

Государственное образовательное учреждение высшего профессионального образования Московский Государственный Университет им. М.В. Ломоносова

Факультет Биоинженерии и Биоинформатики

*На правах рукописи*

**ГОЛОВИН Андрей Викторович**

**КОНФОРМАЦИОННАЯ ДИНАМИКА  
НУКЛЕИНОВЫХ КИСЛОТ ПРИ  
ВЗАИМОДЕЙСТВИИ С ЛИГАНДАМИ**

Специальность 02.00.10 —биоорганическая химия

Диссертация на соискание учёной степени

доктора химических наук

Москва – 2014

# Содержание

<b>1 Введение . . . . .</b>	<b>9</b>
<b>2 Обзор литературы . . . . .</b>	<b>13</b>
2.1 Задачи и проблемы моделирования биополимеров . . . . .	13
2.1.1 Выбор уровня огрубления модели . . . . .	16
2.1.2 Выбор степеней свободы для огрубления . . . . .	17
2.1.3 Выбор процессов или свойств для крупнозернистого моделирования . . . . .	22
2.1.4 Компромисс между уровнями моделирования . . . . .	24
2.1.5 Заключение . . . . .	25
2.2 Роль сахаро-фосфатного остова в формировании структуры НК . . . . .	27
2.2.1 Структурная биоинформатика ДНК . . . . .	30
2.2.2 Структурная биоинформатика РНК . . . . .	39
2.2.3 Электронная структура сахаро-фосфатного остова . . . . .	42
2.2.4 Квантово-механические расчёты больших и малых систем . . . . .	43
2.2.5 Точность различных методов КМ при исследовании N-гликозидной связи . . . . .	46
2.2.6 Остов ДНК, содержащий мышьяк (As-ДНК) . . . . .	50
2.2.7 Альтернативные варианты остова НК . . . . .	52
2.2.8 Моделирование остова НК методами молекулярной механики . . . . .	54
2.3 Моделирование молекулярной динамики G-квадруплексных ДНК . . . . .	60

2.3.1	Что интересного может рассказать нам моделирование G-ДНК?	62
2.3.2	Как можно сравнить моделирование с экспериментом?	64
2.3.3	Ограничения моделирования МД с классическими силовыми полями	67
2.3.4	Использование ионов в моделировании G-ДНК	69
2.3.5	Исследования взаимодействий G-ДНК с лигандами	72
2.3.6	Перспективы применения моделирования к НК	76
<b>3</b>	<b>Результаты и их обсуждение</b>	<b>80</b>
3.1	Структурные аспекты взаимодействия тм-РНК с рибосомой	80
3.1.1	Разработка подхода к моделированию структурной организации комплексов крупных нуклеиновых кислот	82
3.1.2	Моделирование конформаций тмРНК в комплексах с рибосомой	87
3.1.3	Оценка достоверности упрощённого моделирования	93
3.1.4	Создание автоматического инструмента для упрощённого моделирования	95
3.2	Конформационная динамика взаимодействия макролидных антибиотиков с рибосомой	98
3.2.1	Взаимодействие производных тилозина в рибосомном тоннеле	98
3.2.2	Подход к полноатомному моделированию молекулярной динамики рибосомного тоннеля	101
3.2.3	Взаимодействие пептидных производных 5-О-микаминозилтилонолида с рибосомным тоннелем	103
3.3	Влияние топологии петель на геометрию квадруплексов	110
3.3.1	Особенности структуры гуанин богатых ДНК	110

3.3.2 Сравнение геометрических параметров для известных типов квадруплексов . . . . .	117
3.3.3 Классификация квадруплексных структур . . . . .	127
3.4 Исследование динамики структур аптамера к тромбину . . . . .	132
3.4.1 Аптамер 15-ТВА, ЯМР и РСА конформации. . . . .	139
3.4.2 Почему ЯМР структура более стабильна чем РСА модель? . .	142
3.4.3 Комплексы аптамера 15-ТВА с тромбином . . . . .	144
3.4.4 Структурная динамика G-стебля. Углы закрутки спирали от- ражают структурное напряжение . . . . .	149
3.4.5 Структурная аннотация последовательности 15-ТВА . . . . .	153
3.5 Взаимодействие аптамера 15-ТВА с катионами . . . . .	161
3.5.1 Системы и особенности моделирования взаимодействия кати- онов с 15-ТВА . . . . .	164
3.5.2 Молекулярно механические аспекты связывания катионов с 15-ТВА . . . . .	168
3.5.3 Гибридное молекулярно механическое / квантово- механическое исследование взаимодействия катионов с 15-ТВА . . . . .	188
3.5.4 Изотермическая калориметрия . . . . .	195
3.5.5 Заключение об связывание катионов с аптамером 15-ТВА . .	197
3.6 Ключевые факторы, влияющие на конформационную динамику 15-ТВА, и их применение . . . . .	207
3.6.1 Динамика 15-ТВА и функциональная активность . . . . .	207
3.6.2 Исследование структурного влияния дуплексной части ДНК на G-квадруплекс . . . . .	210
3.6.3 Разработка терапевтического аптамера к тромбину . . . . .	211

3.6.4	Разработка аптасенсора к тромбину . . . . .	212
3.7	Температурная зависимость пути сомоорганизации ДНК шпильки d(GCGCAGC) . . . . .	224
3.7.1	Метод обмена репликами для изучения фазового простран- ства биополимеров . . . . .	224
3.7.2	Шпилечные структуры . . . . .	229
3.7.3	Анализ конформаций . . . . .	230
3.7.4	Моделирование самосборки структуры олигонуклеотида 5'-GCGCAGC-3' . . . . .	231
<b>4</b>	<b>Материалы и методы . . . . .</b>	<b>240</b>
4.1	Материалы и методы . . . . .	240
4.1.1	Алгоритм построения крупно-зернистой модели РНК . . . . .	240
4.1.2	Полноатомное моделирование молекулярной динамики рибо- сомного тоннеля . . . . .	246
4.1.3	Влияние топологии петель на геометрию квадруплексов . . . . .	248
4.1.4	Методы моделирования молекулярной динамики систем с ап- тамером 15-ТВА . . . . .	250
4.1.5	Гибридное молекулярно механическое / квантово- механическое моделирование . . . . .	254
4.1.6	Изотремическая калориметрия . . . . .	256
4.1.7	Исследование структуры аптамера 31-TGT методами ЯМР . . . . .	257
4.1.8	Моделирование конъюгатов 15-ТВА и нанотрубок . . . . .	258
<b>5</b>	<b>Выводы . . . . .</b>	<b>261</b>
	<b>Список литературы . . . . .</b>	<b>263</b>

# Список сокращений

AMBER – пакет программ для моделирования молекулярной динамики; этой аббревиатурой также называют семейство силовых полей PARM.

BFGS – Алгоритм Бroyдена-Флетчера-Гольдфарба-Шанно (итерационный метод численной оптимизации).

CBS – Методы полных базисных наборов.

CCSD – Теория связанных кластеров, золотой стандарт квантовой химии.

CHARMM – Семейство силовых полей на основе пакета программ (Chemistry at HARvard Macromolecular Mechanics).

CNT – Углеродная нанотрубка.

COSMO – Модель неявно заданного растворителя в квантовой химии.

COSY – Корреляционная спектроскопия ЯМР.

CPMD – Пакет программ с реализацией плосковолновой версии теории функционала плотности.

DFT – Теория функционала плотности.

FRET – Форстеровский резонансный перенос энергии.

GAFF – Силовое поле для описания малых молекул, совместимое с AMBER.

GAMESS – Программа пакета вычислительной квантовой химии.

GBSA – Метод неявного учёта растворителя.

GPU – Графический процессор (вычислительное устройство).

GROMACS – Пакет программ для моделирования молекулярной динамики.

LINCS – Алгоритм линейного расчёта ограничений.

MLD – мРНК-подобный домен тмРНК.

MSD – Среднеквадратичное смещение.

NMR, ЯМР – Спектроскопия ядерного магнитного резонанса.

NOE – Эффект Оверхаузера.

NOESY – Спектроскопия ядерного эффекта Оверхаузера.

NPT – Изотермическо-изобарический ансамбль.

OMT – 5-О-микаминозилтилонолид.

OPLS – Семейство силовых полей.

PDB – Банк данных 3D-структур белков и нуклеиновых кислот.

PDBID – Идентификатор записи в банке PDB.

PME – Вариант расчёта электростатических взаимодействий на основе суммирования по Эвальду.

PNA – Пептидо-нуклеиновые кислоты.

RDF – Функция кругового распределения.

REMD – Моделирование молекулярной динамики с обменом репликами.

RMSD, СКО – Среднеквадратичное отклонение.

SPC, TIP3P, TIP4P – Модели воды для учёта растворителя явным образом.

TBA – Тромбин-связывающий аптамер

TLD – тРНК-подобный домен тмРНК.

ДНК – Дезоксирибонуклеиновая кислота.

ДСК – Дифференциальная сканирующая калориметрия.

НИВЦ – Научно-исследовательский вычислительный центр.

ППЭ – Поверхность потенциальной энергии

ПТЦ – Пептидилтрансферазный центр

РНК, RNA – Рибонуклеиновая кислота.

РСА – Рентгеноструктурный анализ

тмРНК – Транспортно-матричная РНК



# Глава 1

## Введение

Сложная трёхмерная организация биологических макромолекул определяет их биологическую функцию. Благодаря успехам современной структурной химии, мы теперь знаем, что пространственные структуры как белков, так и нуклеиновых кислот многообразны. В то же время, принципиальное ограничение современных методов состоит в том, что с их помощью можно выявить и охарактеризовать на атомном уровне только самые представленные, т.е. наиболее энергетически выгодные в данных условиях конформации молекул и их комплексов. Кроме того, при описании процессов, в которых участвуют биологические макромолекулы, в большинстве случаев оперируют только структурой исходных соединений и продуктов их превращений, в то время, как и механизм, и динамика происходящих взаимодействий зачастую остаются “за кадром”.

Молекулярное узнавание биологических макромолекул достигается за счёт нековалентных взаимодействий, сопряжённых с конформационными переходами, которые либо предшествуют образованию комплекса, либо происходят пост-фактум. Такие процессы описываются либо механизмами конформационного отбора, либо гипотезой индуцированного соответствия.

Выявление путей, по которым происходит тот или иной процесс, – нетривиальная задача даже для анализа взаимодействия биомакромолекул с низкомолекулярными лигандами. В последнее время стало понятно, что в большинстве исследованных систем реализуется путь конформационного отбора. По этой причине первой задачей каждого нового исследования становится поиск и описание набора конформационных состояний, которые может принимать макромолекула. Следует заметить, что путь конформационного отбора зачастую предполагает выбор таких малопредставленных состояний, которые чрезвычайно сложно зарегистрировать экспериментально. Большинство современных исследований, направленных на описание конформационного профиля биомакромолекул, относятся к пептидам и белкам. В то же время, структура и функция нуклеиновых кислот (НК) все больше привлекает внимание исследователей, в связи с развитием нового направления терапии – применения олигонуклеотидов в качестве лекарственных препаратов для лечения заболеваний человека; кроме того, сами нуклеиновые кислоты рассматриваются теперь как возможные мишени для разработки лекарственных средств. Однако экспериментальные подходы определения конформационного профиля НК, так же, как и белков, имеют ограничения, связанные с представленностью конформаций, ключевых для функционирования НК. Данную проблему помогают решить современные методы молекулярного моделирования, которые позволяют получить информацию о природе и кинетических переходах между конформациями исследуемой макромолекулы.

Эта работа посвящена разработке и применению методов молекулярного моделирования к различным по размерам и сложности системам, в которых ключевую роль играют нуклеиновые кислоты. Выбор систем для исследования определялся современными возможностями вычислительной техники; по мере

перехода к системе меньшего размера, увеличивали детализацию моделирования. Так, крупные системы рассматривали с помощью упрощённых моделей, в то время, как структуры сравнительно коротких олигонуклеотидов анализировали гораздо более детально, благодаря возможности применения методов молекулярной и квантовой механики.

В ходе работы разработан новый подход к компьютерному моделированию структуры больших супрамолекулярных комплексов, основанный на упрощённом представлении нуклеотидов и аминокислот. На основе этого подхода впервые предложены структуры комплексов тмРНК с рибосомой на разных этапах элонгации трансляции. Высокое качество моделирования подтверждено результатами по уточнению структуры тмРНК, которые получены в экспериментах *in vitro*. Наличие пространственных ограничений стало ключевым фактором, который позволил использовать упрощённые модели при сохранении точности моделирования. На основании данных по моделированию молекулярной динамики комплексов производных тилозина с рибосомным тоннелем, впервые предсказана ингибирующая активность этих соединений по отношению к рибосоме *Escherichia coli*.

Используя нетривиальный подход (полноатомное локальное окружение антибиотика в рибосоме), нам удалось показать, что эффективность ингибирования элонгации трансляции производными тилозина обусловлена образованием сети водородных связей, которые позиционируют альдегидную группу лактонного кольца антибиотика для формирования ковалентной связи с реакционноспособными группами рибосомного белка и 23S рРНК в рибосомном тоннеле.

В работе предложена классификация малых G-квадруплексных ДНК по топологии петель на основе геометрического описания структуры самого квадруплексного тяжа. С помощью этой классификации впервые сделаны предпо-

ложения о влиянии топологии молекулы на структурную стабильность и динамическое поведение G-квадруплексных структур.

Применение метода моделирования молекулярной динамики к минимальному 15-звенному квадруплексу ДНК впервые позволило продемонстрировать, что латеральные петли могут оказывать на него как стабилизирующее, так и дестабилизирующее влияние, в зависимости от их длины и последовательности.

При исследовании влияния катиона на стабильность квадруплекса было показано, что его диффузия в центральную полость квадруплекса может быть сложным разнонаправленным процессом, в ходе которого возможны значимые структурные перестройки в G-квадруплексе. Эффективное хелатирование катионов в центре минимального 15-звенного квадруплекса определяется действием латеральных петель: подвижность петель уменьшает вероятность диссоциации комплекса аптамер-катион.

Эти результаты имеют не только фундаментальную научную значимость, но и были использованы для дизайна потенциального лекарственного препарата антитромботического действия на основе ДНК-аптамера к тромбину, который получил название RA-36. В экспериментах *in vitro* и *in vivo* соединение RA-36 проявляет высокую антикоагуляционную активность и низкую токсичность. В настоящее время препарат проходит доклинические испытания для регистрации в качестве лекарственного средства. Знания динамики структуры аптамера к тромбину использованы в дизайне сенсоров для определения концентрации тромбина. Сенсор представляет собой электропроводник на основе углеродных нанотрубок. Впервые предложен механизм работы такого сенсора. На текущий момент проводится оптимизация опытных образцов сенсоров к тромбину.

## Глава 2

# Обзор литературы

### 2.1 Задачи и проблемы моделирования биополимеров

С тех пор, как экспериментальные исследования веществ и химических процессов стали массовым явлением, стала актуальна разработка теоретических моделей, описывающих эти процессы. Создание математического аппарата дифференциальных уравнений несколько веков назад позволило кратко формулировать модели и аналитические решения для самых простых систем. Таких, как: идеальный газ для веществ в газовой фазе, ячейка с гармоническими осцилляторами, модель вещества в твёрдой фазе или модели Изинга для взаимодействующих спинов. С появлением компьютеров в прошлом веке, использование вычислительных подходов позволило строить теоретические модели значительно большей сложности, тем самым расширяя область их применения. Вычислительная мощность даёт возможность повысить точность моделей за счёт более точного представления базовых физических процессов. Такое развитие событий привело к высокой точности предсказаний, которые широко ис-

пользуются в технических наук, таких, как самолётостроение или инженерное проектирование.

Хотя теоретическая и вычислительная химия быстро развивается в течение последних нескольких десятилетий, точность и применимость этих подходов по-прежнему весьма ограничены. Это связано с рядом факторов: 1) степень свободы разных объектов, которыми могут быть одновременно электроны, ядра, атомы и молекулы; 2) взаимодействие между этими частицами регулируется законами квантовой механики, а именно, уравнениями Дирака или Шредингера или в особых случаях - классическими уравнениями движения; 3) при температуре, большей 0 К, поведение частиц регулируется уравнениями статистической механики (Бозе-Эйнштейна, Ферми-Дирака); 4) электростатические взаимодействия имеют значимый эффект на значительных расстояниях; 5) временная шкала различных химических процессов может легко охватывать 15 порядков, и 6) энергия или изменения свободной энергии химических процессов могут быть очень малы, по сравнению с полной энергией взаимодействующих частиц. Эти факторы значительно усложняют разработку моделей в химии, что приводит к ощущению, что этот процесс - скорее искусство, чем наука.

Каждая модель включает в себя выбор основных степеней свободы, взаимодействий, которые регулируют движение вдоль этих степеней свободы, метода для генерации конфигураций по выбранным степеням свободы. Непростой задачей является поиск баланса между точностью и требуемой компьютерной мощностью для описания интересующего процесса на основе физически осмысленных механизмов.

В химии можно выбрать различные уровни моделирования (Таблица 2.1), т.е., с участием различных типов степеней свободы. Самое точное описание включает в себя свойства ядер, атомов и электронов, и этот уровень часто назы-

Уровень	Частицы	Размер, нм	Количество степеней свободы	Уменьшение количества взаимодействий	Уменьшения вычислительных затрат
1	Нуклоны+электроны	$10^{-6}$	$N_n^{\geq 3}$		
2	Ядра+электроны	$10^{-6}$ - $10^{-5}$	$N_e^{\geq 3}$	10–100	$>10^3$
3	Атомы	0.03-0.3	$N_a^{\geq 3}$	10–100	$>10^3$
4	Супраатомные частицы	0.5-10	$N_b^{\geq 1-2}$	2-5	2-25
5	Супрамолекулярные частицы	0.5-10	$N_b^{\geq 1-2}$	2-10	2-100

Таблица 2.1: Характерные размеры частиц на разных уровнях теории моделирования, а также шкала вычислительных затрат, в зависимости от числа нуклонов ( $N_n$ ), электронов ( $N_e$ ), атомов ( $N_a$ ) или частиц ( $N_b$ ). Супраатомными частицами здесь и далее называются частицы, состоящие из нескольких атомов.

вают квантовой химией. Если исследователь не заинтересован в разрыве/формировании химических связей или в описании возбуждённых состояний молекулы, то можно устранить электронные степени свободы из модели и рассматривать только атомы. Таким образом, на втором этапе точная модель переходит в упрощённую с удалением электронных степеней свободы.

Процедура так называемого “огрубления” (т.е., объединения нескольких степеней свободы в одну) может быть применена между любыми двумя уровнями моделирования и, следовательно, любую модель в химии можно рассматривать как упрощённую модель, по отношению к удалённым (подвергшимся “огрублению”) степеням свободы. Примером такого огрубления является так называемое “крупнозернистое моделирование”, в котором одна частица представляет более, чем один тяжёлый атом (не водород). Если эти атомы принадлежат одной молекуле, то такая модель является супраатомной, молекулярной

или крупнозернистой моделью. Если частицы включают в себя более одной молекулы, то такая модель является супрамолекулярной.

Целью этой главы является исследование способов огрубления при описании химических систем, а также поиск ошибок, которые могут появиться при анализе таких моделей. Так как упрощение, вплоть до 3-го уровня, сильно зависит от природы рассматриваемого вещества, то обсуждение этого типа огрубления для НК можно найти в следующих главах обзора. Здесь мы ограничимся обсуждением огрубления уровней моделирования от атомных и выше (уровни 3-5 в таблице 2.1). Это означает, что все объекты моделирования рассматриваются в терминах классической (ньютоновской) механики. Из-за большого разнообразия публикаций о разных объектах и подходах, в обзоре будет приведена общая классификация подходов, и будут упомянуты варианты их использования. Также будут описаны результаты тестирования этих подходов. Такая классификация прояснит выбор модели для выбранной задачи в настоящей работе.

### **2.1.1 Выбор уровня огрубления модели**

В зависимости от целей исследования, можно выделить множество различных уровней моделирования. Начиная с уровня кварков, систему можно последовательно огрубить системы до уровня галактик [1]. В химии наиболее часто используются следующие уровни упрощения (Таблица 2.1):

1. Нуклоны и электроны.
2. Ядра и электроны.
3. Атомы.
4. Супраатомы или объединённые атомы.



## 5. Супрамолекулы

Также можно выделить взаимодействия и законы, которые направляют движение частиц на разных уровнях огрубления (1-5):

1. Сильные взаимодействия, законы Паули и Кулона.
2. Законы Паули и Кулона.
3. Кулоновские взаимодействия, взаимодействия Ван-дер-Ваальса, ковалентные связи.
4. Кулоновские взаимодействия, взаимодействия Ван-дер-Ваальса.
5. Кулоновские взаимодействия, взаимодействия Ван-дер-Ваальса.

Взаимодействия уровней 1 и 2 описываются законами квантовой механики, а взаимодействия уровней 3-5 можно описать законами классической статистической механики. Вычислительная сложность определяется количеством степеней свободы, количеством частиц или участков взаимодействия вместе с уравнениями движения.

Огрубление, переводящее систему с уровня 2 на уровень 3, может быть сделано разнообразными способами, но при этом возникает значимое количество проблем, в то время, как огрубления выше 3-го уровня проводить проще. Это вызвано малой совместимостью принципов классической и квантовой механики.

### 2.1.2 Выбор степеней свободы для огрубления

Любое огрубление подразумевает устранение степеней свободы. Это неизбежно ведёт к уменьшению применимости модели. Например, при перехо-

де от системы, состоящей из ядер с электронами (уровень 2), к атомной системе (уровень 3) становится неприемлемым процесс релаксации электронно-возбужденных состояний молекул. Как правило, огрубление приводит к потере точности модели, хотя для определённых типов моделей и её свойств это не так. Например, свойства жидкой воды при температуре окружающей среды выше 0 К более точно описывается моделью SPC [2] (уровень 3), чем неэмпирической моделью на уровне 2, построенной с помощью теории функционала плотности (DFT) [3]. Это вызвано ограниченной точностью функционалов, используемых при расчёте DFT. Выбор степеней свободы для огрубления зависит от свойств и фазового состояния вещества. Можно составить список условий, которые должны быть выполнены при огрублении для сохранения физически корректной системы. Этот список расположен в порядке увеличения влияния огрубления на точность и скорость вычислений. Степени свободы, которые можно огрубить:

1. не должны быть существенными для процесса или не должны определять свойства исследуемого объекта.
2. требуют большого объёма вычислений, и при их огрублении польза от ускорения вычислений будет компенсировать потерю точности.
3. должны удовлетворять условию, что взаимодействия, регулирующие эти степени свободы, в значительной степени не пересекаются со взаимодействиями, регулируемыми другими степенями свободы системы. Это означает, что частотные компоненты движения по степеням свободы, которые огрубляют, не должны пересекаться с другими частотами движения.
4. должны быть заменены простым, эффективным представлением этого взаимодействия для регулирования поведения степеней свободы, которые остаются в системе.

Желательно, чтобы все перечисленные пункты были учтены при описании системы. Обсудим пример огрубления для перехода с уровня 3 на уровень 4: использование так называемого объединённого атома [4] Рассмотрим алифатические группы  $-CH$ ,  $-CH_2$ - и  $-CH_3$ -группы как объединённые атомы. При объединении атомов водорода с углеродом число нековалентных взаимодействий существенно снижается (в случае липидов - почти в 10 раз, условие 2). Происходит потеря дипольных взаимодействий в  $CH$ -группах и Ван-дер-Ваальсовых взаимодействий между атомами водорода, которыми можно пренебречь (условие 1). Внутренние движения внутри фрагмента  $-CH_n$  в значительной степени не связаны с движением других атомов, и энергия торсионного угла с участием атомов H этой группы может быть включена в соответствующие параметры для торсионного угла вокруг связи C-C (условия 3-4). Итак, все четыре условия для соответствующего огрубления, так или иначе, встречались в этом случае. Если необходимо знать положения огрублённых атомов, например, при расчёте таких величин, как ядерный эффект Оверхаузера (NOE) или параметр порядка ( $S_{CH}$ ), то положение атомов водорода можно легко восстановить, зная позиции атомов углерода и его окружение [5].

Другой пример огрубления – использование геометрических ограничений для малых молекул, которые не имеют внутримолекулярных степеней свободы вращения вокруг связей. Это, например, такие растворители, как вода, метанол или хлороформ или ограничения на длину связи в биополимерах [4]. Последнее часто используется при моделировании молекулярно-биологических объектов, потому что они удовлетворяют условиям с 1 по 4 [4] и позволяют при использовании алгоритмов SHAKE[6], LINCS [7] или других аналогичных методов получить прирост вычислительной производительности в четыре раза.

Примером огрубления, которое не удовлетворяет условиям 3 и 4, является использование неявной модели растворителя. Это попытка имитировать эффект растворителя с помощью функции, которая зависит только от координат частиц растворённого вещества. Если растворителем является вода, то это приводит к серьёзным искажениям поверхности потенциальной энергии растворённого вещества. Хотя движения большой молекулы могут происходить в масштабах от фемтосекунд до миллисекунд, а времена релаксации молекулы воды - порядка пикосекунд, однако их движения в пико- и наносекундных временных интервалах не являются независимыми. Таким образом, условие 3 не выполняется для этих процессов.

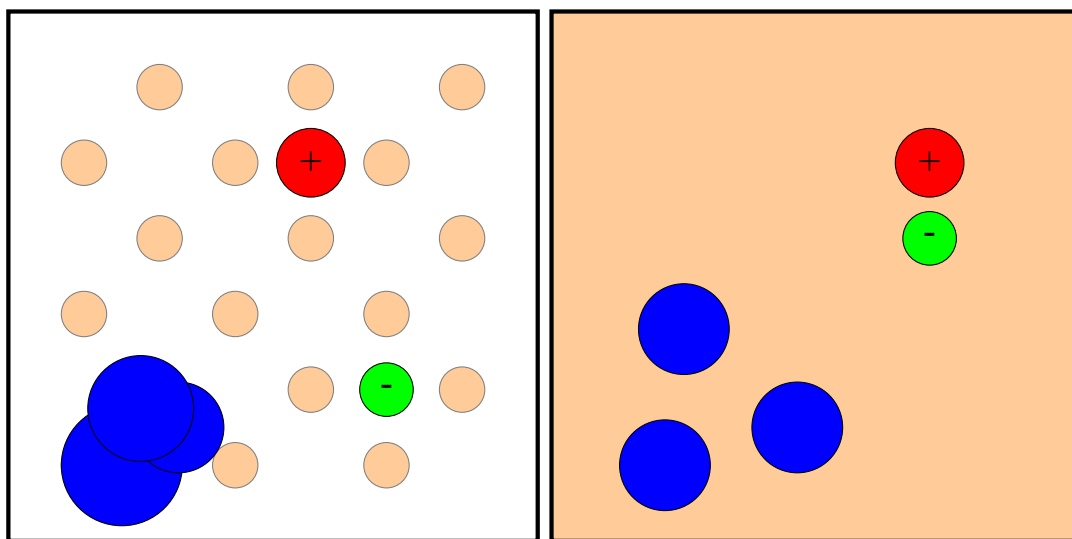


Рисунок 2.1: Иллюстрация гидрофобного эффекта и негативного влияния неявного способа расчёта растворителя (огрубление растворителя). Слева: явно заданный растворитель, справа: неявно заданный растворитель. Растворитель показан оранжевым цветом, гидрофобные частицы - синим, и катионы - красным, анионы - зелёным.

В явно описанном растворителе (оранжевые частицы в левой части рисунка 2.1) неполярные частицы агрегируют, а электростатическое взаимодействие между ионами уменьшается, что приводит к растворению ионной пары. Так называемые гидрофобные, или неполярные частицы характерны тем, что их взаимодействие с водой слабее, чем взаимодействие молекул воды между собой.

Это приводит к исключению воды из контакта с гидрофобными частицами и их последующей агрегации. Ионы с противоположными зарядами лучше взаимодействуют с водой, чем друг с другом, что приводит к растворению ионных пар. Очевидно, что “гидрофобный эффект” растворения ионных пар не может быть правильно воспроизведён, на основании координат частиц и описании свойств растворителя (правая часть рисунка 2.1). Эффективное взаимодействие между атомами растворённого вещества и растворителем, а именно, их энтропия является сложной функцией распределения молекул растворителя вокруг молекулы. Таким образом, условие 4 трудно достижимо при использовании неявного растворителя [8].

Огрубление с уровня 3 до уровня 4 для биомолекул является проблемным из-за неоднородности биомолекул. Инвариантность загрубляемой группы лежит в основе подхода к крупнозернистому моделированию, что в значительной степени справедливо для однородных полимеров и неверно для биополимеров, которые состоят из различных по природе остатков и участвуют в разных типах взаимодействий. В процессе огрубления геометрия и баланс между различными взаимодействиями должны быть сохранены, во избежание потери характеристических особенностей этих молекул [8]. Кроме того, энтропия играет значительную роль в молекулярно-биологических процессах. Это означает, что потеря энтропийной составляющей в процессе огрубления должна быть компенсирована. С другой стороны, сокращение вычислительных затрат между уровнями 3 и 4 не очень значительно, по сравнению с переходом между другими уровнями (Таблица 2.1).

Эти соображения приводят к выводу, что огрубления от уровня 3 для уровня 4 неэффективны для биополимеров (белки, ДНК, РНК и сахара). Ограниченное снижение количества мест взаимодействий приводит к потере существенных

характеристик таких молекул, как с точки зрения внутримолекулярных взаимодействий, так и с точки зрения взаимодействий с растворителем и энтропии. Только липиды, которые имеют относительно длинные однородные алифатических хвосты, могут быть в состоянии сохранить основные характеристики амфифильных молекул при огрублении от уровня 3 до уровня 4. Из-за обилия липидов в мембранах снижение вычислительных затрат может восполнить потерю точности. Поскольку использование явно заданного растворителя необходимо для правильного описания гидрофобного эффекта, а расчёт взаимодействий в растворителе потребляет львиную долю вычислительной мощности, то огрубление растворителя в супрамолекулярные частицы кажется перспективным подходом. Для воды огрубление от уровня 3 до уровня 5 должно сохранить термодинамические и диэлектрические свойства, а также возможность образования водородных связей практически полностью [6]. Это не получится, если представить воду как частицу с потенциалом Леннарда-Джонса без заряда [9; 10]. Огрубление степеней свободы растворителя в моделировании биополимеров имеет хорошие шансы на то, чтобы удовлетворить условия с 1 по 4-ое, но зависит от того, как моделируются взаимодействия конкретных загрубленных частиц.

### **2.1.3 Выбор процессов или свойств для крупнозернистого моделирования**

Сам факт загрубления предполагает потерю информации и, возможно, излишнее упрощение взаимодействий между частицами. В принципе, невозможно воспроизвести все измеряемые свойства вещества при любом уровне загрубления системы. В таком случае возникает вопрос: какие свойства представляют

интерес и каким образом они должны быть представлены, а от каких свойств можно отказаться? Кроме этого, огрубление также может ограничить применимость модели к конкретному фазовому состоянию или некоторой области фазовой диаграммы вещества [11]. Например, модели для жидкой воды, которые обычно используются на уровне 3 моделирования (огрубления), хорошо воспроизводят термодинамические, диэлектрические, структурные свойства и водородные связи молекул воды в жидкой фазе при нормальных условиях. Но эти же модели неприменимы для газовой фазы или жидкой фазы при высокой температуре и давлении. Следующие уровни огрубления тем более ограничены в применении вокруг той точки фазового пространства, для которой эта модель была построена.

В системах с биополимерами исследователей чаще всего интересуют следующие свойства молекул:

1. молекулярная структура растворённого вещества или структура растворителя;
2. термодинамические свойства, такие, как теплота парообразования, плотность, избыток свободной энергии или поверхностное натяжение. Также можно определить второстепенные термодинамические величины, которые характеризуют ответ системы на изменение фазовых свойств, например, теплоёмкости, изотермической сжимаемости или коэффициента теплового расширения;
3. диэлектрические свойства, в частности, статическая диэлектрическая проницаемость растворителя, которая явно изменяет эффективность электростатических взаимодействий;

4. динамические свойства, такие, как диффузия, вязкость и другие молекулярно-динамические особенности.

Модели для растворителя на уровне 3 должны, по крайней мере, воспроизвести структуру жидкости или смеси жидкостей, а также их термодинамические свойства (теплоту парообразования, плотность и диэлектрическую проницаемость). Модели на уровне 4, параметризованные по данным эксперимента или мелкозернистого моделирования, не могут реалистично имитировать структуру жидкости, например, с точки зрения геометрии взаимодействий, таких, как водородные связи. Для надмолекулярных моделей растворителей (уровень 5) сравнение теплоты парообразования и выигрыш свободной энергии с экспериментальными данными не является однозначным [12]. Таким образом, растворители на уровне 5 должны, по крайней мере, воспроизвести термодинамические свойства поверхностного натяжения, плотности и значения диэлектрической проницаемости.

#### **2.1.4 Компромисс между уровнями моделирования**

Как упоминалось ранее, процесс огрубления может уменьшить полезность модели с разными последствиями:

1. Уменьшение диапазона состояний фазового пространства, где эта модель может быть применена.
2. Снижение возможности переноса параметров модели между сходными объектами.
3. Снижение точности определения различных свойств системы.



4. Возможно изменение физического смысла для конкретного свойства или процесса, что приводит к некорректному процессу в крупнозернистой модели.
5. Сокращение энтропийного вклада или потенциальной энергии в общую энергию системы может привести к некорректному балансу между этими величинами.

Важно, чтобы наименьшее количество вышеописанных пунктов было реализовано в модели. Суммарная потеря полезности модели за счёт перечисленных последствий должна быть компенсирована большим увеличением вычислительной эффективности крупнозернистой модели.

### **2.1.5 Заключение**

В связи с быстрым ростом и доступностью вычислительных мощностей, стало популярным использование вычислительного моделирования для описания поведения сложных биомолекулярных систем. Тем не менее, разработка модели с требуемым разрешением, будь то субатомная, атомная, молекулярная или надмолекулярная модель, – это нетривиальная задача. Модель должна отражать физические механизмы, которые определяют поведение биополимеров и процессы, в которых они участвуют.

Для огрублённого моделирования необходимо выбирать набор степеней свободы учитывать особенности взаимодействия. После калибровки параметров взаимодействия можно имитировать эволюцию системы. В зависимости от качества выбора и физической корректности приближений, моделирование может привести, в лучшем случае, к пониманию деталей интересующего механизма, а, в худшем случае, – к набору состояний, не соответствующих действи-

тельности. Во избежание ошибок, а таких примеров достаточно в литературе, следует учитывать четыре условия, которые должны быть удовлетворены в процессе огрубления. В частности, условия 3 и 4, как правило, нарушаются при разработке крупнозернистой модели при недостаточном внимании к детальному анализу влияния огрублённых степеней свободы на оставшиеся. Если эта связь между степенями свободы значима, то выбор соответствующей функциональной формы и параметров для описания этой связи на грубом уровне может привести громоздкому описанию этой связи. При объединении моделей различного разрешения для достижения значимых результатов надо добиться совместимости их термодинамических свойств. Будет ли цель построения таких моделей достигнута, зависит от особенностей интересующей системы. К сожалению, часто бывает трудно разработать надёжную вычислительную модель, которая имеет биологический смысл, а не является просто набором красивых изображений.

## 2.2 Роль сахаро-фосфатного остова в формировании структуры НК

В большинстве известных живых систем генетическая информация заложена в последовательности ДНК, длина которой превышает минимальную кодирующую часть. Эти последовательности ДНК образуют общеизвестную структуру двойной спирали, поддерживаемой каноническими взаимодействиями пар нуклеотидов АТ и GC. Другие молекулы, которые участвуют в экспрессии генетической информации, взаимодействуют с определёнными участками ДНК с особой последовательностью и структурой [13; 14]. Наряду с канонической двойной спиралью, ДНК может формировать альтернативные структурные формы, например квадруплексные, топологическое устройство которых может быть очень разнообразно. Естественно, что в балансе сил, влияющих на структуру таких форм, могут появляться новые факторы, такие, как хелатирование катионов металлов азотистыми основаниями [15].

Молекулы рибонуклеиновой кислоты (РНК) синтезируются на основе выбранных участков геномной ДНК в ходе транскрипции. Долго считалось, что РНК имеет исключительно вспомогательную роль при передаче информации от ДНК до белкового синтеза. Сегодня становится понятно, что РНК - это ключевой игрок в экспрессии генов и в её регуляции. Молекулы РНК образуют, как правило, одноцепочечные структуры, структурное и функциональное разнообразие которых очень велико. Молекулы РНК образуют взаимодействия внутри одной цепи, образуя коротки двуспиральные участки с характерной А-формой дуплекса. Считается, что формирование этих участков является основой для образования полной структуры РНК [16]. После образования двуспиральных участков конечная форма РНК формируется за счёт большого количества раз-

нообразных неканонических взаимодействий. Большие РНК образуют идеально организованные иерархические структуры, и в ходе функционирования они могут претерпевать структурные перестройки. Эти свойства, необходимые для функционирования, оттачивались в ходе миллиардов лет эволюции до самых “незначительных” деталей [17—22].

Химическое разнообразие в структуре нуклеиновых кислот (НК) представлено только азотистыми основаниями, в то время как сахаро-фосфатный остов - это монотонная составляющая НК, представленная насыщенными связями и, как результат, имеющая значительную конформационную свободу, обеспечиваемую совместным вращением торсионных углов (Рисунок 2.2). Это наблюдение может привести к ложному ощущению, сахаро-фосфатный остов играет исключительно пассивную роль в процессе формирования структуры. Мнение, ориентированное на роль азотистых оснований, предполагает, что взаимодействие оснований друг с другом является определяющим при формировании третичной структуры [13; 14; 23—27]. Являясь ароматическими, азотистые основания остаются плоскими во всех возможных случаях, однако пары оснований, динуклеотиды их комбинации проявляют значимое разнообразие конформаций. Описание неперiodической двойной спирали математически сложно достижимо и нуждается в наборе параметров для описания смещения пар нуклеотидов относительно идеального случая [28; 29]. Естественно предположить, что свойства ДНК, определяемые последовательностью, возникают исключительно благодаря стекинг-взаимодействиям между соседствующими парами нуклеотидов. И действительно, большое количество данных по исследованию конформации В-формы ДНК указывает на то, что сахаро-фосфатный остов - это пассивный элемент структуры, который ограничивает конформационное пространство для оснований [26].

Так же, как и для ДНК, очевидно, что свойства РНК при образовании структуры диктуются не только каноническими взаимодействиями в двуспиральных участках, но и множеством неканонических взаимодействий между азотистыми основаниями во всех частях молекулы [18—21].

На сегодняшний момент можно найти классификацию взаимодействий пар нуклеотидов, которая явным образом включает в себя описание контактов с 2'-ОН группой рибозы [18]. Этот подход может быть расширен до специфического описания взаимодействий азотистых оснований с фосфатной группой остова с помощью водородных связей [21]. Учёт этих взаимодействий приводит к появлению нескольких сотен разных парных взаимодействий, которые могут быть разделены на 12 основных семейств [30]. Предполагается, что форма пар нуклеотидов внутри двуспиральных элементов и основные “третичные” парные контакты сохраняют консервативность последовательности и структуры в ходе эволюции, в то время как мутации, меняющие форму НК, вероятнее всего, приводят к изменению функции [17—20].

Особенность РНК состоит в возможности образования стабильных взаимодействий остова с азотистым основанием [18; 21]. Таким образом, существует определённое равновесие между влиянием азотистых оснований и сахарофосфатного остова, на которые в свою очередь, могут влиять внешние факторы. Этими факторами являются: взаимодействие с белками и другими лигандами, гидратация, хелатирование катионов и т.д. Очевидно, что наше представление о том, что азотистые основания играют решающую роль в образовании структуры НК и в её динамике, является упрощённым пониманием баланса энергетических вкладов в структуру нуклеиновых кислот. К сожалению, роль вклада остова НК в этот баланс на сегодня изучена слабо. Тем не менее, уже сейчас можно считать, что влияние остова на структуру НК значимо[31]..

Описание роли остова затруднено как экспериментально, так и теоретически. В то время, как методом РСА можно достаточно точно (с разрешением до  $2.5\text{\AA}$ ) определить положение азотистых оснований и атомов фосфора, установить конформацию сахарных остатков достаточно сложно. Это, в основном, определяется тем, что атом фосфора гораздо лучше регистрируется. Определение конформации остова также затруднено в экспериментах по ЯМР-спектроскопии [32]. Описание сахаро-фосфатного остова вычислительными методами квантовой химии является сложной задачей. Описание остова в терминах молекулярной механики имеет ограниченную точность по причине того, что точечные заряды не зависят от конформации, и расчёт только парных взаимодействий в силовых полях не полностью описывает энергии гибкого и заряженного остова НК [31].

## 2.2.1 Структурная биоинформатика ДНК

Каждый нуклеотидный остаток ДНК состоит из трёх ковалентно связанных частей: остатки фосфорной кислоты и дезоксирибозы образуют сахаро-фосфатный остов, а азотистое основание присоединено к остатку сахара через N-гликозидную связь. В случае ДНК азотистые основания представлены четырьмя плоскими ароматическими основаниями: G, A, T, C. В двуспиральной молекуле ДНК два тяжа направлены антипараллельно по отношению друг к другу. Самыми распространёнными являются три формы двуспиральной ДНК: A, B и Z. Два тяжа образуют две разные по размеру бороздки, которые называют большой и малой. Большинство ДНК в геноме находится в B-форме, которая часто называется канонической формой ДНК. В этой структуре спираль делает оборот каждые 3.4 нм, и расстояние между прилежащими азотистыми основа-

ниями составляет 0.34 нм. Это расстояние определяется Ван-дер-Ваальсовыми взаимодействиями [33]. Таким образом, примерно 10 нуклеотидов образуют один виток спирали. В растворе, содержащем высокую концентрацию спирта или высокую ионную силу, В-форма ДНК переходит в А-форму. Связывание белка также может изменить форму ДНК. А-форма ДНК делает виток каждые 2.8-3.1 нм, и на один виток приходится примерно 11 нуклеотидов. По сравнению с В-формой ДНК, А-форма имеет более короткую и компактную структуру, в которой основания расположены под углом к центральной оси спирали. Расстояние между основаниями остаётся неизменным и составляет 0.34 нм.

В то время, как А- и В-формы ДНК образуют правозакрученную спираль, Z-ДНК имеет левозакрученную спираль. Один её виток составляет 4.6 нм и содержит 12 нуклеотидных остатков. В основном, ДНК, богатые GC-парами при высокой концентрации этанола или высокой ионной силе склонны образовывать такую форму. Биологическая роль Z-ДНК ещё недостаточно изучена. Наряду с упомянутыми формами, ДНК может образовывать мультитяжевые структуры, шпильки и прочие элементы вторичной структуры [34].

Конформация сахаро-фосфатного остова описывается шестью торсионными углами:  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  и  $\chi$  (Рисунок 2.2).

Так как все связи с остатке сахара являются насыщенными, то кольцо сахара не является плоской структурой. Конформация кольца определяется оптимальными значениями торсионных углов, и, естественно, зависит от того, какие заместители находятся рядом с каждой рассматриваемой связью. В разных внешних условиях эти заместители могут располагаться как по одну сторону от кольца, так и по разные. Ожидается, что все пять торсионных углов, которые описывают состояние сахара, не являются полностью независимыми друг от друга. Таким образом, можно представить описание конформации сахара с

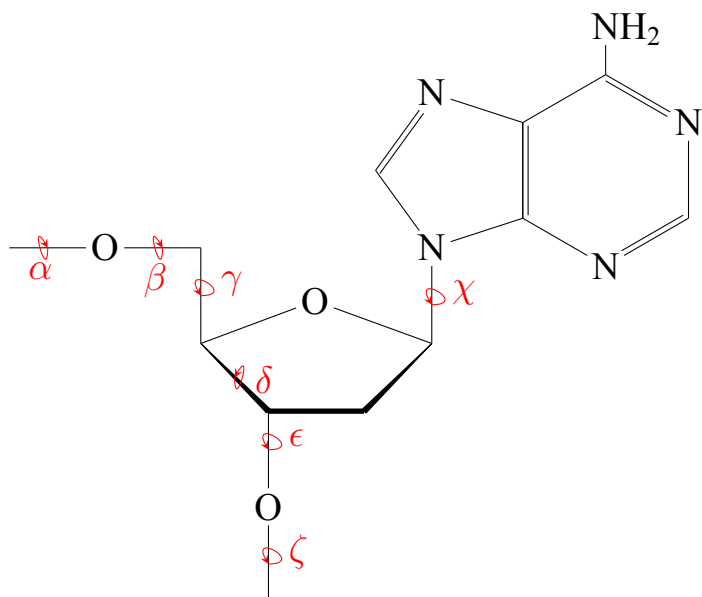


Рисунок 2.2: Торсионные углы ДНК.

помощью всего двух переменных: сдвиг фаз угла псевдотворения максимальное отклонение от плоскости кольца [35]. Угол псевдотворения  $P$  определяет точное положение структуры на пути псевдотворения от выбранной стандартной конформации, для которой, по всеобщему соглашению  $P=0^\circ$ . Оптимальная конформация сахара для нуклеиновых кислот определяется максимально возможным положительным значением торсионного угла  $C(1')-C(2')-C(3')-C(4')$ . В таком случае степень отклонения фуранозного кольца от плоского состояния описывается максимальным отклонением изгиба кольца от плоскости кольца ( $t_m$ ), которое для дуплексов ДНК и РНК варьирует в диапазоне от  $30^\circ$  до  $50^\circ$ .

При описании конформации сахарного кольца часто применяют термины “эндо” и “экзо”. “Эндо” означает, что рассматриваемые атомы находятся по одну сторону кольца от  $C5'$ , и наоборот, “экзо” означает, что интересующие атомы находятся по другую сторону от плоскости кольца относительно  $C5'$ . В А-форме РНК и А-форме ДНК каноническим состоянием сахара является  $C3'$ -эндо, которое соответствует углу  $P=18^\circ$ . Состояние сахара в  $C2'$ -эндо конформации типично для В-формы ДНК и характеризуется углом  $P\sim 162^\circ$ .



Идеальная структура (структура сравнения) с углом  $P=0^\circ$ , имеет смешанную C3'-эндо/C2'-экзо конформацию. Область значений углов от  $90^\circ$  до  $270^\circ$ , которая включает в себя C2'-эндо конформации, известна как южная (S) область значений, а для значений от  $270^\circ$  до  $90^\circ$ , которая включает C3'-экзо конформации, принято называть северной областью (N).

Так как конформация сахара также зависит от значения угла  $\delta$ , то существует математическая зависимость (при идеальной  $sp^3$  гибридизации атомов углерода) между  $P$ ,  $t_m$  и  $\delta$  [36]:

$$\delta = 1.09 \times t_m \times \cos(P + 144^\circ) + 120^\circ [35].$$

N-гликозидная связь, которая связывает азотистое основание и остаток сахара, описывается углом  $\chi$ . Итак, ясно что, для полного описания состояния сахара необходимы все семь торсионных углов каждого нуклеотида.

Локальное состояние структуры ДНК является результатом баланса сил между оптимальным положением азотистого основания и оптимальным состоянием сахара-фосфатного остова. Основным источником информации о его конформации - это данные РСА. Кроме исследований индивидуальных молекул, был также проведён биоинформатический анализ всех известных на тот момент данных о структуре ДНК. В результате были обнаружены общие черты в конформации остова для разных молекул. Очевидно, если качество РСА слишком низко, то даже использование статистических методов не позволит избежать ошибочных значений. Поэтому надо учитывать в любом подобном анализе качество входных данных, при этом стараясь избежать потери редкой информации, которая недостаточно представлена в доступной выборке. Другой особенностью подобных данных является влияние кристаллической упаковки на структуру ДНК. Результаты такого исследования приведены на рисунке 2.3.

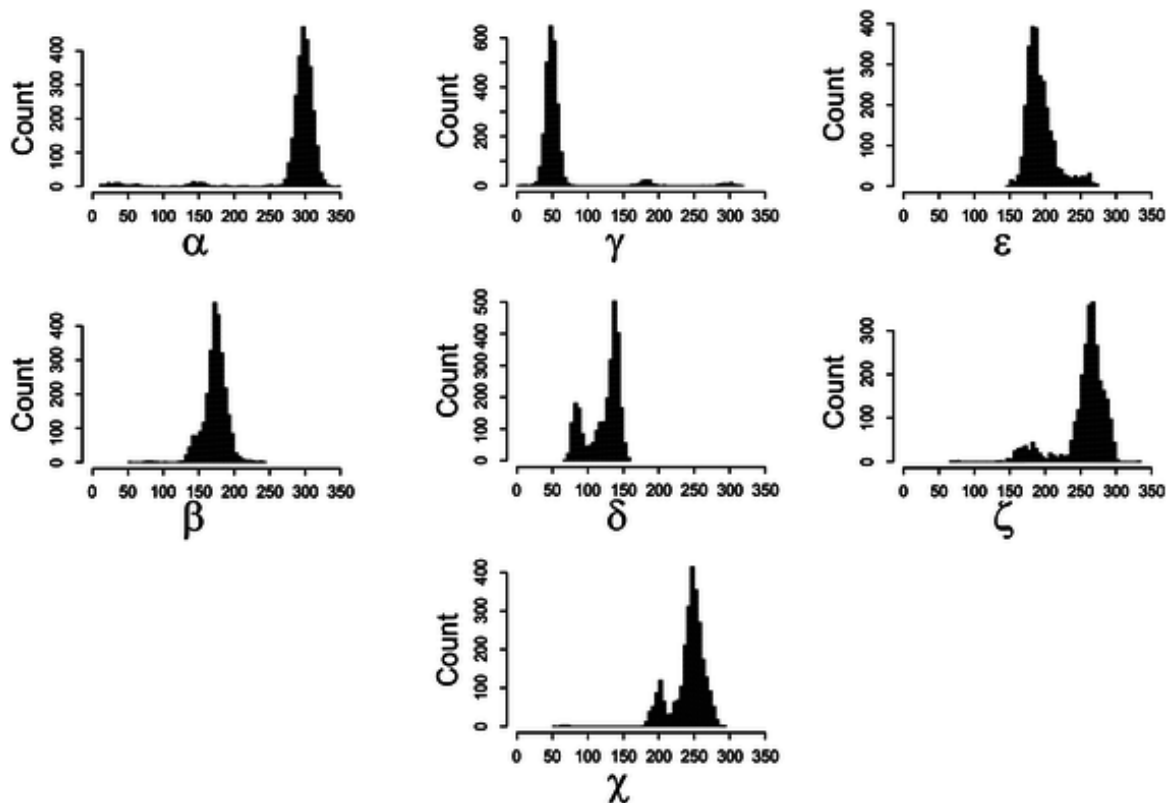


Рисунок 2.3: Гистограммы распределения торсионных углов в ДНК с В-формой и А-формой для структур с разрешением ниже 1.9Å.

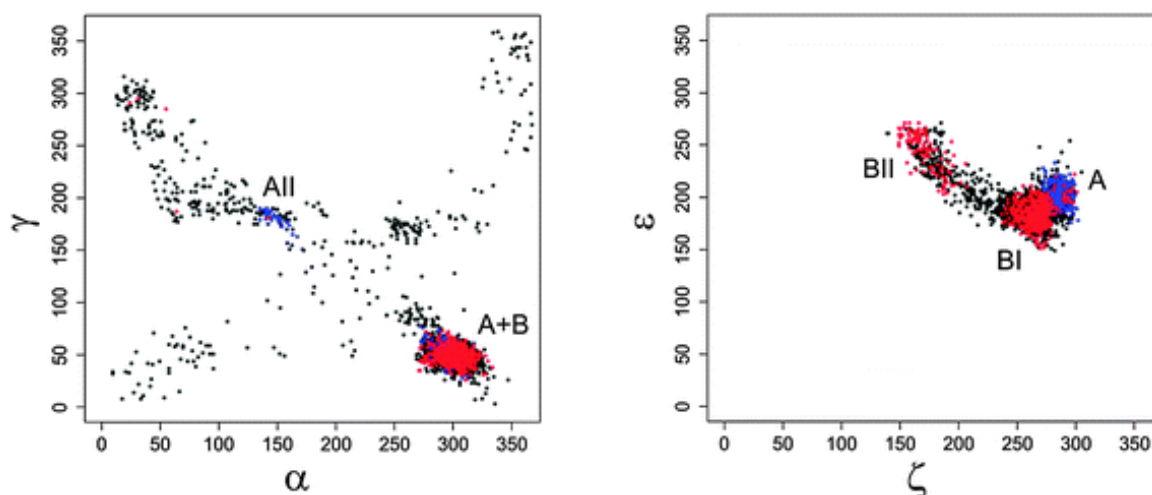


Рисунок 2.4: Свободная структура В-формы ДНК отмечена красным, а в комплексе с лигандами - чёрным. Синим отмечена А-форма ДНК. Основные конформации отмечены буквами [37].

Видно, что отдельные торсионные углы имеют некие разрешённые зоны, в то время как некоторые области совершенно не “заселены”. Здесь можно провести аналогию с картами Рамачандрана для белков. Основное отличие, на-

верное, состоит в том, что значения углов хорошо коррелируют друг с другом (Рисунок 2.5). Существует взаимосвязь между разными торсионными углами остова НК, которая характеризуется специфической комбинацией значений. Рассмотренные геометрические семейства конформаций остова ДНК отражают внутреннее предпочтение для некоторых значений углов и топологические требования, выдвигаемые разными формами ДНК.

Широкое разнообразие конформаций ДНК было изучено методами статистического анализа значений торсионных углов в структурах на основе данных РСА [37; 38]. В часто упоминаемом исследовании Свозила и коллег [37] рассматривается 8000 индивидуальных нуклеотидов примерно из 450 структур ДНК, разрешение которых не превышало 1.9Å. Для того, чтобы учесть корреляцию в значениях торсионных углов соседних нуклеотидов, также был проведен анализ динуклеотидных блоков. Данные для таких блоков определяли как индивидуальный признак, т.е. независимо от формы и топологии ДНК (например А- или В-формы ДНК). Как и ожидалось, анализ выявил основные конформеры, но также были найдены конформеры, которые одновременно имели признаки как А-, так и В-формы ДНК. В-форма ДНК представлена в двух основных конформерных формах: VI и VII. Каноническая конформация VI наиболее представлена среди известных данных. В обеих формах сахарное кольцо имеет С2'-эндо-конформацию. В свободном состоянии VI-форма ДНК имеет две разделённые области значений углов  $\chi$  и  $\epsilon$  (Рисунок 2.5). В то время, как угол  $\chi$  принимает значения около 260°, угол  $\epsilon$  обеспечивает транс-состояние (180°). Любопытно, что в VII-форме ДНК происходит симметричное изменение значений углов  $\chi \sim$  на 170° и  $\epsilon \sim$  - на 245°.

При образовании комплексов с лигандами переход из формы VI в форму VII перестаёт быть дискретным (Рисунок 2.5). Кроме этого, появляются конформе-

ры, которые не могут быть отнесены ни к А-, ни к В-форме ДНК. Эти конформации можно описать как однонуклеотидное состояние в А-форме, в то время, как остальные нуклеотиды представлены в В-форме ДНК; некоторые нуклеотиды могут находиться в некоем транзитном состоянии из А- в В-форму [37].

Каноническая А-форма ДНК, часто называемая АI, описывается состоянием сахара С3'-эндо ( $\epsilon \sim 80^\circ$ ): угол  $\chi$  принимает значения около  $300^\circ$ . Значения углов в следующем нуклеотиде (далее такие углы будут иметь индекс +1) составляют:  $\alpha \sim 300^\circ$ ,  $\beta \sim 180^\circ$ . Следующая основная конформация - это АII, при которой углы  $\alpha_{+1}$  и  $\gamma_{+1}$  принимают транс-конформацию. Также существуют конформеры, похожие на А-форму, и для них характерно значение угла  $\chi \approx 200^\circ$ . Конформация сахара принимает О4'-эндо-конформацию (угол  $\delta \sim 100^\circ$ ), что является переходным состоянием из С2'-эндо в С3'-эндо-конформацию.

Обе А- и В-формы ДНК разрешают движения в углах  $\alpha$  и  $\gamma$ , подобно колечатому валу, при этом они не влияют на ход основной цепи. Конформацию АII с транс-состоянием углов  $\alpha$  и  $\gamma$  можно найти в свободных молекулах. В то время, как для В-формы ДНК подобные состояния углов можно найти только в комплексе с лигандами и белками [39].

Z-форма ДНК - это очень интересный вариант конформации ДНК и серьезный тест для моделирования в классических силовых полях [40]. Все остатки гуанозина имеют син-конформацию вокруг угла  $\chi$ . В то время, как динуклеотид 5'-CpG-3' может иметь только одну конформацию - Z - , динуклеотид 5'-GpC-3' может иметь конформации ZI и ZII. Можно сказать, что истинным повторяющимся элементом в Z-ДНК является динуклеотид. Для комбинации "ZI/ZII" можно утверждать, что для остатка гуанозина  $\epsilon=290^\circ$ , то для следующего остатка цитидина угол  $\beta$  должен составлять  $240^\circ$  (форма ZI). Если для гуанозина угол

$\epsilon=65^\circ$ , то для цитидина  $\beta_{+1}=165^\circ$  (форма ZII). Надо отметить, что угол  $\beta$  не влияет на угол  $\epsilon$  последующего нуклеотида.

Ограничения на выборку, накладываемые разрешением структур РСА, приводят к тому, что некоторые важные топологические элементы остова не представлены. Например конформации нуклеотидов, где угол  $\gamma$  находится в транс-состоянии, что характерно для первого нуклеотида в петлях квадруплексных структур [37; 40].

Поступает всё больше сообщений о том, что остов ДНК вносит значимый вклад в свойства В-формы ДНК, в зависимости от последовательности [23; 41; 42]. Необходимость детального понимания вариативности структуры ДНК становится всё более насущной, так как до сих пор не найден универсальный код соответствия для ДНК-белковых взаимодействий [43; 44]. Дальнейший прогресс в нашем понимании влияния конформаций на функцию ДНК может быть связан с переходом на рассмотрение тетрануклеотидных контекстов в последовательности, количество которых составляет 136 штук. С другой стороны, большинство структурных данных для свободной ДНК получено для исключительно коротких последовательностей, что, в свою очередь, сужает горизонт анализа. Естественно, нет никаких гарантий, что конформации, найденные методом РСА, представляют выборку свойственных для живой клетки конформаций.

Белки узнают специфические сайты узнавания на ДНК, используя две стратегии, которые принято называть “прямое” и “непрямое” считывание. В случае прямого считывания, в последовательности ДНК узнаётся специфический паттерн с образованием определённых контактов азотистых оснований с аминокислотами на поверхности взаимодействия ДНК с белком. Непрямое считывание подразумевает узнавание сайта связывания, благодаря специфическим

изменениям гибкости или подвижности дуплекса ДНК, изменению электростатического потенциала и т.д. Обе модели узнавания приводят к изменениям в структуре ДНК, вплоть до изгиба дуплекса. Вероятно, что обе стратегии дополняют друг друга. Узнавание последовательности ДНК - очень сложный процесс. Трудно использовать простой код для описания процесса узнавания последовательности ДНК белками даже внутри одного класса ДНК-связывающих белков [43—45].

Надо отметить, что биоинформатический анализ (т.е. анализ данных, полученных экспериментально) не обнаружил конформеров ДНК, содержащих нуклеотидов с углом  $\gamma$  в транс-области. Тем не менее, такие конформации были найдены при моделировании молекулярной динамики с использованием ранних версий силового поля семейства parm. Накопление таких конформаций приводило к потере дуплексной структуры В-формы ДНК [31]. Это сравнение указывает на эффективность биоинформатического анализа для исследования стабильных конформаций ДНК, и он может быть использован как метод сравнения для всех остальных методов. Отсюда также следует, что незначительное изменение в свойствах остова может привести к значимым изменениям в геометрии ДНК. В ходе биоинформатического анализа было обнаружено 126 индивидуальных кластеров конформаций, из которого были отобраны 20 наиболее значимых [37]. Некоторые из кластеров содержат менее 10 индивидуальных представителей. Такие малонаселённые кластеры могут появляться из-за ограничений, связанных с экспериментальным определением структуры, и, вероятно, такие конформации не представлены в природных условиях. Тем не менее, они могут представлять важную информацию о состоянии торсионных углов в неканонических структурах. С другой стороны, значимое количество динуклеотидов не попало ни в один из 126 кластеров. Возможно, будущие квантово-

механические исследования смогут прояснить эту картину. В свою очередь, биоинформатические данные могут быть использованы как внутренний контроль для всех последующих исследований вычислительного характера.

## 2.2.2 Структурная биоинформатика РНК

Большая сложность конформационных состояний остова РНК приводит к широкому набору возможных структур РНК. Полное описание хода цепи РНК в целом затруднено, т.к. редко удаётся получить качественные данные для установления структуры [46—48]. Структуры больших РНК представляют основную часть данных о структуре РНК, но для таких объектов разрешение РСА ограничено 2.5-3.5 Å. На сегодняшний день нет высокоточных данных о многих важных конформациях РНК. В то время, как для остова ДНК удалось достичь разрешения в 1.9Å, то для основного набора данных из структур рибосомы это значение не превышает 2.4Å. Часто встречаются ошибки при расшифровке данных РСА, когда ориентация азотистого основания определяется как син-, в то время, как последующие данные РСА указывают на анти-. Таким образом, много индивидуальных состояний остова РНК могут быть неверно определены или получены в ходе оптимизации структуры, а не из экспериментальных данных. Конечно, можно исключить предположительно ошибочные данные из рассмотрения, но это внесёт предвзятость в выборку. Возможно, такие конформации характерны для напряжённых частей РНК. С другой стороны, если нуклеотид имеет несколько равнопредставленных конформаций, то среднее арифметическое положение, которое получается после анализа электронной плотности, вероятно, будет бессмысленным. Структурно-биоинформатический анализ ориентируется на те молекулы, структуры которых статичны. Современные методы

определения структуры детектируют наиболее представленные конформации, но, если представленных конформаций несколько, то определение структуры затруднено, и исследователи склоняются к тем данным, которые представляются им более рациональными. Стоит отметить, что для РНК именно конформационное равновесие играет ключевую роль в выполнении её функции.

Попробуем описать наиболее представленные конформации остова РНК. В РНК наиболее встречающийся набор значений торсионных углов  $\epsilon \sim 290^\circ$  и  $\alpha_{+1} \sim 295^\circ$ , как в канонической А-форме РНК [46; 48]. Углы  $\alpha$  и  $\gamma$  показывают чистое тримодальное распределение с пиками в  $295^\circ$  и  $55^\circ$ , соответственно. Угол  $\beta$  имеет широкое распределение Гаусса с пиком в  $180^\circ$ . Угол  $\epsilon$  имеет один пик негауссовской формы при  $180^\circ$ , а большинство выпадающих значений относится к явным ошибкам определения структуры [46]. Угол  $\delta$  имеет чистое бимодальное распределение, соответствующее С3'-эндо и С2'-эндо-конформации сахара. Для угла  $\epsilon$  наблюдается широкое распределение, что соответствует богатому разнообразию известных структур РНК. Вращение вокруг N-гликозидной связи имеет два основных пика негауссовской формы в значениях  $200^\circ$  (анти-) и  $60^\circ$  (син-).

Внимательное исследование и анализ экспериментальных данных на наличие ошибок является критически важным этапом при построении выборки для биоинформатического анализа конформаций остова РНК. Проблема большого количества степеней свободы в динуклеотидах РНК может быть решена упрощением анализа или уменьшением выборки, но при этом неизбежна потеря данных. Подход, использующий два псевдоторсионных угла на нуклеотид (подобен углам в диаграмме Рамачандрана для белков), может быть удобен для работы с нефилтрованными выборками экспериментальных данных. Применение этого подхода к анализу расчётных данных приводит к потере информации, что



не позволяет эффективно сравнивать экспериментальные данные и расчётные [49].

Ряд исследователей анализирует полноатомное конформационное пространство. Метод, основанный на анализе распределения данных каждого угла [50], был использован для построения т.н. структурного словаря. Этот анализ не предполагает корреляций между значениями углов, что приводит к потере важной информации. Пространство множества углов также анализировалось методом кластеризации “без внешних параметров”, но этот подход также не привёл к успешному решению проблемы [51]. Мюррей с коллегами [47] предложил полезный и физически взвешенный приём описания сопряжения сахар-сахар одним параметром. Реализация подхода состоит в описании углов  $\delta$  и  $\delta_{+1}$ , что является естественным представлением геометрии вокруг фосфатной группы (Рисунок 2.2). Они провели тщательный анализ данных для больших РНК, при этом данные фильтровали с помощью предлагаемого подхода. Также проводили фильтрацию по значениям В-фактора и пространственным пересечениям Ван-дер-Ваальсовых сфер. В результате были выделены 42 конформера остова РНК.

В работе Шнейдера и коллег систематизация проводилась на основе 14 торсионных углов динуклеотида включая угол вокруг N-гликозидной связи. При анализе были рассмотрены 2841 индивидуальных динуклеотида из большой субъединицы рибосомы из *Haloarcula marismortui*. 2010 динуклеотидов были отнесены к А-форме РНК, в то время, как оставшиеся 831 были определены как неканонические. Ожидается, что фокус кластерного анализа был направлен на эти 831 динуклеотида. Оказалось, что 510 из них нельзя было осмысленно классифицировать. Их геометрия могла быть как ошибочной, так и представлять редкие топологические явления. В ходе работы были классифицированы

18 конформаций, не относящихся к А-форме РНК. Кластеры с малой заселённостью не имели значимого влияния.

Результаты от разных исследователей [47; 48; 52] позже были объединены в один номенклатурный список [46]. Была разработана программа по автоматическому анализу и отнесению структуры РНК. Из 46 семейств конформаций остова РНК наиболее представлена классическая А-форма РНК с 4637 индивидуальными представителями. Следующей по встречаемости является конформация А-формы РНК, подобная коленчатому валу, с 275 индивидуальными представителями.

В заключение хочется отметить, что описание остова РНК является крайне сложной проблемой. Анализ известной структуры РНК или результатов вычислительных экспериментов не должен проводиться без сравнения с известными данными. Остов РНК представляет исследователям много интересных вопросов, которые могут быть исследованы с помощью трудоёмких, длительных, но высокоточных квантово-механических расчётов.

### **2.2.3 Электронная структура сахаро-фосфатного остова**

Сахаро-фосфатный остов имеет достаточно сложную электронную структуру, и делокализация электрона затрудняет его теоретическое описание. Благодаря эффектам гиперконъюгации, некоторая часть электронной плотности делокализована по всему остову, и степень делокализации зависит от конформации остова. Некоторые квантово-химические представления могут воспроизвести это явление с разными успехом. Методы с эффективным учётом корреляционных эффектов [53] недооценивают электронную плотность для ковалентных связей, располагая излишек вдоль всего остова. Высококачественные подходы

могут описывать эти эффекты аккуратно. Ожидаемо, что зависимость эффекта гиперконъюгации от конформации остова сильно затрудняет описание остова в классических силовых полях. Эффект гиперконъюгации возникает из донорно-акцепторных взаимодействий локальных орбиталей в положении С1' и С3'. Эти эффекты зависят от конформации остова и могут быть описаны только точными квантово-химическими расчётами. Этот факт, к сожалению, делает практически невозможным аккуратное описание распределения электронной плотности с помощью классических силовых полей.

#### **2.2.4 Квантово-механические расчёты больших и малых систем**

Исследование остова нуклеиновой кислоты квантово-механическими методами требует внимательного выбора модельной системы и методов расчёта. Здравый смысл предлагает для изучения модельную систему с максимальным размером. Тем не менее, выбор модели является сложным вопросом, и часто необходимо изучать небольшие системы. Одной из очевидных причин является то, что, чем меньше система, тем качество применяемых расчётов может быть выше. С другой стороны, вычислительная стоимость высококачественных методов квантовой механики стремительно растёт с числом атомов в системе [53]. Тем не менее, эта проблема постепенно отступает, в связи с быстрым развитием аппаратного и программного обеспечения. Истинная причина, которая затрудняет исследование крупных системы, - это сложность поверхности потенциальной энергии (ППЭ).

Квантово-механические (КМ) подходы не позволяют накопить данные о конформационном ландшафте, доступные в моделировании молекулярной ме-

ханики. В КМ-расчётах часто используется оптимизация геометрии, располагающая структуру в ближайшем минимуме ППЭ. На сегодняшний момент нет инструментов для эффективного исследования ППЭ. С ростом числа атомов также растет количество локальных минимумов энергии для двугранных углов. Исследования нуклеиновых кислот методами молекулярной механики показывают, какой удивительно богатой может быть конформационная динамика остова НК практически во всех диапазонах времён наблюдения [54—56].

В то время, как в расчётных экспериментах система становится всё больше, это ещё далеко до размера реального биополимера и, кроме того, в таких расчётах не учитывается влияние растворителя. Без учёта растворителя в системе исследователь сталкивается с большим количеством как внутримолекулярных взаимодействий, так и других структурных явлений, которые не встречаются в реальной жизни. К примеру, часто в таких взаимодействиях участвуют концевые группы, 2'-гидроксил, анионный фосфат. Например, при исследовании системы сахар-фосфат-сахар было обнаружено, что учёт стекинг-взаимодействий в НК не является полезным, так как энергия этих взаимодействий была переоценена, что привело к неверным значениям энергии остова [57]. Спонер и коллеги попытались использовать иминометиленовую группу вместо азотистых оснований, которая имитировала ароматические азотистые основания. К сожалению, даже такая модель была непригодна для изучения энергетического профиля остова. Вероятной причиной являлось то, что результат оптимизации геометрии зависел от начальной ориентации  $-NCH_2$ -групп. Существование различных способов взаимодействия метиленовых групп с остовом привело к невозможности выделения вкладов в энергию, характерных для остова [57]. Таким образом, ни азотистые основания, ни их простые аналоги не стоит использовать при изучении энергии конформации остова НК. Тем не менее, если задача предпо-

лагает взаимодействие оснований с остовом, то они должны учитываться явно, или объектом изучения должна являться N-гликозидная связь.

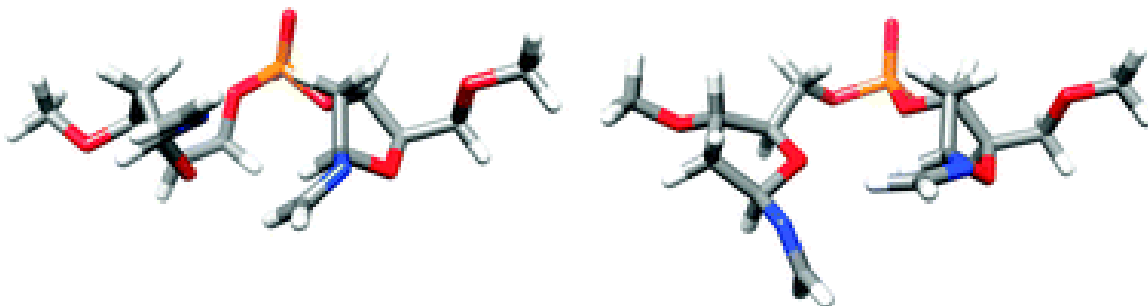


Рисунок 2.5: Два варианта взаимодействия иминометиленовых групп между собой: стекинг-взаимодействие (слева) и T-образное взаимодействие (справа).

Синим цветом обозначены атомы азота, серым - углерода, красным - кислорода, оранжевым - фосфора. [37].

Заряд остова затрудняет перенос в реальные системы результатов, полученных в газовой фазе путём КМ-вычислений. Даже один анионный заряд фосфатной группы вызывает трудности, в то время, как между двух или более зарядов возникает электростатическое отталкивание. Кроме того, очевидно, что сольватация меняет электронную структуру остова НК [58]. Самый лучшим вариантом подавления чрезмерного электростатического отталкивания является использование моделей с введением неявного растворителя, хотя точность таких расчётов значительно ниже, по сравнению с вычислениями в газовой фазе. Протонирование анионных атомов кислорода является химически бессмысленным. Возможна нейтрализация заряда одновалентными катионами рядом с атомами кислорода, но это также неоптимально, так как большую часть времени фосфаты не взаимодействуют с одновалентными катионами в растворе. Данные моделирования показывают, что примерно 10% времени катионы напрямую связаны с фосфатной группой [59].

Использование слишком малых систем также приводит к неточной геометрии остова. Специфической проблемой РНК является 2'-гидроксил рибозы: его

наличие усложняет конформационное пространство цепи РНК, так как он может быть легко вовлечён во внутримолекулярные взаимодействия. Тем не менее, в неканонических сегментах РНК группа 2'-ОН очень часто участвует в ключевых молекулярных взаимодействиях в качестве донора в водородной связи [17—21]. Таким образом, исследования остова РНК часто включают окружающий контекст, для обеспечения реалистичной картины. Например, недавнее исследование небольшой модельной системы [60] указывает на то, что группы 2'-ОН в неканонических регионах РНК предпочтительно взаимодействуют с соседними O3'. Тем не менее, релевантность этого утверждения трудно проверить, так как применённые расчёты не учитывают ни корреляцию между торсионными углами остова, [46; 48], ни образование водородных связей между основаниями [17]. В действительности, группы 2'-ОН в неканонических регионах часто выступают в качестве доноров в критических водородных связях. Этот пример иллюстрирует ограниченность небольших модельных систем. Стоит обратить внимание на то, что проблемы с возникающими в газовой фазе взаимодействиями и с наличием заряда у остова влияют также на качество описания нуклеиновых кислот в молекулярной механике.

### **2.2.5 Точность различных методов КМ при исследовании N-гликозидной связи**

На протяжении многих лет наиболее распространённым методом в расчетах КМ был метод Хартри-Фока (HF), это метод пренебрегает электронной корреляцией. Даже в настоящее время он используется некоторыми исследователями для оптимизации геометрии, за которой затем следует расчёт энергетических параметров с учётом корреляционной энергии. При разработке классических

силовых полей семейства “Amber”: ff94, ff98, ff99 [61], КМ-расчёты проводились именно таким образом.

Недавние исследования для улучшения параметризации силовых полей нуклеиновых кислот использовали уже более эффективные методы. Последнее усовершенствование поля CHARMM основывается на уровне теории MP2/6-31+G\* для оптимизации геометрии, а расчёт энергии проводился на уровне теории RI-MP2/cc-pVTZ [62]. Методы DFT (теория функционала плотности) и LMP2 (локальная теория возмущений Мюллера-Плессет) были применены для улучшения потенциала торсионных углов  $\alpha$  и  $\gamma$  для силового поля AMBER (версия parmbsc0) [31].

В то время, как участилось применение современных методов квантовой химии, а именно, DFT, часто встречается его использование без оценки возможных ошибок. Точность важна, в особенности для описания остова нуклеиновых кислот, где небольшие ошибки могут накапливаться вдоль цепи биополимера. Ошибка в несколько десятых долей ккал/моль в потенциале торсионного угла  $\chi$  может полностью деградировать структуру РНК [63; 64].

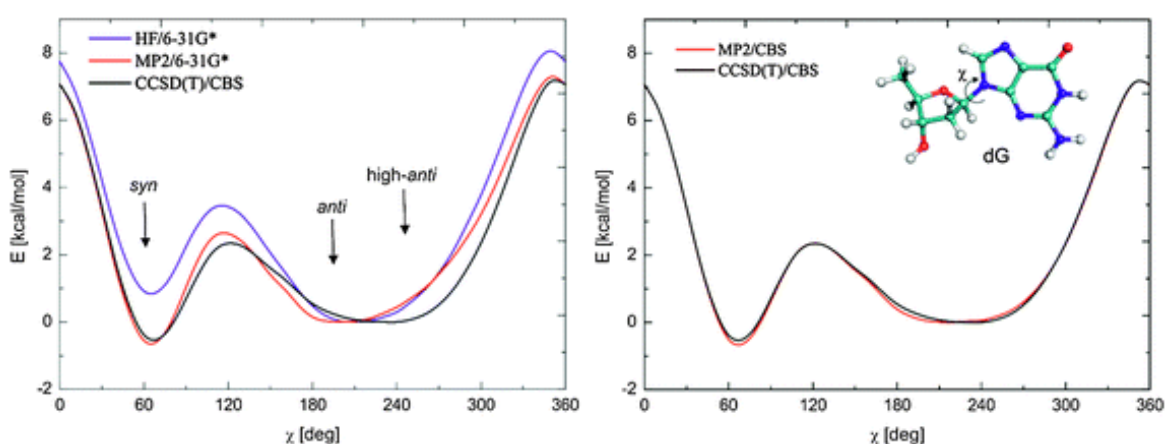


Рисунок 2.6: Профиль энергий торсионного угла  $\chi$  дезоксигуанозина (dG). Слева находятся профили, полученные со следующими уровнями теории: HF/6-31G\* (синий), MP2/6-31G\* (красный), а также высокоточный CCSD(T)/CBS (чёрный). Справа - профили, полученные с уровнями теории MP2/CBS (красный) и CCSD(T)/CBS (чёрный). [64].

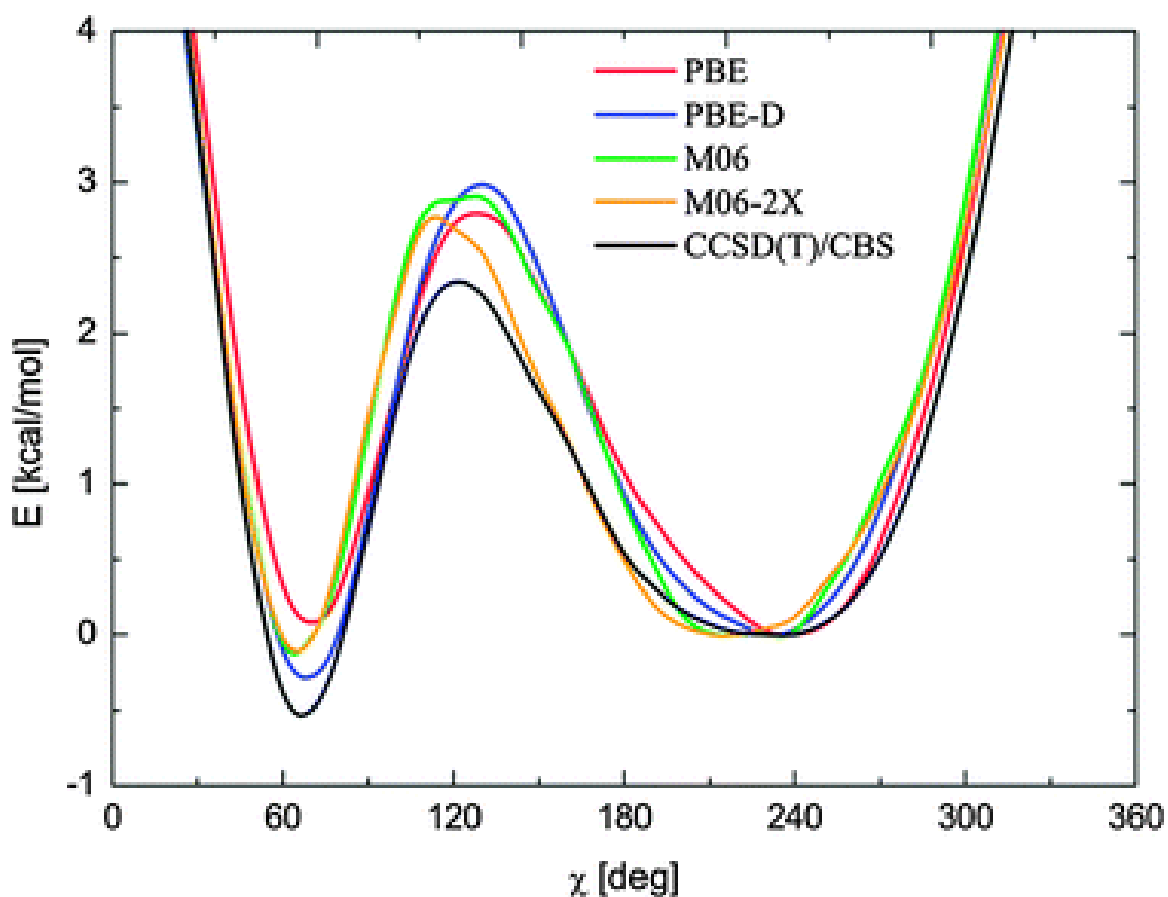


Рисунок 2.7: Профиль энергий вращения угла  $\chi$  для dG на основе различных функционалов DFT и базиса 6-311++G(3df, 3pd): PBE (красный), с коррекцией дисперсии PBE-D (PBE/D-1.06-23) (синий), M06 (зелёный), M06-2X (оранжевый), и профиль сравнения CCSD(T)/CBS (чёрный).

Рисунки 2.6 и 2.7 иллюстрируют результаты по исследованию энергии при вращении вокруг угла  $\chi$  для дезоксирибонуклеозидов [64]. Подход HF/6-31G\* даёт ненадежные результаты как для энергий, так и для геометрии. Значительные ошибки встречаются для относительных значений энергии в важных точках на ППЭ, в основном, в син-/анти- минимуме и в анти-области. Геометрия глобального минимума получилась довольно неточной, и наблюдалось существенное изменение оптимального значения угла  $\chi$  в сторону анти-области. Большая часть информации о конформациях, использованная для построения текущих силовых полей, была получена с помощью HF/6-31G\*. Так как метод HF сравним по требованиям к вычислительной мощности с GGA и другими



функционалами DFT, то нет никаких оснований для использования этого метода.

Уровень теории MP2 с небольшим базисом 6-31G\* значительно улучшает данные, полученные методом HF/6-31G\*. Однако в некоторых важных частях графика зависимости энергии от значения угла уровень ошибки все ещё неприемлемы (Рисунок 2.7).

Современные методы DFT работают лучше, чем HF/6-31G\*, однако ошибки относительных значений энергии для син/анти областей все ещё присутствуют. Примечательно, что описана неточно именно нижняя высота барьера. По данным Спонера и коллег, PBE87 принадлежит к числу самых точных функционалов GGA для анализа конформаций молекул, однако он по-прежнему дает весьма значимые ошибки для рассматриваемой системы. Частичного улучшения можно достичь путём уточнения PBE с помощью эмпирической коррекции дисперсионной составляющей (PBE-D) [65]. Таким образом, ни один из методов DFT, протестированных в работе Спонера, не может быть рекомендован для получения точных профилей энергии торсионных углов нуклеозидов.

Метод MP2/CBS предоставляет достаточную точность профиля угла  $\chi$ , с ошибкой не более 0,1 ккал/моль, близкую к точности данных, полученных высокоточным методом CCSD(T)/CBS (Рисунок 2.7).

Итак, расчёты профиля угла  $\chi$  показали, что уровни теории KM ниже, чем MP2/CBS, и не могут быть рекомендованы для точных расчётов энергии, а для самых затратных CCSD(T) коррекция не требуется. Современные методы DFT позволяют нам изучать гораздо более крупные фрагменты НК, состоящие вплоть до нескольких нуклеотидов с учётом полярного растворителя. Тем не менее, перед их использованием важно учитывать коррекцию для дисперсионных взаимодействий и полезно проверять важные точки методом MP2/CBS

для репрезентативного набора геометрий. Следует отметить, что требования к уровню теории в расчётах остова отличаются от таковых для азотистого основания, где уровень MP2/CBS не является достаточным, и вычисления MP2/CBS могут корректироваться на уровне CCSD(T)-вычислений даже при использовании небольшого поляризованного базисного набора. Напротив, для расчёта образования пар оснований уровень MP2/CBS широко признан достаточно точным, хотя поправки CCSD(T) могут ещё улучшить результаты. Различные части нуклеотида имеют разные требования к уровню теории для достижения результатов разумного качества.

Отметим также, что точность расчётов КМ является лишь одним из аспектов в вычислениях, связанных с остовом НК. Одинаково важны как выбор адекватной модели, так и правильный учёт сольватационных эффектов. Показано, что учёт эффектов растворителей значительно облегчает использование данных КМ при описании НК в классическом силовом поле [64; 66].

### **2.2.6 Остов ДНК, содержащий мышьяк (As-ДНК)**

Возможности использования вычислений КМ для изучения остова нуклеиновых кислот, были проиллюстрированы в недавней дискуссии о возможности жизни на основе мышьяка. Было высказано предположение, что бактерия GFAJ-1 использует арсенаты в месте фосфатов в остове ДНК [67]. С самого начала это предложение было окружено большим количеством противоречий, частично из-за отсутствия информации о структуре и стабильности сложных эфиров соединений мышьяка.

Для того что бы мышьяк заменил фосфор в ДНК, остов на основе мышьяка должен иметь соответствующие структурные, термодинамические и кинетиче-

ские свойства. Расчёты КМ показали, что мышьяк мог бы быть правдоподобной заменой фосфора, исходя из термодинамической и геометрической точек зрения. В расчётах использовали анализ Natural Bond Orbital (NBO) для оценки главных стереоэлектронных эффектов, действующих в регулярной ДНК-цепи и в её As-варианте [68]. Оптимизированные геометрии и эффект сверхсопряжения вдоль  $C3'-O3'-X-O5'-C5'$  ( где  $X = P, As$ ) проявляют поразительное сходство для арсенатов и фосфатов.

Таким образом, арсенат был бы идеальной заменой для фосфата в основной цепи ДНК как по структурным, так и по термодинамическим соображениям, что было показано и другими [69]. Однако, как было отмечено Младеком и коллегами, этого не достаточно, чтобы утверждать о возможности жизни на основе мышьяка, поскольку эти данные не дают никакой информации о гидролитической стабильности остова As-ДНК. Ранее исследования показали, что простые арсенат-эфиры гидролизуются на много порядков быстрее, чем их фосфатные аналоги [70—72].

Более поздние результаты поддержали гидролитическую инертность As-ДНК, спекулируя на том, что затрудненный доступ из-за полинуклеотидной цепи может привести к резкому снижению скорости гидролиза остова [67]. Таким образом, были проведены эксперименты по оценке скорости гидролиза серии фосфорных и арсенатных эфиров с алкильными группами возрастающей сложности.

Оказалось, что энергии активации незначительно зависят от сложности органической части сложных эфиров, а вот разность между энергиями активации фосфатов и арсенатов существенна: 15 ккал/моль как в газовой фазе, так и в растворителе [70]. Это предполагает, что сложные эфиры мышьяка могут гидролизироваться на 7-10 порядков быстрее, чем фосфатные эфиры, и эта разница

не зависит от размера окружающих их алкильных групп. Кроме того, сравнение уровней энергии в газовой фазе и в растворе показывает, что разница в скоростях гидролиза между фосфатными и арсенатными эфирами справедлива и для менее полярных условий. Неподходящие гидролитические свойства вряд ли позволят построить стабильную As-ДНК. При решении подобных вопросов квантово-химические исследования с помощью скромных усилий могут обеспечить чёткое химическое представление об интересующем объекте [70]. В настоящее время доминирует мнение, что результаты в исходной экспериментальной работе были интерпретированы неверно, а именно: не существует ДНК с атомами мышьяка в остове [73].

### **2.2.7 Альтернативные варианты остова НК**

Создание искусственной жизни является одной из самых далёких и желанных целей современной молекулярной биологии и химии. В связи с этим, значимые усилия были направлены на создание альтернативных генетических систем на основе химических модификаций остова НК. Модифицированные нуклеиновые кислоты также могут быть полезным для фармацевтических целей [74]. На сегодняшний день синтезирован ряд модификаций остова нуклеиновых кислот, тем не менее, ни одна из них не была способна самовоспроизведению. В какой-то степени все эти модификации требуют помощи со стороны генетического аппарата на основе НК [75]. Возможно, самым большим прорывом в этой области была работа Пиньейру и коллег, которые впервые разработали генетическую систему, способную к дарвиновской эволюции на основе синтетического полимера (1,5-дегидрохекситол нуклеиновая кислота, HNA), содержащего информацию [76].

Ранее обсуждалась As-ДНК, которую можно рассматривать в качестве простого альтернативного варианта остова, построенного по аналогу ДНК. Альтернативные переносчики генетической информации могут также включать химические модификации сахарной части остова. Действительно, простые сахара могут быть компонентами для “информационных” полимеров. Есть примеры по построению остова на основе триозы (TNA) [77] и гликоля (GNA) [78]. Так как TNA и GNA более просты, чем НК, предположили, что они играли роль “информационного” полимера в мире до появления НК. Другом классе искусственных остовов представлен циклогексановыми производными для замены остатков сахара, например, HNA [79] и ANA [80]. Несмотря на это, пептидные нуклеиновые кислоты (PNA) [81; 82] представляют собой наиболее известный класс альтернативных “информационных” полимеров с модифицированным остовом.

Несмотря на растущую популярность альтернативных вариантов остова НК, вычислительные исследования этих систем ещё далеки от систематичности [83; 84]. Например, в исследовании, сочетающем кристаллографические и квантово-химические подходы, показано, что ANA:РНК дуплексы стабилизируются образованием водородной связи внутри одной цепи с участием атомов O2' из альтрозной части одного остатка и атома кольца O4' следующего остатка [83]. Однако углубленное изучение альтернативных вариантов остова ДНК требует атомистического моделирования с подходящими силовыми полями, которые предназначены для каждого отдельного варианта остова. Очевидные сложности, возникающие в силовых полях для описания ДНК и РНК, которые подчеркиваются в многочисленных частях этого обзора, указывают на необходимость создания хорошо сбалансированного силового поля для альтернативных видов остова. Несмотря на огромные усилия, которые были предприняты для

параметризации ДНК и РНК, а также на доступность однозначных экспериментальных данных, их всестороннее тестирование выявило тот факт, что силовые поля всё ещё далеки от совершенства. В принципе, аналогичные усилия должны быть направлены на подготовку надёжных силовых полей для нуклеиновых кислот с модифицированным остовом. Примечательно, что почти все расчётные исследования, опубликованные для вариантов остова, сосредоточились на PNA, как на объекте с относительно простой параметризацией [85—87].

## **2.2.8 Моделирование остова НК методами молекулярной механики**

Наиболее распространённый вычислительный подход к изучению нуклеиновых кислот - это моделирование молекулярной динамики (МД) в явно заданном растворителе. НК описывается классическими эмпирическими потенциалами (силовые поля). МД может раскрыть информацию обо всех аспектах эволюции структуры во времени, в том числе, можно наблюдать позиции всех молекул воды и ионов. Так как тепловые колебания важны для функционирования биополимеров, то методы моделирования МД незаменимы.

Даже самые точные вычисления, с точки зрения описания потенциальной энергии, могут привести к нереалистичным результатам при недостаточности выборки. Кроме того, даже при достижении времени наблюдения в микросекундном диапазоне масштаб выборки остаётся недостаточным для многих приложений. Нельзя ожидать, что простое уточнение расчётов (современные подходы DFT) является оптимальным решением для изучения биологических систем [53].

Тем не менее, самый важный вопрос для современного молекулярного моделирования нуклеиновых кислот - это качество силовых полей. Благодаря квантово-химическим расчётам, у нас есть возможность детально изучить ключевые аспекты работы силовых полей. Силовые поля - это аналитические функции, связывающие структуру с потенциальной энергией. Они состоят из набора периодических функций для двугранных углов и гармонических функций для длин связей и валентных углов. Молекулярные взаимодействия в силовых полях аппроксимируются функцией Леннарда-Джонса. Электростатические эффекты основаны на постоянных точечных зарядах, локализованных на центрах атомов. Параметры силового поля не реагируют на изменения в геометрии молекулы и изменения окружающей среды.

В силовых полях многие эффекты (все виды поляризации и переноса заряда) не учитываются, по определению. Тем не менее, вклад таких взаимодействий может быть учтён эффективно, но неточно. Для этого приходится оптимизировать параметры уравнений силовых полей во время процедуры подгонки, и, в основном, оптимизации подвергаются параметры для описания торсионных углов. Эти параметры не имеют прямого аналога в природе и используются для окончательной настройки силового поля, оказывая большое влияние на результаты моделирования [64].

Все последние уточнения силовых полей для нуклеиновых кислот были основаны на модификации параметров торсионных углов для сахаро-фосфатного остова [31; 62; 64; 88]

Любое изменение точечных зарядов в любом случае требует полномасштабной репараметризации для двугранных углов. Точные данные из расчётов КМ играют двойную роль. С одной стороны, они обеспечивают как можно более точными данными для описания остова. С другой стороны, физически правиль-

ные данные затем используются для параметризации функции торсионных углов, не имеющих прямого физического смысла, минимизируя различия между описаниями КМ и молекулярной механики (ММ). В то время, как потенциал торсионного угла формально является внутримолекулярным свойством, его оптимизация учитывает и основные межмолекулярные взаимодействия. Однако некоторые специфические межмолекулярные взаимодействия либо неправильно учтены, либо полностью игнорируются в описании силовых полей. Таким образом, любое силовое поле, в конечном счете требует тщательного тестирования в моделировании.

Исходное силовое поле, предложенное Корнеллом и коллегами, - "AMBER" (версии parm94 или ff94 ), широко используется для моделирования нуклеиновых кислот [89]. Описание торсионных углов в этом силовом поле было несколько раз пересмотрено. Модификации силовых полей ff98 и ff99 содержат незначительные изменения описания торсионного угла вокруг N-гликозидной связи и вокруг связей типа O-C-C-O для увеличения степени закрутки спиральной B-формы ДНК [90]. Важным изменением была модификация силового поля parmbsc0 (Барселона), где было пересмотрено описание торсионного угла  $\gamma$ , который ранее необратимо переключался в транс-состояние, что приводило к необратимому разрушению структуры B-формы ДНК [31] (Рисунок 2.8).

Недавно была предложена модификация силового поля parm $\chi_{OL3}$ , которая была специально разработана для РНК на основе поля parmbsc0 [64]. Parm $\chi_{OL3}$  предотвращает необратимую дегградацию РНК с образованием неспиральных двуцепочечных структур из-за патологического смещения угла  $\chi$  в состояние анти- (Рисунок 2.8). Parm $\chi_{OL3}$  был параметризован с самым высоким уровнем теории с учётом электронной корреляции и с использованием модели раствори-



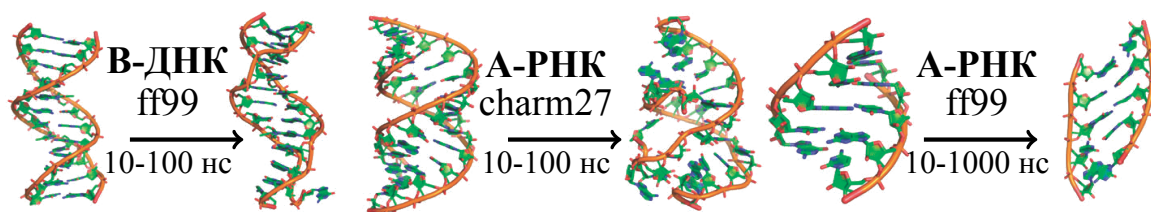


Рисунок 2.8: Деформации канонических молекул НК при использовании неадекватной версии силового поля. А) Прогрессивная деградация В-формы ДНК (CGCGAATTCGCG, дуплекс), в связи с медленным накоплением необратимых транс-состояний в версии силового поля ff99. Б) Дестабилизация дуплекса РНК (UA)<sup>6</sup> в поле CHARMM27. В) Деградация спиральной формы дуплекса РНК (GCGCG) в ff99, в связи с неточностью описания угла вокруг N-гликозидной связи. Эта проблема была исправлена в parm $\chi_{OL3}$ .

теля COSMO для учета сольватации НК. На сегодняшний момент рекомендуется использовать комбинацию силовых полей ff99 + parmbsc0 для ДНК и ff99 + parmbsc0 + parm $\chi_{OL3}$  для РНК. Также было уточнено описание потенциала угла  $\chi$  для ДНК (parm $\chi_{OL4}$ ) для улучшения моделирования неканонических молекул ДНК с нуклеотидами в син-конформации, таких, как антипараллельные G-квадруплексные ДНК [40]. И, хотя ff98, ff99 и parm $\chi_{OL4}$  можно рассматривать всего лишь как попытки тонкой настройки, однако parmbsc0 и parm $\chi_{OL3}$  имеют большое значение для достижения стабильности канонических и неканонических структур ДНК.

Расчёты КМ помогли также в уточнении силовых полей семейства CHARMM [25] для нуклеиновых кислот. Затем были созданы и испытаны несколько версий силового поля. При описании РНК в силовом поле CHARMM была пересмотрена параметризация торсионных углов 2'-гидроксильного протона рибозы для того, чтобы избежать неустойчивости РНК при моделировании МД (CHARMM36) [88] (Рисунок 2.8). Тем не менее, даже в новой версии в некоторых тестах по-прежнему наблюдалась некоторая нестабильность структуры

РНК [91], так что необходимо приложить некоторые дополнительные усилия, для более точного описания РНК в полях семейства CHARMM.

Наряду с упомянутыми ранее модификациями, были предприняты другие попытки оптимизации полей семейства “AMBER”, которые ещё недостаточно испытаны и обычно не используются [92; 93]. Эти модификации связаны с уточнением описания угла  $\chi$ .

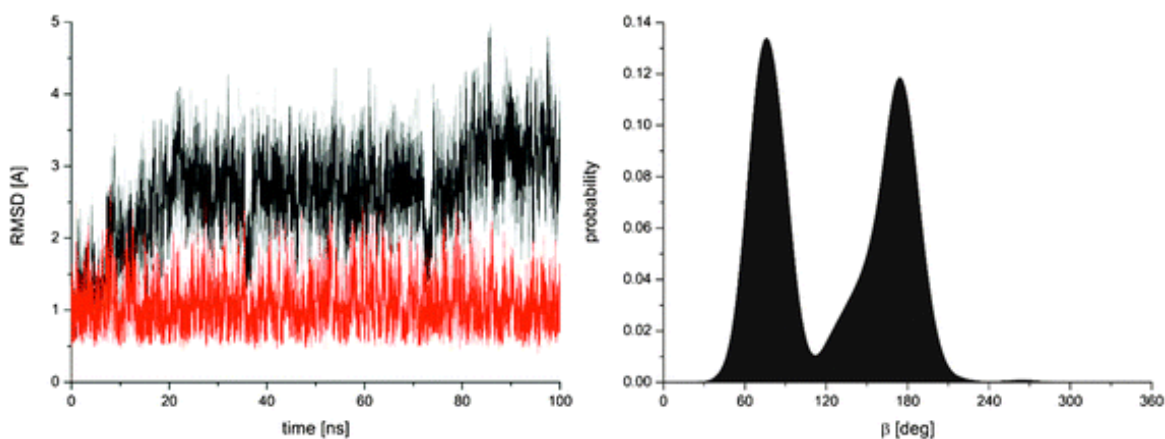


Рисунок 2.9: Моделирование МД канонической А-формы дуплекса РНК (GCACCGUUGG)<sup>2</sup> (PDBID:1QC0) с использованием недавней версии поля AMBERTOR [94]. Слева: среднеквадратичное отклонение (RMSD) остова. Чёрной линией обозначено моделирование с AMBER99TOR, красной линией - ff99 + parmbsc0 + parm $\chi_{OL3}$ . Справа: распределение значений торсионного угла  $\beta$  при моделировании AMBER99TOR. Высокая заселённость малых значений  $\chi$  не согласуется с кристаллографическими данными и показывает, что неудачная репараметризация может ухудшить баланс сил из-за непредсказуемых побочных эффектов. Это может быть связано как с неточностью первичных расчётов КМ, так и с неточностью последующей подгонки.

Одной из главных причин успеха силового поля “AMBER” является его элегантная модель для описания электростатических взаимодействий. Частичные атомные заряды, которые сами по себе не имеют большого физического смысла, пригодны для воспроизведения наблюдаемой в КМ картины. Этот подход известен как подгонка электростатического потенциала (ESP) вокруг небольших фрагментов нуклеиновых кислот. Так как ESP яв-

ляется важным фактором, определяющим электростатическую часть молекулярных взаимодействий, то описание точечного заряда с учётом ESP является оптимальным в пределах этого силового поля. Этот подход также хорошо себя показал для описания стэкинга и других взаимодействий с участием азотистых оснований [95]. Тем не менее, последние исследования показали, что молекулярно-механическое описание гетероатомов переоценивает энергию стэкинг-взаимодействий [96]. Это может иметь значимые последствия для достоверности моделирования, так как может привести к систематической излишней стабилизации структур со стэкингом в моделировании. К сожалению, проблему нельзя исправить в нынешнем виде силового поля. Для гибкого и заряженного сахаро-фосфатного остова подход ESP менее оправдан, так как использование постоянного набора зарядов не позволяет точно описывать поверхность потенциальной энергии остова.

## 2.3 Моделирование молекулярной динамики G-квадруплексных ДНК

Гуаниновые квадруплексы (G-ДНК), возможно, наиболее важные неканонические структуры ДНК. Основной структурной единицей G-ДНК является плоский квартет циклически связанных водородными связями гуанинов. Несколько последовательных квартетов, образующих стопочную структуру, называют тяжом G-ДНК, в центральную полость которого часто выстраиваются одновалентные катионы. На рисунке 2.10 показан простой тяж с двумя квартетами, стабилизированных одним катионом  $K^+$ . Молекула G-ДНК может состоять из четырех, двух или одной молекул олигонуклеотидов, которые соединяются вместе и образуют тяж G-ДНК в различных комбинациях с параллельными и антипараллельными ориентациями цепей.

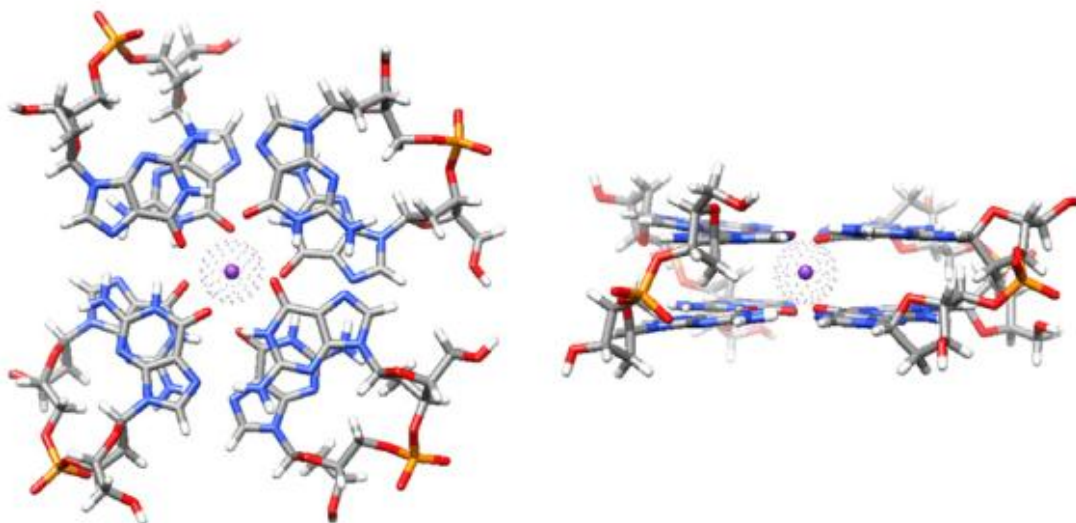


Рисунок 2.10: Двуквартетный тяж G-ДНК с антипараллельно направленными цепями. Слева представлен вид сверху, справа - вид сбоку. Стабилизирующий катион изображён фиолетовой сферой.

В природе наиболее актуальны варианты с одноцепочечной (мономерной) топологией. Под топологией G-ДНК понимают способ организации петель при

образовании квадруплекса. Мономерным и димерным квадруплексам, состоящим из одной или двух последовательностей ДНК, нужны одноцепочечные петли для изменения хода цепи. Структура петель играет важную роль в биологической функции G-ДНК. Хотя архитектура G-ДНК выглядит на первый взгляд простой, на самом деле она представлена огромным набором топологических вариантов, которые часто являются чрезвычайно чувствительными к последовательности оснований и окружению [15; 97—100]. Даже концевые остатки нуклеотидов в последовательности ДНК могут иногда влиять на общую топологию G-квадруплекса [101].

Первые расчёты по моделированию G-ДНК с использованием соответствующего силового поля и протоколов моделирования были опубликованы в 1998 году [102], и примерно в то же самое время появились данные о моделировании МД для I-квадруплексов (квадруплексы на основе инозина). Эти исследования G-ДНК были выполнены с помощью пакета программ AMBER [89] и с использованием первого поколения силовых полей (ff94). Первое исследование было основано на внушительной для того времени длине траектории в 20 наносекунд. Исследование выявило несколько результатов, которые остаются полностью актуальными до сих пор (что, однако, не всегда справедливо для области моделирования МД). Например, была обнаружена уникальная жёсткость квадруплексного ствола и показано, что тяж заметно дестабилизирован в отсутствие катиона в центральной полости, а также была обнаружена возможность гидратации центральной полости при отсутствии катиона. Авторы работы подчеркнули, что тяж может быть достаточно стабильным с ненасыщенной катионами центральной полостью.

Исследование показало глубокие структурные конфликты в рентгеновской структуре G-ДНК с латеральными петлями  $(G_4T_4G_4)_2$ , поддерживая альтерна-

тивную топологию с диагональными петлями, которая была предсказана на основании данных ЯМР (и позже, наконец, была подтверждена методом РСА) [103; 104]. Моделирование продемонстрировало несовершенство описания катионов внутри стебля. Первое исследование показало, что тщательно выполненные и проанализированные вычислительные эксперименты по моделированию МД предоставляют интересную информацию о G-ДНК.

### **2.3.1 Что интересного может рассказать нам моделирование G-ДНК?**

При оценке результатов моделирования важно понять, что именно этот метод делает и каковы его основные ограничения и приближения. МД в явном растворителе является, по сути, системой с одной молекулой или с одной копией молекулярного комплекса, растворенной в водном растворе. Большинство изученных систем начинается с точной стартовой геометрии (набор координат атомов), как правило, полученных из экспериментальных данных с высоким разрешением: зачастую из результатов РСА и ещё чаще – из данных структурного ЯМР. Качество стартовой геометрии критически влияет на последующее моделирование по двум причинам. Во-первых, время моделирования может быть недостаточно для того, чтобы преодолеть энергетические барьеры для перехода в более выгодную конформацию структур. Во-вторых, качество теоретической модели биополимера (силового поля) может быть недостаточно точным, чтобы найти правильную структуру.

В то же время моделирование может быть полезно для обнаружения неточностей в экспериментальных структурах, о чём, как правило, свидетельствуют

быстрые структурные изменения в ходе молекулярной динамики, когда система имеет высокое стартовое значение потенциальной энергии.

Классической задачей для моделирования является исследование известных экспериментальных структур и их ближайших аналогов (например, полученных путём замены оснований, изменением протонированного состояния и т.д.) [105—108]. Более амбициозной задачей является попытка использовать имитационное моделирование для предсказаний новых структур, а в крайних случаях – для имитации пути самосборки молекулы. Однако есть соблазн исключить такие значимые вычислительные затраты полностью, так как они в целом более рискованны и требуют глубокого понимания того, какие приближения можно применять. Есть много исследований в современной литературе, в которых авторы недостаточно хорошо ориентируются в возможностях современных методов моделирования. Такие работы часто опубликованы в высокорейтинговых журналах, так как результаты выглядят, на первый взгляд, более привлекательными, по сравнению с исследованиями, которые выполняются с применением более консервативных подходов.

Начиная со стартовой геометрии, биомолекулярная система, окружённая молекулами воды и ионами, проходит 1-1000 и более наносекунд молекулярной динамики. Метод имитирует подлинные тепловые флуктуации исходной структуры. Затем исследователи пытаются изъять из накопленных данных полезную информацию об исследуемом объекте, однако количество этих данных по-прежнему очень мало, по сравнению с явлениями в реальном мире. С появлением новых мощных аппаратных средств, таких, как графические процессоры (GPU), или даже специализированного оборудования, временной масштаб моделирования биомолекул будет постепенно смещаться в микросекундную область или даже миллисекундную [109; 110].

## 2.3.2 Как можно сравнить моделирование с экспериментом?

Многие полагают, что моделирование в растворе нацелено на то, чтобы имитировать эксперимент. Однако условия моделирования являются специфическими и не соответствуют ни одному эксперименту. Типичный эксперимент моделирования молекулярной динамики содержит одну биомолекулу, например, тяж G-ДНК, и его окружают конечным количеством растворителя и ионов при введении периодических граничных условий. Размер ячейки с растворителем выбирают так, чтобы он был достаточно мал для минимизации затрат на расчёт растворителя, но, в то же время, необходимо имитировать условия, при которых воспроизводятся свойства чистого растворителя вдали от растворённого вещества. К сожалению, при относительно высокой концентрации G-ДНК (растворённого вещества), часто используемой в эксперименте, 5-50 мМ, невозможно поместить одну молекулу в ячейку, а значит, нельзя удовлетворить последнему условию.

Периодически размещённые в ячейках молекулы G-ДНК, очевидно, не могут взаимодействовать друг с другом. Например, при высокой концентрации, молекулы G-ДНК не могут агрегировать, хотя должны, если, конечно, несколько молекул G-ДНК не присутствуют в одной и той же ячейке. Это указывает на искусственность моделируемой системы и может приводить к появлению артефактов в моделировании [111]. Помимо вопросов о конечном размере, высокой концентрации растворённого вещества и периодичности, есть также потенциальные проблемы с влиянием солей на структуру. Важно подчеркнуть, что моделирование должно служить в качестве автономного дополнения к экспериментам. Например, структуры РНК “изгиб-поворот” [112—114] или внутренние петли РНК 5'-UAA / 5'-GAN [115; 116] являются важными строительными



блоками, которые периодически встречаются в рибосоме, и их стабильность обеспечивается также дополнительными взаимодействиями РНК с окружением или даже с белками. Таким образом, их равновесное состояние в изолированном виде, определённое методом ЯМР, отличается от того, что происходит в рибосоме, и, следовательно, динамика конформации в изолированном состоянии имеет ограниченный биохимический смысл. В то же время, в отличие от экспериментальных подходов, моделирование может аккуратно исследовать свойства переходов таких РНК из состояний, близких к биологически значимым. С другой стороны, если разворачивание структуры предотвращено достаточно высоким энергетическим барьером, то результат моделирования не приведёт к более стабильной структуре. Таким образом, моделирование собирает полезные данные только о близких альтернативных структурах. Для данного случая не стоит проводить аналогии между моделированием в растворе и экспериментом. Моделирование таких структур должно быть связано с исходным состоянием этого элемента в составе рибосомы. То же самое можно сказать о квадруплексной ДНК: есть все основания использовать имитационное моделирование для микросостояний G-ДНК, которые нельзя обнаружить в экспериментальных исследованиях, но они могут участвовать в формировании квадруплексов [117].

Кроме того, можно исследовать ту топологию G-ДНК, которая, вероятно, наименее представлена в типичных условиях эксперимента. Целью такого исследования может быть понимание причин, по которым эта конформация не является представленной. Надо понимать, что слабо представленные конформации не регистрируются экспериментальными методами. С другой стороны, при оценке применимости моделирования, надо признать, что результаты эксперимента могут быть неадекватными из-за условий проведения эксперимен-

та. Структуры, полученные методом РСА, могут в определённой степени быть подвержены эффекту кристаллической упаковки [118—120].

Для определения структуры методом ЯМР иногда вводят изменения в последовательность или используют другие химические модификации [121; 122]. Это редко обсуждается, но термодинамика и кинетика небольших систем ДНК может заметно меняться при присоединении меток, например, хромофоров в экспериментах с FRET. Это становится очевидным при рассмотрении чувствительности структурной динамики ДНК к самым тонким изменениям последовательности [123]. Собственно, моделирование может быть использовано для исследования влияния присоединения хромофора на структуру ДНК [124].

Стандартный подход к моделированию может быть представлен следующим образом: создаётся начальная геометрия (обычно берётся из эксперимента, возможно, с изменениями), биополимер погружают в виртуальную ячейку с водой при наложении периодических граничных условий, устанавливают температуру моделирования (с использованием алгоритмов-“термостатов”), а затем позволяют системе двигаться, благодаря тепловым флуктуациям. После окончания моделирования проводят анализ накопленных данных. Таким образом, моделирование может рассматриваться в качестве альтернативного инструмента, в дополнение к экспериментальным методам. Моделирование позволяет задавать конкретные вопросы о поведении молекул и, конечно, отвечать на них. Как было отмечено выше, нужно быть очень консервативным и осторожным при попытке построения модели, не основанной на экспериментальных данных атомарного разрешения, даже если модель строится для части молекулы. В хорошо выполненном вычислительном исследовании всегда должен присутствовать параграф с оценкой ограничений применимости результатов модели-

рования, а также и краткий обзор соответствующей литературы; а результаты не должны быть представлены как результат реального эксперимента.

### **2.3.3 Ограничения моделирования МД с классическими силовыми полями**

Основным преимуществом метода моделирования МД является уникальный уровень детализации всех аспектов эволюции структуры биополимера во времени, с определённым приближением её энергии. В том числе в МД, доступна информация обо всех молекулах воды и ионах. Эту информация невозможно в настоящее время получить из эксперимента. Тем не менее, МД имеет два хорошо известных и значительных ограничения. Во-первых, время наблюдения в моделировании очень мало, по сравнению со временем протекания большинства соответствующих молекулярно-биологических процессов. Например, “расплетание” пары оснований, происходит в микросекундном диапазоне [123; 125; 126], а самосборка структуры G-ДНК происходит в диапазоне времён от микросекунд до нескольких дней [127].

Короткие времена наблюдения подразумевают, что при расчёте доступна лишь малая часть фазового пространства (выборки) молекулы вокруг исходной геометрии. Если исходные структуры не являются представленными в реальной смеси конформеров, то такая выборка может привести к неточным результатам. Ранее уже подчёркивалось, что такой подход является рискованным способом применения моделирования, т.е. моделирование МД предполагаемой гипотетической модели структуры с малой выборкой приведёт к метастабильным конформациям, близким к стартовой геометрии. Во-вторых, увеличение мощности вычислительных средств только подчёркивает неточности, допущенные

при построении силовых полей, хоть они и являются простыми, аналитически-ми функциями, связывающими структуру с потенциальной энергией.



Рисунок 2.11: Схематическое изображение координаты конформационного равновесия. В коротких моделированиях МД выборка будет сделана рядом с исходной структурой. Если это экспериментальная структура, то такая выборка обеспечит точное представление о равновесных флуктуациях. Если применить более длительное время моделирования и/или усовершенствованные методы по сканированию фазового пространства, то можно преодолеть энергетические барьеры перехода либо к непредставленным, либо к представленным структурам. Однако, если непредставленные структуры из-за неточности силового поля имеют более низкую энергию, то переход к представленной структуре становится менее вероятен.

Если силовые поля неточны, то, при наличии достаточной вычислительной мощности, результаты моделирования приведут к непредставленным или неправильным геометриям (рисунок 2.11). Когда временные рамки моделирования были ограничены сотнями пикосекунд, исследователи даже с неточным описанием электростатики и примитивными силовыми полями получали “стабильное” поведение системы [128; 129]. С использованием параллельных вы-

числений в программах для молекулярной динамики стало ясно, что силовые поля нуждаются в улучшении, равно как и в более точном расчёте удалённых электростатических взаимодействий. Обсуждение современных достижений в описании нуклеиновых кислот приведены в части обзора, посвящённой сахаро-фосфатному остову НК.

### **2.3.4 Использование ионов в моделировании G-ДНК**

Структура нуклеиновых кислот, как известно, чувствительна к ионным условиям и должна в моделировании быть чувствительной к концентрации и типу ионов. Однако прямое сравнение ионных условий между экспериментом и моделированием не является осмысленным. Есть существенные различия между тем, как учитываются ионы при моделировании, и как это происходит в эксперименте.

Во-первых, в силовом поле ионы аппроксимируются как сферы с Ван-дер-Ваальсовым радиусом с постоянным точечным зарядом в центре, что отличается от реальности. Во-вторых, современные расчёты МД производятся в маленьких ячейках с растворителем при периодических граничных условиях и с учётом электростатики по методу Эвальда (в реализации PME). Ячейки большого размера в настоящее время не используются, так как они сильно замедляют моделирование. В небольшой ячейке распределение ионов на гранях ячейки не воспроизводит распределение ионов в фазе [130]. Сайты связывания ионов на поверхности биополимера достаточно хорошо описываются силовыми полями, но обмен между разными сайтами может быть искажён из-за малого размера ячейки моделирования. Однако эти эффекты до сих пор не определены количественно. То же самое можно сказать и о возможных артефактах, вызыва-

емых методом РМЕ. ДНК имеет заряд -1 на каждом остатке фосфата, поэтому даже чистая нейтрализация заряда в ячейке средних размеров приводит к концентрации катионов  $\sim 0,2-0,3$  М. Добавление избытка соли, соответственно, увеличивает концентрацию катионов. Формальная концентрация исследуемого вещества тоже велика и составляет  $\sim 0,01$  М, но молекулы данного вещества не имеют контактов друг с другом из-за использования периодических граничных условий. Большинство исследований по моделированию НК имеют короткие времена наблюдения и, по большей части, остаются в окрестности одного конформационного макросостояния, и они менее чувствительны к качеству об-счёта ионов, чем моделирование самосборки структуры.

Большинство работ по моделированию НК, опубликованных до сих пор, было сделано либо с набором нейтрализующих одновалентных катионов (обычно  $\text{Na}^+$  или  $\text{K}^+$ ) или с добавлением  $\sim 100-200$  мМ моновалентной соли для достижения физиологической ионной силы. На сегодня нет сообщений о том, что моделирование МД в нейтрализующих и избыточных солевых условиях приводит к разным конформациям, но есть проблемы для параметров отдельных ионов или для баланса между параметрами разных ионов [131]. Однако, чаще всего, экспериментально найденные сайты связывания ионов хорошо воспроизводятся, как при моделировании каталитического кармана HDV рибозима [59; 132].

На текущий момент при использовании только нейтрализующего количества катионов результаты несильно зависят от описания ионов. На самом деле, ДНК-дуплекс недостаточно чувствителен к соли: на это указывает то, что структура дуплексной ДНК не изменяется при высоких концентрациях соли ( $\sim 5$  М, неопубликованные данные Читмана и коллег). Также не наблюдается существенных различий в моделировании, когда система не содержит мобильных противоионов, и нейтрализация ячейки достигается формальным распределе-

нием компенсаторных зарядов среди всех молекул воды или за счёт применения модификации метода Эвальда с распределением виртуальных зарядов по решётке. Последнее наблюдение несколько удивительно, тем не менее, структура ДНК дуплекса (CGCGAATTCGCG)<sub>2</sub> остаётся нативной при времени наблюдения до 750 нс. Ожидается, что эта структура должна денатурировать при такой неестественно низкой ионной силе. С другой стороны, структура РНК без нейтрализующих катионов сразу теряет стартовую конформацию и денатурирует при моделировании МД (Читман, неопубликованные данные). На сегодняшний день все ещё отсутствует общее мнение на том, как хорошо модель точечного заряда иона может воспроизводить влияние соли на структуру и динамику НК.

Очевидно, что G-ДНК является особенной системой, где ДНК-ионные взаимодействия имеют решающее значение, так как структура содержит ионы в центральной полости стебля. Есть несколько известных проблем в изучении молекул G-ДНК, которые связаны с приближениями при описании ионов. Поведение ионов при моделировании (а также данные КМ-вычислений) указывает на то, что радиус ионов является слишком большим, чтобы оптимально описывать взаимодействия ионов внутри стеблей [133]. Это уменьшает вероятность связывания Na<sup>+</sup> в плоскости квартета и может привести к случайному выталкиванию K<sup>+</sup> из полости внутри стебля. Было предположено, что сокращение радиуса K<sup>+</sup>, который находится внутри стебля, может быть использовано в моделировании G-ДНК. Также вероятно, что комплексообразование катионов несколько недооценено в современных силовых полях. Все эти вопросы могут быть решены при сравнении расчётов в силовых полях и данных КМ о взаимодействиях между катионами и гауниновыми квартетами [133].

Одной из явных проблем является потеря катионом экспериментально установленного сайта связывания в структуре, состоящей из стебля и петли в G-

ДНК  $(G_4T_4G_4)_2$ . Это, пожалуй, единственный задокументированный случай, когда экспериментально установленный сайт связывания одновалентного катиона не очень хорошо воспроизводится в моделировании. Многие катион-связывающие сайты в РНК были правильно воспроизведены или даже предсказаны в моделировании МД. Считается, что, в целом, описание поведения ионов во внутренней полости G-ДНК происходит адекватно. Самым ярким артефактом является недавно обнаруженная кластеризация (кристаллизация соли) катионов и анионов при определённой неподходящей комбинации параметров катионов и анионов. Этого можно избежать, если использовать новейшие параметры для описания ионов [134]. Эта проблема никогда не возникала при моделировании, в котором катионы использовались только для нейтрализации, и для такого моделирования можно использовать традиционные параметры. Недавние работы показали, что результаты моделирования РНК в определённой степени зависят от модели воды, но для G-ДНК подобных данных ещё нет.

### **2.3.5 Исследования взаимодействий G-ДНК с лигандами**

Квадруплексные структуры встречаются в промотерах онкогенов, а также в теломерах. Малые молекулы могут стабилизировать или изменять структуру G-квадруплекса, влияя на его функцию, поэтому молекулы, связывающие G-квадруплекс, активно исследуются как потенциальные противораковые агенты. Ряд недавних публикаций подчёркивает роль G-ДНК в качестве мишени для терапии рака [135; 136]. Активное внимание исследователей к комплексам квадруплекс-лиганд привело к появлению специальных выпусков журналов, посвящённых этой теме [137]. Компьютерный дизайн и/или дизайн на основе структуры активно применяется для поиска и оптимизации лигандов, способ-



ных специфически взаимодействовать с заданными G-квадруплексными мишенями. Эти подходы включают в себя как сравнительное моделирование и структурный анализ [138; 139], так и более детальные исследования с применением докинга [140—142], моделирование молекулярной динамики и моделирование с вычислением изменения свободной энергии [143—157]

В последних обзорах подробно обсуждают использование методов моделирования как G-ДНК, так и G-РНК [158; 159]. Методы моделирования могут зачастую улучшить наше понимание принципов взаимодействия квадруплекса и его лиганда. Например, многие лиганды G-ДНК эффективно взаимодействуют с квартетом за счёт стэкинга, однако такие лиганды неспецифичны. Методы моделирования могут помочь в разработке молекул, специфичных к последовательности и структуре, используя различия в структуре бороздок квадруплекса [160]. Общей чертой подобных исследований является использование силовых полей семейства AMBER, в том числе, их модификации “GAFF” для описания лигандов.

При оценке достоверности вычислений для комплексов G-ДНК с лигандами всегда надо иметь в виду несколько вопросов: 1) качество исходной структуры (способ расчёта очевиден, когда доступна надёжная структура, полученная методом РСА, в то время, как результаты докинга могут привести к нереалистичной картине); 2) качество описания силового поля для лиганда; 3) каковы ожидаемые цели? (в то время, как качественные структурные данные вполне достижимы, а качественные расчёты свободной энергии, скорее всего, нет).

Выше была показана необходимость детальной настройки силовых полей для НК для предотвращения структурной деградации нативных структур ДНК и РНК. Качественная параметризация силовых полей для лигандов также может быть весьма непростой задачей. Есть преимущество в том, что в данном

случае исследователь имеет дело с небольшой молекулой, поэтому небольшие ошибки в параметризации мономера не накапливаются, как это происходит с силовыми полями биополимеров. Относительно удобно параметризовать плоские ароматические системы, что справедливо для изрядной части G-ДНК-связывающих лигандов. С другой стороны, параметризация надёжного силового поля для сложных лигандов, таких, как аминогликозиды, будет затруднено, если вообще возможно.

Поляризация составляющие взаимодействия важны для многозарядных лигандов. Так, водородная связь с участием атома галогена не может быть точно описана моделью точечного заряда из-за анизотропной плотности заряда на нём [161; 162], и это только некоторые проблемы в параметризации лигандов. Параметры для описания Ван-дер-Ваальсовых взаимодействий можно позаимствовать из силовых полей для биополимеров. Параметры для описания длины связей и углов слабо связаны с нековалентными взаимодействиями и могут быть получены из экспериментальных данных и высококачественных вычислений КМ. Заряды для атомов могут быть получены, исходя из рассчитанной для лиганда поверхности электростатического потенциала (RESP), на уровне теории КМ HF/6-31G\* [163; 164]. Вычисление частичных зарядов для атомов может быть затруднено для сложных молекул. И, в итоге, исследователи вновь сталкиваются со сложной задачей по описанию торсионных углов. Единственная возможность провести строгое описание торсионных углов – это затратные высококачественные КМ-вычисления.

При моделировании, где биополимер описывается в силовом поле семейства “Amber”, принято использовать специализированное для описания малых молекул силовое поле GAFF [165]. GAFF является расширением силового поля “Amber” с параметрами для широкого спектра органических молекул: в ос-

новном, для молекул, схожих с известными лекарствами. Оригинальная версия этого поля использует 33 основных и 22 специальных типа атомов, включая атомы H, C, N, O, S, P, F, Cl, Br и I, таким образом, покрывая основную часть химического разнообразия. Был разработан алгоритм для автоматизированной процедуры для соотнесения типа атома, заряда и параметров силового поля для широкого круга органических соединений [166]. Эта удобная процедура ни в коем случае не заменяет подробную ручную работу по параметризации соединения. Для точного моделирования можно использовать такие параметры в качестве отправной точки и тщательно проверять их с помощью точных КМ-вычислений, а при необходимости – уточнять. Так как органические молекулы сложны, а силовые поля являются по природе эмпирическими приближениями, то кажется нереалистичным создание автоматизированной процедуры для качественной параметризации органических соединений [95]. Следует также отметить, что расчёты КМ, которые использовались при построении силового поля GAFF в 2004 году, сегодня уже не удовлетворяют условию высококачественного расчёта.

Вопрос параметризации также справедлив для моделирования химических модификаций в составе нуклеиновых кислот как в азотистых основаниях, так и в сахаро-фосфатном остове. Более подробно этот вопрос рассмотрен в главе, посвящённой остову НК.

Многие исследователи пытаются вычислить изменение свободной энергии при образовании комплекса G-ДНК-лиганд, и для этого чаще всего применяется популярный подход ММ-PBSA. Этот подход использует значительные упрощения (используется растворитель в неявном виде), и маловероятно, что заметное улучшение точности данного подхода будет достигнуто в обозримом будущем. Одним из основных ограничений всех методов ММ и КМ, которые используют

неявное описание растворителя, – это возникновение ошибки при оценке энергии сольватации, которая зависит от атомных радиусов. Атомные радиусы – это эмпирически подобранные параметры, которые не соответствуют “реальным” физическим свойствам. Неточность определения энергии сольватации напрямую вносит ошибку в расчёты изменения свободной энергии связывания. Лучшие результаты при использовании неявного описания растворителей были достигнуты, когда параметры оптимизировали для определённого класса молекул и молекулярных комплексов при наличии экспериментальных данных. В качестве примера можно привести исследования Ли и коллег. Они сообщили, что, согласно методу PBSA, изменение свободной энергии при связывании G-ДНК с перилен-производным (Tel03) составляет  $\sim -35$  ккал/моль [146]. Это значение сильно отличается от экспериментального для связывания подобных соединений ( $\sim -10$  ккал/моль). Это указывает на то, что метод MM-PBSA может задавать правильные тенденции, но вычисленные абсолютные значения изменения свободной энергии часто завышены. В другой работе Аргвала и коллег были рассчитаны восемь различных комплексов G-ДНК с телометсатином, и изменение свободных энергий составило от -1 до -16 ккал/моль [151]. Такой диапазон значений указывает на неприемлемый уровень ошибки в результатах расчёта.

### **2.3.6 Перспективы применения моделирования к НК**

Как долго надо наблюдать за НК в моделировании МД, чтобы получить биохимически осмысленные данные? Первые публикации сообщали о диапазоне времён наблюдения 0,5-5 нс. Даже такие короткие траектории моделирования могут иногда предоставить полезные данные. К примеру, в приведённой работе была показана высокая жёсткость квадруплексного тяжа. Тем не менее, с

ростом производительности компьютеров и накоплением опыта, предполагается, что минимальная длина отдельных траекторий должна быть порядка 25-50 нс. В случае, когда исходная структура уже достаточно релаксирована, моделирование может правильно отражать многие её структурные особенности. Чтобы увидеть большую амплитуду движений и найти разные микросостояния, которые разделены небольшими энергетическими барьерами, продолжительность наблюдения должна быть в диапазоне времён 100-1000 нс для каждой траектории. Принято считать, что моделирование со временем наблюдения 1 нс может позволить преодолеть энергетические барьеры порядка 5 ккал/моль, а при времени наблюдения 100 нс преодолеваются барьеры не более 8 ккал/моль. Поскольку разные системы имеют разную сложность конформационного пространства, как с точки зрения термодинамики, так и с точки зрения кинетики, то нельзя заведомо сказать, сколько времени необходимо наблюдать за эволюцией системы. Для каждого конкретного случая решение принимается индивидуально. Для определённых целей множественное (несколько траекторий) моделирование с длиной траектории 25 нс является достаточным, в то время, как для других систем микросекундный масштаб времени будет недостаточен. Хорошим советом является предложение работать на пределе современных вычислительных возможностей. Чем дольше наблюдения, тем выше вероятность того, что интересные события из динамики системы не будут пропущены. Очевидно также, что, чем дольше происходит наблюдение за системой, тем более вероятно, что исследователи будут сталкиваться с несовершенством используемых силовых полей.

Лучше иметь набор нескольких (скажем, 5-10) траекторий с достаточным временем наблюдения, чем полагаться только на одну очень длинную траекторию. Хотя моделирование МД основано на детерминистическом описании,

оно в действительности имеет стохастическую природу. Опыт показывает, что с увеличением длины моделирования вероятность наблюдения структурного изменения может уменьшаться. Единственная траектория в самом начале может быть испорчена редким необратимым событием, что приводит к потенциально неверной интерпретации результатов. В то время, как множественные повторения приведут к накоплению ценных результатов. Например, при исследовании роли мутации в локальном окружении можно обнаружить статистически достоверные данные об изменении свойств в удалённой части молекулы [167].

Очевидно, что даже после самых современных улучшений силовые поля остаются далеки от совершенства. Вполне вероятно, что увеличение выборки поможет в будущем найти дополнительные артефакты в силовых полях, которые на данный момент скрыты из-за больших значений энергетического барьера между стартовой “нативной” конформацией и потенциально артефактным состоянием. Надо быть особенно осторожным с методами моделирования, которые ускоряют накопление данных, а именно: метод обмена репликами, метадинамика и направленная МД. Такие способы, с одной стороны, крайне полезны. С другой стороны, их применение не является тривиальным. Во-первых, эти методы всегда основаны на значительных приближениях, в дополнение к тем, которые используются в стандартном моделировании МД. Направленная МД применима к простым путям, таким, как переход из А-формы в В-форму ДНК, но более сложные изменения могут легко выйти за пределы применимости метода [167]. Приближения, основанные на насыщении энергией так называемых коллективных переменных, не всегда в полной мере признаются сообществом исследователей. Для неподготовленного исследователя сложной задачей является в полной мере понять значение таких расчётов, если авторы публикаций не сформулировали это явно и точно [168]. Всегда существует некоторая

погрешность, которая возникает при попытке решить проблему выборки. Тем не менее, это очень ценные методы, но они не смогут заменить обычное моделирование МД, которое остаётся золотым стандартом.

Несмотря на имеющиеся ограничения, точность хорошо откалиброванных силовых полей для нуклеиновых кислот достаточно высока, несмотря на их простоту. Учитывая трудности при разработке поляризуемых силовых полей и высокие требования достаточно точного описания в КМ, вероятно, что современные силовые поля будут доминировать в исследованиях по моделированию НК в течение следующего десятилетия. Существуют дополнительные возможности для дальнейшего улучшения современных силовых полей, и было бы справедливо признать, что ограниченная гибкость и нефизический характер описания НК в простой функциональной форме, определённо не позволяют создать идеальное силовое поле. Таким образом, рано или поздно возможности для улучшения силового поля через торсионные углы будут исчерпаны.

КМ-вычисления могут в будущем дополнить атомистическое моделирование, хотя они всегда будут отставать в производительности от ММ-моделирования, что сказывается на размере систем и широте выборки. Довольно простой и распространённый подход к моделированию биополимеров – это расчёт структуры методами КМ в газовой фазе, как это было сделано для оценки энергии взаимодействия G-квартетов с ионами [169—171]. Хотя такие расчёты должны быть значительно более точными, чем данные из моделирования МД, трудно экстраполировать эти результаты на полные и сольватированные системы нуклеиновых кислот. Для этого необходимо учитывать всю молекулу НК и её окружение в КМ-расчётах, а это возможно при использовании современных подходов к гибриднему квантово-механическому/молекулярно-механическому моделированию.

## Глава 3

# Результаты и их обсуждение

### 3.1 Структурные аспекты взаимодействия тм-РНК с рибосомой

Основной целью данной работы был поиск подходов к пониманию механистических особенностей функционирования нуклеиновых кислот и определения требований к динамике структуры нуклеиновой кислоты при её функционировании. Наиболее известной и распространённой органеллой клетки, где нуклеиновая кислота играет ключевую роль, является рибосома. Рибосома – это РНК-белковый ансамбль, в структурной организации и функционировании которого РНК играют ключевую роль – при этом зачастую в ходе белкового синтеза, выполняемого рибосомой, ассоциированные с ней РНК меняют свою конформацию. Самым ярким примером подобной РНК считается тмРНК. Предположительно, эта РНК не только изменяет взаимное расположение своих доменов в ходе выполнения предписанной ей функции, но вообще теряет некоторые элементы вторичной структуры. Рибосома – это гигантский рибозим [172], который способен функционально взаимодействовать с широким кругом соедине-



ний – от низкомолекулярных антибиотиков до крупных биополимеров (мРНК, тРНК, белковые факторы и т.д.). Относительно недавно открыт механизм транс-трансляции. Это сложный, хорошо организованный процесс, который переключает синтез полипептидной цепи с мРНК на специальный участок (MLD - домен тмРНК сходный с мРНК) специализированной транспортно информационной РНК (тмРНК) [173; 174]. Бактериальные клетки используют транс-трансляцию для возобновления активности рибосом после ареста, который происходит при трансляции с повреждённой мРНК, направляя повреждённую мРНК и продукт её трансляции в системы деградации [175]. Во время своего функционирования тмРНК взаимодействует с несколькими белками и рибосомой.

Следует сказать, что на момент начала запланированной работы структурные данные по этим комплексам были представлены только структурой псевдоузлов, входящих в тмРНК [176], и структурой комплекса белка SmpB с фрагментом тРНК-подобной части тмРНК [177]. К тому моменту имелись лишь предварительные сведения о строении инициаторного комплекса “тмРНК-SmpB-70S рибосома” [178; 179]. Относительно недавно Мур и Сайер (Moore & Sauer) [180] предложили гипотетический механизм, с помощью которого тмРНК может различать активные рибосомы от неактивных. Однако особенности строения тмРНК в фазах элонгации и терминации не рассматривались ими. Ранее в лаборатории Донцовой О.А. разработан метод для выделения комплексов тмРНК с рибосомой на разных этапах транс-трансляции [181]. Остановка транс-трансляции достигалась расположением сигнала терминации трансляции в соответствующем месте тмРНК и инактивации фактора терминации, или высвобождения, RF2. Используя этот подход, Бугаева и коллеги с помощью метода химического зондирования пробинга и криоэлектронной спектроскопии исследовали структуру тмРНК. С точки зрения данной работы особенный инте-

рес вызывают значимые структурные изменения, которые претерпевает тмРНК в ходе своего функционирования в комплексе с рибосомой. Используя подходы молекулярного моделирования и накопленные экспериментальные данные, мы разработали метод, который позволил представить расположение и приближительную структурную организацию тмРНК в комплексе с рибосомой на разных этапах транс-трансляции.

### **3.1.1 Разработка подхода к моделированию структурной организации комплексов крупных нуклеиновых кислот**

При поиске решения поставленной задачи был проведён анализ возможностей современных инструментов для моделирования третичной структуры РНК. Точная оценка возможностей и ограничений современных подходов необходимая ступень для выбора оптимального подхода к моделированию структуры тмРНК в комплексе с рибосомой. За последние годы разработано несколько подходов для предсказания третичной структуры РНК. Большинство из них основано на экспериментальных данных, а также на имеющейся информации о вторичной структуре РНК. Нередко используется гипотетическая или предсказанная вторичная структура РНК, что в случае ошибки приводит к совершенно неверному предсказанию третичной структуры этой макромолекулы. Число известных трехмерных структур РНК, организованных в базы данных PDB [182], SCOR и RNABase [183], постоянно растет, что позволяет проверять модели трехмерной структуры РНК и повышать их качество. Используя приёмы гомологичного моделирования, можно использовать известную структурную информацию для соответствующего участка в составе крупной НК неизвестной структуры.

На сегодняшний день только один из методов моделирования, iFoldRNA, предполагает автоматизированный процесс предсказания трехмерной структуры РНК [184]; однако серьезное ограничение iFoldRNA – размер молекулы РНК не превышает 150 остатков, а это существенно меньше интересующего нас объекта.

Необходимо отметить программный пакет ERNA-3D, который может создавать трехмерное представление структуры РНК исходя из известной вторичной структуры [185]. Эта программа, исходя из спиральных участков вторичной структуры РНК, автоматически создает спиральные участки ее третичной структуры в виде А-формы двойной спирали. В первой версии программы одноцепочечные участки моделировались с помощью итерационных поворотов сахарофосфатного остова. В современной версии одноцепочечные участки и мотивы моделируются по гомологии с известными структурами высокого разрешения базы SCOR. Если для одноцепочечной последовательности нет гомолога в базе SCOR, то эти нуклеотиды могут быть расставлены в пространстве вручную.

Сравнительный анализ последовательностей в программе ERNA-3D был использован для определения расположения трехмерного остова маленького домена сигнальной молекулы РНК, включающей псевдоузел. Программа ERNA-3D а использована и для моделирования тмРНК *Escherichia coli*, *Bacillus anthracis* и *Caulobacter crescentus* [186]. Полученные модели обладают всеми необходимыми функционально-структурными свойствами, но не дают представления о расположении тмРНК на рибосоме. MANIP – программа, собирающая известные фрагменты РНК в сложную трехмерную архитектуру [187]. Для ее работы необходима база данных, содержащая трехмерные фрагменты известных структур. Алгоритм автоматически определяет и показывает возможные

водородные связи между нуклеотидами. Программу MANIP использовали при моделировании структуры РНКазы Р [188].

Каркас программной системы Nucleic Acid Builder (NAB) описывает нуклеиновые кислоты как иерархическую структуру и может быть использован для конструирования как двухцепочечных, так и одноцепочечных РНК – вплоть до нескольких сотен нуклеотидов [189]. На основе данных по специфическим взаимодействиям и по известным трехмерным структурам можно определить набор пространственных ограничений для любых структурных элементов, например, для псевдоузлов. Полученные модели могут быть улучшены с помощью встроенных инструментов оптимизации геометрии молекулы и моделирования молекулярной динамики.

Программа MC-Sym использует символьные вычисления и численные значения для построения трехмерных структур РНК. Основу этой программы составляют структурные данные, которые записаны как набор символов представлений домена. Численные значения используются для улучшения символьной модели [190]. С помощью этого подхода изучена специфичность рибосвитча SAM-I к SAM [191].

В общем случае – чтобы привести структурные данные в соответствие с экспериментальными результатами – все модели, полученные с помощью описанных выше программ, требуют дополнительной проверки и оптимизации вручную. Окончательные модели обычно оптимизируют как по геометрии, так и по молекулярной динамике. Резюмируя краткое рассмотрение известных подходов к моделированию структуры РНК, необходимо отметить, что все они имеют серьезные ограничения, связанные с размером изучаемых РНК, а также сложностями, возникающими при моделировании комплексов РНК-РНК и РНК-белок. Первые попытки проводить моделирование комплексов рибосомы

с тмРНК с использованием полноатомного описания всех участников процесса выявили неэффективность доступных вычислительных ресурсов и алгоритмов. Это ограничение было преодолено при использовании сильно упрощённого описания нуклеотида. Вместо описания всех атомов и их возможных попарных взаимодействий мы решили описывать нуклеотид одной частицей. Центр этой частицы приходится на положение атома фосфора. Для описания хода сахарофосфатного остова используется гармонический осциллятор с равновесным значением длины связи  $5.6\text{\AA}$ , а заряд частицы составляет  $-1$ . Очевидно, что такой набор параметров не описывает ни водородные связи, ни стэкинг-взаимодействия. Для восполнения потерянной при упрощении информации мы использовали метод дистанционных ограничений. Этот подход позволяет сохранить относительное взаимное расположение остатков, используя набор попарных ограничений на расстояние между частицами. В основу метода дистанционных ограничений заложена информация по вторичной структуре РНК, из которой достаточно однозначно определяется большая часть пространственных ограничений. В первую очередь это касается областей стебля вторичной структуры РНК, которые в нативных условиях образуют А-форму двойной спирали. Это также распространяется на участки последовательности, для которых известна третичная структура – для тмРНК это псевдоузлы рК1, рК2, рК3 и рК4. Что касается петель и выпетливаний, то, в случае отсутствия взаимодействий с их участием, это достаточно подвижные элементы третичной структуры, что явно следует из результатов ЯМР-исследований. В случае моделирования взаимодействия тмРНК с рибосомой их конформация, вероятно, будет определяться взаимным расположением элементов структуры, для которых можно построить дистанционные ограничения, а именно: стебли и псевдоузлы. Важной особенностью метода дистанционных ограничений считается

возможность производить моделирование петель с использованием известных структур гомологов, имеющихся в базе данных PDB, и эта функция была нами использована для сохранения геометрии места контакта тмРНК с белком SmpB. Способ описания молекул с помощью упрощенной (“cores grain”) модели в первую очередь уменьшает необходимое время расчетов, так как длительность моделирования молекулярной механики квадратично зависит от количества атомов в системе. Для моделирования всего нуклеотида мы используем лишь один атом фосфора, что дает выигрыш во времени расчета примерно в 1500 раз. Кроме того, подобное упрощение позволяет быстро выявлять ошибки подхода и оптимизировать алгоритм, при этом число пространственных ограничений уменьшается тоже, как минимум, в 1500 раз.

Таким образом, после выбора подхода моделирования необходимо было выбрать описание для рибосомы. Здесь мы решили тоже упростить описание как нуклеотидов, так и аминокислотных остатков. Так как из экспериментальных данных следует, что при связывании тмРНК конформация рибосомы не изменяется, мы поставили запрет на движение частиц, описывающих рибосому. Как и в случае с тмРНК, остатки РНК и белков в рибосоме описывались одной частицей с зарядом, соответствующим остатку, а для моделирования примерной формы поверхности использовали модифицированные параметры потенциала Ленарда-Джонса (нуклеотид:  $\sigma=16\text{\AA}$ ,  $\epsilon=5.000$  кДж/моль; аминокислота:  $\sigma=8\text{\AA}$ ,  $\epsilon=1.000$  кДж/моль), эти параметры примерно описывают средний радиус нуклеотида и аминокислоты.

### **3.1.2 Моделирование конформаций тмРНК в комплексах с рибосомой**

В экспериментальной части работы, которая выполнена Бугаевой Е.Ю и коллегами были выделены комплексы рибосомы с тмРНК. В работе использован набор тмРНК, которые имели такие различия в первичной структуре, которые позволяли выделять комплексы тмРНК-рибосома, соответствующие разным этапам трансляции открытой рамки считывания тмРНК. Данные, суммирующие описание как вторичной структуры тмРНК, модификаций тмРНК, так и результаты химической модификации тмРНК в комплексах с рибосомой, приведены на Рисунке 3.1. Как видно из Рисунка 3.1, вторичная структура тмРНК представлена набором псевдоузлов и протяжённых стеблей с внутренними петлями. Для стеблей мы создавали дистанционные ограничения на основе известных данных для двухтяжевой РНК. При описании псевдоузлов использованы модели структур индивидуальных псевдоузлов, полученные на основании данных криоэлектронной микроскопии [192] или ЯМР-анализа [176]. Исходя из результатов по химической модификации, можно предполагать, что все стабильные элементы вторичной структуры сохраняются в исследуемых комплексах.)

Для структурной интерпретации полученных экспериментальных данных по конформации тмРНК в рибосоме на разных этапах трансляции мы провели моделирование предложенным выше способом. Для описания структуры рибосомы использованы координаты атомов, полученные на основании тех данных рентгено-структурного анализа (РСА), где указан путь мРНК в рибосоме [193]. Мы предположили, что кодируемая тмРНК открытая рамка считывания должна располагаться там же, где находится мРНК. Моделью тмРНК-2 мы назвали такое состояние тмРНК когда: стоп-кодон находится в А-сайте рибосо-

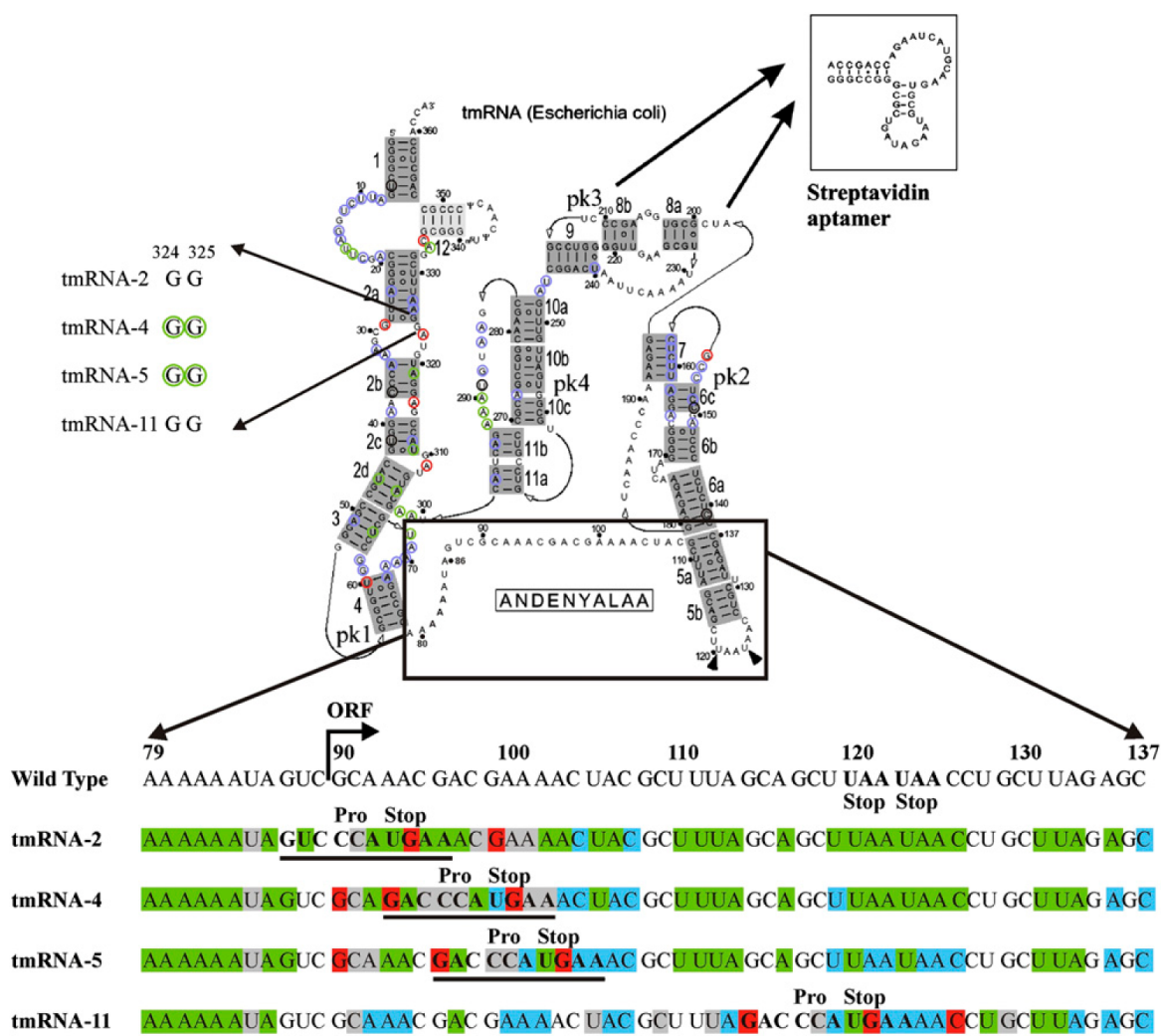


Рисунок 3.1: Описание защит остатков тмРНК в комплексах с рибосомой.

Вторичная структура тмРНК построена по данным [rnp.uthct.edu/rnp/tmRDB/tmRDB.html](http://rnp.uthct.edu/rnp/tmRDB/tmRDB.html). Псевдоузел pk3 (нуклеотиды U212-A239) заменен на аптамер к стрептавидину. Общие эффекты от связывания с рибосомой для всех комплексов отмечены зелёным кругом (защита) и красным кругом (экспонирование). Синим кругом отмечены остатки, на которые образование комплекса не влияло. Чёрными кругами отмечены остатки, склонные к деградации. Нуклеотиды G324 и G325, доступность которых изменяется в комплексах, показаны слева. Область A79-C137 детально представлена внизу рисунка.

мы, а предыдущий кодон вместе с тРНК занимает Р-сайт. В этом случае часть тмРНК должна находиться в Е-сайте рибосомы и при этом сохранять структуру, подобную акцепторному стеблю тРНК. Положение стебля L1 заимствовано из работы Harms и коллег [194]. В ходе оптимизации геометрии было принято, что только тРНК- и мРНК-подобные части тмРНК неподвижны в пространстве, а



все остальные области располагаются в пространстве произвольно. В стартовой структуре мы расположили остатки тмРНК по кругу, концы которого замыкались на неподвижных элементах. Формирование известных структурных элементов вторичной структуры происходило в ходе оптимизации геометрии под действием дистанционных ограничений (Рисунок 3.2).

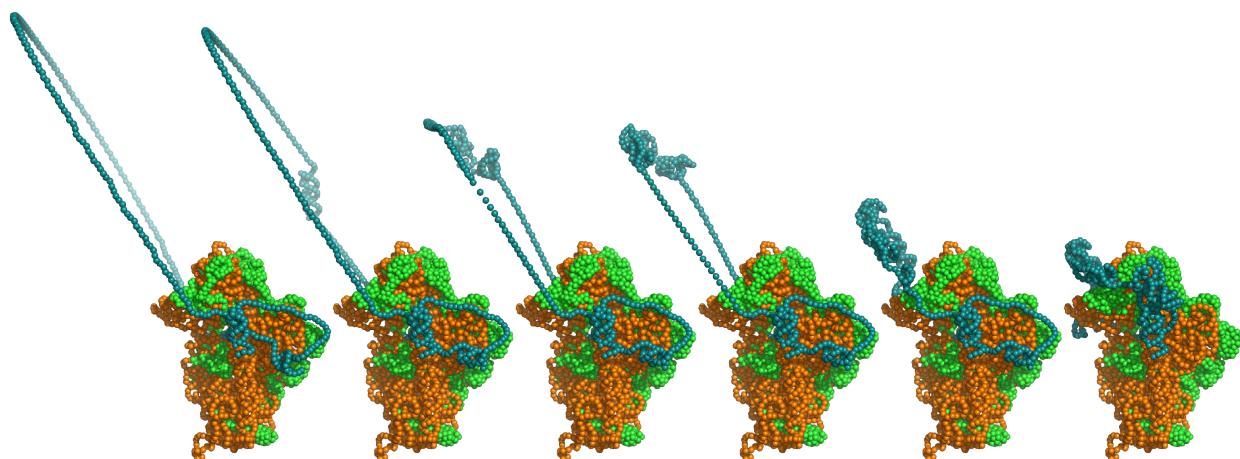


Рисунок 3.2: Этапы сборки структурных элементов тмРНК вокруг 70S рибосомы. Для лучшего восприятия 50S субчастица скрыта от наблюдателя. Тёмно-бирюзовым цветом отмечены нуклеотиды моделируемой тмРНК, оранжевым – рибосомной РНК, зелёным – аминокислотные остатки белков.

Полученная в результате моделирования структура тмРНК-2 внутри рибосомы согласуется с данными по химической модификации, полученными как Бугаевой Е.Ю. с коллегами, так и другими авторами (Рисунок 3.3). Белок SmpV занимает свое положение на тРНК-подобной части тмРНК (TLD) – как это показано Гутманом и др. на основании данных PCA [177]. Область TLD тмРНК находится в E-сайте рибосомы. Защищенные остатки A79–A86 расположены рядом со входом в мРНК-связывающий сайт. Псевдоузел рК1 расположен рядом с E-сайтом в месте высвобождения тРНК. Оказалось, что между рибосомными субчастицами, со стороны стебля белка L1, недостаточно места для одновременного размещения TLD вместе с белком SmpV и псевдоузлом рК1. Стебли 2a, 2b и 2c поддерживают расположение домена TLD в E-сайте рибосомы. Стебель

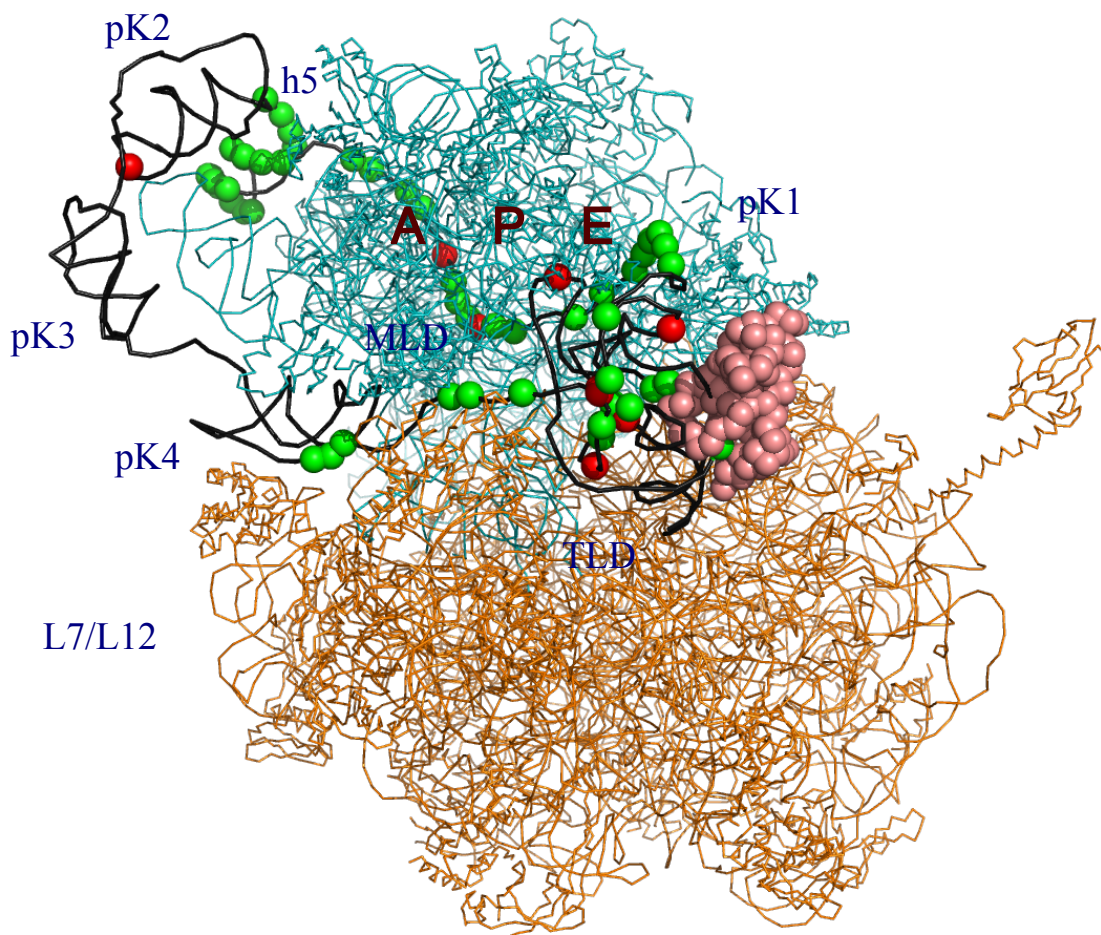


Рисунок 3.3: Модель структуры тмРНК-2 в комплексе с рибосомой (вид сверху). 30S субчастица окрашена синим, 50S субчастица – зелёным. тмРНК обозначена чёрной линией, которая отражает ход сахарофосфатного остова. Белок SmpV показан розовыми сферами. Нуклеотиды тмРНК, которые защищены рибосомой от химической модификации, отмечены зелёными сферами. Остатки, доступность которых увеличивается при образовании комплекса, показаны красными сферами. Структурные элементы тмРНК подписаны текстом.

2d образует связку между псевдоузлом pK1 и аркой, которая опоясывает область головы 30S субчастицы, и располагается на плече. Арка образована псевдоузлами pK4, pK2 и pK3. Стебель 5 расположен рядом со входом в тоннель для мРНК и может легко расплетаться при продолжении движения рибосомы по открытой рамке считывания мРНК-подобной части тмРНК. При расплетании стебля 5 образующегося избытка длины тмРНК достаточно для того, чтобы арка не смещалась со своей позиции на плече 30S субчастицы рибосомы. Кодон

“возобновления” находится в Р-сайте рибосомы и второй кодон в рамке считывания в А-сайте. Из полученной модели видно, что все остатки одноцепочечных участков тмРНК взаимодействуют с рибосомой и расположение остатков с увеличенной доступностью для химической атаки соответствует возможным нарушениям цепи РНК.

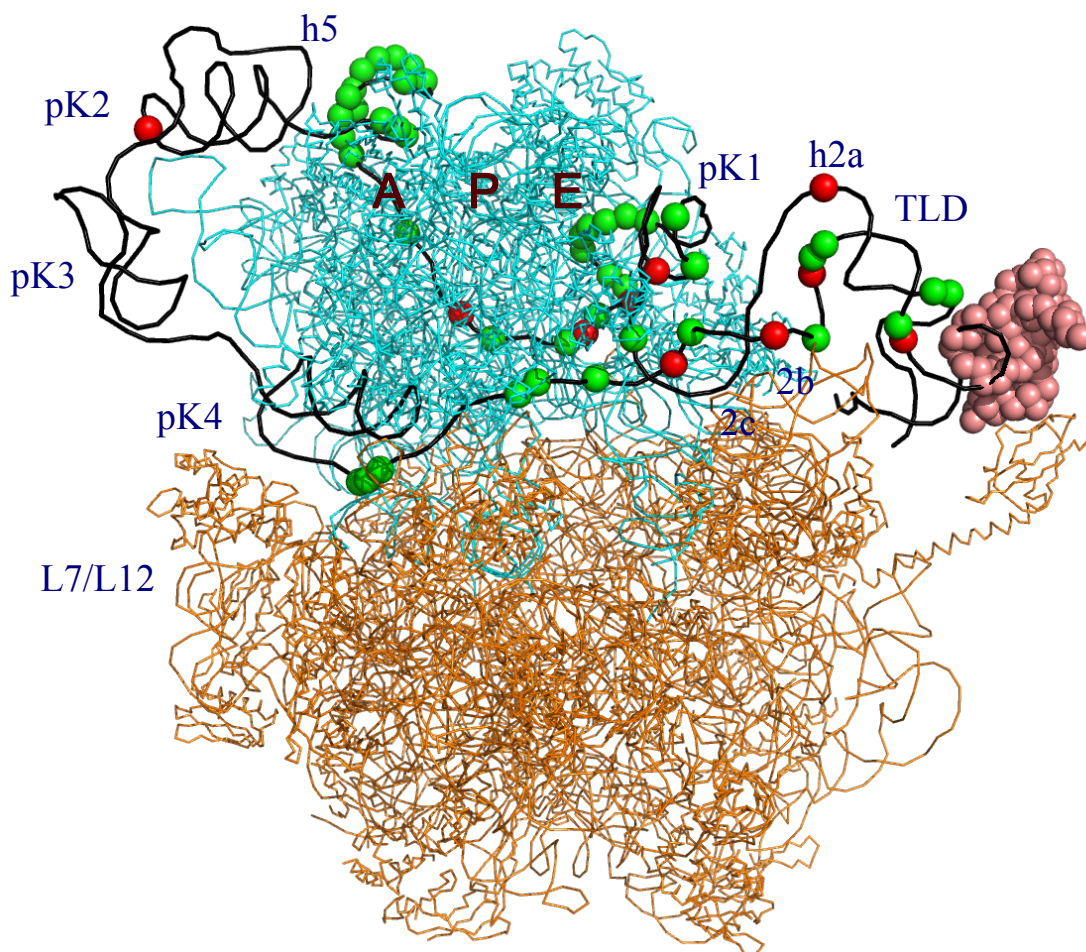


Рисунок 3.4: Модель структуры тмРНК-4 в комплексе с рибосомой (вид сверху). 30S субчастица окрашена синим, 50S субчастица – зелёным. тмРНК показана чёрной линией, которая отражает ход остова. Белок SmpV показан розовыми сферами. Нуклеотиды тмРНК, которые защищены рибосомой от химической модификации, помечены зелёными сферами. Остатки, доступность которых увеличивается при образовании комплекса, показаны красными сферами. Структурные элементы тмРНК подписаны текстом.

Таким же способом нами получена модель комплекса тмРНК-4 с рибосомой (Рисунок 3.4). Как и в случае модели тмРНК-2, можно говорить о полном соответствии полученной модели с данными химической модификации и криоэлект-

кронной микроскопии. На этом этапе транс-трансляции область TLD тмРНК вместе с белком SmpB перемещается, покидает зону трансляции и располагается на “платформе” 30S субчастицы. Этот домен тмРНК не зафиксирован на поверхности рибосомы и может занимать практически любую позицию, не пересекаясь с деацелированной тРНК, которая покидает рибосому через E-сайт. Стебли 2a, 2b и 2c, с одной стороны, соединяют TLD и псевдоузел рК4, который является частью арки, а с другой стороны, петля A79–A86 поддерживается стеблем 2d. Этот стебель, благодаря дегидратации, может быть легко расплетён при движении рибосомы, оставаясь при этом недоступным для химической модификации – как это выявлено при сравнении паттернов модификации этого участка в комплексе с рибосомой и в свободном состоянии. Защищённая петля A79–A86 прочно удерживается в мРНК-связывающем сайте рибосомы [193], что приводит к изменению конформации первого нуклеотида в кодоне возобновления и как результат к увеличению его доступности для химической модификации. Арка, которая может быть визуализирована с помощью криоэлектронной спектроскопии состоит из трёх псевдоузлов. Как и в случае тмРНК-2, эта арка окружает голову 30S субчастицы, начиная с плеча. Модель предполагает движение арки вокруг 30S субчастицы, что согласуется с данными криоэлектронной микроскопии (Рисунок 3.5).

Спираль 5 расположена рядом со входом в тоннель для мРНК. Третий и четвертые кодоны открытой рамки считывания тмРНК расположены соответственно в P- и A-сайтах, что согласуется с данными РСА [193]. Как и ранее, все защищенные в комплексе остатки тмРНК, находящиеся в одноцепочечных областях, вовлечены во взаимодействие с рибосомой. Расположение остатков, доступность которых увеличивается при формировании комплекса, соответствует возможным нарушениям РНК-цепи. В разработанной нами модели принято во

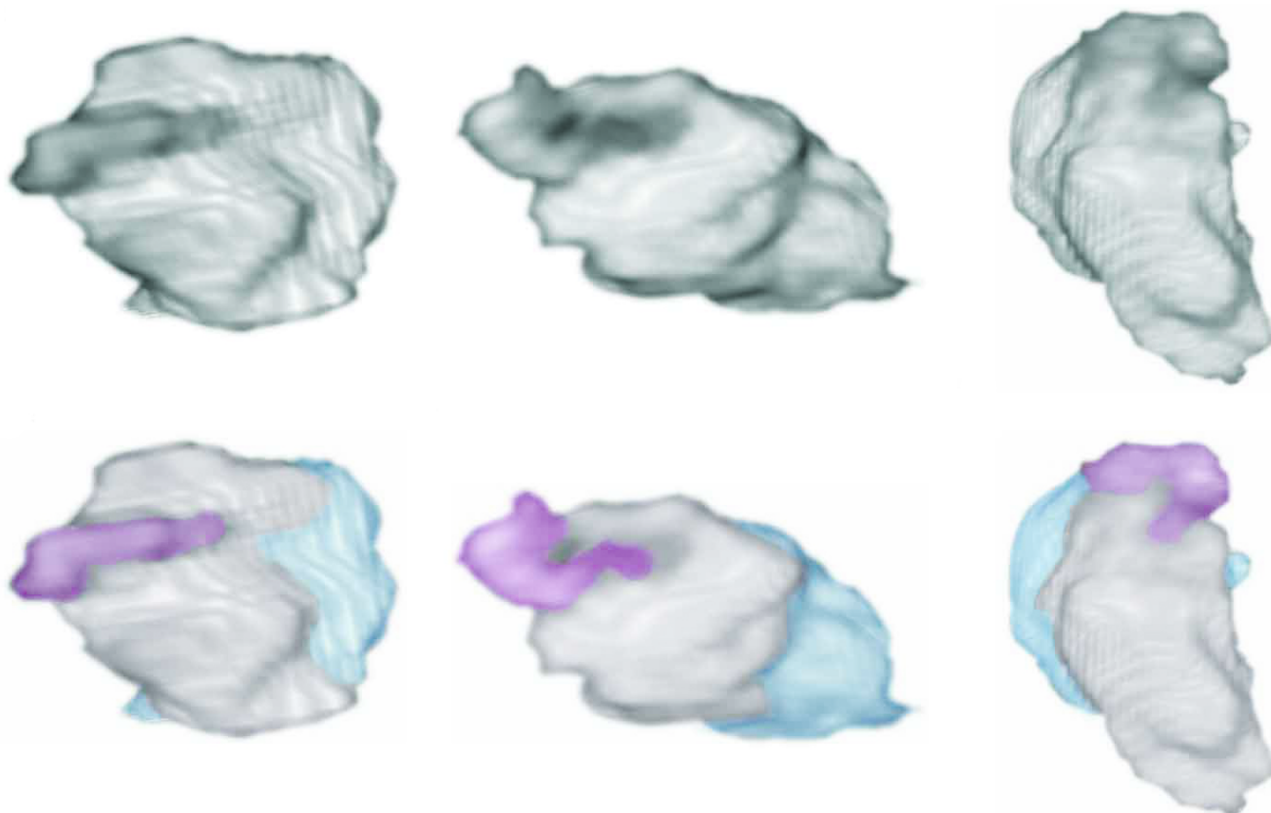


Рисунок 3.5: Трёхмерная реконструкция комплекса тмРНК-4 с рибосомой на основе данных криоэлектронной микроскопии. Три индивидуальные частицы с разным положением арки показаны со стороны 30S субчастицы.

внимание участие белка SmpV на всех этапах функционирования тмРНК – в соответствии с экспериментальными данными, полученными ранее [175; 195; 196]. Однако есть информация о том, что две молекулы белка SmpV могут связываться с рибосомой на пре-инициаторном этапе [179; 197], стимулируя распознавание неактивной рибосомы.

### 3.1.3 Оценка достоверности упрощённого моделирования

После публикации вышеизложенных результатов вышла работа [198], в которой на основе данных метода криоэлектронной микроскопии предложена полноатомная модель тмРНК в комплексе с рибосомой. Сравнительный анализ структуры этого комплекса с построенной нами моделью выявил достаточно

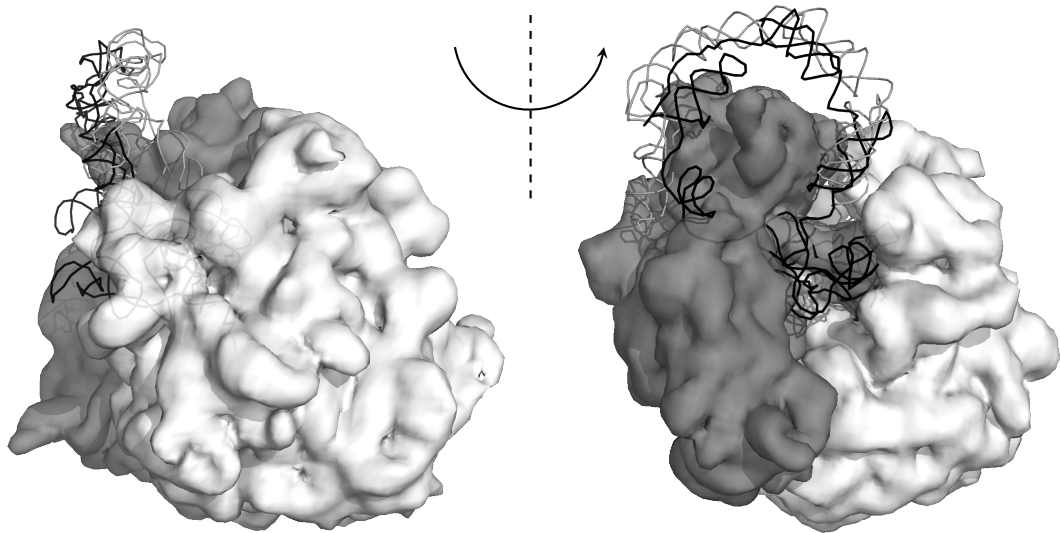


Рисунок 3.6: Пре-инициаторный комплекс рибосомы: модель и экспериментальные данные. Серым отмечены субчастицы рибосомы. Положение тмРНК по данным криоэлектронной микроскопии показано серыми линиями, на основе проведенного моделирования – чёрным.

высокую точность моделирования для той области тмРНК, которую мы называем аркой. и наименьшую – при расположении коротких двуспиральных участков, локализация которых определятся стэкинг-взаимодействиями во внутренних петлях молекулы. Как уже упоминалось, мы пренебрегли такого рода взаимодействиями для того, чтобы иметь принципиальную возможность построить модель такого крупного комплекса как тмРНК-рибосома.

Необходимо отметить, что для моделей, которые рассмотрены в этой части работы, данных по структуре ещё не было. Следует сказать, что переход к упрощенному моделированию позволяет достаточно точно локализовать в составе модели крупные структурные элементы, однако, для детального понимания атомистических механизмов, которые определяют функционирование крупных НК, необходимо использовать полноатомное моделирование.

### 3.1.4 Создание автоматического инструмента для упрощённого моделирования

В этой главе представлены результаты по созданию инструмента моделирования, который может быть унифицирован и распространён для широкого использования. Нами разработан метод и запущен веб-сервер для моделирования трёхмерной структуры малых РНК на основании данных по вторичной структуре. Как и ранее, мы оперировали методом пространственных ограничений, который позволяет достаточно однозначно определять параметры ограничений. Это касается двуспиральных участков вторичной структуры РНК, которые в нативных условиях образуют А-форму двойной спирали. Для петель и выпетливаний можно использовать данные по конформации петель с подобной последовательностью в известных структурах. Важным условием востребованности такого инструмента является возможность получения полноатомных моделей без дополнительного программного обеспечения. Первые варианты нашего сервиса (<http://dualopt1.cmm.msu.ru>, Рисунок 3.7) были реализованы на основе языка Perl, а молекулярно-механистическая часть высчитывалась с помощью программного пакета Gromacs. Предложена модульная схема составления пакета программ, которая позволяет использовать отдельные функции (подпрограммы) независимо друг от друга. Разбиение на модули позволяет создавать гибкую программную среду, которая легко обновляется и дополняется по мере развития. Подробности алгоритма работы сервиса можно найти в главе «Материалы и методы».

## Анализ результатов тестирования сервиса

Цель создания сервиса – попытка частично ликвидировать разрыв в количестве данных между вторичными и третичными структурами РНК. Для реализации этой цели разработан многоэтапный подход к построению модели РНК с упрощенным описанием нуклеотида, на основе которого и создан автоматизированный метод для моделирования трехмерной структуры РНК. В качестве иллюстрации возможностей метода проведено моделирование 5S рРНК, тмРНК *E. coli*, а также проанализирована его производительность. Полученные модели соответствуют всем необходимым структурным требованиям, следующим из описания вторичных структур и информации по структурам этих двух РНК. Это свидетельствует о приемлемом качестве полученных моделей. Однако следует подчеркнуть, что этот подход помогает построить модель на основе известных данных и практически не обладает предсказательной способностью. Возможность получения дополнительной информации из моделирования подобного рода возникает только при указании пользователем дополнительной информации, достаточной для формирования структуры РНК близкой к нативному состоянию. Несмотря на вышеизложенные ограничения, последние модификации нашего сервиса позволяют выполнять широкий спектр операций при моделировании третичной структуры РНК и восстанавливать упрощенные модели до полноатомного состояния.



# RNA Tertiary Structure Modelling

All operations to build your own model you may find in small [movie](#).

Two steps to fold coarse-grained model of RNA tertiary structure

Step #1:

- fill in RNA secondary structure form with CT format (also with pseudoknots);
- select helices auto-stacking gap (see description below);
- select short-distance tertiary interactions prediction (triples, "backside pentariples");
- choose online representation of output structures;
- push "Finish Step #1" button to obtain secondary and intermediate\* structures.

\* intermediate structure is the structure with minimum of "free energy" contains formed helices and stacking, but without tertiary interactions

### RNA secondary structure form

- RNA should be shorter then 400 nucleotides!
- You can paste any part of CT file into form, output structure numeration will start from the number 1. All interactions with unrepresented parts of the molecule will be discarded. So try to paste structural independent domains.
- programm algorithm could process H-type pseudoknots!
- secondary structure of RNA contains pseudoknots may be not plane
- input CT should not contain tertiary interactions, please check it
- for more details about CT format, look [RNA Structure Formats](#)
- to convert bracket (Vienna) notation into CT, use [converter](#) (dose not support pseudoknots)
- some useful RNA secondary structure databases:
  - [RNA SSTRAND](#)
  - [Comparative RNA Web Site](#) (require easy registration)

Paste you structure in CT format here, or use "Sample CT"

### Auto-stacking gap

- GAP is a maximum distance in nucleotides of one strand between two helices to perform coaxial stacking:
- if one helix could generate several stacking forms with other helices, it will form none; in that case please use manual stacking (Step #2).

### Include short-distance tertiary interactions prediction

- rRNA like triples A:(A:U) ; A:(G:C) ; A:(C:G) :

- "backside pentariple" U:(C:G)(U:A) :

Yes       No

### Represent output structures in browser

- required Java®:
- may retard page downloading

Yes       No

Start first folding:

Versions: web interface v.1.0, software v.1.1  
 © 2008 Gleb Speshilov, Lomonosov Moscow State University

Рисунок 3.7: Изображение стартовой страницы сервиса для моделирования трехмерной структуры малых РНК на основании данных по вторичной структуре.

## **3.2 Конформационная динамика взаимодействия макролидных антибиотиков с рибосомой**

### **3.2.1 Взаимодействие производных тилозина в рибосомном тоннеле**

При исследовании комплекса тмРНК с рибосомой нами показано, что упрощенное моделирование при достаточном количестве экспериментальных данных может привести к реалистичной модели третичной структуры. Хорошее совпадение модели с экспериментальными данными обусловлено тем, что в этом подходе крупные элементы структуры распределены в ограниченном объеме, так что внутренние стерические ограничения, принятые в рассматриваемой системе, приводят к разумному результату. В то время как в природе, такая структурная организация формируется широким разнообразием нековалентных взаимодействий.

Используя разработанный подход, мы попытались предсказать активность производных антибиотика, который связывается с рибосомным тоннелем. Важной составляющей этой системы является не сам антибиотик, а собственно рибосомный тоннель – нас интересует потенциальная подвижность нуклеотидов, которые находятся на поверхности тоннеля. На сегодняшний день есть только общие представления о функционировании остатков 23S рРНК в рибосомном тоннеле. Следует отметить, что в отличие от предыдущей задачи, в этой модели нас интересует подвижность отдельных оснований, а не остова РНК. Это обусловлено тем, что, согласно данным РСА, в районе рибосомного тоннеля подвижности остова 23S рРНК чрезвычайно низкая. Очевидно, что для решения поставленной задачи необходимо максимально точно оценить возмож-

ные нековалентные взаимодействия между нуклеотидами (в случае же с тмРНК мы определяли местоположение не оснований, а крупных элементов вторичной структуры).

Итак, кратко опишем объект исследования. Полипептидная цепь, синтезируемая в пептидилтрансферазном центре рибосомы (ПТЦ), продвигается по рибосомному тоннелю (РТ), или вообще не зашифровывать), по мере того, как к цепи присоединяются новые остатки. Рибосомный тоннель пронизывает тело 50S субчастицы насквозь – от ПТЦ до выхода из рибосомы, где растущая полипептидная цепь встречает направляющие факторы и шапероны. Основная роль рибосомного тоннеля – это направление синтезируемой белковой цепи от ПТЦ к выходу из рибосомы. Известны и факты, подтверждающие взаимодействие боковых групп аминокислот вновь синтезируемого белка с остатками 23S рРНК. В некоторых случаях такое взаимодействие настолько значимо, что приводит к остановке синтеза и появлению альтернативных продуктов экспрессии некоторых генов [199—201]. На сегодня известны две области рибосомного тоннеля, где происходит связывание полипептидной цепи с остановкой трансляции. Первый сайт состоит из остатков 23S рРНК, которые относятся к так называемому внешнему окружению ПТЦ, второй сайт расположен дальше по ходу тоннеля и организован не только нуклеотидами 23S рРНК, но и аминокислотными остатками рибосомных белков L4 и L22 [200]. Эти два регуляторных участка ассоциированы с процессом трансляции, перекрываются с сайтами связывания антибиотиков-макролидов, которые затрудняют выход растущей цепи белка, блокируя тоннель. Макролиды – это члены большого семейства природных и синтетических антибиотиков, которые действуют на рибосому. Они содержат макролактонное кольцо, состоящее из 14, 15 или 16 атомов, к кольцу присоединен один или более остатков природных сахаров [202]. Как только появились

данные РСА по структуре рибосомы, тут же стали пристально изучать комплексы рибосомы с макролидами [203; 204]. В результате этих исследований получена масса информации по взаимодействиям между макролидами с рибосомой. В частности показано, что остатки сахара в макролиде располагаются вдоль тоннеля, в то время как лактонное кольцо контактирует с большим участком тоннеля, практически полностью занимая его полость в [205]. Недавно показано, что некоторые нуклеотиды ингибируют рибосому при синтезе пептида с определенной последовательностью [206]. Понимание этого сложного явления – как антибиотик взаимодействует не только с рибосомой, но и с продуктом трансляции – пока находится на самом начальном этапе. Решение этой задачи даст начало новому направлению антибиотикотерапии – разработке класса высокоспецифичных ингибиторов трансляции, которые будут подавлять синтез только полипептида, содержащего последовательность, характерную исключительно для белка-мишени определенного патогена.

Ранее под руководством академика РАН А.А. Богданова был синтезирован ряд пептидных производных тилозина [207—209]. Идея их синтеза заключается в том, чтобы с помощью пептидной части воспроизвести растущую цепь, а макролидную использовать как “якорь” для позиционирования пептида в рибосомном тоннеле. Такие конъюгаты оказались удобным инструментом для исследования регуляторного участка рибосомного тоннеля [209]. Недавно синтезирован ряд таких производных тилозина, в которых остаток триптофана находится на разном расстоянии от макролидного кольца. Здесь триптофан выбран неслучайно – предполагается, что именно остатки триптофана играют ключевую роль в функционировании пептидов TnaC и SecM как ингибиторов трансляции, связывающихся в рибосомном тоннеле.

### 3.2.2 Подход к полноатомному моделированию молекулярной динамики рибосомного туннеля

Как упоминалось выше, полноразмерное моделирование молекулярной динамики рибосомы пока остается очень затратным подходом. Поставив перед собой конкретную задачу – на атомарном уровне получить информацию по взаимодействиям рибосомного туннеля с антибиотиками, мы для ее решения выбрали такой метод моделирования молекулярной динамики, который позволяет наблюдать за эволюцией системы во времени. В связи с тем, что начальное положение всех частей антибиотика не было определено, сначала необходимо привести систему в состояние структурного равновесия и только потом проводить анализ. По предварительным результатам стало понятно, что для достижения равновесия может понадобиться длительное время наблюдения, что не позволит при моделировании МД учесть все атомы рибосомы и растворителя ( $\approx 10^6$  штук атомов).

Ранее в литературе описан подход [210; 211], суть которого состояла в следующем. Вокруг интересующего “объекта” создавали сферическое окружение из остатков компонентов рибосомы. Атомы, которые находятся на краю сферы, подвергали позиционным ограничениям, т.е. их движение было запрещено. После этого сферическая система “помещалась” в тригональную ячейку и “заполнялась” водой. В этой системе объем рибосомы, заполненный зарядами, замещён на гораздо менее значимую с точки зрения электростатических взаимодействий воду. Надо признать, что такой подход не только далек от реальности, но и не оптимален.

Мы предложили строить не сферическое окружение, а кубическое ячейке кубической формы, что позволяет более удачно воспроизводить электростати-

ческие свойства окружения, характерные для рибосомы. Следует принять во внимание, что в трансляции белка часто задействованы не отдельные рибосомы, а полисомы, и наш подход наиболее удачно описывает окружение, свойственное полисомам, где вероятно высока плотность отрицательного заряда из остова РНК и относительно не много воды. Немаловажно и расположение катионов, нейтрализующих заряд РНК. Известно, что при сборке рибосомы из структурных компонентов требуются ионы магния. Однако того числа ионов магния, которое обнаружили при разрешении структуры рибосомы методом РСА [212], явно недостаточно для нейтрализации отрицательного заряда РНК. Известно, что есть определённая геометрия расположения фосфатных групп между двумя близлежащими нуклеотидами НК – когда образуется так называемый “магниевый мостик”. Нами разработан инструмент, который позволяет располагать катионы магния для образования таких мостиков в любой заданной системе, содержащей остатки НК. После насыщения системы катионами магния оставшийся избыток отрицательного заряда нейтрализовали катионами натрия.

Теперь можно было приступать непосредственно к моделированию. Для этой цели мы выбрали структуру рибосомы *E. coli*, которая получена методом РСА с разрешением 3.1Å (PDBID: 3OFR)[212]. Учитывая, что в составе рРНК прокариотических рибосом есть модифицированные основания, мы дополнили исследуемую структуру информацией по модификации оснований 23S рРНК [213]. Позиции новых атомов оптимизированы методом наискорейшего спуска, который позволяет удалить высокоэнергетичные (слабые) контакты, не затрагивая геометрию системы. Местоположение тилозина, как и его производных, в рибосоме *E. coli* неизвестно, поэтому его определяли по суперпозиции большой субчастицы рибосом *E. coli* и *H. marismortui* (PDBID: 1K9M, разрешение

3Å) [205], а все производные тилозина располагали по аналогии. Следует отметить, что при суперпозиции субчастиц с особой тщательностью рассмотрены консервативные остатки A2100, A2101, G2102, A2103 и A2538. Полученные структуры ориентировали так, чтобы рибосомный тоннель был направлен параллельно оси X. Все остатки рибосомного тоннеля, хотя бы один атом которых входил в кубическую ячейку, содержащую тилозин, были экстрагированы. Длина грани ячейки составила 70 Å. Возможные пустоты в структурах заполнялись водой, описываемой моделью tip4p, а остатки биополимеров – силовым полем parm99sb [214]. Таким образом, подготовленные системы подвергали стандартным шагам подготовки к моделированию МД. Так, на остатки, которые находились на расстоянии меньше 1 Å, накладывались позиционные ограничения. Время наблюдения за каждой системой составляло 500 нс. Конфигурации системы сохранялись в файл каждые 60 пс – каждая такая конфигурация в последствии будет называться кадром и кадром траектории.

### **3.2.3 Взаимодействие пептидных производных 5-О-микаминозилтилонолида с рибосомным тоннелем**

Мы уже говорили о том, что моделирование МД использовали для описания взаимодействия рибосомного тоннеля с некоторыми макролидами и растущей полипептидной цепью [210; 211]. По отношению макролидам тилозинового семейства этот подход ранее не применялся. Выше мы аргументировали выбор основных ограничений, которых придерживались при построении модели: экстрагированное из большой субчастицы рибосомы кубическое окружение рибосомного тоннеля, сильно затрудненное движение остатков, атомы которых на-

ходятся на краю выборки, и неограниченная свобода движения для остальной части системы.

Для подтверждения релевантности выбранной стратегии мы сначала исследовали динамику комплекса ‘‘тилозин–50S субчастица *H. Marismortui*’’, структура которого известна (PDBID: 1k9m). Среди множества доступных параметров, описывающих динамику системы, мы выбрали среднеквадратичное смещение (MSD) лактонного кольца относительно стартового (экспериментального) положения после суперпозиции системы относительно рибосомной части. Этот параметр отличается высокой чувствительностью. Согласно выбранному описанию окружения тоннеля, ожидается, что в ходе моделирования рибосомная часть не совершает ни поступательных, ни вращательных движений. Таким образом, именно смещение ключевого фрагмента макролида – лактонного кольца – будет индикатором нестабильности комплекса. Комплекс тилозина с рибосомой стабилизирован одной ковалентной связью, образованной боковой группой A2062 23S рРНК и альдегидной группой лактонного кольца. В полученных траекториях моделирования МД для тилозина не обнаружено значимых сдвигов в течение всего временного интервала наблюдения а системой (250 нс). Так что показана применимость используемого подхода для моделирования в микросекундном диапазоне.

Полученные данные по изменению MSD для всех изучаемых систем приведены на Рисунке 3.8. Значения MSD для тилозина и Вос-Gly-ОМТ выходят на равновесные на уровне  $0.01 \text{ нм}^2$ , что соответствует смещению атомов всего на  $1 \text{ \AA}$ . Это значение меньше РСА-разрешения структуры рибосомы, что позволяет говорить о высоком качестве построенной модели. Значения MSD для Вос- $\gamma$ -аминобутирил-ОМТ и ОМТ выше, и соответствуют  $0.03 \text{ нм}^2$  и  $0.06 \text{ нм}^2$ . Смещение тилозина меньше, чем его производного Вос-Gly-ОМТ, что обуслов-



лено присутствием в молекуле тилозина двух дополнительных остатков сахара, которые образуют водородные связи с компонентами рибосомного туннеля. В частности (Рисунок 3.9) мицинозы образует три очень стабильные водородные связи с остатками G748 и A751 23S рРНК рибосомы *E. coli*. ОМТ, в отличие от его производного, образует всего одну стабильную водородную связь с компонентами туннеля, что, вероятно, и является основной причиной его относительно высокой подвижности в модели.

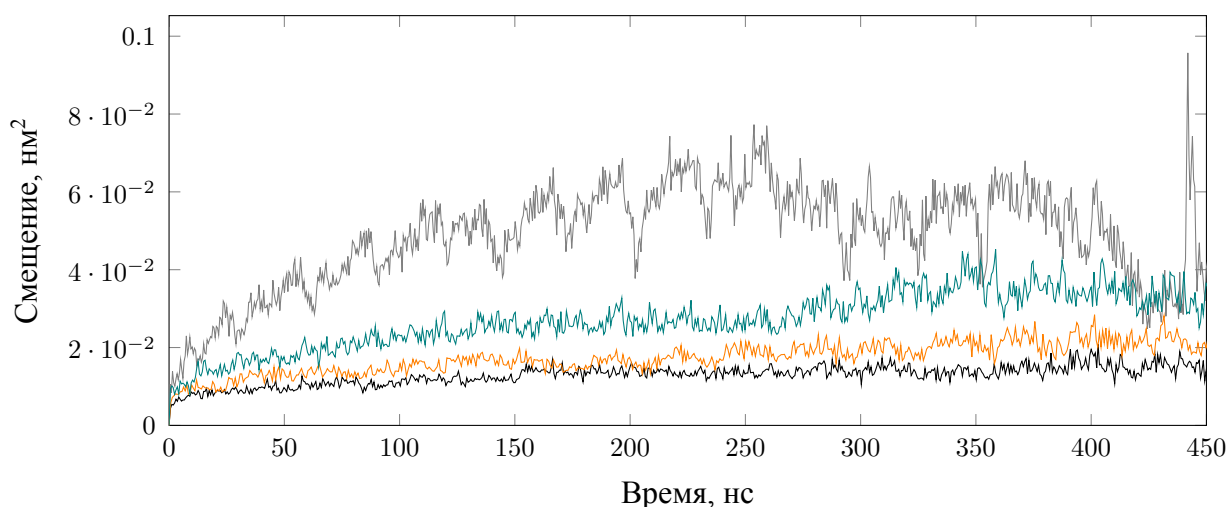


Рисунок 3.8: Среднеквадратичное смещение лактонного кольца тилозина (чёрная линия), ОМТ (серая линия), Woc-Gly-ОМТ (оранжевая линия) и Woc- $\gamma$ -аминобутирил-ОМТ (зелёная линия) производных в рибосомном туннеле 50S субчастицы рибосомы *E. coli*.

Надо отметить, что значения MSD для Woc-Gly-ОМТ сравнимы с таковыми для тилозина и значительно ниже, чем у ОМТ и Woc- $\gamma$ -аминобутирил-ОМТ. Это наблюдение указывает на эффективность сети водородных связей, образуемых звеном Woc-Gly-. Действительно, нами обнаружено, что карбонильная группа Gly образует две водородные связи: одна с аминок группой Lys90 из белка L22, вторая с нуклеотидом A752 23S рРНК (Рисунок 3.9). Результаты моделирования согласуются с общеизвестным фактом о способности карбонильного кислорода образовывать две водородные связи, выступая при этом в качестве акцептора. Для оценки стабильности водородных связей мы рассчитывали относительное

число кадров траектории с рассматриваемой водородной связью. Полученная доля в дальнейшем будет указываться в процентах, принято считать, что водородная связь стабильна, если её доля составляет более 30%. Стабильность, рассчитанная для вышеуказанных водородных связей, образуемых карбонильной группой Gly, составила 42% (Lys90) и 58% (A752).

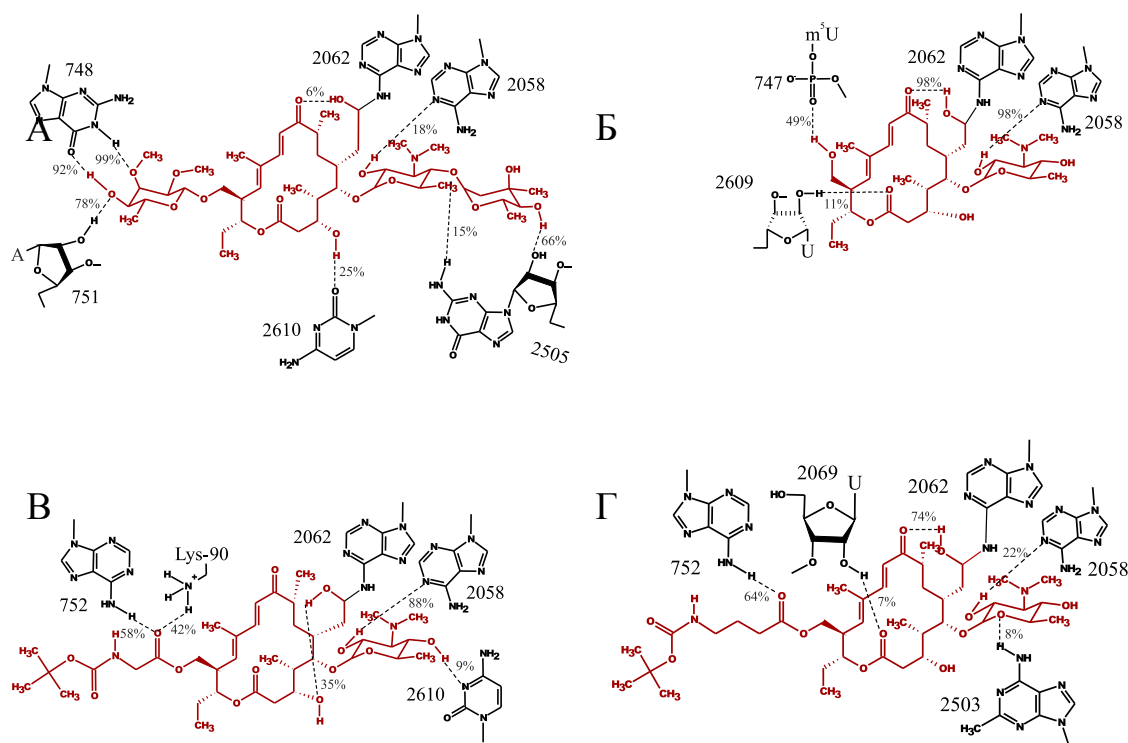


Рисунок 3.9: Структуры тилозина (А), ОМТ (Б), Вос-Gly-ОМТ (В) и Вос- $\gamma$ -аминобутирил-ОМТ (Г) и их водородные связи с остатками 23S рРНК и лизином 90 из белка L22 в рибосомном тоннеле 50S субчастицы рибосомы *E. coli*.

В отличие от Вос-Gly-ОМТ, Вос- $\gamma$ -аминобутирил-ОМТ может формировать только одну стабильную водородную связь с А752 (Рисунок 3.9) и не образует водородную связь с остатком Lys90 из-за высокой подвижности прилегающей Вос- $\gamma$ -аминобутирил группы. Следует заметить, что нет явных свидетельств о вовлечении Вос-группы в гидрофобные взаимодействия. Хотелось бы отметить, что гидрофобные взаимодействия эффективны в том случае, когда по-

движность молекул воды на поверхности контакта отличается от подвижности в водной фазе. Для рибосомного тоннеля сложно говорить о водной фазе вообще, так как практически все молекулы воды структурированы, а это значит, что вклад гидрофобных взаимодействий здесь незначителен. Из Рисунка 3.9 видно, что все рассматриваемые соединения образуют водородную связь гидроксильной группы карбиноламинного мостика с атомами кислорода основания A2062 и/или атомами кислорода O3 и O9 из макролидного кольца. Роль таких водородных связей в стабилизации комплекса антибиотик-рибосома отнюдь не очевидна.

Сравнивая активность рассмотренных соединений в экспериментах по ингибированию трансляции *in vitro*, мы обнаружили, что соединение Вос-Gly-ОМТ превосходит по эффективности ОМТ. Замена большой гидрофобной Вос-группы на меньшую гидрофильную формил-группу не приводит к изменению активности, что указывает на правильность выводов, сделанных на основе результатов моделирования.

Мы уже отметили, что антибиотики тилозинового ряда, благодаря наличию альдегидной группы в при С6, обладают способностью формировать ковалентную связь с компонентами рибосомного тоннеля [205]. Давно известно, что для таких антибиотиков как тилозин, ОМТ, десмикозин и прочих аналогов восстановление этого карбонила до гидроксила приводит к значительному резкому снижению их ингибиторной активности. Как показано А.А. Богдановым с коллегами, восстановление альдегидной группы тилозина приводит к тому, что тилозин теряет способность конкурировать с эритромицином и становится очень слабым ингибитором трансляции.

Соответственно, надо подчеркнуть, что для эффективного формирования этой ковалентной связи необходимо правильным образом ориентировать аль-

дегидную группу лактонного кольца в положении C20 и амино группу A2062 23 рРНК. Именно на этапе формирования водородных связей обеспечивается точная пространственная ориентация участков взаимодействующих молекул. Электростатические, ван-дер-Ваальсовы и гидрофобные взаимодействия не имеют геометрических детерминант, их влияние на пространственную структуру макромолекул обусловлено их совместным действием и поэтому при изменении геометрии их влияние ослабевает. Водородным связям для эффективного вклада в общую энергию связывания требуется достаточно жесткие геометрические ограничения, которые позволяют специфически ориентировать молекулу антибиотика для успешного протекания реакции образования ковалентной связи. Кинетика этого сложного двухступенчатого процесса уже исследована [215].

Мы моделировали состояние так называемого “конечного комплекса” с рибосомой, структурно подобной *H. marismortui* по данным PCA. А это не позволяет *a priori* утверждать, что результаты моделирования будут соответствовать данным экспериментов, проведенных на рибосомах *E. coli*. Тем не менее, сравнивая ингибирующую активность соединений в ряду тилозин-ОМТ-производные в зависимости от числа водородных связей, образуемых, согласно модели, каждым из них с областью тоннеля, можно предложить разумное объяснение полученных экспериментальных данных. Особо следует отметить высокую активность N-AcGly-содержащих соединений. Интересно, что именно атомы, входящие в остов пептида, способны образовывать две стабильные водородные связи как с белком L22, так и с 23S рРНК. Известно, что нуклеотид A752 образует уотсон-криковскую пару с U2609, и эта пара является ключевым структурным и функциональным элементом рибосомного тоннеля [212]. Она вовлечена во взаимодействия как с “тормозящим” пептидом, так и макро-

лидными антибиотиками [210; 216—219]. С другой стороны, функциональная значимость Lys90 рибосомного белка L22 так же хорошо известна. Его мутации приводят к множественным эффектам ассоциированных с активностью “тормозящего” пептида TnaC.

Таким образом, выявлен новый потенциальный сайт нековалентного взаимодействия остова растущего пептида с поверхностью рибосомного тоннеля. Возникающие здесь взаимодействия регулируют скорость прохождения через рибосомный тоннель синтезируемого пептида, аминокислотный состав и конформация которого и определяют эффект этой регуляции. Вероятно, эти контакты могут влиять на созревание белка и его фолдинг на выходе из рибосомного тоннеля. В этой части работы показано, что низкоамплитудные движения азотистых оснований рибосомы играют ключевую роль в связывании антибиотиков. На основании полученных результатов можно говорить о высокой точности метода полноатомного моделирования, который позволяет детально воспроизвести геометрию водородных связей, необходимых, в данном случае, для позиционирования альдегидной группы тилозиновых производных внутри рибосомного тоннеля. Важно отметить, что значимых изменений в направлении сахаро-фосфатного остова не выявлено – как и предполагалось при построении модели. Следует признать, что для полного описания и определения динамики движения остова НК рибосома не самый удобный объект, так как придется учитывать движение всех атомов этого огромной РНК-белковой машины - и это очень затратные расчёты.

## 3.3 Влияние топологии петель на геометрию квадруплексов

### 3.3.1 Особенности структуры гуанин богатых ДНК

Известно, что в различных условиях тяжи гуанозиновых олиго- и полинуклеотидов способны агрегировать друг с другом. Нередко такая агрегация стабилизируется присутствующим в растворе одновалентным катионом с подходящим радиусом типа калия или натрия. С помощью РСА показано, что такие полигуаниновые последовательности образуют структуру с новым типом укладки ДНК - четырехцепочечную спираль [220—222], где четыре гуаниновых основания из разных цепей образуют плоскую структуру, удерживаемую G-G попарными водородными связями (Рисунок 3.10). Такие структуры отличаются высокой стабильностью и называются гуаниновыми G-квартетами, или G-тетрадами. Каждый G-квартет скреплен в сумме восемью водородными связями, образованными взаимодействием “уотсон-криковской стороны” одного гуанинового основания с “хугстиновской стороной” другого.

К НК, которые потенциально могут образовать G-тетрадный мотив, приковано внимание многих исследователей. Это связано с тем, что гуанин-богатые последовательности чрезвычайно широко представлены во всех открытых на данный момент геномах. Такие мотивы обнаружены и в промоторных областях, и в сайтах переключения в составе последовательности иммуноглобулиновых генов, и в “горячих точках” рекомбинации, и др.[223].

G-квартеты обнаружены также на концах ДНК эукариотических хромосом, т.е. в теломерах [224]. Последовательность ДНК теломер представляет собой тандемные повторы коротких блоков поли-G, которые иногда включают в себя

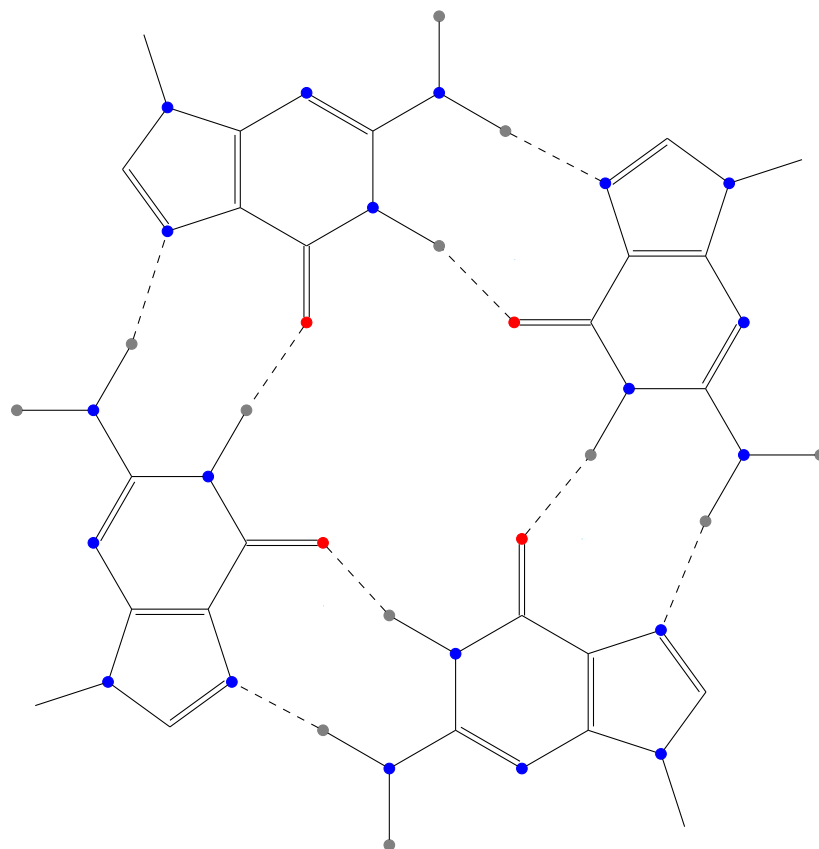


Рисунок 3.10: Схема G-квартета. Представлены четыре остатка гуанина из копланарной четверки квартета. Каждый нуклеотид одновременно выступает и донором (атомы N1 и N2), и акцептором (атомы O6 и N7) водородных связей.

аденин и/или тимин:  $G_n T_n, G_n T_n G_n, G_n A_n$  или  $(TTAGGG)_n$ . Тип повтора является видоспецифичным. Так, повтор  $(TTAGGG)_n$  характерен для млекопитающих,  $TTAGG$  для насекомых, а  $TTTAGGG$  – повторы большинства растений. В теломере ДНК находится в комплексе со специфическими белками. Теломеры защищают информацию на концах генома от от “недорепликации” и других повреждений, которые могут случиться в ходе рекомбинации или при воздействии нуклеаз.

ДНК теломер в соматических клетках человека в среднем составляет 8–10 т.п.н. Терминальные 100–200 нуклеотидов с 3'-конца представляют собой однотяжевой “хвост” [225], конформационно ничем не ограниченный. В живых клетках этот “хвост” ассоциирован с белком POT1 [226], в отсутствие же этого белка однотяжевая теломерная ДНК способна складываться и димеризоваться, формируя четырехцепочечные шпильки, которые могут стабилизироваться формированием G-тетрад [227; 228]. Другой способ стабилизации такой ДНК — формирование внутримолекулярных G-квартетов путем многократного складывания. Такие G-квартет-содержащие структуры называются квадруплексами или тетраплексами [229]. Помимо теломер, G-квадруплексные последовательности обнаружены в промоторах ряда онкогенных и опухолеассоциированных генов, таких как *k-ras* [230], *c-kit* [231] и *bcl2*[232].

Таким образом, перспективным направлением терапии онкологических заболеваний может стать дизайн ингибиторов экспрессии таких генов с помощью специфичных к квадруплексам агентов типа порфирина TMPyP4 [233; 234]. G-квадруплексы могут ,быть образованы короткими олигонуклеотидами с последовательностью  $G_m X_n G_m X_o G_m X_p G_m$ , где  $m$  – число остатков гуанина в G-блоке, которые обычно и задействованы в образовании G-тетрад.  $X_n, X_o$  и  $X_p$  могут



быть комбинацией любых нуклеотидов, включая G; такие участки формируют петли между G-тетрадами.

Некоторые последовательностей аптамеров тоже относятся к квадруплексным ДНК, т.е. в растворе способны формировать G-квадруплексы. Аптамеры – это короткие синтетические олигонуклеотиды или пептиды, способные специфически узнавать свою мишень, размер же мишени варьирует в широких пределах – от малых молекул до целых клеток. Так что аптамеры считаются аналогами моноклональных антител [235; 236].

Уже идентифицированы G-квадруплексные аптамеры, мишенью которых является тромбин [237] и STAT-3[238]. Известны G-квадруплексные аптамеры с антираковыми свойствами – они проходят сейчас стадию клинических испытаний; механизм их действия связан с белком нуклеолином, участвующем в процессинге РНК [239]. Формирование квадруплексной структуры из заданной квадруплексной последовательности может происходить различными путями – и зависит этот путь от числа G-блоков. Четыре отдельных цепи могут ассоциировать друг с другом с образованием межмолекулярного G-квадруплекса (Рисунок 3.11).

Внутримолекулярные тетраплексы образуются из одноцепочечной молекулы в результате сложной пространственной укладки нуклеотидной цепи (Рисунок 3.12). Фундаментальный строительный элемент всех квадруплексных структур – G-квартет. Квартеты в структуре квадруплекса располагаются друг над другом, причем минимум два квартета необходимы для структурной стабильности тетраплекса [240]. Число остатков гуанина в каждом отдельном G-блоке напрямую связано с количеством G-тетрад в окончательно свернутом квадруплексе. Например, в ДНК теломер млекопитающих (тандемный повтор

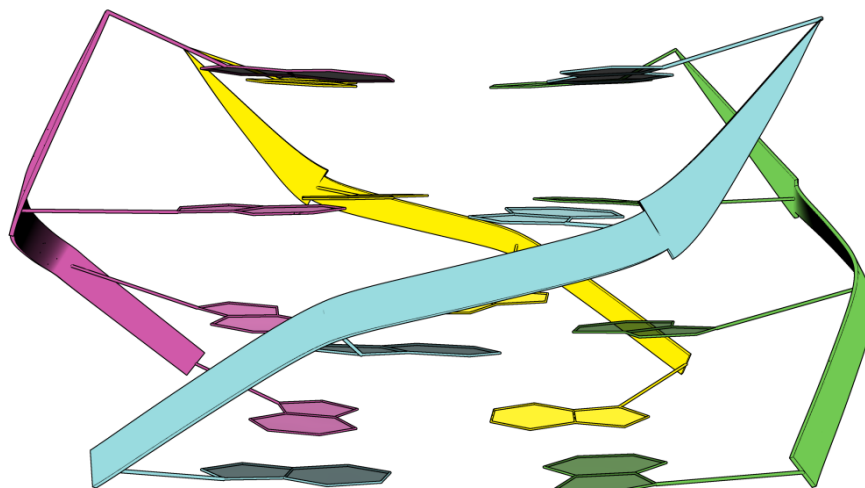


Рисунок 3.11: Четырехтяжевой межмолекулярный параллельный G-квадруплекс.

TTAGGG) квадруплексы, образованные четырьмя такими повторами, имеют три расположенных друг над другом G-квартета [241].

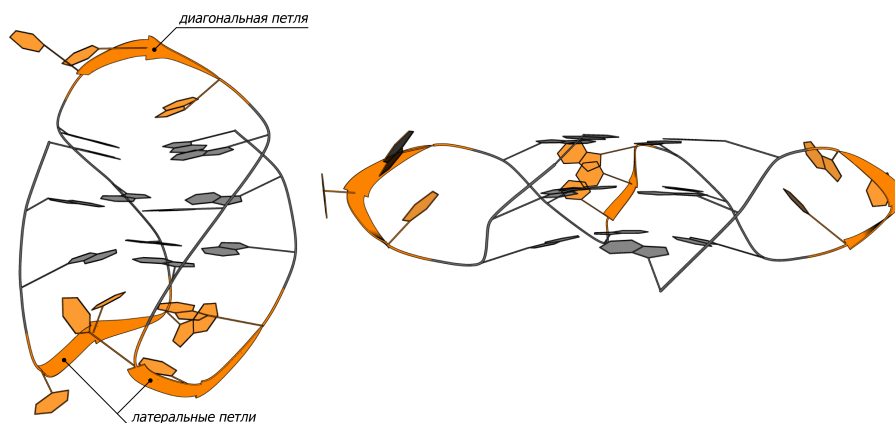


Рисунок 3.12: Сворачивание цепи ДНК в одномолекулярный G-квадруплекс с тремя G-тетрадами.

Два варианта укладки олигонуклеотида  $AGGG(TTAGGG)_3$  различаются направленностью полинуклеотидной цепи в различных тяжах квадруплекса. Слева: топология с петлями латерального и диагонального типов [242]; справа: топология с петлями пропеллерного типа, обращающими ход цепи [243]. Квадруплекс стабилизируют те же самые факторы, что и дуплекс ДНК: стэкинг-взаимодействие оснований, водородные связи, электростатические взаимодей-

ствия и гидратная оболочка. Гидратация сахарофосфатного остова очень важна для стабильности структуры: в упорядоченной гидратной оболочке молекулы воды с помощью разветвленной сети водородных связей объединяют в единое целое основания, сахара и заряженные фосфаты, которые локализованы на внешней поверхности квадруплекса [244—246].

Кроме этих стандартных факторов, стабилизирующих дуплексы ДНК, у квадруплексов есть специфическая, весьма важная, составляющая – большой вклад в стабильность вносит координация Об карбонильной группы катионами [247; 248]. Атомы Об образуют квадрат в каждом квартете, что в квадруплексе дает бипирамидную антипризму с расстоянием  $3,3\text{Å}$  между квартетами [249]. Это отрицательно заряженное пространство между тетрадами должно быть стабилизировано путем координации катиона. На стабильность образованного квадруплекса существенно влияет природа катиона: размер/ионный радиус и заряд [248].

Четыре гуанозинового нуклеозида в составе тетрады могут существовать либо в анти-, либо в син-конформации относительно гликозидной связи – а значит, существует 16 возможных комбинаций. Взаимная ориентация отдельных цепей в квадруплексе влияет на значение торсионного угла вокруг N-гликозидной связи. Например, при параллельной ориентации всех четырех цепей (Рисунок 3.11) все гликозидные углы находятся в анти-конформации.

При антипараллельной ориентации в составе квадруплексов присутствуют как син-, так и анти-гуанины, вне зависимости от того, образован квадруплекс четырьмя цепями или одной.

Различные нуклеотидные последовательности между G-блоками формируют внеспиральные петли. Эти петли бывают трех типов. В параллельном внутримолекулярном квадруплексе необходима петля, соединяющая нижнюю G-

тетраду с верхней, в результате чего получаются петли пропеллерного типа (Рисунок 3.12, справа). Антипараллельные квадруплексы сформированы так, что в них хотя бы одна цепь антипараллельна другим. Такая топология тетрад обнаружена в большинстве определенных на текущий момент бимолекулярных и внутримолекулярных квадруплексных структур.

Помимо пропеллерных в квадруплексах встречается еще два типа петель. Латеральные (боковые) петли соединяют соседние G-тяжи. Две таких петли могут располагаться как на одном, так и на противоположных полюсах молекулы, что соответствует положению “голова к голове” или “голова к хвосту” в бимолекулярных комплексах. Другой тип антипараллельных петель – диагональные петли, соединяющие противоположные G-тяжи (Рисунок 3.12, слева).

Все квадруплексные структуры имеют четыре бороздки – в отличие от двойной спирали, имеющей только две. Бороздки образуются полостями, ограниченными сахарофосфатными остовами. Размеры бороздок очень сильно разнятся, что определяется общей топологией и природой петель, а также гликозидных углов. В квадруплексах с петлями исключительно диагонального или латерального типа бороздки структурно просты, в квадруплексах с петлями пропеллерного типа бороздки имеют более сложный характер. Таким образом, существует ряд структурных переменных (число G-тетрад, их последовательность и длина, тип и направленность петель), которые взятые вместе приводят к большому топологическому и структурному разнообразию квадруплексов. Мы попытались выявить взаимосвязь между структурой и свойствами G-квадруплексных ДНК и определить факторы, влияющие на геометрию квадруплексов.

### 3.3.2 Сравнение геометрических параметров для известных типов квадруплексов

Даже поверхностный анализ многообразия обнаруженных для квадруплексов вариантов структур, присутствующих в базе данных PDB, приводит к заключению о необходимости систематизации и выработке универсальных параметров описания структуры квартетов и их полиморфизма. До настоящего времени такие попытки в литературе не описаны. В качестве структурных характеристик квадруплексов мы выбрали два параметра. Первый – это угол закручивания квадруплекса, т.е. угол поворота между двумя соседними квартетами, соответствующий углу между двумя векторами, проходящими через атомы C1' двух соседних гуанинов (Рисунок 4.1). Этот параметр широко используется для описания структур двойной спирали спирали [250]. Второй параметр более оригинален – он описывает отклонение квартетов от планарности – это расстояние между центрами масс двух квадратов, образованных атомами O6 и N9. Два этих параметра позволяют в достаточной мере описать конформационный полиморфизм и конформационную подвижность структуры квадруплексной ДНК.

#### Четырехтяжевые параллельные квадруплексы

Состав группы (PDBID):

- 1EVM [251], 1EVN [251], 1NP9 [252], 1NZM [253], [254]1O0K - теломерная ДНК (человек)
- 139D [255; 256] – теломерная ДНК (*Tetrahymena*)
- 1EMQ [257] – теломерная ДНК (*Saccharomyces cerevisiae*)
- 1EVO [258] – фрагмент вирусного генома SV40

В рассматриваемом ряду представлен простейший вариант устройства квадруплексных структур. Угол закручивания в таких структурах можно рассматривать как идеальный, поскольку ассоциация четырех тяжей не накладывает никаких структурных ограничений, чего не наблюдается в случае би- и мономолекулярных квадруплексов. Четырехтяжевые параллельные квадруплексы характеризуются широким спектром углов закручивания (Рисунок 3.13) с двумя областями предпочтительных значений: узкой, соответствующей  $21^\circ$ , и более размытой, находящейся в пределах  $27^\circ$ – $34^\circ$ . Планарность квартетов таких структур также варьирует: в большинстве случаев квартеты немного отклоняются от плоскости (на приблизительно  $0.5\text{ \AA}$ ), хотя максимальные значения отклонений могут превышать  $1\text{ \AA}$ . Найденное разнообразие значений параметров иллюстрирует спектр вариаций конформационного полиморфизма четырехтяжевых параллельных квадруплексов, в образование которых при межмолекулярной ассоциации не вносятся никаких структурных ограничений.

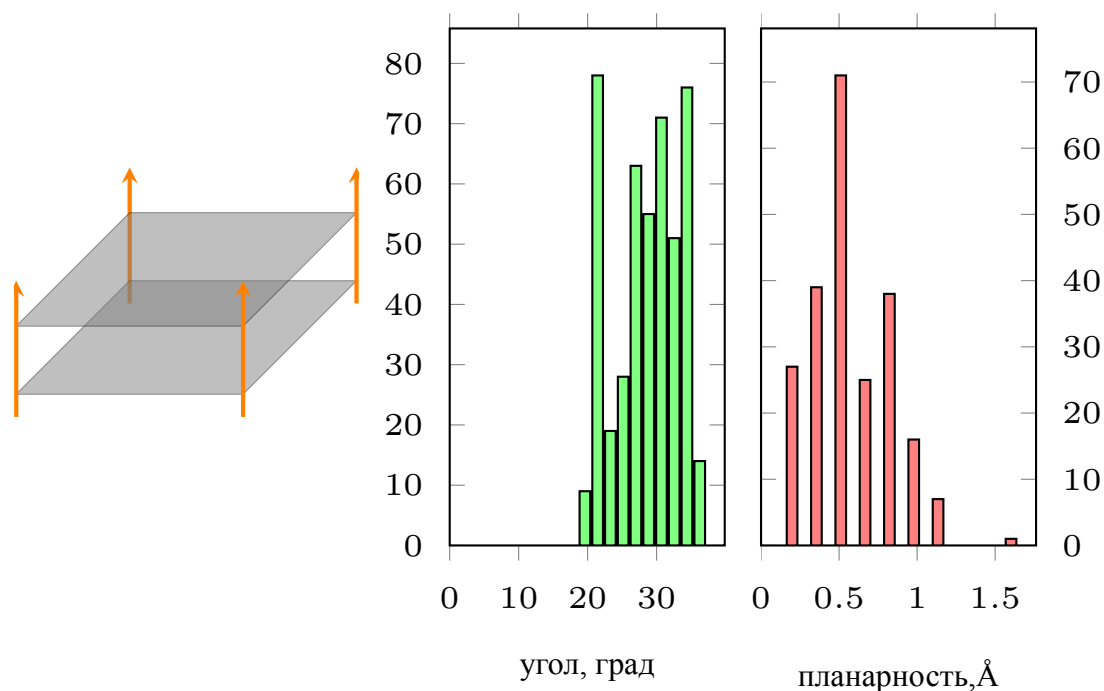


Рисунок 3.13: Схема пространственной организации и геометрические характеристики структур с четырехтяжевыми параллельными квадруплексами.

## Одномолекулярные квадруплексы с топологией цепи (3+1)

Состав группы:

- 2JSK [259][43], 2JSQ [259][43], 186D [260][44], 2GKU[261] [45], 2HY9 [262][46], 2JPZ[263] [47], 2JSL [259][43], 2JSM [259] - теломерная ДНК (человек, Tetrahymena)
- 2F8U [264] – промотор белка BCL2 (человек)

Такие квадруплексы имеют в своем составе две петли латерального и одну пропеллерного типа, обращающую ход полинуклеотидного тяжа в обратную сторону. От предыдущего типа квадруплексов этот отличается только наличием пропеллерной цепи, что сказывается на значениях угла закручивания (Рисунок 3.14). У паттерна распределения углов появляется дополнительное значение с максимумом  $28^{\circ} \pm 4^{\circ}$ . Более того, характерные для квадруплексов с латеральными петлями значения углов закручивания в диапазоне  $10^{\circ} - 20^{\circ}$  сдвинуты на  $2^{\circ} - 4^{\circ}$  в сторону больших значений. Квартеты квадруплексов типа “кресло” жестко планарны – в них планарность фиксируется стэкинг-взаимодействиями с гетероциклическими основаниями латеральных петель.

### Структура типа ‘кресло’

Состав группы:

- 148D[265], 1C32[266], 1C34[266] , 1C35[266], 1C38 [266], 1QDF[267], 1QDH [267], 1RDE [268] – ДНК-аптамер к тромбину
- 2KM3 [269] - теломерный повтор CTAGGG (человек)

Структура типа “кресло” представляет собой мономолекулярный квадруплекс, соединенный тремя петлями латерального типа. Это в достаточ-

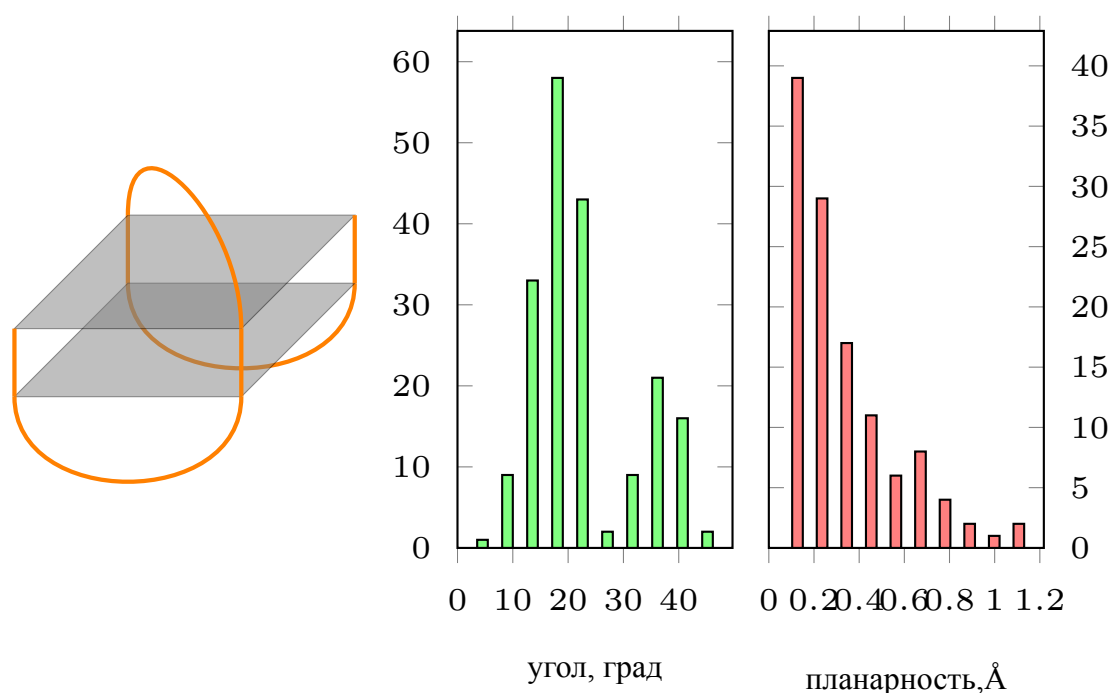


Рисунок 3.14: Схема пространственной организации и геометрические характеристики квадруплексов типа (3+1)

ной степени уникальная структура, представленная лишь двумя молекулами: 15-звенным ДНК-аптамером к тромбину (148D, 1C32, 1C34, 1C35, 1C38, 1QDF, 1QDH, 1RDE) и 22-звенным олигонуклеотидом, образованным теломерным повтором STAGGG (2KM3). Такие структуры характеризуются более низкими значениями углов поворота по сравнению со структурами межмолекулярных четырехтяжевых квадруплексов, у которых петли отсутствуют (Рисунок 3.14). Среднее значение углов закручивания квадруплексов для структур типа “кресло” составляет  $15^\circ$  с отклонением  $5^\circ$ . Квартеты таких структур характеризуются высокой планарностью, скачок значений в области  $0.8\text{--}0.9\text{Å}$  соответствует структурам, разрешенным в растворах с высокой ионной силой. Планарность объясняется тем, что гетероциклические основания латеральных петель образуют стэкинг-взаимодействия с основаниями квартетов и, тем самым, ограничивают их отклонения от плоскости.



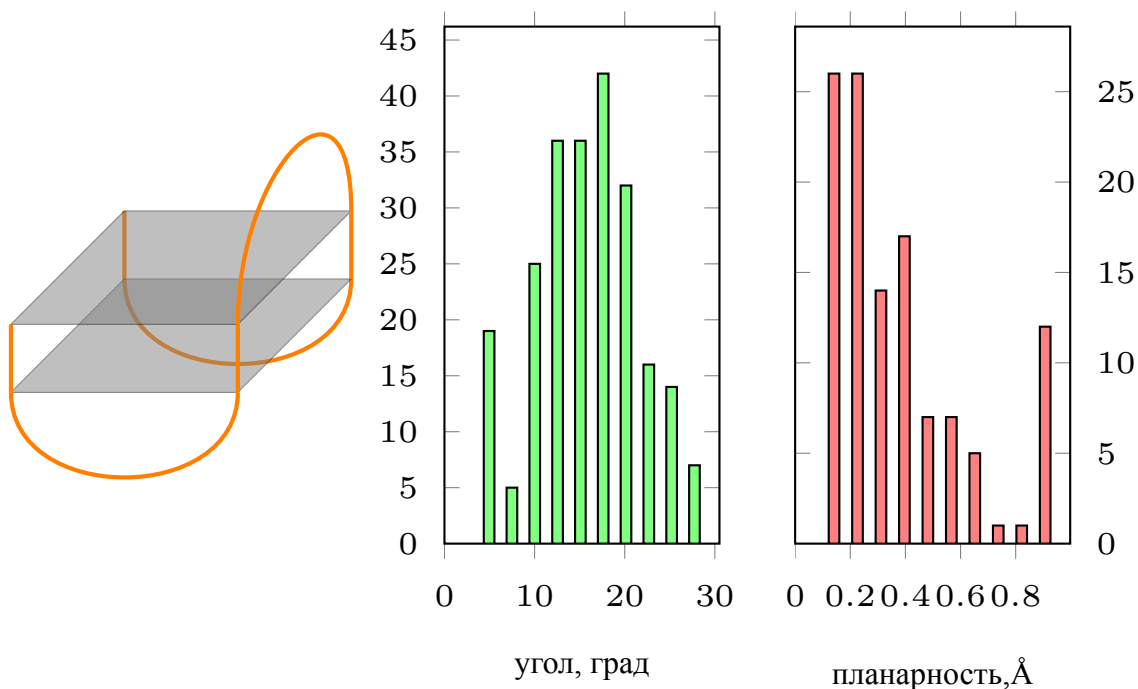


Рисунок 3.15: Схема пространственной организации и геометрические характеристики квадруплексов типа “кресло”

### Структура типа ‘корзина’

Состав группы:

- 2KF8[270], 2ККА[271], 2KOW [272], 143D[273], 230D [274], 201D[275]  
— теломерная ДНК (человек, *Oxytricha*)

Структуры типа “корзина” представляют собой мономолекулярный квадруплекс, соединенный двумя петлями латерального и одной диагонального типа. Как и в предыдущем случае, замена латеральной петли на диагональную приводит к изменению характера распределения углов закручивания — появляется два ярко выраженных плеча, со средними значениями  $18^\circ \pm 4^\circ$  и  $36^\circ \pm 4^\circ$  (Рисунок 3.16).

Таким образом, наличие диагональной петли определяет появление в конформационном ландшафте квадруплексов нового набора структур. Для них характерны большие углы закручивания квадруплекса – эти значения максималь-

ны среди всех, встречающихся в четырехтяжевых параллельных квадруплексах, не имеющих петель. Степень планарности квартетных структур типа “корзина” ниже, чем у рассмотренных ранее (исключительно с латеральными петлями), что, видимо, обусловлено уменьшением стэкинг-взаимодействий гетероциклических оснований диагональной петли с верхним квартетом.

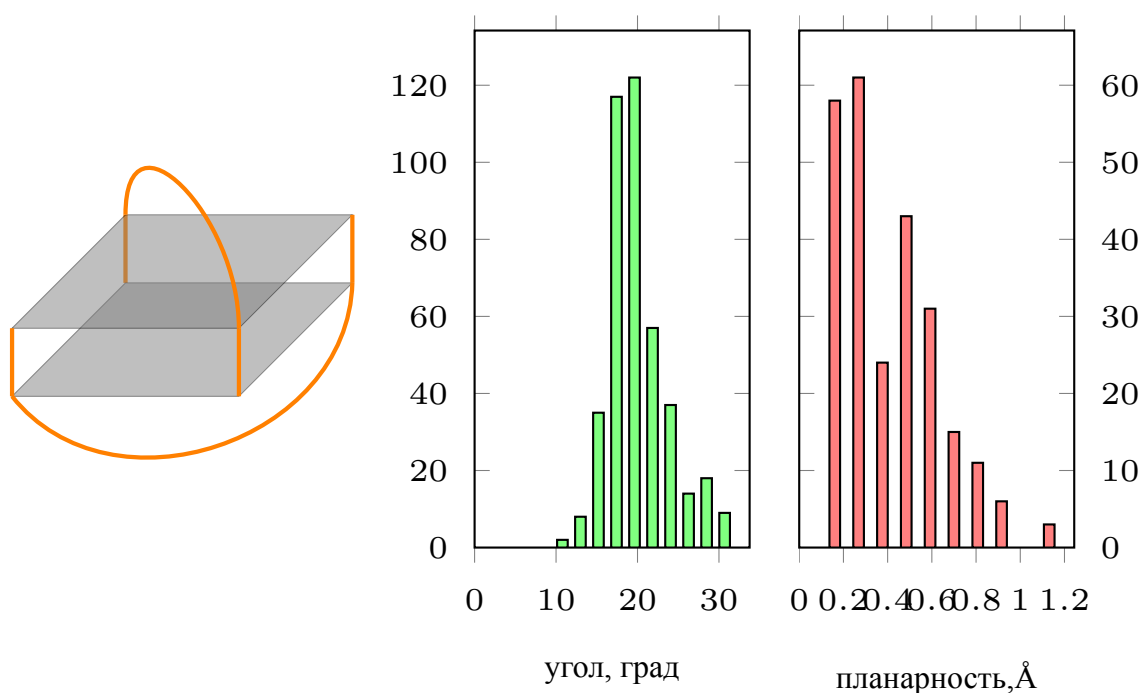


Рисунок 3.16: Схема пространственной организации и геометрические характеристики структур типа “корзина”.

### Одномолекулярные параллельные квадруплексы с пропеллерными петлями

Состав группы:

- 1KF1 [276], 3CDM [277] - теломерная ДНК (человек)
- 1XAV [278], 2A5P [279] – промотор с-MYC (человек)
- 2KQG [280], 2KQH[280], 2KYR[281] – промотор онкогена c-kit (человек)
- 1MYQ [282] – синтетический олигонуклеотид (GGA)<sub>4</sub>

- 1Y8D [283] – аптамер к (ВИЧ-1) интегразе ВИЧ-1

Такие необычные структуры представляют собой внутримолекулярные квадруплексы, все петли которых организованы по пропеллерному типу. а полинуклеотидный тяж трижды меняет направление. Наличие пропеллерных петель жестко определяет структуру квадруплекса: распределение углов закручивания имеет выраженный максимум со значением  $31^\circ \pm 3^\circ$  (Рисунок 3.17). Это значение близко к полученному для 4-тяжевых параллельных квадруплексов, у которых нет петель. Одномолекулярные параллельные квадруплексы с пропеллерными петлями организованы в квартеты с высоким значением планарности. Возможно, что пропеллерные петли обеспечивают оптимальную геометрию квадруплексов с жестко закрепленным сахарофосфатным остовом.

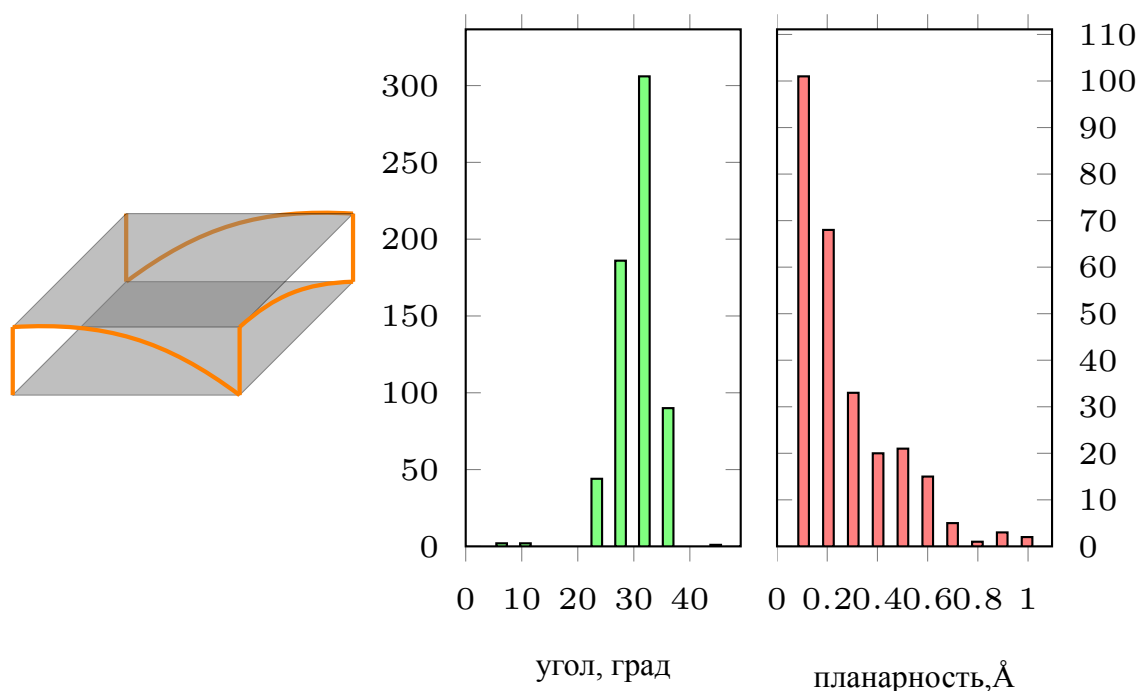


Рисунок 3.17: Схема пространственной организации и геометрические характеристики одномолекулярных параллельных квадруплексов с пропеллерными петлями.

## **Бимолекулярные квадруплексы с латеральными петлями**

Состав группы:

- 1A8N [284], 1A8W [285] – тандемный повтор GGGC
- 1F3S [286] – синтетический олигонуклеотид

В отличие от мономолекулярных квадруплексов, которые образуются путем внутримолекулярного фолдинга, бимолекулярные квадруплексы формируются при димеризации двух сложенных “сами на себя” полинуклеотидных тяжей, которые несут блоки гуанинов. Для этого типа квадруплексов характерна большая длина латеральных петель (5–6 нуклеотидов). Их можно рассматривать как квадруплексы со структурой, промежуточной между мономолекулярными квадруплексами с латеральными петлями и квадруплексами с диагональными петлями. Углы закручивания этих квадруплексов также занимают промежуточное положение в таблице значений - между двумя экстремумами указанных типов структур (Рисунки 3.16,3.17,3.18).

Следует отметить склонность бимолекулярных квадруплексов с латеральными петлями к закручиванию на  $20^\circ \pm 1^\circ$ ,  $27^\circ$ . Квартеты в квадруплексах этого типа жестко планарны – по-видимому, из-за интенсивных стэкинг-взаимодействий с гетероциклическими основаниями петель. Более строгие умозаключения о структуре этих квадруплексов можно будет сделать только после расширения списка представителей – для повышения статистической значимости.

## **Бимолекулярные квадруплексы с диагональными петлями**

Состав группы:

- 156D[287; 288], 1JPQ[244], 1L1H[289], 1QDI[267], 1QDK[267], 3EM2[290], 3EQW[290], 3ERU[290], 3ES0[290], 3ET8[290], 3EUM[290], 2AKG[291], 1K4X[274], 1JRN[244], 2HBN[292], 3EUI[290] — теломерная ДНК (*Oxytricha*)
- 2KAZ [293], 1U64[294], 1LVS[295], 1FQP[296] - синтетические олигонуклеотиды

В молекулы этой группы входят бимолекулярные квадруплексы с диагональными петлями между противоположными углами. Таким образом, эти квадруплексы похожи на структуры типа “корзина”, описанные ранее, у которых диагональная петля заменяет две латеральные петли. Для значений углов закручивания квадруплексов такая замена приводит к исчезновению плеча с максимумом  $35^\circ$ , оставляя только один максимум  $19^\circ \pm 4^\circ$  (Рисунок 3.18). Распределение значений планарности квартетов также похоже на таковое у структур типа “корзина”.

### **Бимолекулярные параллельные квадруплексы с пропеллерными петлями**

Состав группы:

- 1K8P[276], 2HRI[297], 3CE5 [298] – теломерная ДНК (человек)
- 2KYO [299]- промотор онкогена c-kit (человек)
- 1NYD [300], 1EEG[301], 1XCE[302] - синтетические олигонуклеотиды

В этом типе квадруплексов, как и для в его мономолекулярном аналоге, пропеллерные петли жестко ограничивают конформационный полиморфизм квадруплекса: кривая распределения значений угла закручивания имеет максимум при  $31^\circ \pm 3^\circ$  квартеты планарны (Рисунок 3.19).

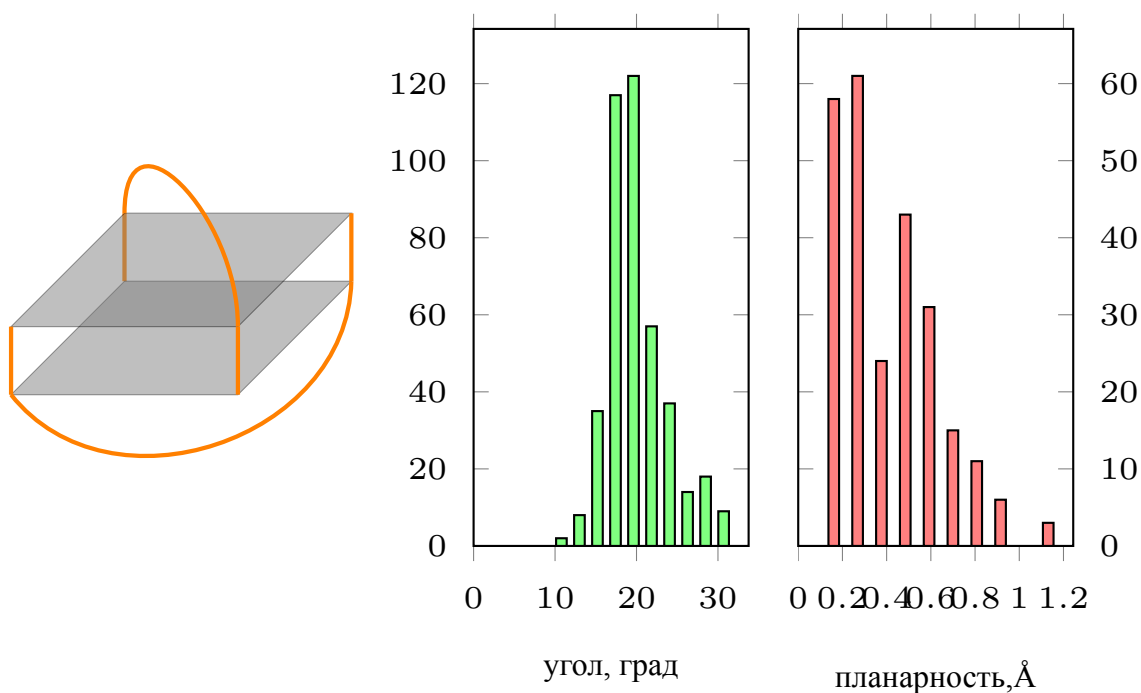


Рисунок 3.18: Схема пространственной организации и геометрические характеристики бимолекулярных квадруплетов с диагональными петлями.

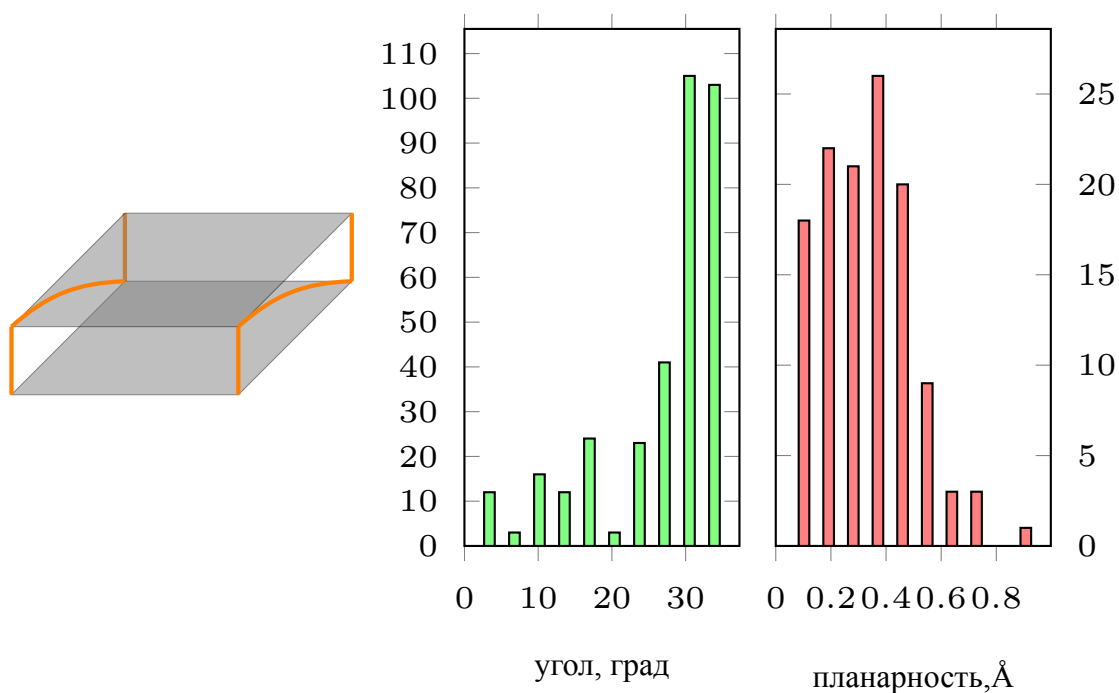


Рисунок 3.19: Схема пространственной организации и геометрические характеристики бимолекулярных параллельных квадруплетов с пропеллерными петлями.

### 3.3.3 Классификация квадруплексных структур

G-квадруплексные структуры локализованы на концах ДНК теломерных участков хромосом и в промоторных областях ряда онкогенных и опухолеассоциированных генов. Это делает квадруплексы привлекательной мишенью для противораковой терапии. Наиболее вероятно, что взаимодействие антираковых препаратов с целевой квадруплексной ДНК происходит за счет стэкинг-взаимодействия с квартетами (терминальными квартетами, либо за счет интеркаляции), с одной стороны, и с бороздками – с другой. Таким образом, детальное знание геометрии этих элементов и причин, способных влиять на нее, абсолютно необходимо для оптимизации поиска новых эффективных противораковых препаратов.

Квадруплексы НК представляют интерес не только как мишени антионкогенных препаратов, но и как структурное ядро терапевтических агентов на основе аптамеров. В частности, это аптамеры к тромбину, к интегразе ВИЧ-1, к опухолевому маркеру – белку нуклеолину, вовлеченному в процессинг РНК. Для оптимизации стабильности и фолдинга этих олигонуклеотидов также необходимо знать природу сил, способных оказывать стабилизирующее и дестабилизирующее воздействие на молекулу аптамера. Для квадруплексов характерно большое топологическое и структурное разнообразие, обусловленное такими переменными, как число G-квартетов, направленность, тип, последовательность и длина петель. Для описания этого конформационного ансамбля мы выбрали два параметра: угол закручивания между двумя соседними квартетами и планарность квартетов.

Показано, что оперирование этими параметрами позволяет охарактеризовать самые сложные квадруплексные структуры. Надо отметить, что под пара-

метром “планарность” имеется в виду обобщение двух явлений. В случае, если квартет симметричный и все основания G образуют водородные связи друг с другом, этим параметром фиксируется именно планарность квартета. В случае разрыва водородных связей симметрия квартета нарушается и тогда параметр фиксирует степень искажения квартета.

Четырехтяжевой межмолекулярный параллельный квадруплекс представляет простейший пример, в котором действует лишь одна из указанных переменных – число G-квартетов. Эта переменная определяет границы, в которых могут существовать такие структуры без ущерба для своей целостности. Границы довольно широки: в терминах угла закручивания они составляют  $19^{\circ}$ – $36^{\circ}$  с двумя областями предпочтительных значений:  $21^{\circ}$  и  $27^{\circ}$ – $34^{\circ}$ . Движения терминальных квартетов таких структур ограничены лишь стэкинг-взаимодействиями с прилежащими квартетами и координационными связями со стабилизирующими катионами – в случае их наличия. Так что G-квартеты таких структур, особенно терминальные, не отличаются строгой планарностью.

Структуры типа “кресло” разительным образом отличаются от 4-тяжевых квадруплексов, в составе которых нет петель. Латеральные петли таких структур ограничивают геометрию квадруплексов, более того, они “уводят” ее в область значений углов закручивания, не свойственных беспетлевому квадруплексам. Если для последних минимальные углы составляют  $19^{\circ}$ , то для структур с петлями исключительно латерального типа область предпочтительных значений находится в области  $15^{\circ} \pm 5^{\circ}$ . Это говорит о том, что латеральные петли, в основном из-за небольшой длины, приносят в структуру квадруплекса существенное напряжение. В то же время, из-за особенностей расположения, вытекающих уже из названия данного типа петель, на терминальных полюсах G-квадруплексов такие петли активно взаимодействуют с краевыми квартетами



квадруплексов за счет стэкинга, что положительно сказывается на планарности квартетов структур типа “кресло”.

Интересно, что добавление петель диагонального или пропеллерного типа к структуре типа “кресло” сразу сказывается на паттерне распределения углов закручивания – область предпочтительных значений углов сдвинута в сторону больших величин, характерных уже для беспетлевых параллельных G-квадруплексов. Кроме того, на паттерне появляется второе плечо, характеризующее область, примыкающую к петле, которая отличается от латеральной. В структурах типа “корзина” хорошо видно, насколько велико напряжение, создаваемое латеральными петлями, располагающимися на одном из полюсов молекулы. Момент, создаваемый этими силами, разворачивает прилегающую к диагональной петле область молекулы, помещая ее в диапазон больших значений углов закручивания,  $36^\circ \pm 4^\circ$ . Это выше допустимой нормы для беспетлевых параллельных квадруплексов, но диагональная петля, за счет своей длины и эластичности, компенсирует эту аномалию.

Необходимо отметить, что такие значения углов закручивания не характерны для структур с петлями исключительно диагонального типа. В таких структурах геометрия квадруплексов довольно жестко ограничена значениями углов  $19^\circ \pm 4^\circ$ , что все же ближе к приемлемым для беспетлевых квадруплексов. Таким образом, можно заключить, что напряжения, создаваемые латеральными петлями, сильнее создаваемых петлями диагонального типа.

Важно, что часто используемый в описании квадруплексов паттерн син/анти-конформаций, в группе бимолекулярных квадруплексов с диагональными петлями представлен наибольшим числом возможных комбинаций. В то же время, именно в этой группе разброс значений углов закручивания квадруплексов наименьший. Это позволяет предположить, что доминирующим фак-

тором влияния на структуру квадруплекса является тип петель. Пропеллерные петли наиболее жестко определяют геометрию квадруплекса. Как в одномолекулярных, так и бимолекулярных тетраплексах с петлями исключительно пропеллерного типа значения углов закручивания находятся в пределе  $31^\circ \pm 3^\circ$ . Даже появление в этих структурах латеральных петель мало влияет на закручивание прилежащих к пропеллерным петлям областей – оно попадает в область значений  $28^\circ \pm 4^\circ$ , а значит вписывается в интервал предпочтительных значений для беспетлевых квадруплексов ( $27^\circ - 34^\circ$ ). Видимо, такие углы закручивания соответствуют оптимальной геометрии G-квартета, поскольку тетрады структур с петлями пропеллерного типа отличаются высокой планарностью от квадруплексов с диагональными петлями.

Нами исследованы все известные на данный момент G-квадруплексные структуры. Их кластеризуют не только по топологии; многие группы представлены последовательностями общего или родственного происхождения. Заметим, что в итоговую версию вошли не все структуры – часть из них обладает слишком специфичной топологией, чтобы составить представительную выборку. Для описания геометрии квадруплексов мы предложили два параметра: угол закручивания тетраплекса и степень планарности G-тетрад – и показали, что основным источником напряжений в структуре квадруплексов являются петли, соединяющие блоки G-тяжей. Петли латерального типа наиболее сильно изменяют геометрию G-квадруплексов, однако, их воздействие компенсируется добавлением в структуру петель диагонального и пропеллерного типа. Петли диагонального типа также достаточно жестко определяют структуру квадруплекса, однако создаваемые ими напряжения не столь велики, как в случае с латеральными петлями. Оптимальной же геометрией характеризуются квадруплексы с петлями пропеллерного типа.

Очевидный интерес для дальнейшего детального исследования динамики целостной молекулы НК представляют небольшие квадруплексы с дестабилизирующими петлями. Такой выбор позволит охарактеризовать весь спектр конформационных превращений молекулы в процессе ее функционирования.

## 3.4 Исследование динамики структур аптамера к тромбину

Аптамеры – это молекулы ДНК или РНК, отобранные из рандомизированных библиотек нуклеотидных последовательностей, проверенных на способность связываться с заранее заданным лигандом. Для полноты картины, надо заметить, что термин “аптамер” распространяется не только на НК – сейчас активно развивается область исследований по пептидным аптамерам. Важно отметить, что размер вариабельной части последовательности, не превышает 60 пар нуклеотидов. Размер лиганда бытне играет роли – это может быть как высокомолекулярное соединение, так и низкомолекулярное. Ключевой момент отбора – стабильность третичной структуры лиганда.

Найдены аптамеры даже к таким гигантским объектам как клетки А размер клетки не сопоставим с размером небольшой нуклеиновой кислоты! В этом случае отбор идёт не на клетку целиком, а на некоторый набор детерминант, доступных для взаимодействия с аптамером, на поверхности клетки. Процедура отбора подразумевает циклическое обогащение смеси и последующую амплификацию “успешных” последовательностей-кандидатов [303]. В связи с тем, что ни одну из исходных библиотек нельзя считать понлй, то и результаты селекции далеки от оптимальных. Несмотря на это, на сегодняшний день найдены десятки последовательностей ДНК и РНК, способных эффективно связываться с заданными мишенями.

Одним из таких достижений стал аптамер к тромбину. Тромбин – это сериновая протеиназа, которая реализует расщепление фибриногена и высвобождение фибрина – основного компонента тромба. Одна из причин успеха связана непосредственно со строением этого белка – наличие на его поверхности двух

зон с положительным зарядом., Так что образование эффективного комплекса между тромбином и отрицательно заряженным остовом НК вполне объясним. Первая плюс-зона тромбина называется экзосайт 1 – здесь происходит связывание фибриногена. Вторая плюс-зона называется экзосайт 2 – тут связывается гепарин.

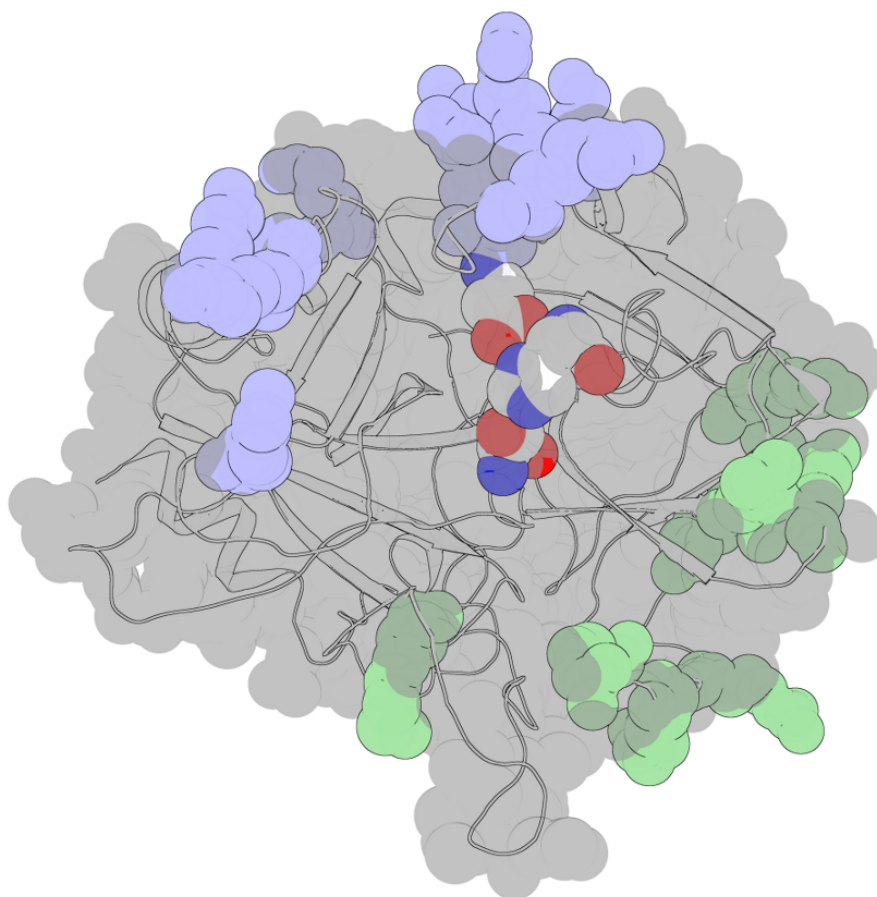


Рисунок 3.20: Третичная структура тромбина. Зеленым отмечен экзосайт 1, голубым экзосайт 2. Каталитический сайт представлен в виде сфер атомов, которые окрашены в следующие цвета: O – красные, C – белые, N – синие.

Известно множество вариантов аптамеров к тромбину [237; 304]. Самым изученным на сегодня является 15-звенный олигонуклеотид с последовательностью d(GGTTGGTGTGGTTGG), также известный как 15-ТВА. Ключевой структурной детерминантой этого аптамера является G-богатая последовательность, которая в растворе образует квадруплексную структуру из двух плоских G-квартетов, расположенных друг над другом. Образующий квадруплекс отно-

сится к мономолекулярному типу, т.е. состоит из одной цепи. Реализации подобной структуры кватреты обеспечивается тремя латеральными петлями. Две из них состоят из двух тиминовых нуклеотидов, а третья содержит последовательность TGT. Несмотря на обилие данных [266; 268; 305; 306] по этому аптамеру и его взаимодействию с тромбином, атомистический механизм узнавания тромбина аптамером не изучен.

На момент выполнения работы были доступны две модели комплекса аптамера с тромбином, которые основаны на данных ЯМР и РСА с [305; 306]. Структура аптамера в этих двух комплексах сильно различалась – несмотря на наличие G-кватретов – и направлением цепи и геометрией петель. Уже на тот момент считалось, что структура, полученная по данным РСА, не отражает реальную картину. Модель комплекса, построенная на основе данных ЯМР, больше соответствовала и значениям электронной плотности, полученным методом РСА, как исходя из величин R-фактора, так и при сравнении электронной плотности модели и экспериментальных значений [307].

Итак, есть два варианта модели третичной структуры комплекса тромбин–аптамер 15-ТВА. Из-за различий в структуре аптамера зона его контакта с тромбином тоже различна. В ЯМР-модели ТТ-петли аптамера взаимодействуют с экзосайтом 1 тромбина, в то время как в РСА-модели TGT-петля аптамера образует контакт с тем же экзосайтом 1 тромбина [306].

Попытки разрешить это противоречие с помощью сайт-направленного мутагенеза аптамера не удалось. Изменения в ТТ-петлях однозначно влияли на эффективность связывания с тромбином, но и мутации в TGT-петле тоже приводили к потере ингибирующей способности аптамера. Стехиометрия взаимодействия тромбина с аптамером также “не лежит на поверхности”. Вокс и Tasset [237; 308] предполагают стехиометрию комплекса как 1 : 1. На основе данных

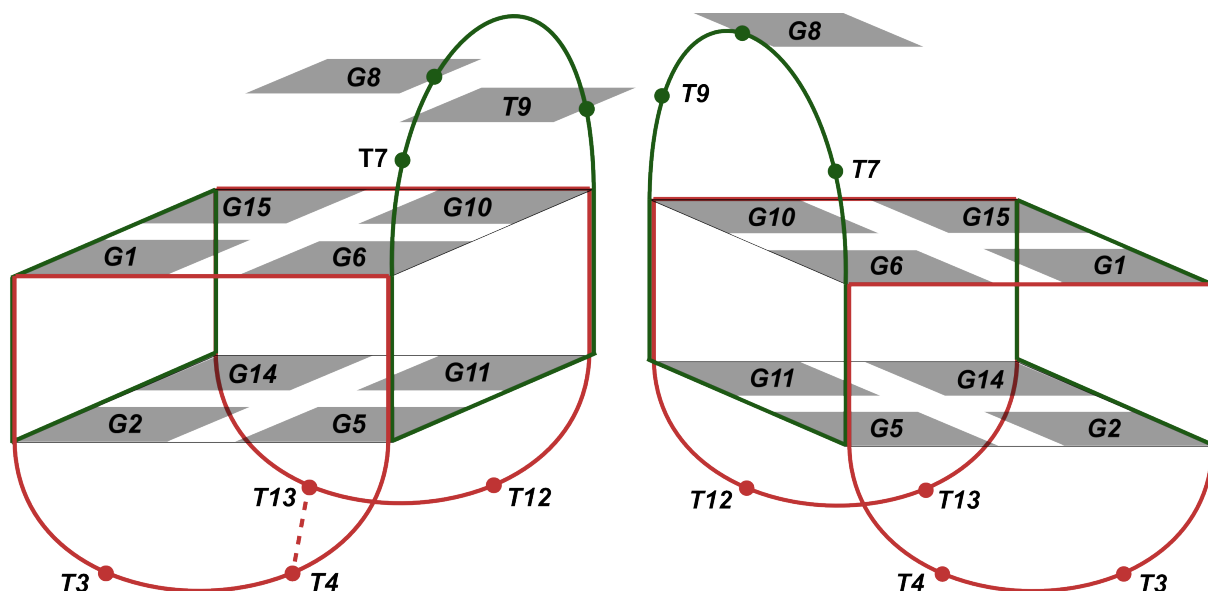


Рисунок 3.21: Слева: Модель на основе ЯМР. Справа: модель на основе PCA. Два G-квартета, верхний квартет образован остатками G1, G6, G10 и G15. Нижний квартет образован остатками G2, G5, G11, G14. Остальные нуклеотиды образуют три латеральных петли: одну TGT и две TT. Противоположные стороны квадруплекса образуют подобные бороздки. Красным выделены малые бороздки, синим выделены большие бороздки. Модели 15-ТВА отличаются как направлением цепи, так и топологией цепей. В ЯМР модели, два азотистых основания G8 и T9 из TGT петли образуют стекинг взаимодействия с верхним квартетом. Азотистые основания участвующие в стекинге отмечены серыми прямоугольниками. Основания T4-T13 образуют TT пару с двумя водородными связями, стабилизируя TT петли. В модели полученной на основе данных PCA, причём основание G8 вступает в стекинг взаимодействие с верхним квартетом. TT петли не взаимодействуют друг с другом.

PCA, Radmanabhan и др. [306] полагают, что 15-ТВА может взаимодействовать как по экзосайту 1, так и через границу кристаллической ячейки с экзосайтом 2 соседней молекулы тромбина. В пользу этой гипотезы свидетельствуют результаты, полученные Pagano и др. [309]. Используя метод изотермической сканирующей калориметрии, они обнаружили комплексы тромбин-аптамер со стехиометрией 2 : 1.

Очень важно, что число нуклеотидов в петлях аптамера невелико – и это уникальная возможность провести структурную аннотацию роли каждого нук-

леотида в формировании структуры аптамера как в свободном состоянии, так и при образовании комплекса с тромбином. Мы использовали моделирование МД в явном растворителе как инструмент для анализа роли “петельных” нуклеотидов в стабильности всей структуры. Нами исследована стабильность положения атомов в зоне контакта для двух вариантов структуры.

Моделирование МД активно используется для исследования структур, содержащих G-квадруплексы [310—315] и стало удобным инструментом для получения атомистической информации о динамических свойствах молекул. Метод МД основан на наблюдении за эволюцией структуры в условиях, максимально приближенных к физиологическим, в том числе и по временным рамкам. Важным моментом в подобного рода моделировании является достоверное описание взаимодействий, в которых участвуют атомы исследуемой молекулы. Совокупность параметров, с помощью которой описывают некий класс молекул, например биополимеры, называют силовым полем. При моделировании на основе силовых полей не учитывается электронная плотность атомов в явном виде, а используются разнообразные аналитические функции, параметры которых адаптируются для воспроизведения экспериментально полученных параметров системы.

Современные силовые поля, такие как parm99 [316; 317] и его модификации, достаточно точно описывают поведение G-квадруплексного стебля. Показано, что диагональные и пропеллерные типы петель моделируются методами МД не так достоверно, как квадруплексная часть структуры [310—312].

В 2007 году была опубликована модификация силового поля parm99 с названием parmbs0 [318]. Это обновленное описание сильно улучшило результаты моделирования В-ДНК при длительном времени наблюдения ториях. Длина траектории (время наблюдения за системой) является важным параметром



моделирования: чем дольше наблюдение за системой, тем более вероятно, что система пройдет все состояния, свойственные равновесным. Для нового поля parm30 показано, что на длительных траекториях это силовое поле хорошо описывает как канонические, так и неканонические структуры НК [318; 319]. Улучшен и алгоритм описания одноцепочечных участков ДНК – таких как петли в 15-ТВА. Однако это улучшение не означает сиюминутного получения идеальных результатов [312].

К началу проведения исследования была доступна информация о двух способах моделирования МД 15-ТВА. Pagano и коллеги показали [309], что моделирование в коротком временном диапазоне (5 нс) дает стабильные структуры как 15-ТВА, так и его модификаций в силовом поле parm98. Результаты моделирования согласовывались с данными ЯМР, которые и были источником по конформации олигонуклеотида. Это означает, что в 5нс-интервале наблюдения изменения структуры не происходило. Отметим, что этого времени недостаточно для наблюдения существенных структурных изменений в молекуле такого размера. Поведение петель согласовывалось с данными ЯМР. В работе, выполненной Jayaram и коллегами [240], показано, что нарушенная структура 15-ТВА возвращается к исходной после 2 нс моделирования в силовом поле OPLS-AA с записями для НК (ранее эти записи были добавлены нами к стандартному полю OPLS: см. <http://rnp-group.genebee.msu.su/3d/ff.htm>).

### **Построение моделей**

Для моделирования МД были выбраны две известные конформации аптмера 15-ТВА. Для анализа конформации по данным РСА были взяты координаты атомов соответствующие аптамеру в комплексе с тромбином (PDBID:1HUT)

[306]. В качестве стартовой конформации по данным ЯМР были использованы координаты аптамера из записи банка PDBID:148D [305].

Для сравнительного анализа влияния петель на динамические свойства структуры использована ЯМР-структура, из которой удалены записи о координатах атомов в петлях. Также исследованы конформации полученных обоими методами структур, в которых изменено исходное положение остатков в TGT-петле. Эти модификации названы соответственно TG(-T), T(-GT) и TG(+T). Знаки '+' и '-' означают изменение положения остатка относительно начальной конфигурации, где '+' – это вовлечение в стэкинг взаимодействие с прилежащим G-квartetом, в то время как, '-' – это разрушение существующего стэкинга с квартетом. Подобные модели получали на основе известных из PDB конформаций аптамера путем вращения остатки нуклеотидов G8 и T9 относительно углов  $\gamma$ ,  $\epsilon$  и  $\chi$  в программе молекулярной визуализации PyMol версии 1.1.

Внесённые изменения требовали оптимизации геометрии структуры в вакууме, что было реализовано модифицированным квази-Ньютоновским алгоритмом BFGS (Broyden–Fletcher–Goldfarb–Shanno)[320] с органичным использованием памяти. Таким образом исключили из рассмотрения слабые (высокоэнергетичные) взаимодействия. Для моделирования комплекса аптамера с тромбином как в ЯМР (PDBID: 1hut), так и PCA (PDBID: 1hao) конформациях, протонирование ряда остатков типа: аспаргат, глутамат и гистидин – проведено согласно экспериментальным данным Ahmed и коллег[321]. Все другие аминокислотные остатки тромбина имели состояние протонированности, соответствующее их pKa в физиологическом растворе. Для комплекса аптамера с тромбином в соотношении 1 : 2 на основе исходных структур комплексов были построены модели, где вторая молекула тромбина “поступала” из соседней ячейки – при этом сохранялась и группа симметрии, и расстояние между аптамером

и тромбином (не более 4Å). Из найденных “соседей” отбирали ту молекулу тромбина, которая взаимодействовала с аптмером по экзосайту 2. Все операции проводили в программе PyMol.

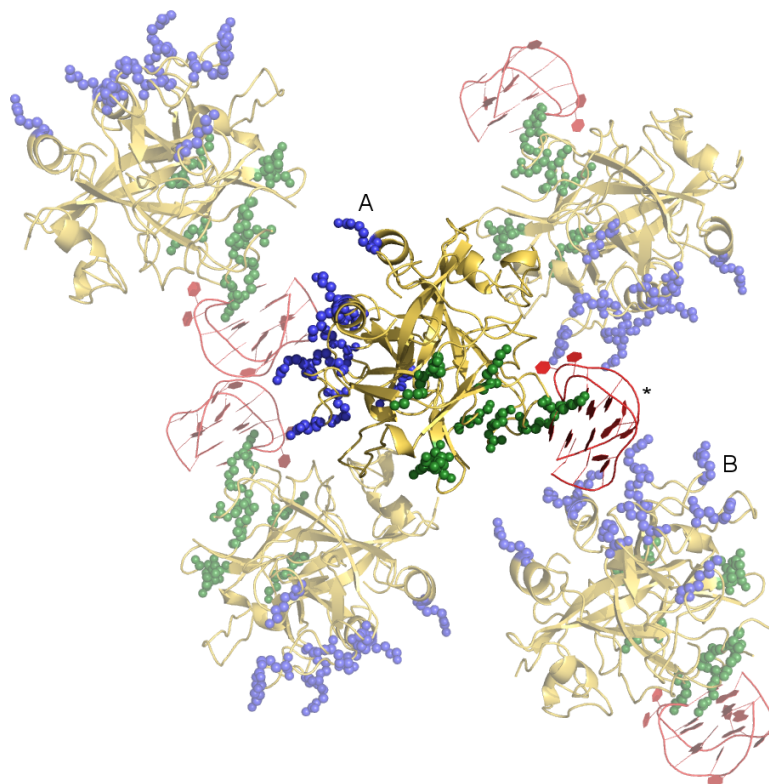


Рисунок 3.22: Модель комплекса тромбин–15-ТВА с соотношением 2 : 1. Взаимодействие тромбина (исходная структура отмечена жёлтым и буквой А) с аптамером (красный) через экзосайт 1 (зелёные сферы). Вторым тромбин отбирали на основании близости расположения экзосайта 2 (отмечен голубыми сферами) к аптамеру. Среди всех возможных “соседей” по кристаллической решётке (отмечены более бледным) только одна молекула тромбина удовлетворяет установленным критериям (отмечена буквой В)

### 3.4.1 Аптамер 15-ТВА, ЯМР и РСА конформации.

При моделировании молекулярной динамики происходит эволюция во времени выбранной конфигурации при учёте кинетической энергии атомов при заданной температуре. Это означает следующее: если конформация не соответствует оптимальной, то при моделировании, скорее всего, она будет значимо из-

меняться. причем с потерей геометрии, соответствующей структурному элементу. Для такой небольшой молекулы как 15-ТВА этим элементом является квадруплекс. В предварительном исследовании нами показано, что наблюдение за аптамером в воде в течение 100 нс не приводит к значительному изменению геометрии квадруплекса, но наблюдается разница в амплитуде движения гуаниновых оснований как в плоскости, так и вне плоскости квартета для РСА структуры. Очевидно, что возможное разрушение структуры квадруплекса лежит за пределами 100 нс–интервала моделирования МД. В расчётах мы сделали упор на наблюдение факта разрушения именно квадруплексной структуры. Выбранная нами длина траектории в 900 нс проверена в серии экспериментов, которые показали, что разрушение структуры квадруплекса происходит при отсутствии стабилизирующего катиона в центре квадруплекса. Как сказано выше, мы использовали два силовых поля для моделирования МД: parm99 и parmbs0. Хотя к началу исследования РСА–структуру считали ошибочной, мы всё-таки расценивали конформацию 15-ТВА в ней как одну из возможных. В ходе моделирования РСА-модель полностью потеряла структуру квадруплекса в обоих силовых полях. В обоих случаях потеря структуры проходила по одному сценарию: T4 и T13, которые принадлежат разным ТТ–петлям, вовлекаются в стэкинг-взаимодействия друг с другом, нарушая планарность нижнего G-квартета. В силовом поле parm99 структура продолжает сохранять элементы квадруплекса вплоть до 190-нс траектории МД, в то время как в силовом поле parmbs0 структура квадруплекса теряется в течении 10 нс (Рисунок 3.23).

ЯМР-структура сильно отличается от РСА-модели. На протяжении всего времени наблюдения, 900 нс, конфигурация аптамера остаётся близкой к исходной в обоих силовых полях – за исключением перестройки в контактах тиминов из ТТ-петель внизу квадруплекса. В силовом поле parmbs0 водородные связи

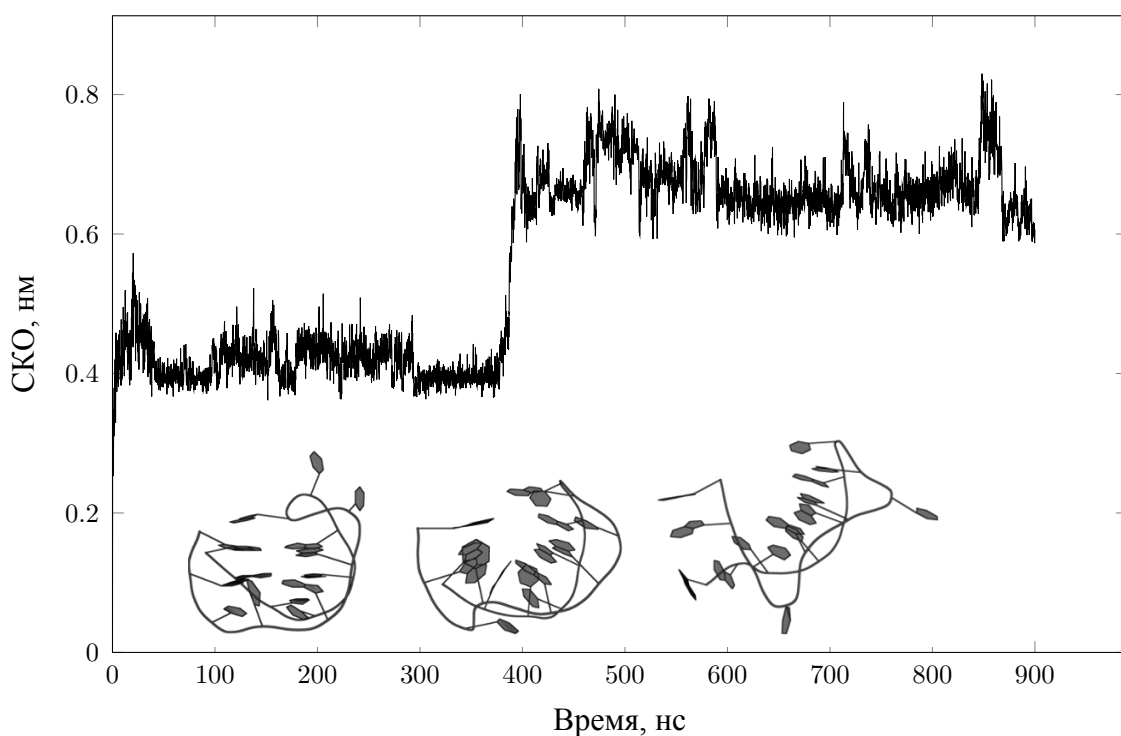


Рисунок 3.23: РСА конформации аптамера в соответствующие моменты траектории на графике временной зависимости СКО всех атомов текущей структуры относительно стартовой. Первое изображение соответствует стартовой структуре.

T4–T13 “переключаются” на T4–T12 и формируется контакт T3–T13. Однако подобная конфигурация TT-петель не очень стабильна и может переключаться обратно – на T4–T13 контакт. Такая подвижность TT-петель может иметь принципиальное значение для образования комплекса с тромбином – атомы оснований T4 и T13 аптамера не задействованы в связывании с белком и потому небольшое изменение их позиции не влияет на этот процесс. В поле parm99 контакты T4–T13 переключаются на T4–T12, в то время как T13 находится в стэкинге с нижним квартетом, а T3 образует выпетливание в сторону растворителя (Рисунок 3.24).

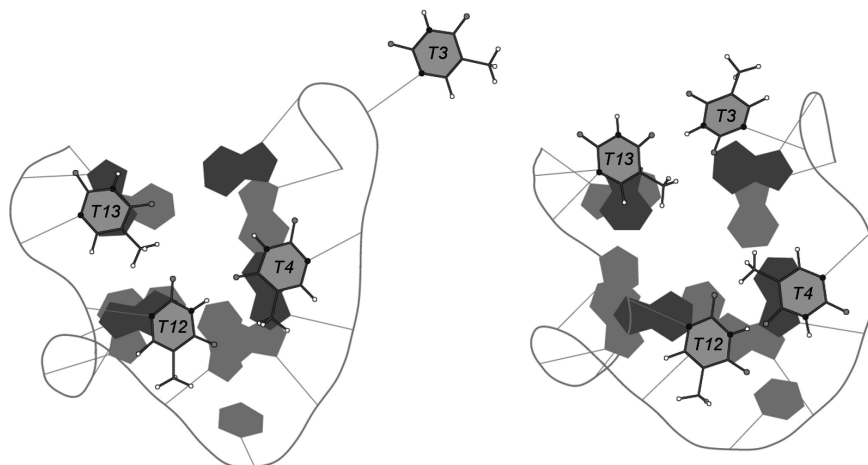


Рисунок 3.24: Вид снизу конечной структуры аптамера после моделирования структуры на основе ЯМР данных. Структурная организация ТТ-петель в ЯМР-модели зависит от используемого силового поля. В случае parm99 (слева) основание Т3 экспонировано в раствор. После нескольких перестроек (через 200 нс) основания ТТ-петель могут организовать такую же структуру, как и в случае силового поля parmbs0 (справа). Далее (через 820 нс) эта структура превращается в ту, которая представлена на рисунке. В поле parmbs0 все тимины склонны к стэкинг-взаимодействию с нижним квартетом и формированию двух ТТ-пар: Т3–Т13 и Т4–Т12. Такая геометрия петель достигается достаточно быстро и остаётся на протяжении всего времени моделирования.

### 3.4.2 Почему ЯМР структура более стабильна чем РСА модель?

Как обсуждалось ранее, интерпретация РСА-комплекса считалась ошибочной – так что потеря квадруплексной структуры при моделировании МД вполне логична. Это стало первым случаем “разрушения” представленной в банке PDB квадруплексной структуры. С одной стороны, моделирование просто предполагает, что одна условно изолированная молекула не имеет стабильной структуры, что нельзя напрямую экстраполировать на эксперимент, так представленность данной конформации может помешать её детекции в растворе методом ЯМР. С другой стороны, стабильность ЯМР-модели в при достаточно

длительном наблюдении позволяет предположить приемлемое качество описания аптамера в силовом поле  $\text{parmbc0}$ . Самое важное здесь то, с помощью моделирования можно понять, какие силы и факторы влияют на формирование структуры аптамера. Понятно, что данные такого рода не получают экспериментальным путём.

Используя разные стартовые конформации, мы попытались определить роль основных структурных элементов аптамера: стебля, петель и их взаимное влияние – на образование комплекса с тромбином. Основным структурным элементом аптамера 15-ТВА – это G-квадруплексный стебель, состоящий из двух квартетов. Согласно ЯМР-модели, этот стебель сверху экранирован от молекул воды двумя основаниями TGT-петли, G8 и T9. Снизу стебель экранирован TT-парой T4–T13. В случае PCA-структуры из TGT-петли только G8 вовлечен в стэкинг-взаимодействие с верхним квартетом, а два основания T взаимодействуют с нижним квартетом, при этом не образуя между собой водородных связей. Таким образом, можно выделить три структурных особенности, которые могут определять квадруплексную структуру аптамера:

- Водородные связи между основаниями T4–T13
- Два основания, G8 и T9, которые вовлечены в стэкинг-взаимодействия с верхним квартетом
- Катион металла между квартетами

ЯМР-модель, по нашим расчётам, обладает всеми этими особенностями, в то время как PCA-модель описывает только вовлечение в стэкинг основания T9 и катион в центре квадруплекса.

Начнем с исследования самой минимальной двухквартетной структуры – вариант аптамера без петель. Ранее проводили исследование [322] четырёх квар-

Система	Исследуемый фактор	Результат
Четырёх тяжевой квадруплекс состоящий из двух G-квартетов с катионом калия в центре.	Влияние петель на стабильность G-стебля	G-стебель был стабилен в течении всех 700 наносекунд наблюдения
Четырёх тяжевой квадруплекс состоящий из двух G-квартетов без катиона в центре*	Роль катиона в стабилизации G-стебля.	G-стебель потерял структуру в начале наблюдения
ЯМР модель с заменой катиона калия на катион натрия.	Влияние параметров описания катиона на поведение стабильной структуры 15-ТВА и её геометрию	Не замечено значительной разницы в поведении структуры в комплексе с разными катионами.
ЯМР модель без катиона внутри квадруплекса	Влияние стабилизирующего иона на структурную стабильность 15-ТВА.	Несмотря на высокую амплитуду флуктуаций, модель сохранила квадруплексную структуру. На 72 нс катион натрия из раствора проник в центр квадруплекса через ТТ петли (Рисунки 3.25,3.24 ) и стабилизировал молекулу до конца наблюдения.
ЯМР модель где основание Т9 ориентировано в раствор и не образует стэкинга с верхним квартетом (TG(-T) ЯМР) РСА модель, где основание Т9 образует стэкинг с G-квартетом (TG(+T) РСА)	Способна ли ЯМР модель сохранить структуру если только G8 из TGT петли образует стэкинг с G-квадруплексом? Может ли стэкинг взаимодействие одного основания сделать РСА модель стабильной?	Аптамер потерял структуру квадруплекса, все квартеты потеряли геометрию, не смотря на наличие пары Т4-Т13 в стартовой структуре Модель проявила гораздо лучшую стабильность и сохраняла квадруплексную структуру в течении 790 нс, в этот момент основания Т4 и Т13 образовали стэкинг взаимодействие, которое привело к нарушению планарности нижнего квартета и квдруплексная геометрия была потеряна

Таблица 3.1: Исследования влияния различных факторов на структурную стабильность 15-ТВА. \* Исходно катиона не было в центральной полости квадруплекса, но в растворе очевидно были катионы способные проникнуть внутрь квадруплекса.

тетных стеблей, но напрямую экстраполировать эти результаты на двухквартетную структуру нельзя из-за очевидного различия в балансе стабилизирующих сил. В таблице 3.1 суммированы результаты наблюдений по влиянию упомянутых факторов на стабильность структуры 15-ТВА.

### 3.4.3 Комплексы аптамера 15-ТВА с тромбином

**Комплекс 1:1.** РСА- структура комплекса аптмера с тромбином была взята из банка PDB (PDB ID: 1hut). Мы полагали, что белок может стабилизировать именно эту конформацию аптамера – т.е. тромбин узнаёт минорную конформацию 15-ТВА и сдвигает конформационное равновесие в сторону РСА-варианта.





Рисунок 3.25: Путь проникновения катиона натрия в центральную полость аптамера в системе, где стартовая структура не связана с катионом. Это изображение соответствует временному интервалу 60-80 нс. Катион (чёрные точки), движется – начало движения в левой части рисунка – вдоль сахарофосфатного остова и проникает внутрь через ТТ-петли. Далее катион проходит нижний квартет и в течение всего времени моделирования занимает позицию в центре квадруплекса. Атомы фосфора в остове отмечены серыми сферами.

Исследование динамических свойств комплексов тромбина с аптамером в РСА- и ЯМР-конформациях проводили с помощью моделирования МД со временем наблюдения 600 нс в силовом поле parmbsc0.

В стартовой конфигурации комплекса с РСА-конформацией аптамера тромбин связывается с TGT- петлей, а ТТ-петли экспонированы в растворитель. Аминокислотные остатки тромбина, способные быть донорами при образовании водородной связи, взаимодействуют не только с нуклеотидными остатками TGT-петли, но и с нуклеотидом ТЗ из ТТ-петли. Интересно, что именно это взаимодействие становится причиной последовательного разрушения структу-

Начальная модель			Конечная модель		
донор	водород	акцептор	донор	водород	акцептор
N3 (T9)	H3 (T9)	OH (Tyr117)	<i>N3 (T7)</i>	<i>H3 (T7)</i>	<i>OH (Tyr117)</i>
N (Ile79)	H (Ile79)	O1P (T9)	<i>N3 (T3)</i>	<i>H3 (T3)</i>	<i>OG (Ser72)</i>
N (Asn78)	H (Asn78)	O1P (T9)	<i>ND2 (Asn78)</i>	<i>HD22(Asn78)</i>	<i>O3' (T7)</i>
NH2 (Arg77A)	HH22 (Arg77A)	O4' (G10)	<i>ND2 (Asn78)</i>	<i>HD22 (Asn78)</i>	<i>O2P (G8)</i>
NH1 (Arg77A)	HH12 (Arg77A)	O4' (G10)	<i>N (Asn78)</i>	<i>H (Asn78)</i>	<i>O2P (G8)</i>
N (Arg77A)	H (Arg77A)	O3' (G8)	<i>N (Tyr76)</i>	<i>H (Tyr76)</i>	<i>O6 (G8)</i>
NH2 (Arg75)	HH22 (Arg75)	N7 (G1)	<b>NH2 (Arg75)</b>	<b>HH22 (Arg75)</b>	<b>O1P (G8)</b>
<b>NH2 (Arg75)</b>	<b>HH22 (Arg75)</b>	<b>O1P (G8)</b>	<b>NH1(Arg75)</b>	<b>HH12 (Arg75)</b>	<b>O1P (G8)</b>
<b>NH1 (Arg75)</b>	<b>HH12 (Arg75)</b>	<b>O1P (G8)</b>	<i>N (Arg75)</i>	<i>H (Arg75)</i>	<i>O4 (T3)</i>
NE (Arg75)	HE (Arg75)	N9 (G1)	<i>OG1 (Thr74)</i>	<i>HG1 (Thr74)</i>	<i>O3' (G1)</i>
NE (Arg75)	HE (Arg75)	N7 (G1)	<i>N (Thr74)</i>	<i>H (Thr74)</i>	<i>O4 (T3)</i>
ND1 (Hys71)	HD1 (Hys71)	O2P (G8)			

Таблица 3.2: Карта водородных связей в комплексе тромбина с РСА конформацией аптамера<sup>a</sup>. <sup>a</sup> Водородные связи которые сохранились в ходе моделирования МД отмечены жирным, вновь образованные отмечены наклонным шрифтом.

Начальная модель			Конечная модель		
донор	водород	акцептор	донор	водород	акцептор
N3 (T12)	H3 (T12)	OE2 (Glu77)	<i>ND2 (Asn 78)</i>	<i>HD22</i>	<i>O3' (T13)</i>
N3 (T12)	H3 (T12)	O (Glu77)	<i>NH2 (Arg 77A)</i>	<i>HH22</i>	<i>O1P (G14)</i>
OG (Ser153)	HG (Ser153)	O4 (T7)	<i>NH2 (Arg 77A)</i>	<i>HH22</i>	<i>O4' (G14)</i>
OH (Tyr117)	HH (Tyr117)	O1P (T13)	<i>NH1 (Arg 77A)</i>	<i>HH12</i>	<i>O (T13)</i>
N (Asn78)	H (Asn78)	O (T13)	<i>NH1 (Arg 77A)</i>	<i>HH12</i>	<i>O5' (G14)</i>
NE (Arg77A)	HE (Arg77A)	O1P (G14)	<i>NH1 (Arg 77A)</i>	<i>HH12</i>	<i>O4' (G14)</i>
<b>N (Tyr76)</b>	<b>H (Tyr76)</b>	<b>O4' (T4)</b>	<b>N (Tyr 76)</b>	<b>H</b>	<b>O4' (T4)</b>
<b>NH2 (Arg75)</b>	<b>HH22 (Arg75)</b>	<b>O (T4)</b>	<b>NH2 (Arg 75)</b>	<b>HH22</b>	<b>O4 (T13)</b>
<b>NH1 (Arg75)</b>	<b>HH12 (Arg75)</b>	<b>O4 (T13)</b>	<b>NH1 (Arg 75)</b>	<b>HH12</b>	<b>O (T4)</b>
NE (Arg75)	HE (Arg75)	O (T4)	<i>NH1 (Arg 75)</i>	<i>HH12</i>	<i>O4 (T13)</i>

Таблица 3.3: Карта водородных связей в комплексе тромбина с ЯМР конформацией аптамера<sup>a</sup>. <sup>a</sup> Водородные связи которые сохранились в ходе моделирования МД отмечены жирным, вновь образованные отмечены наклонным шрифтом.

ры квадруплекса. Паттерн водородных связей комплекса аптамера с тромбином сильно изменилась за время моделирования: их осталось только 17% (детальная информация представлена в Таблице 3.2). В отличие от РСА-конформации в ЯМР-модели аптамер образует контакты с тромбином через ТТ-петли, а ТGT-петля экспонирована в раствор. Как и в случае с моделированием аптамера в свободном состоянии, конечная структура хорошо согласуется со стартовой, за исключением ориентации Т7 (Рисунок 3.26)

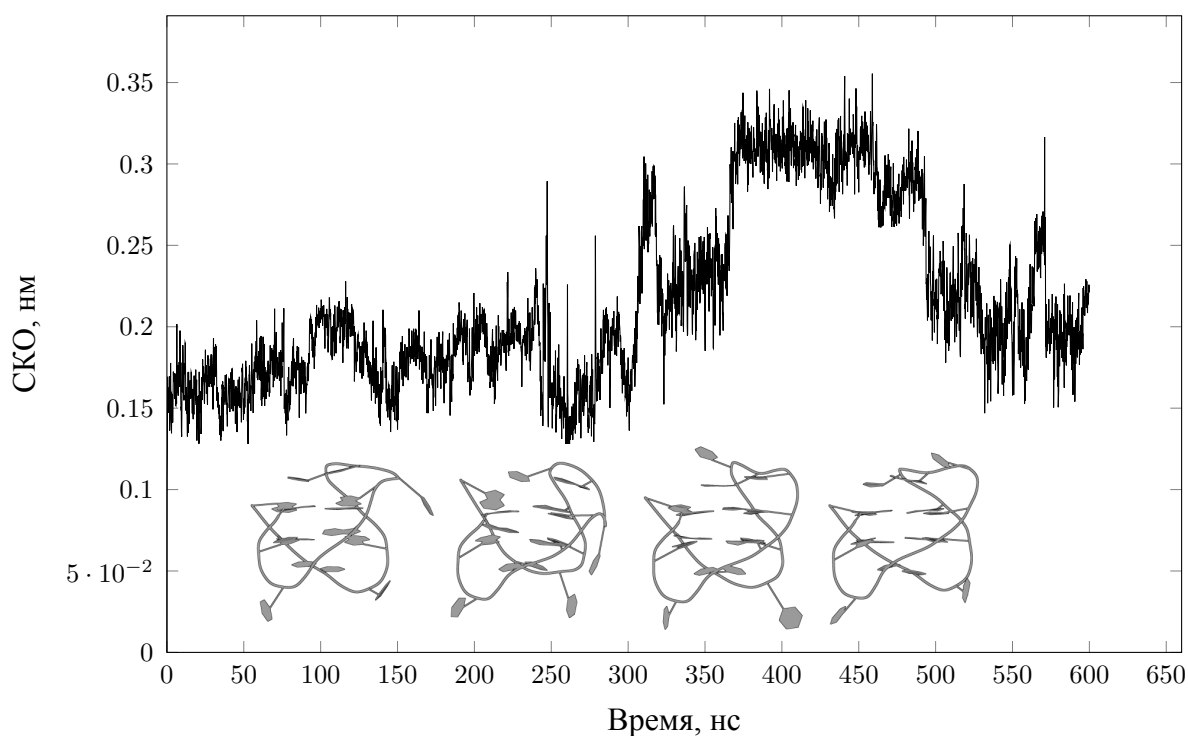


Рисунок 3.26: Динамическое поведение аптамера в ЯМР-конформации в комплексе с тромбином в соотношении 1 : 1. Изображения аптамера соответствуют выбранным моментам на графике СКО. Стартовая конформация аптамера взята из оригинального PDB-файла и использована как структура сравнения при расчёте полноатомного СКО. Изменение значения СКО в районе 400 нс связано с перестройкой TGT-петли: остаток T7 вовлечен в стэкинг с верхним G-квartetом.

В отличие от PCA в ЯМР-модели сохраняется 30% исходных водородных связей с тромбином, а изменение структуры аптамера было незначительным. Интересна роль основания T3, который не представлен в Таблице 3.3. Этот нуклеотид взаимодействует с тромбином через стэкинг с Tug76. Позже, в 2012 году, методом PCA это взаимодействие подтверждено [323]. Необходимо отметить, что большинство водородных связей, образуемых Arg77 с аптамером, также подтверждены данными последнего PCA. Так,  $N\delta$  Asn78 находится на расстоянии  $4\text{\AA}$  от атома 3'-O нуклеотида T13 и потенциально здесь возможно формирование водородной связи – что и было показано нами при моделировании. Таким образом, в ходе моделирования МД комплекса тромбин–15-ТВА нами обнаружены ранее не установленные водородные связи, часть которых позднее

подтверждена методом РСА. Из таблицы водородных связей видно, что остаток Arg77 способен образовывать больше водородных связей, чем Arg75 – но не одновременно. Именно возможность “переключения” водородных связей может приводить к эффективному связыванию аптамера белком. Для макромолекулярного комплекса достаточно выгодно “переключение” водородных связей за счет использования энергии теплового движения атомов биополимеров.

### **Комплекс 1:2.**

При моделировании МД комплекса 15-ТВА с тромбином в соотношении 1 : 2 выявлены незначительные изменения структуры аптамера, причем и в случае РСА, и ЯМР. Для ЯМР-модели изменения структуры были меньше, чем в РСА.

Если мы назовём тромбин из одной кристаллографической ячейки 'А', а вторую молекулу 'В', то в ЯМР-модели, то ТТ-петли аптамера взаимодействуют с экзосайтом 1 белка 'А', а петля TGT – с экзосайтом 2 белка 'В'. Эти взаимодействия не изменяются в течение 600-нс моделирования МД. Нуклеотид Т7 аптамера, который экспонирован в раствор в комплексе 1 : 1, в комплексе 1 : 2 взаимодействует с молекулой тромбина 'В'. а G8 и Т9 эффективно экранируют верхний квартет аптамера от положительно заряженных аминокислот и доноров водородной связи экзосайта 2 белка 'В'.

В случае РСА-модели взаимодействие аптамера с белком инвертировано, т.е. TGT-петля взаимодействует с экзосайтом 1 молекулы 'А', а ТТ-петли с экзосайтом 2 молекулы 'В'. При моделировании МД, практически с самого начала траектории, наблюдалось значимое нарушение планарности квартетов, но это не привело к потере структуры квадруплекса. Множественные контакты ТТ-петель с остатками белка в экзосайте 2 молекулы 'В' стабилизируют эту часть

аптамера, предотвращая деструктуризацию аптамера в рассматриваемом нами диапазоне времён.

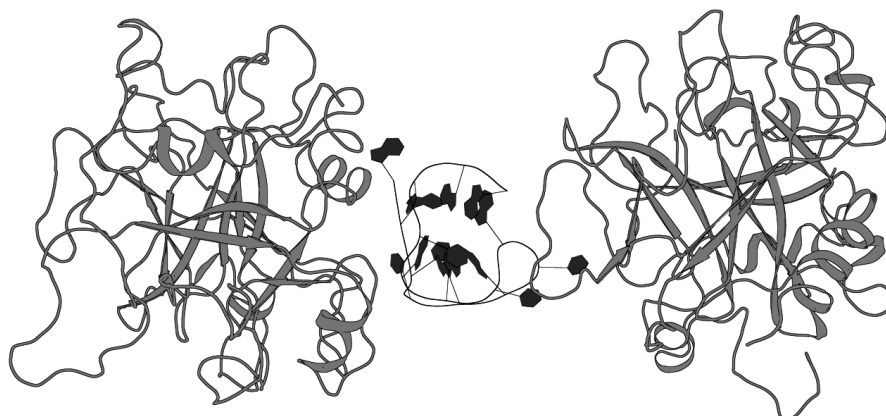


Рисунок 3.27: Структура комплекса тромбин-аптамер (2 : 1) как результат моделирования МД (аптамер в РСА-конформации).

#### **3.4.4 Структурная динамика G-стебля. Углы закрутки спирали отражают структурное напряжение**

Мы выбрали угол закрутки двух прилежащих G-квартетов как ключевой параметр описания самого G-стебля. Этот параметр представлен углом между двумя векторами, которые начинаются и заканчиваются в месте локализации атома C1' прилежащих остатков гуанина. Схема построения векторов приведена на Рисунке 3.28. Мы сравнивали значения угла закрутки, его флуктуации в ходе моделирования динамики 15-ТВА и комплексов тромбин–15-ТВА, а также в беспетлевом двухквартетном квадруплексе. Мы считаем, что беспетлевая структура не имеет напряжений, вызываемых петлями, и отражает оптимальное значение угла закрутки квартетов.

В ходе моделирования МД выявлены существенные различия в значении угла закрутки для аптамера в составе рассматриваемых структур, а также при

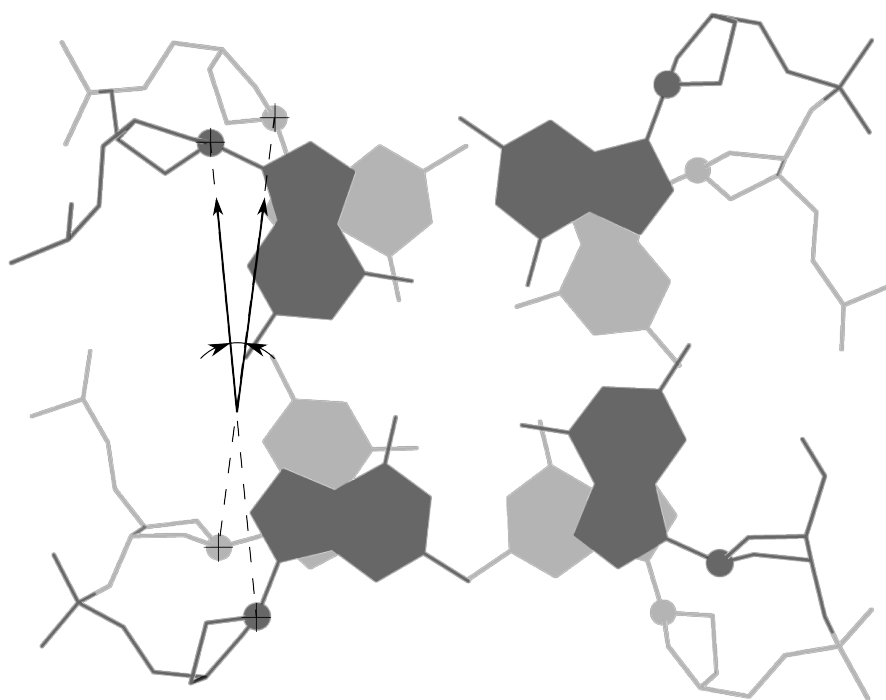


Рисунок 3.28: Схема определения угла закрутки квадруплекса.

сравнении разных структур – что отражает гибкость квадруплекса в планарных движениях (Рисунок 3.29). Также обнаружена разница между углами, относящимся к разным бороздкам квадруплекса. Используя как объект сравнения структуру без петель, мы оценили вклад петель и влияние белка на значения угла закрутки квадруплекса аптамера.

Заметим, что диапазон значений углов закрутки, который можно найти в опубликованной ЯМР-структуре (горизонтальные синие линии на Рисунке 3.29), значительно отличается от значений моделирования МД свободного аптамера (чёрные линии на Рисунке 3.29). Этот факт, вероятно, отражает неточности в параметрах силового поля, которые сдвигают значения угла закрутки относительно экспериментальных значений – как это известно для дуплексных ДНК в В-форме. Тем не менее относительное изменение значений, при переходе от системы к системе, эффективно отражает влияние петель и комплексообразования с белком. С другой стороны, диапазон углов закрутки для

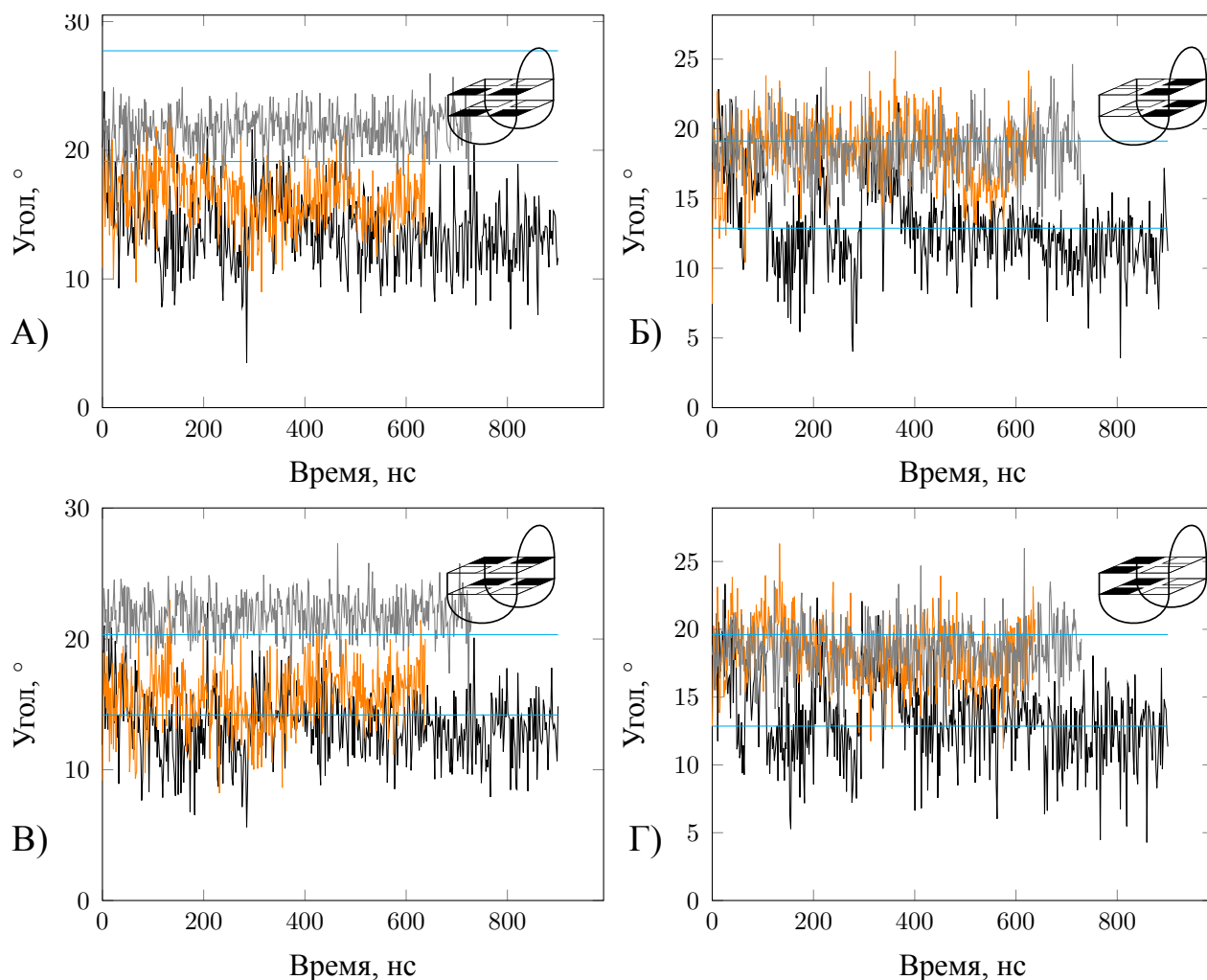


Рисунок 3.29: Временная зависимость для угла закрутки между квартетами.

Угол измерялся для разных систем в разных бороздках, что отражено в схематических вставках. Данные подвергнуты сплайн-интерполяции. Серые линии – это угол между квартетами в беспетлевой структуре. Черные линии – это ЯМР-конформация 15-ТВА в свободном состоянии. Коричневые линии – это комплекс ЯМР-конформации с тромбином (в соотношении 1 : 1). Две горизонтальные синие линии ограничивают диапазон значений, которые можно обнаружить в 12-ти моделях ЯМР-структуры (PDB ID:148d).

ЯМР-модели 15-ТВА может быть не точен, так как при оптимизации структуры на основе ЯМР-ограничений моделируется отжиг структуры от 1000°К до 75°К после оптимизации геометрии в наборе программ X-PLOR 3.1. Кроме того, ЯМР-структура (Рисунок 3.30) имеет выраженную деформацию плоскостей квартетов (непланарность), что может отражать неточность представления ЯМР-данных.

Проведенное нами моделирование МД дало структуру аптамера с остатками квартетов, расположенными в одной плоскости. Мы сравним здесь только изменения значений углов при переходе от беспетлевой структуры к структуре аптамера и к комплексу аптамера с тромбином. В целом, петли влияют на угол закрутки квадруплекса – и петли 15-ТВА заметно ограничивают гибкость квадруплекса. В результате мы наблюдаем значительное уменьшение угла закрутки и уменьшение амплитуды флуктуаций (чёрные и серые линии на Рисунке 3.29). Значения углов закрутки для беспетлевой структуры и ЯМР-модели 15-ТВА практически не пересекаются.

Полагая что, моделирование МД в диапазоне 100 нс по кинетике Арениуса может покрывать фазовое пространство молекулы с различием в энергии 7–8 Ккал/моль от минимума свободной энергии, можно оценить влияние петель. И это влияние значимо. Интересно, что образование комплекса с тромбином уменьшает негативное влияние петель на углы закрутки квадруплекса. Влияние тромбина на углы закрутки в основном распространяется на большие бороздки квадруплекса. Аптамер в комплексе с тромбином характеризуется значениями углов закрутки, более близкими к беспетлевой структуре, чем к структуре свободного аптамера 15-ТВА. В комплексе с аптамером тромбин компенсирует влияние петель на большую бороздку, в то время как подобного влияния на малую бороздку не наблюдается. ТТ-петли являются единственным структурным элементом, способным влиять на угол закрутки малой бороздки; и именно ТТ-петли претерпевают значительные изменения в ходе моделирования (Рисунок 3.30). Вероятно, что именно это напряжение, создаваемое ТТ-петлями в стартовой конформации, и является причиной неустойчивости РСА-модели аптамера, так же как и ЯМР-структуры TG(-Т), где Т9 удален из стэкинга с верхним квартетом.



Для проверки этого предположения две модифицированные модели были подвергнуты МД-моделированию. Первая модель – это ЯМР-структура после моделирования МД, где основание нуклеотида Т9 переориентировано из стэкинг взаимодействий с верхним квартетом в сторону растворителя. Вторая модель – это, тоже ЯМР-конформация после моделирования МД, но с переориентированием не только Т9, но и соседнего нуклеотида G8. Таким образом, верхний квартет полностью лишился стэкинг-взаимодействий с TGT-петлей.

В обеих структурах TT-петли имели ранее уравновешенную в МД конформацию – для минимизации влияния на напряжённость квадруплекса. Обе модифицированные структуры успешно сохраняли структуру квадруплекса в ходе моделирования. Основания, которые были переориентированы, восстановили стэкинг-взаимодействия с верхним квартетом.

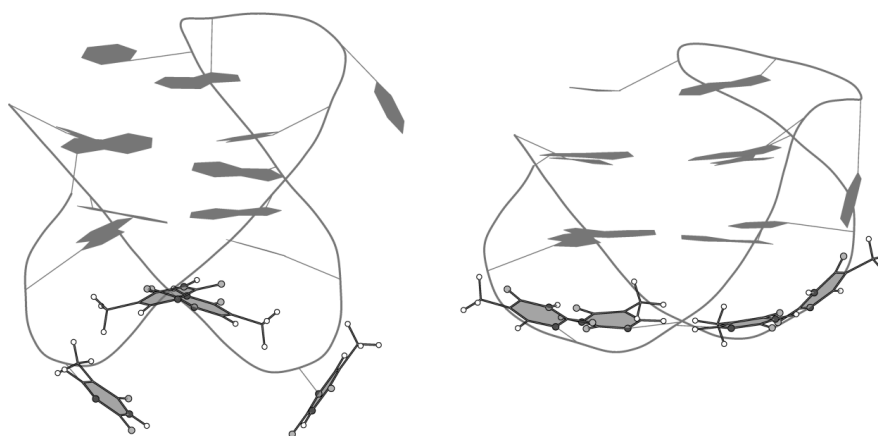


Рисунок 3.30: Структура аптамера 15-ТВА. Слева изображена ЯМР-структура 15-ТВА из банка PDB (PDB ID: 148d), справа – после моделирования МД для конформации слева. Тимины из TT-петель обведены чёрными линиями.

### 3.4.5 Структурная аннотация последовательности 15-ТВА

Тромбин, связывающий аптамер (15-ТВА), – это интригующий пример взаимодействия белка с ДНК, содержащей G-квадруплекс. В дополнение к зало-

женной отбором возможности связываться с тромбином и, следовательно, к высокой вероятности получить “статус” лекарственного препарата, 15-ТВА представляет важный объект исследования основных физико-химических свойств G-квадруплексов ДНК. Олигонуклеотид 15-ТВА содержит минимально возможное количество G-квартетов – всего два A беспетлевых квадруплексов с двумя квартетами не регистрировали до этого – и тут встает вопрос: как петли стабилизируют структуру 15-ТВА? Структурная стабильность G-квадруплексов напрямую зависит от присутствия катиона во внутренней полости, однако очевидно, что положение катиона внутри полости непостоянно, так что между аптамером и раствором должен происходить обмен катионами. Итак, молекула должна существовать какое-то время без катиона, а так как предположительно место связывания только одно, то формирование стабильной структуры со всего двумя квартетами достаточно удивительно. 15-ТВА содержит три коротких одноцепочечных петли, которые сохраняют сближенность цепей квадруплекса и одновременно необходимы для связывания с тромбином. Кроме очевидной энтропийной составляющей, определяющей сборку квадруплекса из одной или четырех цепей, эти петли могут оказывать прямое действие на G-квадруплексный стебель, и это действие может быть и стабилизирующим, и дестабилизирующим. Дестабилизирующий эффект может быть связан с длиной петель: петля может быть слишком короткой для того, чтобы поддерживать оптимальную геометрию квадруплекса. Стабилизирующее действие может оказывать дополнительный стэкинг и экранирование молекул воды, которые способны заменить любую водородную связь внутри биополимера без потери энтальпии системы.

Наряду с общими нетривиальными свойствами, понимание структуры 15-ТВА осложнено взаимоисключающими экспериментальными данными. Это

результаты по электронной плотности аптамера в комплексе с тромбином и данные ЯМР по взаимной сближенности некоторых протонов для свободного аптамера в растворе. Основное различие между этими данными – это ориентация основной цепи– ее ориентация в одном методе является зеркальным отражением в другом. Кроме этого есть различия в данных по положению и способу координации стабилизирующего катиона. Marathias и Bolton [266] предположили, что с 15-ТВА связываются два катиона калия, но есть информация [324—326] о том, что с 15-ТВА связывается только один катион. О связывании аптамера с тромбином тоже есть масса несогласующихся данных.

Моделирование МД относится к достаточно разработанным подходам для выяснения механизмов взаимодействия одновалентных катионов сНК, включая G-квадруплексные ДНК. Но надо признать, что есть очевидные ограничения в применении молекулярной механики для этих целей [311; 312]. Например, простое парное взаимодействие в силовом поле не может описать поляризацию электронного облака, что является ключевым эффектом при координации катионов. Из-за этого ограничения, мы решили не исследовать детально разницу между катионами калия и натрия при образовании комплекса с 15-ТВА или оценивать число катионов, одновременно связанных с аптамером. Надо отметить, что в ходе работы мы не наблюдали одновременного связывания двух и более катионов с аптамером. В принципе, желательно исследовать взаимодействие 15-ТВА с катионами с использованием гибридного метода молекулярной механики/квантовой механики [327; 328]. Эти методы к сожалению весьма вычислительно затратны и ограничивают время наблюдения в течении десятка пикосекунд.

Даже продуманное до мельчайших деталей исследование системы в газовой фазе методами стандартных методов квантовой механики не даст полного пред-

ставления о балансе сил в 15-ТВА. Но моделирование МД позволяет оценить влияние заряда в центре квадруплекса на его стабильность.

Мы сравнили динамику 15-ТВА в тех конформациях, которые соответствуют данным РСА и ЯМР по комплексу с тромбином (т.е. с различным ходом основной цепи), в комплексе с белком и в отсутствие петель. Общее время наблюдения превышало 12 мкс, что для такой небольшой молекулы как 15-ТВА считается значимым.

Исследование разбили на два этапа. Сначала были проанализированы две известные структуры аптамера 15-ТВА, а затем их реализация во всех известных моделях комплекса с тромбином. В результате моделирования МД показано, что в растворе, в свободном состоянии, может быть устойчива только структура аптамера, описываемая ЯМР, а предлагаемая методом РСА, даже при увеличении стэкинг-взаимодействия, теряет структуру квадруплекса. Установлена основная причина таких различий – разница во внутримолекулярных взаимодействиях, которые могут либо стабилизировать, либо разрушать структуру G-квадруплекса. Стабилизирующая роль, в основном, играет TGT-петля и катион внутри квадруплекса. Оценить относительный вклад каждого компонента пока не представляется возможным. ЯМР-модель аптамера оказалась достаточно стабильной даже в отсутствие катиона внутри квадруплексного стебля – более того, в ходе моделирования показано, что такой аптамер захватывает катион из раствора, что полностью стабилизирует его квадруплекс. Так называемая беспетлевая версия 15-ТВА без катиона в общем-то нестабильна, но за время моделирования ее структура сохраняется.

Итак, рассмотренные структуры 15-ТВА можно расположить в следующий ряд по уменьшению стабильности: ЯМР-модель > беспетлевая модель > РСА-модель. Из этого следует, что главный эффект на стабилизацию структуры

15-ТВА оказывают три петли (ЯМР-структура). Нами показано, что при искусственном ослаблении структуры 15-ТВА выведением нуклеотида Т9 (TGT-петля) из стэкинга с верхним квартетом, структура сильно дестабилизируется – даже при наличии катиона в центре квартета. Это указывает, но не доказывает, важную роль стэкинга нуклеотида Т9 с квадруплексом. Нельзя исключить и вероятность, что искусственное вмешательство в структуру аптамера может приводить к дестабилизации, не связанной непосредственно со стэкингом. И этот фактор в ходе моделирования МД выявить трудно. В ходе динамики дестабилизированной структуры ее разрушение напрямую зависит от подведения ТТ-петель. Когда основание Т9 не вовлечено в стэкинг с верхним квартетом, TGT-петля не может сбалансировать дестабилизирующее напряжение от ТТ-петель. Заметим, что более точное описание взаимодействия с катионом может несколько изменить баланс сил, хотя данные по мутагенезу 15-ТВА косвенно подтверждают сделанные нами выводы о роли TGT-петли.

Кроме того, полученные нами результаты свидетельствуют о важности стэкинг-взаимодействий между квадруплексом и минимум двух нуклеотидов TGT-петли, а именно: G8 и Т9. Только нуклеотид Т7 TGT-петли оказался не вовлечен во внутримолекулярные взаимодействия. По все видимости, его роль чисто структурная – создание достаточной длины петли, соединяющей тяжи квадруплекса. Ранее Smirnov и коллеги [329], а также Griffin и др. [330] выяснили, что оптимальная длина этой петли составляет три нуклеотида. Нами выявлено дестабилизирующее действие ТТ-петель на G-квадруплексный стебель. Один их факторов – изменение угла закрутки квадруплекса (Рисунок 3.29). Видимо, длины ТТ-петель недостаточно для оптимальной геометрии квадруплекса и это вызывает напряжение в структуре. Заметим, что в ходе моделирования ЯМР-структуры 15-ТВА ТТ-петли существенно меняют конформацию (Рису-

нок 3.30). Это значит, что в исходной модели в ТТ-петелях было напряжение, которое релаксировало в ходе моделирования МД. Может быть два объяснения наблюдаемому эффекту. Первое – стартовая структура не идеальна и в ходе моделирования происходит ее оптимизация. Второе – геометрия петель в квадруплексе может неточно воспроизводиться приближениями для построения силового поля, как это ранее показано в литературе [312]. Это означает, что в ходе моделирования МД происходит оптимизация структуры аптамера относительно стартовой геометрии, но это улучшение имеет смысл в терминах используемого силового поля. Интересно отметить, что при моделировании ЯМР-структуры с выведенным из стэкинга основанием Т9 при релаксированном состоянии ТТ-петель структура аптамера оказалась высокостабильной и в ходе моделирования нативное положение нуклеотида Т9 восстановилось. Аналогичный результат получен и при моделировании системы, где одновременно изменялись положения G8 и Т9. Таким образом, оптимизация геометрии ТТ-петель приводит к общей стабилизации структуры аптамера 15-ТВА.

ТТ-петли также играют критическую роль при связывании с тромбином. Если РСА-конформация аптмера верна, то связывание с тромбином происходит при участии ТGT-петли. Несмотря на множество контактов, которые образует ТТ-петли с белком и которые могли бы стабилизировать квадруплекс, дестабилизирующее действие этих петель четко выражено. С другой стороны, в случае ЯМР-модели 15-ТВА, его ТТ-петли взаимодействуют с тромбином и этот контакт уменьшает напряжение G-стебля аптамера по сравнению с динамикой свободного соединения в растворителе. Это видно по данным для значений угла закрутки квадруплекса (Рисунок 3.29). Из этого следует, что ЯМР-модель лучше отражает структуру комплекса 15-ТВА с тромбином [331]. При моделировании МД комплексов аптамер-тромбин в стехиометрии 1 : 2 показано, что эти струк-

туры стабильны и имеют право на существование. В комплексе 1 : 2 с аптамером в РСА-конформации показана роль ТТ-петель на взаимодействие с тромбином. Множественные контакты ТТ-петель с белком по экзосайту 2 предотвращают разрушение структуры аптамера в течение всего времени наблюдения (600 нс). Однако Tsiang и коллеги показали, что мутации в экзосайте 2 не изменяют ингибирующей активности аптамера, что свидетельствует о том, что аптамер не взаимодействует с экзосайтом 2 тромбина [332].

G-стебель в ЯМР-структуре 15-ТВА находится в явном напряжении из-за ТТ-петель (значимо изменены углы закрутки квадруплекса по сравнению с беспетлевой структурой). По результатам моделирования можно предположить, что ТГТ-петля стабилизирует квадруплекс, в то время как ТТ-петли вносят напряжение в структуру. Baldrich и др. показали [331], что при связывании тромбина наблюдается уменьшение структурного напряжения, вызванного ТТ-петлями в молекуле аптамера. Необходимо подчеркнуть, что все сделанные нами умозаключения строятся только на анализе моделирования МД, результаты которого хорошо согласуются с экспериментальными данными. Мы не предпринимали попыток провести моделирование с расчётом свободной энергии, так как у нас нет подходящего прямого метода для расчётов с приемлемой точностью. В 2010 году в проводимом нами исследовании использовано время наблюдения, которое на 1-2 порядка превосходило известные способы моделирования МД для G-квадруплексных ДНК. Считается, что достоверность МД напрямую зависит от длины траектории. Итак, если обобщить полученные при МД-моделировании результаты, то можно сделать следующие выводы.

1. В целом, петли стабилизируют структуру квадруплекса в составе 15-ТВА.

Однако ТТ-петли, с одной стороны, способствуют сближенности тяжёлой

квадруплекса, с другой – дестабилизируют его структуру. И это связано с их длиной. Влияние TGT-петли только стабилизирующее.

2. TT-петли участвуют в узнавании тромбина и при связывании белка уменьшается структурное напряжение в G-квадруплексе, приближая его геометрии к оптимальному состоянию.
3. Согласно данным моделирования МД, ЯМР-структура аптамера 15-ТВА лучше отражает реальную структуру его комплекса с тромбином и экспериментальные данные, чем РСА-структура.

Полученные результаты будут полезны при конструировании биосенсоров и дизайне лекарственных препаратов на основе аптамера 15-ТВА.



### 3.5 Взаимодействие аптамера 15-ТВА с катионами

Мы уже обсуждали, что правильно упакованный мономолекулярный G-квадруплекс без катиона в центральной полости может быть стабильным при моделировании МД в явно заданном растворителе в течение временного интервала, достаточного для захвата катиона щелочного металла из раствора. Вопрос о координации катиона в аптамере 15-ТВА тоже обсуждался. На момент исследования данные о положении и числе катионов, способных связаться с аптамером, были противоречивы. Мы детально исследовали координацию катионов в 15-ТВА и механизмы проникновения катионов во внутреннюю полость. Кроме того исследованы механизмы взаимодействия одновалентных катионов с G-квадруплексами. Ранее в экспериментальных данных выявлено, что длина и состав петель влияют на структурную стабильность G-квадруплексов и эффективность их сборки [333—336], Однако полного понимания факторов, которые направляют самосборку квадруплексной структуры, не было. Известно, что формирование квадруплексной структуры зависит от заряда и размера катиона, который связывается в центральной полости стебля. Так, катион натрия можно обнаружить в двух положениях: в центральной полости между квартетами и в плоскости одного из квартетов [337; 338].

Катионы большего размера такие как  $K^+$  и  $NH_4^+$  не могут находиться в плоскости кварта по стерическим соображениям – так что они стремятся расположиться симметрично между прилегающими квартетами – по центру. Большинство известных на сегодня G-квадруплексных структур получены методом ЯМР, возможности которого сильно ограничены для получения информации о положении катионов металлов. В целом современные экспериментальные методы дают ограниченную информацию о динамическом поведении катионов,

связанных в центральной полости квадруплекса. Моделирование МД становится удобным инструментом для исследования G-квадруплексных структур [117; 133; 143; 149; 339—346].

При МД-моделировании общая картина взаимодействия катионов с G-квадруплексными ДНК достаточно реалистична [347]. В ЯМР-исследовании структуры 15-ТВА в присутствии катионов стронция выявлено наиболее вероятное место связывания этого катиона – центральная полость G-стебля и позиция между плоскостями квартетов. При такой геометрии связывания катион координируется восемью атомами кислорода, Об, азотистых оснований гуанинов квадруплекса [348]. В другой работе исследовали взаимодействие с аптамером двух катионов калия [349; 350]. Один катион располагается между верхним квартетом и TGT-петлей (в дальнейшем мы будем называть это место “верхний сайт связывания калия”, или “верхний сайт”) или между нижним квартетом и TT-петлями (далее “нижний сайт связывания калия”, или “нижний сайт”).

В 2009 году Плавец и коллеги прямо показали место связывания катиона  $^{15}\text{NH}_4^+$  с 15-ТВА, используя ЯМР и другие спектроскопические методы [351]. Полагая, что квадруплекс в присутствии катионов аммония и калия имеет одну и ту же топологию, авторы исследовали различия в положении нуклеотидов петель. Оказалось, что структура петель мало зависит от природы выбранных катионов. Было показано, что катион аммония связывается в центральной полости квадруплекса и между квартетами (в дальнейшем мы будем называть это место связывания “центральным”). Кроме того был обнаружен сигнал, который соответствовал связыванию катиона в нижнем сайте – его интенсивность была в 10 раз слабее, чем сигнала из центрального сайта. Рядом с TGT-петлей катион аммония не зарегистрирован. Амоний может образовывать как водородные связи с карбонильными атомами азотистых оснований гуанина, так и электро-

статические контакты. Это накладывает очевидные ограничения на взаимное расположение доноров и акцепторов водородной связи. Катион калия образует как координационные, так и электростатические взаимодействия, требования которых к геометрии расположения участников менее жесткие. В дополнение, методом плавления с ультрафиолетовой детекцией, было показано, что аптамер к интегразе ВИЧ (Т30923), который образует сходный с 15-ТВА по топологии квадруплекс, вероятно связывает три катиона калия во всех трёх сайтах: нижнем, центральном и верхнем [352]. Вся эта информация не позволяет даже оценить вероятное положение катиона калия в структуре 15-ТВА.

Нами проведено исследование атомистической картины взаимодействия катиона металла с аптамером 15-ТВА с использованием методов МД в явно заданном растворителе (более 30 индивидуальных траекторий), гибридного метода молекулярной динамики/квантовой механики (МД/КМ) и экспериментальной валидации методом изотермической калометрии (ИКТ). В результате показано наличие трех сайтов связывания катиона в молекуле 15-ТВА. Два из них, верхний и нижний, не могут удерживать катион продолжительное время, так что он “засасывается” в центральный сайт между G-квартетами. Петля TGT-действует как ворота, которые направляют движение катиона во внутреннюю полость и после этого помогают заблокировать катион внутри полости в центральном сайте. Оказалось, что связывание катиона может происходить совершенно разными способами и зависит это от природы катиона. В исследовании показано, что для связывания катиона молекулой 15-ТВА структура квадруплекса должна находиться либо в собранном, либо в близкой к собранной форме. Катион стабилизирует собранную структуру аптамера и направляет близкую к собранной структуру к конечному состоянию. Можно думать, что основные ре-

зультаты, полученные в ходе этого исследования, скорее всего справедливы и для других квадруплексных структур.

### **3.5.1 Системы и особенности моделирования взаимодействия катионов с 15-ТВА**

#### **Классическое моделирование МД**

Прежде всего следует сказать, что все вычислительные эксперименты были разделены на логические части. Первый блок экспериментов относился к ЯМР-структуре 15-ТВА (PDB ID: 148d) [353]. Что бы увеличить область рассматриваемого фазового пространства, мы рассчитали 8 траекторий МД: 3 траектории с катионами натрия и 5 с катионами калия (Таблица 3.4). В этих траекториях исходно не было катионов в комплексе с аптамером 15-ТВА – только катионы в растворителе. Известно, что катионы натрия – слабые стабилизаторы квадруплексной структуры 15-ТВА. Но мы решили проводить моделирование с катионами натрия и калия.

Описание катионов на основе простых парных взаимодействий далеко от идеала [354]. По сравнению с данными квантово-химических расчётов простое описание в силовом поле неверно оценивает энергию связывания и преувеличивает степень крутизны функции Леннарда-Джонса в той части, где описывается отталкивание [347]. Так, при моделировании взаимодействия с центральной полостью квадруплекса эффективный размер катиона калия должен находиться между известными сегодня параметрами для катионов натрия и калия. Используемое в силовых полях семейства *parm* описание катиона калия дает завышенный размер катиона.

Второй набор траекторий призван ответить на вопрос: как изменится взаимодействие с катионами при удалении TGT-петли из ЯМР-структуры аптамера. Проведено семь вычислительных экспериментов для увеличения покрытия фазового пространства системы. Одна система рассчитывалась с катионом внутри квадруплекса и 10 – без катиона внутри аптамера, способного стабилизировать структуру (Таблица 3.4).

Третий набор систем призван помочь установить роль TT-петель. TT-петли были удалены из ЯМР-структуры 15-TVA и набор траекторий, аналогичный второму, был рассчитан. Как и ранее, пакет программ GROMACS [355] использовался для расчёта и анализа траекторий МД. Основные параметры моделирования были те же, что приведены ранее. Необходимо отметить, что в ходе моделирования МД двух систем с катионами калия мы наблюдали образование кристаллической структуры KCl. Это известный артефакт используемого силового поля, вызванный дисбалансом между описанием воды, катионов и анионов [134]. Несмотря на образование кристалла соли, избыток свободных катионов калия был достаточен для нейтрализации общего заряда системы – так что эта конфигурация могла бы быть использована в моделировании.

При исследовании связывания ионов натрия проблем, связанных с образованием кристаллов не возникало. Мы не ожидаем, что связывание катиона с аптамером может значительно измениться из-за использования модифицированных параметров для KCl в рамках расчётов по классической схеме парных взаимодействий, т.е. параметры катионов используемые в литературе приводят к идентичным результатам (рисунок 2 и обсуждение в статье Фадрны и коллег [347] ). Надо отметить, что эффективный размер катиона в силовом поле также определяется глубиной потенциальной ямы в функции Ленорда-Джонса,

системы для моделирования и начальные конфигурации	длина траектории, нс	количество траекторий	количество атомов				
			ДНК	Вода	K <sup>+</sup>	Na <sup>+</sup>	Cl <sup>-</sup>
15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	300	3	488	22356	–	25	11
15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	900	1	488	22356	–	25	11
15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: K <sup>+</sup> , Cl <sup>-</sup> )	300	2	488	22356	25	–	11
15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: K <sup>+</sup> )	300	3	488	22444	14	–	–
без TGT 15-ТВА со стабилизирующим Na <sup>+</sup> в центре (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	300	1	390	26540	–	29	19
без TGT 15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	60	10	390	26536	–	29	19
без TT 15-ТВА со стабилизирующим Na <sup>+</sup> в центре (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	60	1	358	17068	–	16	8
без TT 15-ТВА без стабилизирующего Na <sup>+</sup> (ионы: Na <sup>+</sup> , Cl <sup>-</sup> )	60	10	358	17068	–	16	8

Таблица 3.4: Выбранные системы для моделирования.

так как глубина и радиус частицы компенсируют друг друга в отталкивающей части функции.

### Гибридные молекулярно динамические/квантово-механические расчёты

Стартовая структура для моделирования связывания 15-ТВА с катионом калия была взята из банка PDB (PDBID:1c35 [350]) – в этой структуре с аптамером связано 2 катиона калия. Для моделирования структуры с катионом бария использована структура из записи PDBID:1rde [348] – эта структура получена для комплекса аптамера с катионом стронция (нет структурных данных по комплексу 15-ТВА с барием). Согласно данным спектроскопии кругового дихроизма, комплексы аптамера с катионами бария и стронция имеют очень схожие параметры. Так как исходно предполагалось, что результаты моделирования будут подтверждаться экспериментально, мы решили отказаться от исследования стронция, так как этот элемент накапливается в тканях и таким образом токсичен для живых организмов. Так что использование бария в этой части работы продиктовано необходимостью провести корреляцию между теоретически по-

лученными и экспериментальными данными. Построены пять систем для моделирования с учётом электронной плотности (Рисунок 3.31).

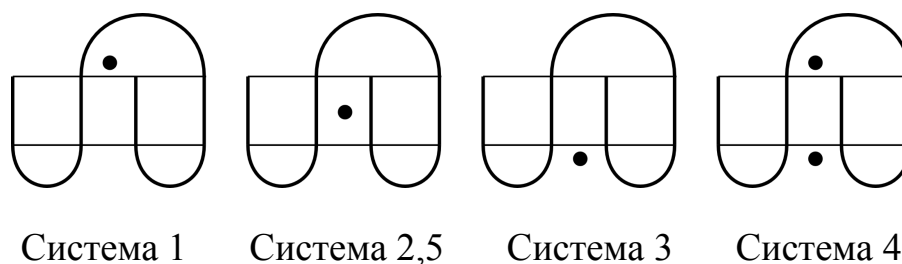


Рисунок 3.31: Системы для моделирования комплекса 15-ТВА с катионами методом МД/КМ.

Верхнее положение катиона (системы 1 и 4) было рассчитано как среднее арифметическое всех позиций катиона калия рядом с верхним квартетом (остатки G1, G6, G10 и G15) из опубликованной геометрии 1c35. Центральное положение катиона (системы 2 и 5) – это геометрический центр фигуры, образуемой положениями карбонильных атомов кислорода азотистых оснований квадруплексных гуанинов. Нижнее положение катиона (системы 3 и 4), как и верхнее, рассчитано, как среднеарифметическое для положений катионов калия, близких к нижнему квартету (остатки G2, G5, G11 и G14) в структуре 1c35. В системах 1–4 использован катион калия, а в системе 5 катион бария.

Молекулярно-механистическая часть системы описывалась силовым полем parmbs0 [31], а квантово-механическая – плоско-волновым вариантом теории функционала плотности с использованием псевдопотенциалов (PW-DFT) [356] со использованием формализма поляризованного спина [357] и функционала PW91 [358]. Взаимодействие между валентными электронами и оставшейся частью атома описывалось с помощью сверхмягких VDB-псевдопотенциалов (псевдопотенциалы Ван-дер-Билта) [359]. Квантовая часть системы состояла из атомов оснований всех нуклеиновых кислот кроме T7, и рассматриваемого катиона. Разделение механической и квантовых частей проводилось по N-

гликозидной связи с введением замыкающих водородов для насыщения оборванной связи. Общий размер квантовой системы составил 178 атомов (179 для системы 4 с двумя катионами калия). Ван-дер-Ваальсовы взаимодействия, которые плохо описываются теорией функционала плотности, корректировали с помощью аналитического потенциала Гримми [360]. Ячейка каждой системы была наполнена водой, представленной моделью TIP4P. Общий заряд системы нейтрализовался, так же как и в случае обыкновенных систем, заменой молекул воды на соответствующие ионы. Вновь добавленные молекулы воды и ионы уравнивались вокруг неподвижного комплекса 15-ТВА с катионами в течении 100 пс.

Для подготовленных систем было запущено моделирование МД/КМ с использованием программных пакетов GROMACS и CPMD и специального интерфейсного кода [361] для эффективного описания электростатических взаимодействий между квантовой и механическими системами. Длительность наблюдения за системами составляла 1.5 пс с шагом интегрирования 0.12 фс. Температура системы, 310 К, контролировалась с помощью алгоритма Нуза-Хувера [362]. Использование сверхмягких псевдопотенциалов позволило установить радиус обрезания плоских волн на значение  $30 R_u$ . Квантовая система представлена кубической формой с длиной ребра  $40 R_u$  и волновая функция квантовой системы описывалась с помощью 90000 плоских волн.

### **3.5.2 Молекулярно механические аспекты связывания катионов с 15-ТВА**

Вероятность проникновения и предполагаемая траектория движения катиона к месту связывания в 15-ТВА мы изучали методом моделирования МД с



описанием взаимодействий согласно силовому полю parmbs0. Мы моделировали полную ЯМР-структуру аптамера 15-ТВА в водном растворе, содержащем соли NaCl, KCl или просто катионы калия (без ионов Cl<sup>-</sup>) для нейтрализации заряда основной цепи аптамера. Во всех случаях стартовое положение катионов было случайным и квадруплексная часть аптамера не была связан ни с катионом. Во всех восьми вычислительных экспериментах (Таблица 3.4) молекула аптамера 15-ТВА “захватывала” катион из раствора и связывалась в центральном сайте. Мы отметили, что первоначальное связывание катиона с аптамером является Скорость-лимитирующей стадией процесса была первая стадия связывания. Рассмотрим детально механизм захвата каждого типа катионов.

### **Захват катиона натрия**

Так как радиус катиона натрия меньше радиуса катиона калия, логично предположить, что на пути его продвижения в центральный сайт квадруплекса и меньше препятствий. В двух из трёх случаев катион натрия сначала связывался с верхним связывающим сайтом. В одном случае первичное связывание происходило в районе ТТ–петель, после чего катион проникал через пору между ТТ-петлями в центральный сайт. В рассмотренных нами траекториях МД первичное связывание происходило на 1,112 и 61 нс. Время, затраченное катионом натрия на связывание в верхнем сайте, составило 1 и 4 нс, в то время как в нижнем сайте катион провёл 7 нс. Для проникновения катиона в центральную полость через верхний сайт необходимо соблюдение нескольких условий (Рисунок 3.32): 1) азотистые основания G8 и T9 должны изменить своё положение, чтобы дать доступ к карбонильным атомам кислорода гуаниновых оснований верхнего квартета; 2) азотистые основания верхнего квартета должны разойтись для того, чтобы сформировать пору достаточного размера для свя-

зывания катиона в плоскости квартета. Только при выполнении этих условий катион натрия может переместиться в центральный сайт связывания внутри квадруплексного стебля и занять там стабильное положение.

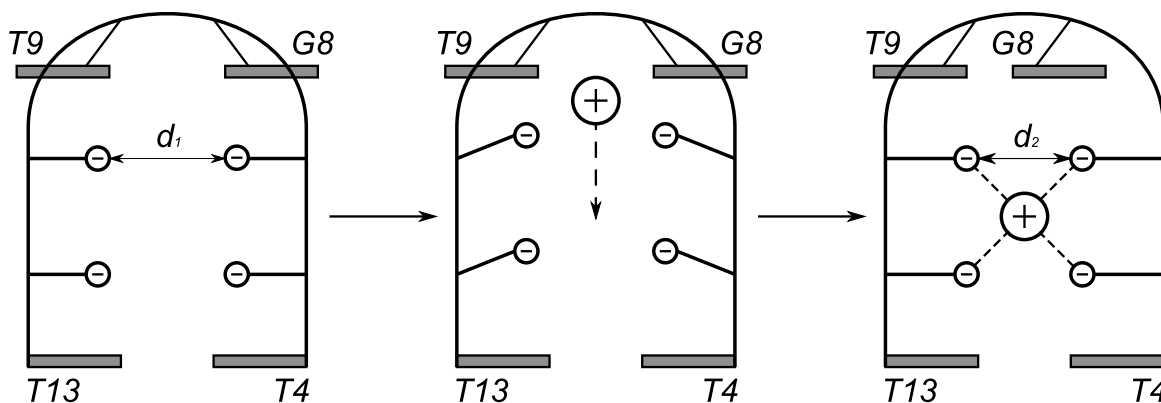


Рисунок 3.32: Схема движений азотистых оснований аптамера 15-ТВА при проникновении катиона в центральный сайт связывания через верхнюю часть структуры.

Динамика проникновения катиона натрия внутрь центральной полости через верхнюю часть структуры представлена на Рисунке 3.33. Площадь поры, которая образуется в центре верхнего квартета, зависит от положения катиона. Если катион находится в центральном сайте, то размер поры уменьшается относительно того состояния, когда катион находится вне центрального сайта. В рассматриваемой системе проникновение катиона в центральную полость произошло на 112 нс. Это стало возможным благодаря смещению основания G8 и появлению доступа к карбонильным атомам кислорода (Рисунок 3.33, нижний график). Катион преодолевает двое ворот: первые формируются остатками TGT-петли, а вторые G-основаниями верхнего квартета. После проникновения катиона в центральный сайт связывания все ворота закрываются, “запирая” тем самым катион внутри аптамера. Поверхность поры, образуемой остатками гуанина, становится меньше и амплитуда колебаний снижается (Рисунок 3.33, верхний график после 112 наносекунды). Подобным изменениям подвергается

и G8: центр масс азотистого основания приближается к центру масс азотистых оснований верхнего квартета, т.е. G8 накрывает пору в верхнем квартете (Рисунок 3.33, нижний график после 112 наносекунды).

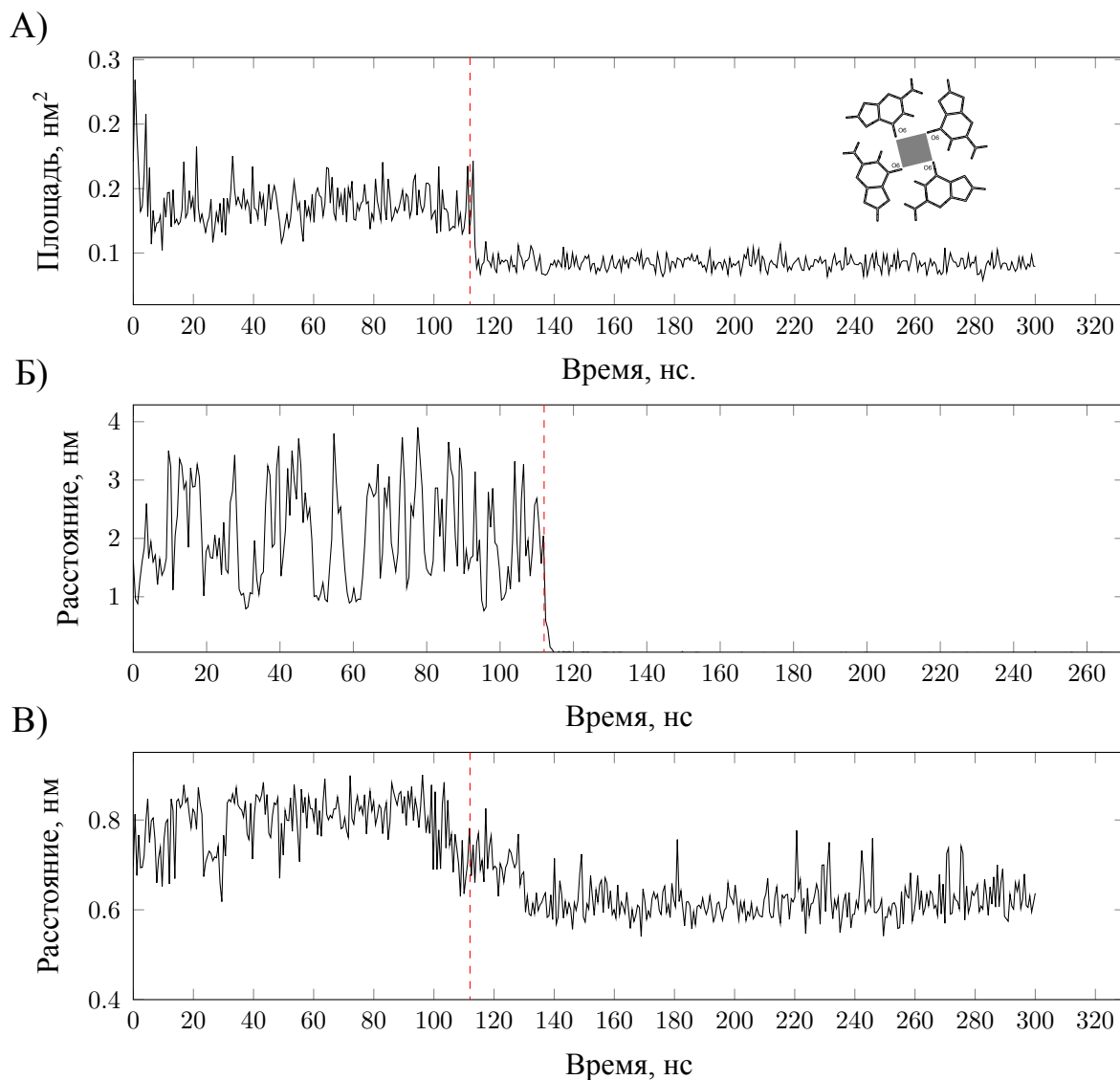


Рисунок 3.33: Динамика структуры аптамера 15-ТВА при проникновении катиона натрия в центральную полость квадруплекса через верхнюю часть структуры. (А) Площадь фигуры, образуемой атомами Об верхнего квартета. Схема приведена в верхней правой части графика – серым отмечена фигура, для которой вычислялась площадь. (Б) Расстояние между катионом натрия и центром масс восьми атомов Об, принадлежащих квадруплексному стеблю, т.е. центральному сайту связывания. (В) Расстояние между центром масс азотистого основания остатка G8 и центром масс атомов Об верхнего квартета квадруплекса. Вертикальная прерывистая линия отмечают момент проникновения катиона в центральный сайт связывания. Аналогичная хема для Т9 не показана.

Очень похожий процесс происходит при проникновении катиона натрия в аптамер 15-ТВА через нижнюю часть структуры, то есть через ГТ петли (Рисунок 3.34).

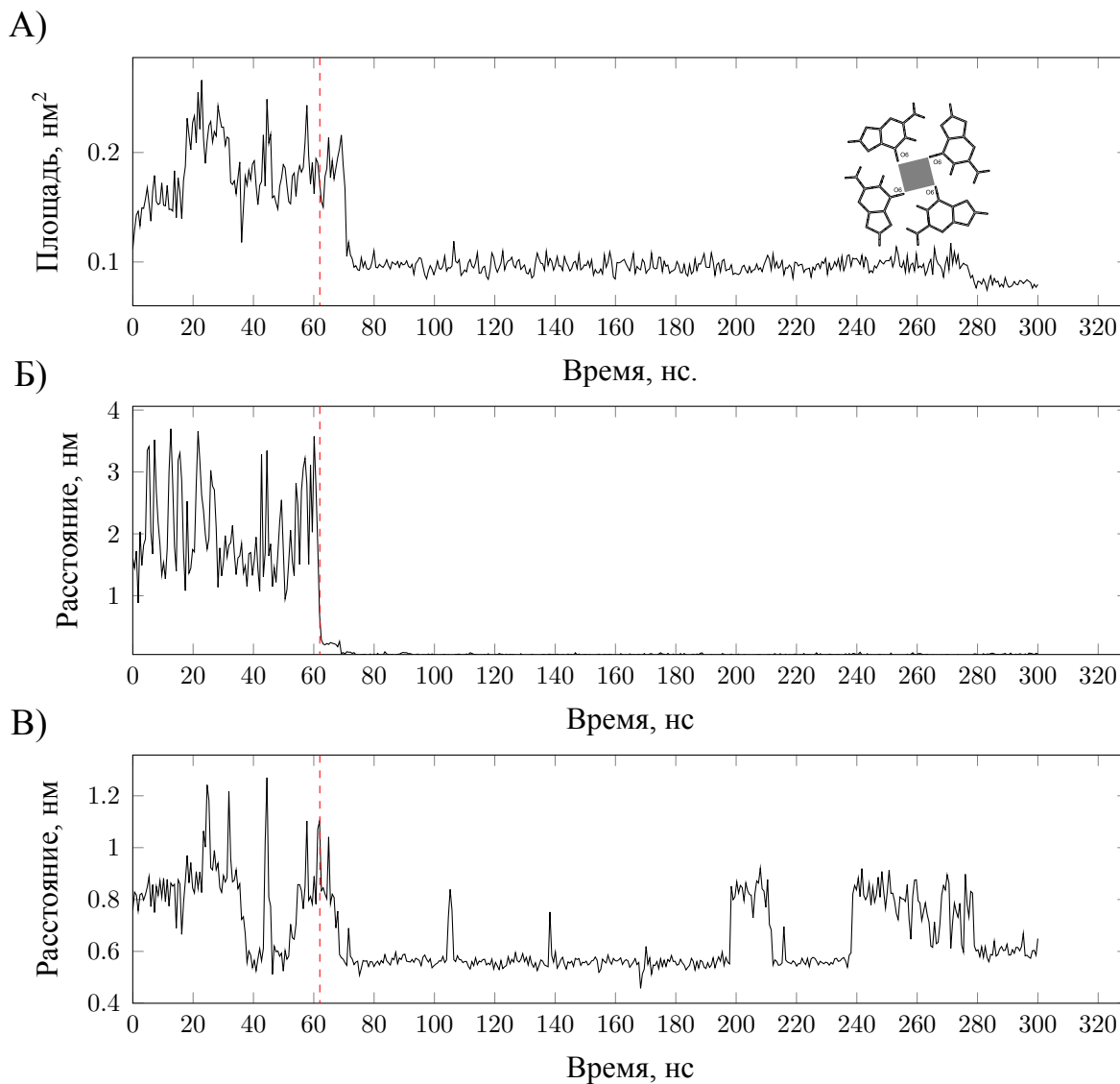


Рисунок 3.34: Динамика структуры аптамера 15-ТВА при проникновении катиона натрия в центральную полость квадруплекса через нижнюю часть структуры. Б) Расстояние между катионом натрия и центром масс восьми атомов Об, принадлежащих квадруплексному стеблю (центральному сайту связывания). А) Изменение площади фигуры, образуемой атомами Об верхнего квартета. В) изменение расстояния между центрами масс азотистых оснований Т4 и Т13. Проникновение катиона внутрь квадруплекса наблюдалось на 62-й нс: прерывистая линия опускается до минимальных значений, в этот момент все ворота были открыты, после того, как катион занял своё положение в центральном сайте, ворота захлопнулись.

## Механизм дегидратация катионов при проникновении катионов в центральный сайт аптамера

Моделирование молекулярной динамики позволяет следить шаг за шагом за процессом десольватации катиона во время его движения к центральному сайту связывания аптамера. Для визуализации этого процесса мы следили за значениями круговой функции распределения плотности (RDF) воды вокруг изучаемого катиона. Нормальное распределение воды вокруг катиона получено, но основе участка траектории длиной 1 нс, где катион не взаимодействовал с аптамером. Функция RDF имеет чёткий пик на расстоянии  $2.4\text{Å}$ , плечо пика заканчивается на расстоянии  $3.2\text{Å}$  от центра катиона. Суммарное число молекул, которое попадает в сферу указанного радиуса, можно считать числом молекул в первой сфере окружения катиона. Мы выбрали этот параметр как показатель степени гидратированности катиона (Рисунок 3.35).

Для расчёта изменения числа молекул воды вокруг катиона во время процесса связывания его с аптамером мы разбили траектории на фрагменты по 100 пс. Для каждого фрагмента мы рассчитывали и значения функции RDF, и число молекул воды в окружении катиона. Для визуализации процесса дегидратации построены зависимости изменения числа молекул воды вокруг катиона от расстояния до центрального сайта связывания. На Рисунках 3.36, 3.37 отображен процесс дегидратации катиона натрия при проникновении через TT-петли и TGT-петлю соответственно.

Исходно в растворителе с катионом взаимодействует около 6 молекул воды, когда катион начинает приближаться к TGT-петле, он постепенно теряет три молекулы. Проходя TGT-петлю, катион теряет ещё одну молекулу воды. После этого катион продолжительное время движется около поры, образуемой верх-

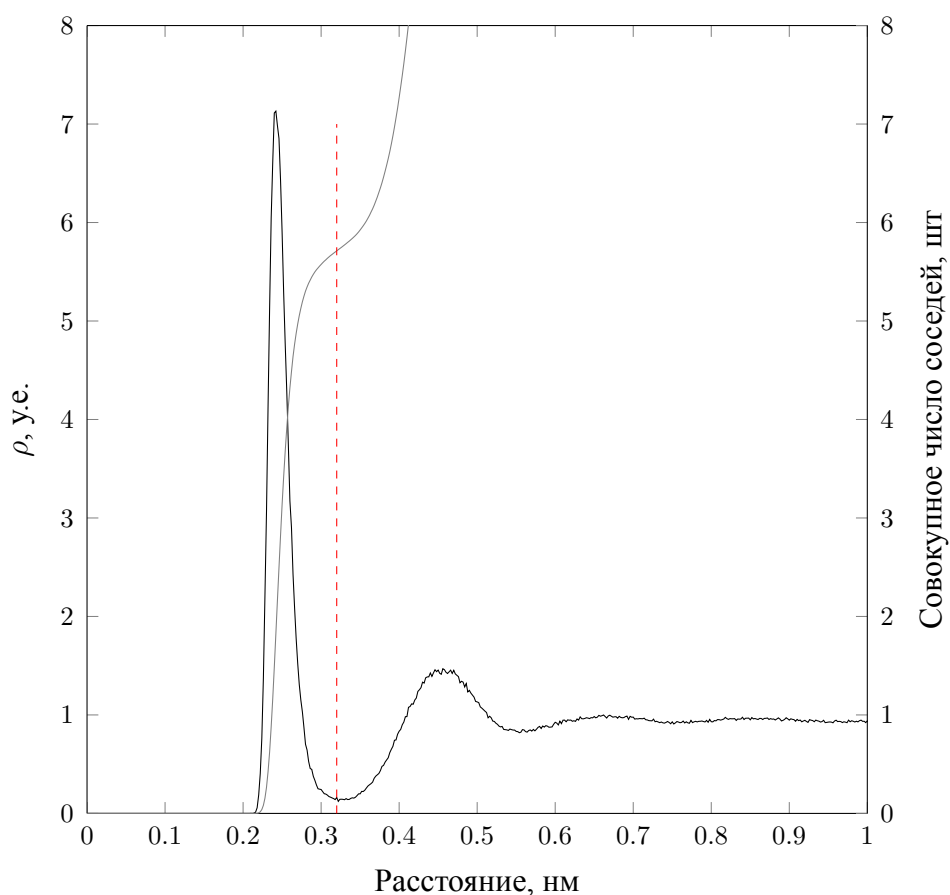


Рисунок 3.35: Гидратация катиона натрия в водном растворе. Чёрной сплошной линией отмечена функция кругового распределения плотности воды вокруг катиона натрия в растворе. Серой сплошной линией отмечено изменение суммарного количества молекул воды в сфере заданного радиуса с центром в месте положение катиона. Пунктирная линия отображает выбранный нами радиус для подсчёта молекул воды вокруг катиона.

ним квартетом, и во время этого движения постепенно теряет две оставшиеся молекулы воды – и только потеряв всю воду, катион проваливается внутрь квадруплекса. В случае движения через ТТ-петли катион продолжительное время флуктуирует рядом с ТТ-петлями и нижним квартетом квадруплекса, во время этого движения он теряет 2-3 молекулы воды из своего окружения. Приближение к плоскости нижнего квартета приводит к потере ещё двух молекул, а последняя теряется при прохождении плоскости нижнего квартета.

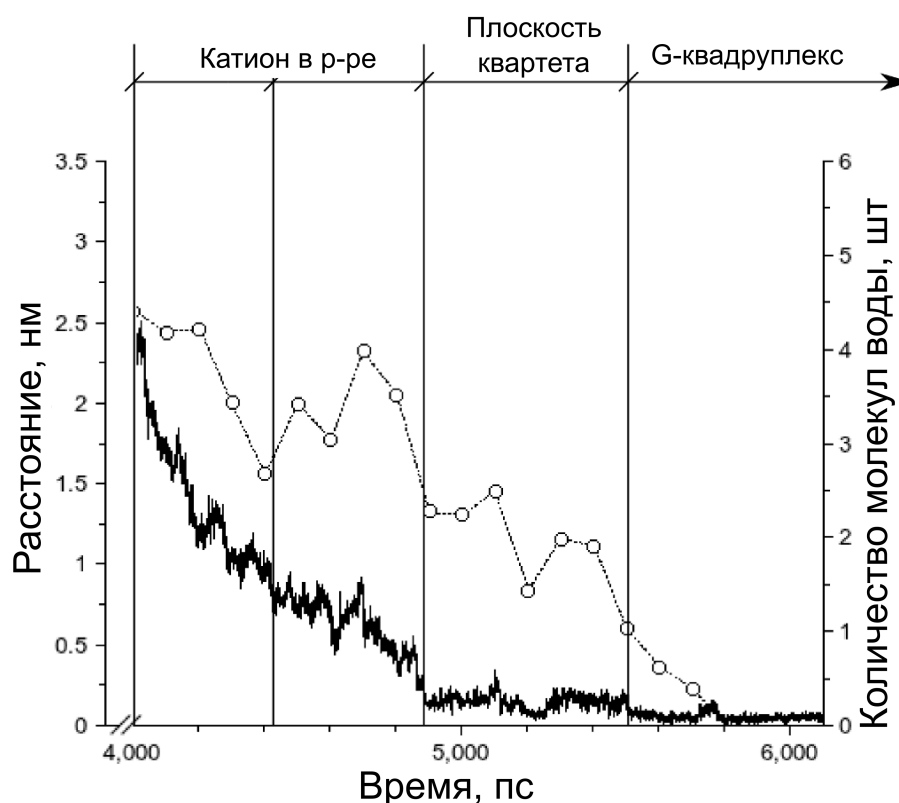


Рисунок 3.36: Проникновение катиона натрия в центральный сайт через ТТ-петли. Сплошной линией отмечено расстояние от катиона до центрального сайта связывания. Пунктирной линией отмечено число молекул воды вокруг катиона.

### Моделирование взаимодействия катиона калия с аптамером 15-ТВА

Системы, в которых моделировалось взаимодействие катиона калия с аптамером 15-ТВА, содержали лишь то количество катионов калия, которое было необходимо для нейтрализации заряда системы, обусловленного отрицательным зарядом сахарофосфатного остова ДНК. В двух случаях из трёх катион сначала связывался с нижним сайтом. Это наблюдалось на 10-й и 44-й нс наблюдения. Катион провёл в одном случае 4, а в другом 1 нс в нижнем связывающем сайте, после чего проник в центральный сайт связывания. В третьей траектории аптамер не смог захватить катион из раствора до 149-й нс наблюдения. К этому времени структура квадруплекса стала разрушаться, в результате чего основания G1 и G2 вышли из состава G-стебля (Рисунок 3.38). В этот момент катион

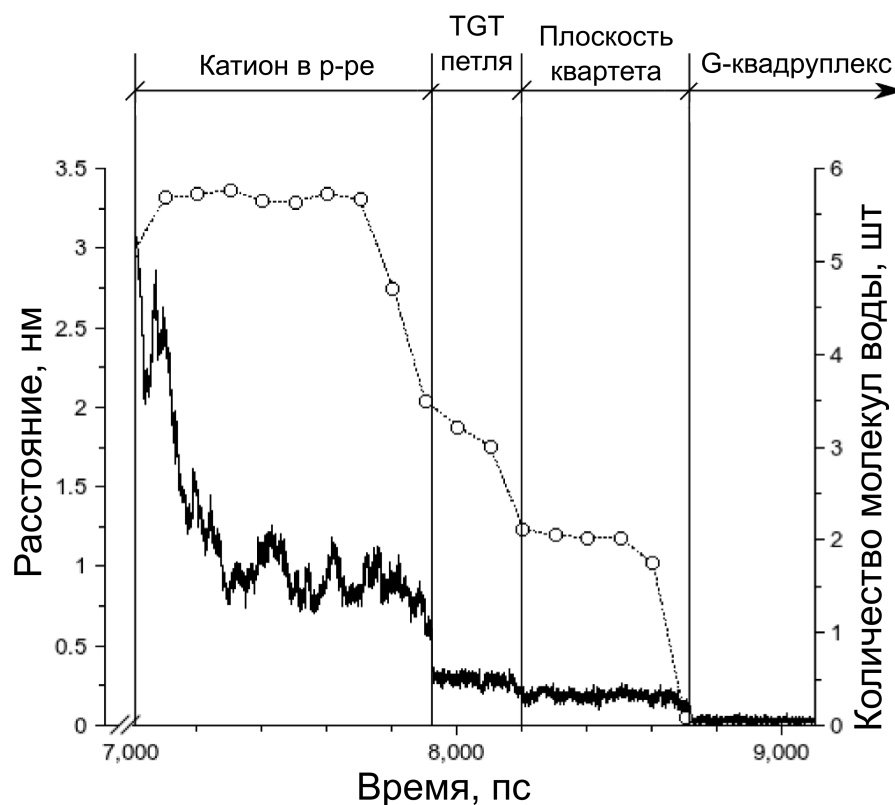


Рисунок 3.37: Проникновение катиона натрия в центральный сайт через TGT-петлю. Сплошной линией отмечено расстояние от катиона до центрального сайта связывания. Пунктирной линией отмечено число молекул воды вокруг катиона.

калия попадает внутрь квадруплекса через образовавшуюся брешь, связываясь в верхнем сайте – в этот момент его структура отличается от всех обсужденных ранее. После короткого пребывания в верхнем сайте ( $\approx 1$  нс) катион перемещается в центральный сайт. Последующее восстановление структуры аптамера происходит достаточно медленно из-за нового взаимодействия G1-T9, которое сформировалось одновременно с первичным захватом катиона. К 156-й нс наблюдения полного восстановления структуры не произошло: изменена конформация остатка G1 из-за вращения вокруг связи C3'-O3' и значение торсионного угла  $\epsilon$  с  $174^\circ$  до  $68^\circ$  (Рисунок 3.39).

Квадруплекс полностью не восстановил структуру до конца времени наблюдения. Хотя траектория длиной 300 нс считается продолжительной – при срав-



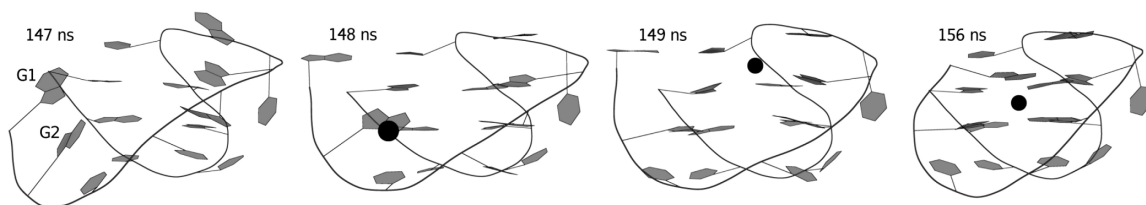


Рисунок 3.38: Проникновение катиона калия в структурно дестабилизированный квадруплекс.

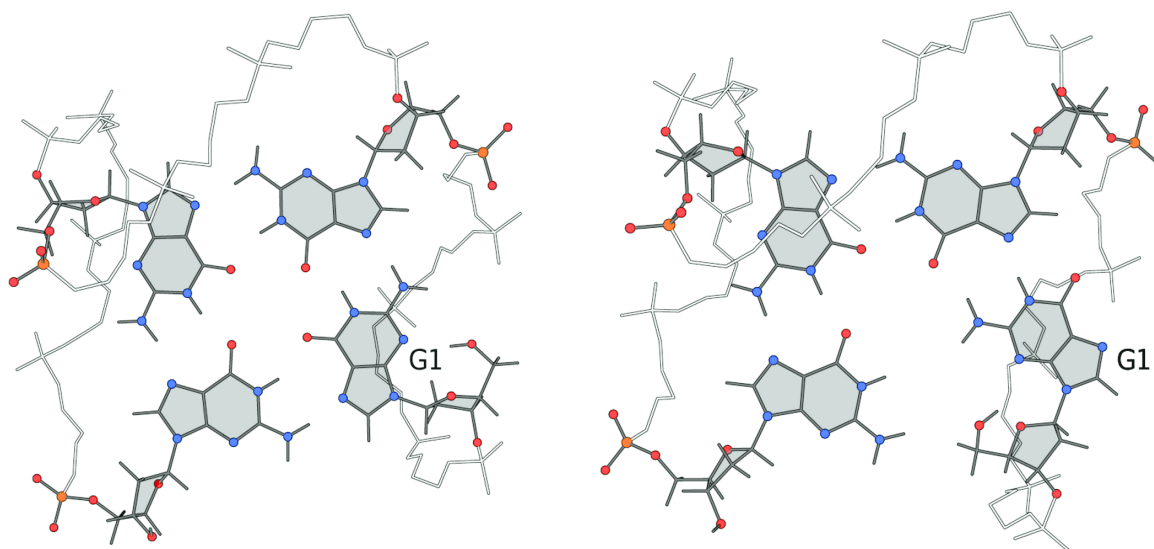


Рисунок 3.39: Изображение вращения остатка G1 вокруг торсионного угла  $\epsilon$ .

нении с используемыми в большинстве аналогичных исследований, – этого времени недостаточно для покрытия всех конформационных переходов в молекулах такого размера. Поведение систем с KCl в растворе с 15-ТВА в целом согласовывалось с наблюдениями, описанными выше. В обеих системах наблюдали изменения структуры 15-ТВА при связывании катиона калия. В первом случае, как в системе, обсужденной ранее, остаток G1 изменял конформацию, вращаясь вокруг связи C3'-O3', и торсионный угол изменил значение с  $174^\circ$  до  $74^\circ$ . Это привело к увеличению площади поры между гуанинами верхнего квартета. Вторые ворота были открыты, в то время как первые (конформация TGT-петли) закрывались, как только катион попадал во внутреннюю полость (Рисунок 3.40). на Рисунке 3.40 показано, что основной элемент первых ворот – это

основание G8 петли TGT, в то время как движения остатка T9 не меняются в ходе связывания катиона с аптамером (серая линия на Рисунке 3.40).

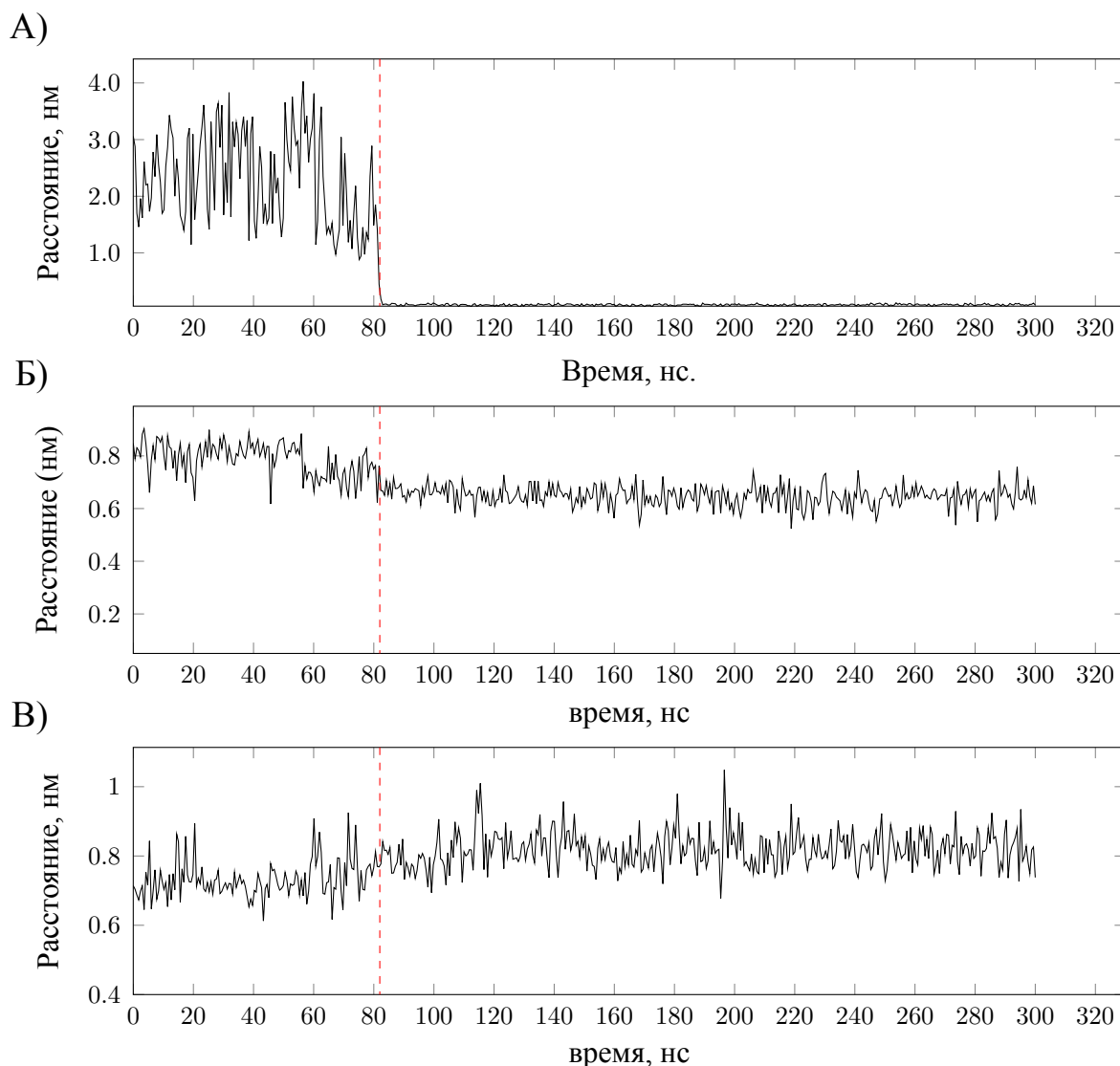


Рисунок 3.40: Поведение TGT петли при связывании катиона калия из раствора. Прерывистая линия отражает расстояние между центрами масс атомов Об гуанинов квадруплекса и катионом калия. Сплошная чёрная линия это расстояние между центрами масс остатка G8 и восьми атомов Об гуанинов квадруплекса. Сплошная серая линия это расстояние между центрами масс остатка T9 и восьми атомов Об гуанинов квадруплекса. Проникновение катиона во внутреннюю полость квадруплекса (прерывистая линия достигает нулевых значений при 82 наносекунде наблюдения) коррелирует с движением G8 но не T9. Катион калия не был замечен в плоскости квартета.

Во второй системе, с KCl, аптамер не захватывает ни одного катиона калия из раствора до 129 нс, после чего структура аптамера теряет организацию

(Рисунок 3.41). G-стебель разделяется на две части и расстояние между ними достигает  $10\text{\AA}$ . В результате образуется большая пора, в которую катион калия с лёгкостью проникает между плоскостями квартетов (Рисунок 3.41).

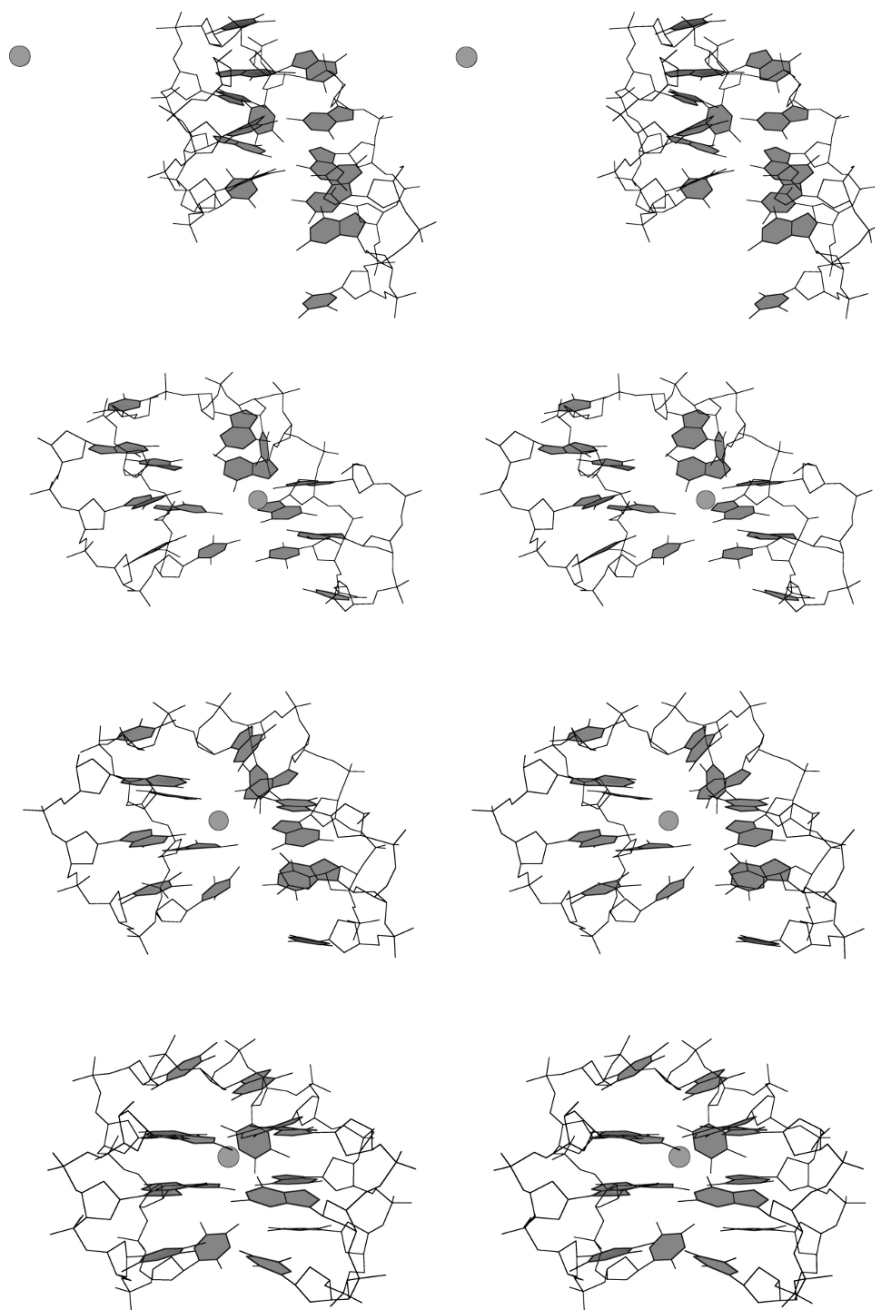


Рисунок 3.41: Стереорезультат проникновения катиона калия в центральный связывающий сайт через пору, образуемую при изменении структуры квадруплекса (стерео-изображение). Время наблюдения (сверху вниз): 128, 129, 130 и 131 нс.

После этого квадруплекс полностью восстанавливает нативную структуру менее чем за 1 нс и остается стабильным до конца наблюдения.

### **Вклад петель связывание катионов**

Описанные выше наблюдения показывают, что TGT-петля играет важную роль во взаимодействиях аптамера с катионами, находящимися в растворе. Для более детального изучения роли TGT-петли мы провели моделирование МД аптамера 15-ТВА без TGT-петли в водном растворе с NaCl. Из 11 расчетных систем одна система в начальной конфигурации содержала катион в центральном сайте связывания, остальные 10 содержали катионы только в растворе. В четырех системах из десяти модифицированный аптамер потерял структурную организацию в течение первых 10 нс наблюдения и тем самым возможность захвата катиона. В тех случаях, когда захват катиона происходил очень быстро (0–6 нс), аптамер сохранял структуру квадруплекса и, захватив катион, был стабилен до конца наблюдения (60 нс) – как и система с катионом в центре квадруплекса изначально. Таким образом, TGT-петля вносит важный вклад в стабильность 15-ТВА. С одной стороны, отсутствие TGT может ускорять захват квадруплексом катиона из раствора, так как верхний квартет становится полностью доступен для молекул растворителя. С другой стороны, удаление TGT-петли сильно снижает структурную стабильность квадруплекса без катиона в центральном сайте связывания. Мы считаем, что второй фактор явно превалирует над первым, так как по результатам моделирования вероятность распада структуры ещё до захвата катиона составляет 40%.

Аналогичные алгоритмы использованы и при исследовании роли TT-петель в захвате катиона: TT-петли удаляли из молекулы 15-ТВА и проводили моделирование МД по захвату катиона в растворе с NaCl. Из 11 построенных си-

Система, катион	Связывающий сайт	$N_n b$ время для 1:1, нс	$N_b$ время для 2:1, нс	вероятность
15-TBA, Na <sup>+</sup>	нижний	898.7	55.2	0.06
15-TBA, Na <sup>+</sup>	нижний	898.7	5.31	0.0059
без TGT 15-TBA, Na <sup>+</sup>	верхний	57.0	38.8	0.68
без TGT 15-TBA, Na <sup>+</sup>	верхний	59.8	36.8	0.62
без TGT 15-TBA, Na <sup>+</sup>	верхний	59.8	1.7	0.028
без TGT 15-TBA, Na <sup>+</sup>	верхний	59.8	0.4	0.067
без TGT 15-TBA, Na <sup>+</sup>	нижний	59.5	2.6	0.044
без TGT 15-TBA, Na <sup>+</sup>	верхний	59.5	13.0	0.22
без TGT 15-TBA, Na <sup>+</sup>	верхний	59.5	2.8	0.047
без TT 15-TBA, Na <sup>+</sup>	нижний	58.8	8.2	0.14
без TT 15-TBA, Na <sup>+</sup>	нижний	56.3	19.3	0.34
без TT 15-TBA, Na <sup>+</sup>	верхний	57.7	13.4	0.23
без TT 15-TBA, Na <sup>+</sup>	нижний	57.7	3.3	0.057
без TT 15-TBA, Na <sup>+</sup>	нижний	57.0	14.4	0.25

Таблица 3.5: Вероятность существования комплекса аптамера с двумя катионами натрия, когда один из катионов находится в центральном связывающем сайте. Таблица представляет все индивидуальные явления когда наблюдалось связывание двух катионов ( [1]+(1) ).

стем одна исходно несла катион в центральном сайте связывания, в остальных катион присутствовал только в растворе. Потеря структуры квадруплекса наблюдалась только в двух случаях из десяти. Если катион захватывался структурированной ДНК, то квадруплексная структура сохранялась до конца времени наблюдения. Во всех стабильных системах захват катиона происходил в диапазоне 2–6 нс наблюдения.

В используемой временной шкале не наблюдали обмена катиона, связанного с аптамером, на катион из раствора, однако обнаружены промежуточные структуры, в которых катион был связан в верхнем (7 раз) или в нижнем (5 раз) связывающем сайте, в то время как структура аптамера была близка к нативной и другой катион находился в центральном связывающем сайте. В таких структурах, по видимому из-за отталкивания положительных зарядов, катион в центральной полости был смещен в сторону плоскости квартета, противоположного тому, с которым связался второй катион. Статистика о продолжительности этого явления приведена в Таблице 3.5. Через некоторое время катион верхнего/нижнего связывающего сайта диссоциировал от квадруплекса в раствор.

Подробный анализ траекторий нативного аптамера 15-ТВА выявил только одно событие, когда второй катион связывается аптамером: при связывании катиона в нижнем сайте относительная вероятность найти второй катион, связанным в нижнем сайте составляла 0.06. Для системы без ТТ-петель обнаружено одно событие, когда произошёл полный обмен катионов за связывание с квадруплексом: на 45-й нс траектории МД один катион из раствора приблизился к верхнему связывающему сайту (Рисунок 3.42), что привело к выталкиванию катиона из центрального связывающего сайта. В течение очень короткого промежутка времени наблюдается образование комплекса 1 : 2 между ДНК и катионами, причем очень похожий на тот, который описывали Marathias и Bolton [349]. Новый катион входит в центральную полость и через некоторое время (3.5 нс) катион, который исходно находился в центральном связывающем сайте, переходит из комплекса с ДНК в раствор.

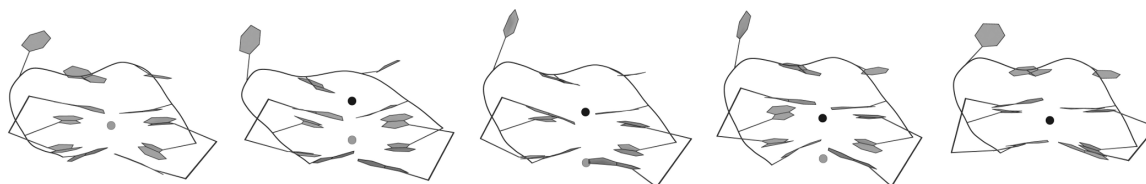


Рисунок 3.42: Обмен катионами натрия между модифицированным аптамером 15-ТВА (без ТТ-петель) и солевым раствором. На рисунке представлены конфигурации системы в ходе моделирования. Исходно связанный катион отмечен серой сферой, а “приходящий” – черной сферой.

Таким образом, впервые методами моделирования показана возможность обмена катионами между квадруплексной ДНК и раствором только за счёт теплового движения. Ранее при моделировании МД по профилям свободной энергии [117; 344; 363] было высказано предположение, что квадруплекс в свободном состоянии всегда связан с катионом. Результаты этой работы расширяют наше понимание общей картины связывания квадруплексами катионов. По-видимому, короткие или смещённые квадруплексы способны вытолкнуть

связанный с ними катион при кратковременном связывании второго катиона из раствора. И этот механизм чисто электростатический по сути – отталкивание двух катионов, лежащих на одной линии, приводит к сначала к смещению связанного катиона и затем к выходу его в растворитель. Однако связывание катиона из раствора выходом другого в раствор не уменьшает вероятность дестабилизации квадруплекса во время этого процесса.

### **Анализ занятости состояний связанных катионов и оценки энергий**

Основная цель моделирования МД – это исследование динамики структуры выбранной системы. Конечно, очень важно дополнить общую картину данными по изменению свободной энергии. К сожалению, метод моделирования МД не позволяет следить за относительным изменением общей свободной энергии во время расчёта. Надо отметить, что общая энергия системы сильно флуктуирует, в то время как кинетическая энергия системы корректируется алгоритмом термостатирования. Это является основной причиной, почему нельзя рассчитывать изменение свободной траектории во время моделирования. Кроме того, очень важна конфигурация системы, а именно, фазовое пространство для микросостояний системы.

Моделирование МД основано на гипотезе эргодичности, где постулируется, что среднее статистическое по ансамблю равно среднему значению во времени для системы – что можно отслеживать в ходе моделирования МД. Например, набор состояний положения катионов в очень длинной траектории должен достаточно точно описывать обмен катионами между разными микросостояниями и, следовательно, может использоваться для расчёта разницы свободной энергии между микросостояниями. Такой подход расчёта изменения свободной энергии является единственно правильным с физической точки зрения.

Есть методы, которые позволяют следить за свободной энергией системы не напрямую, а приблизительным образом, используя анализ конформаций после расчёта. Такие подходы экстрагируют координаты исследуемого объекта из траектории, удаляют молекулы воды и заменяют их описанием растворителя в неявном виде, например, с помощью теории Пуассона-Больцмана. С переменным успехом этот подход активно использовался для исследования G-квадруплексных ДНК некоторое время назад [339; 364]. Подобные расчёты вносят дополнительные значительные приближения к уже сделанным в ходе моделирования МД – в итоге получаемые значения содержат значительные систематические и случайные ошибки.

Учитывая все эти ограничения, мы решили не использовать подобные методы для исследования изменения свободной энергии в ходе взаимодействия аптамера 15-ТВА с катионами, а следить за некоторыми компонентами свободной энергии, к примеру, за энергией взаимодействия ДНК и катиона, на основе приближений силового поля. Такие расчёты проводились ранее, например Bolton и коллегами [266]. Такой расчёт не учитывает как эффект растворителя, так и энтропийный фактор, и результат не обязательно коррелирует с изменением свободной энергии. Если не учитывать затраты на дегидратацию, то получаемые значения могут быть сильно завышены или даже быть противоположны по знаку эффекта. В литературе можно найти убедительные примеры того, как современное неявное описание растворителя (метод Пуассона-Больцмана) приводит к противоположному знаку значений изменения свободной энергии по сравнению с расчётами, не учитывающими растворитель или использующими устаревшие модели описания растворителя [364; 365].

Тем не менее, мы провели расчёты, результаты которых позволили расширить понимание механизма взаимодействия катионов с квадруплексными ДНК.



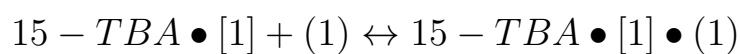
Относительные энергии разных состояний находят из населённости этих состояний, которая, в свою очередь, может быть определена из очень длительных траекторий молекулярной динамики. Таким образом, нам нужно было набрать достаточное число конформаций для каждого состояния. чтобы понять, как катион связывается в центральном сайте и покидает квадруплекс много раз. К сожалению, нам не удалось достичь такого времени наблюдения, чтобы детально изучить этот механизм, но интересные факты все-таки получили.

Мы задали следующие состояния взаимодействия катиона с квадруплексом: [0] – квадруплекс, не содержащий катион в центральном сайте, [1] – квадруплекс содержащий один катион в центральном сайте, [0]+(1) – квадруплекс с одним катионом в нижнем или верхнем сайтах связывания, [1]+(1) – квадруплекс с двумя катионами, один из которых в центральном сайте, и (1)+(1) – квадруплекс с двумя катионами в верхнем и нижнем сайтах.

Мы наблюдали множество раз, как квадруплекс захватывая катион из состояния [0] переходит в состояние [1], но ни разу не наблюдали обратного перехода – из [1] в [0], – несмотря на относительно длительное наблюдение за состоянием [1]. Мы регистрировали быстрый обмен катионами, т.е. переход из [1] в [1] через состояние [1]+(1). Несмотря на то, что из этих данных нельзя получить точное значение свободной энергии, совершенно очевидно, что состояние с катионом в центральном сайте, [1], гораздо выгоднее безкатионного, [0]. Из-за быстрого перехода катиона вовнутрь квадруплекса из верхнего и нижнего сайтов связывания, т.е. для пути  $[0] \rightarrow (1) \Rightarrow [1]$ , мы не можем оценить разницу в свободных энергиях между [0] и (1).

Очевидно, что состояние (1) гораздо менее выгодно, чем [1], так как конфигурация (1) не может конкурировать с состоянием [1]. Однако мы можем оценить вероятность существования случаев [1]+(1). Это может быть сдела-

но подсчётом временных интервалов связывания второго катиона с верхним или нижним сайтом аптамера, т.е. времени жизни состояния (1). В Таблице 3.5 приведены периоды времени времени и вероятности ( $P_b$ ) для состояния [1]+(1) наблюдаемого в разных траекториях. Зададим  $N_b$  как период времени в который наблюдалось состояние [1]+(1), а  $N_nb$  это общее время для состояний [1] и [1]+(1) в рассматриваемой траектории. Другими словами мы рассматриваем химическую реакцию между аптамером и катионом.



$15 - TBA \bullet [1]$  - это аптамер с катионом в центральном сайте связывания, а  $15 - TBA \bullet [1] \bullet (1)$  - это аптамер связанный с двумя катионами, один в центральной полости, а другой в верхнем или нижнем сайте связывания.

В таблице 3.5 приведены значения времени пребывания второго катион в аптамере, имеющем катион в центральном сайте (состояние [1]+(1)). Отметим, что не во всех траекториях в центральной полости исходно присутствовал катион, поэтому требовалось некоторое время для захвата первого катиона. Это время не учитывалось при подсчёте вероятностей. Следует напомнить, что все системы, где не было либо ТТ-, либо ТGT-петель, строились без катиона внутри и не все из них сохранили квадруплексную структуру, а системы, ее потерявшие, не рассматривались. Общее время наблюдения таких траекторий соответствовало 420 и 540 нс, т.е. в сумме это 940 нс, что приводит к соотношению  $\approx 0.2$  состояний [1]+(1)/[1] или  $\approx 0.15$  для отношения [1] + (1)/[1] + [1] + (1).

В системах, где удалены некоторые петли, доступный квартет более вероятно взаимодействует с катионом. Для систем с нативной молекулой 15-ТВА мы не детектировали состояний типа [1]+(1) в течение времени наблюдения

более 300 нс, хотя состояние [1] сохранялось в течении 726 нс. Тем не менее нам удалось обнаружить состояние [1]+(1) с общим периодом порядка 60 нс в траекториях, по длине близких к микросекунде.

Хотя все эти данные трудно назвать сходящимися, грубые прикидки показывают, что значение  $([1]+(1)/[1])$  для целой молекулы 15-ТВА лежит в диапазоне 0.01–0.1. Следует отметить, что концентрация аптамера при моделировании составляла около 0.007 М, а катиона около 0.18 М.

Для катиону калия мы не детектировали состояние [1]+(1), хотя. надо сказать, размер выборки и время наблюдения были малы (см. описание захвата катиона калия выше) и мы не проводили моделирования беспетельных вариантов 15-ТВА, где квартеты более доступны для связывания второго катиона.

Итак, для проверки значений изменения свободной энергии, полученных экспериментально, нам бы понадобилось наблюдать за системой приблизительно в 100 раз дольше, т.е. примерно 3 мкс. Рассчитанная нами вероятность состояния [1]+(1) достаточно хорошо согласуется с данными, полученными Trajkovski и др. [351] для катиона аммония при исследовании методом ЯМР. Авторы обнаружили слабое связывание второго катиона в нижнем сайте связывания. К сожалению, из экспериментальных данных сложно точно отличить состояние [1]+(1) от (1). На основе полученным нами результатов мы все-таки полагаем, что рнчь шла о состоянии [1]+(1), т.е. о комплексе со стехиометрией 1 : 2.

Таким образом, при моделировании МД выявлено, что, конфигурация [1] гораздо стабильнее всех проанализированных нами. Однако связывание второго катиона, когда первый находится в центральном сайте связывания (состояние [1]+(1)), энергетически возможно. Как отмечалось ранее, точность предсказан-

ной населённости состояния  $[1]^{+}(1)$  может сильно зависеть от выбора параметров силового поля.

### **3.5.3 Гибридное молекулярно механическое / квантово-механическое исследование взаимодействия катионов с 15-ТВА**

В исследовании связывания катиона методами классической механики катион захватывался аптамером с образованием промежуточного комплекса, связываясь в нижнем или верхнем сайте. Однако моделирование МД в классических силовых полях не позволяет учитывать поляризацию электронных облаков, которая однозначно описывает координационные взаимодействия катиона и его хелаторов. По этой причине мы решили прибегнуть к помощи гибридного МД/ММ метода для оценки относительной стабильности разных позиций катиона калия в комплексе с аптмером 15-ТВА. Несмотря на непродолжительное время наблюдения за системой, доступное для используемых вычислительных ресурсов, гибридный метод МД/ММ может быть использован как дополнительным независимый тест, который дополнит данные классической молекулярной механики.

Пять систем (Рисунок 3.31) были построены для моделирования МД/КМ: три из них, 1–3, содержали по одному катиону калия (соответственно в верхнем, среднем и нижнем сайте связывания); система 4 включала два катиона калия в верхнем и нижнем сайтах с и система 5 – катион бария в центральном сайте связывания. Катион бария был выбран как катион сравнения – он образует комплекс с аптамером, очень схожий с комплексом для стронция. Для по-

следнего известно, что катион в нем связан в центральном сайте аптамера со стехиометрией 1 : 1.

В двухсистемах, где катион калия находился в верхнем сайте связывания (системы 1 и 4), в ходе моделирования он смещался в центральный сайт. В системе 4 это происходило параллельно с выталкиванием катиона калия из нижнего сайта связывания в сторону раствора. В системе 3 положение катиона калия в нижнем сайте связывания было также нестабильно—после некоторых флуктуаций катион смещался в сторону центрального сайта (Рисунок 3.43).

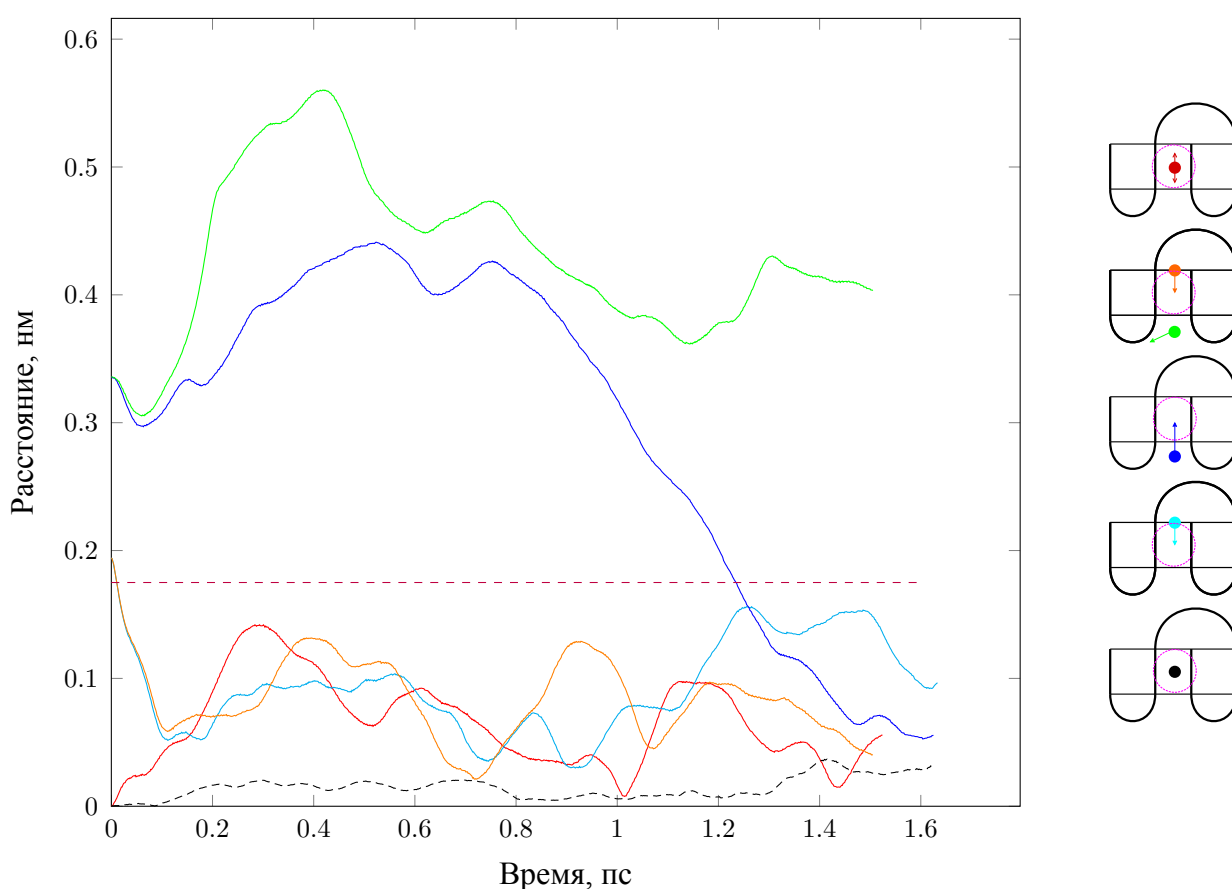


Рисунок 3.43: Поведение катионов в различных комплексах с 15-ТВА при моделировании МД/КМ. Слева: Расстояние между катионом и геометрическим центром квадруплекса. Граница центрального сайта связывания отмечена розовой пунктирной линией. Если расстояние близко к 0 это значит, что катион находится в центре центрального связывающего сайта. Справа: схематическое изображение движений катионов. Цвета катионов соответствуют графику расстояний.

Только системы 2 и 5, в которых катионы калия и бария исходно находились в центральном сайте связывания, были устойчивы (Рисунок 3.43). Во время моделирования катион калия совершал периодические колебания внутри центрального сайта связывания – от верхнего к нижнему квартету и обратно. В отличие от калия, катион бария оставался практически неподвижным в центре полости между квартетами. Положение первого пика парной функции радиального распределения (PRDF) карбонильных атомов кислорода  $O_b$  вокруг катиона калия составляет  $2.7\text{\AA}$ , что очень близко к соответствующему значению при исследовании *ab initio* динамики сольватации катиона калия в метаноле [366] (Рисунок 3.44).

Среднее число хелаторов, которые попадают в окружение катиона калия и соответствует первому пику PRDF, равно 5.2. Это же значение описывает среднее координационное число калия внутри квадруплекса. Как упоминалось, катион бария практически не смещается от стартовой позиции. Параметры его координации схожи с экспериментальными данными [367], среднее координационное число равно 8, а расстояние до хелаторов колеблется около  $2.8\text{\AA}$ .

Движение катиона калия к центру квадруплекса сопровождается оптимизацией общей энергии квантовой подсистемы (Рисунок 3.45). Так как классическое силовое поле достаточно приближенно учитывает энергию донор-акцепторных взаимодействий, то мы ожидали, что анализ значений энергии полученных для них МД/КМ расчётов может дать реалистичные значения энергии взаимодействия катиона с квадруплексом. Для вычисления энергии индивидуальных компонентов мы использовали тот же теоретический уровень и рассчитали относительное изменение общей энергии квантовой системы как функцию от времени и представили на графике вместе с расстоянием до центрального сайта связывания.

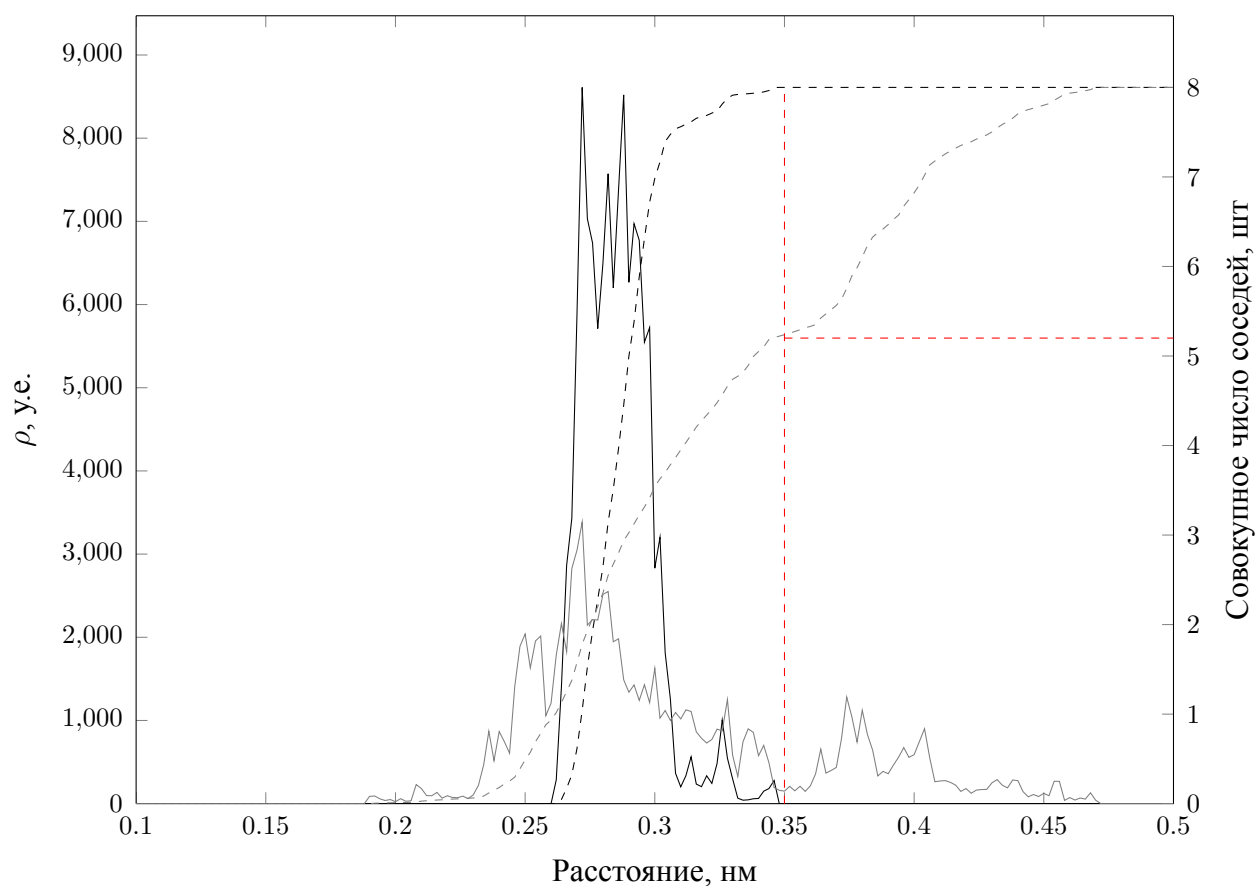


Рисунок 3.44: Динамика катионов калия и бария в центральном сайте связывания при моделировании МД/КМ. Парная функция радиального распределения между атомами кислорода Об и катионом калия (красная сплошная линия) или бария (чёрная сплошная линия). Позиция катиона бария внутри ДНК стабильна и расстояние от бария до хелаторов флуктуирует в небольшом диапазоне (от  $2.69\text{\AA}$  до  $2.88\text{\AA}$ ). Позиция катиона калия менее чёткая, наиболее выраженный пик приходится на расстояние в  $2.71\text{\AA}$ . Среднее координационное число катиона калия равно 5.2.

Относительное изменение энергии рассчитывали как:

$$\Delta E = E_{QM} - E_{QM} - E_{QM}^{K^+}$$

В рассматриваемом случае это 'комплекс' это квантовая подсистема с заглушками в виде атомов водорода, описанная ранее. Аптамер это такая же система как и комплекс, но без катиона калия.  $K^+$  это катион калия в ячейку такого же размера как и системы 'комплекс' и 'аптамер'. Размер ячейки важен так как

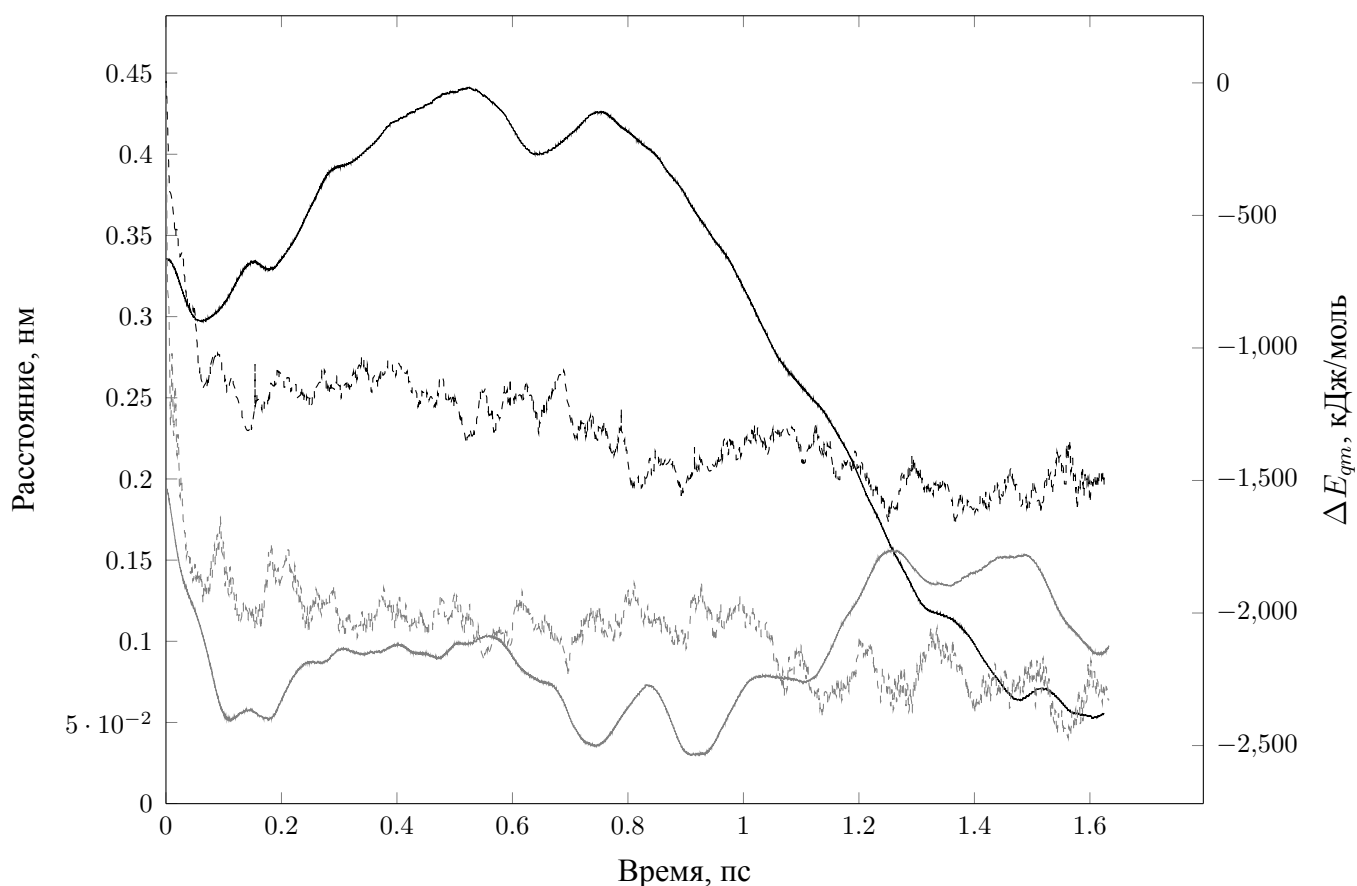


Рисунок 3.45: Изменение относительных значений энергии комплексов катионов калия с 15-ТВА во время моделирования МД/КМ. Расстояние между катионом и геометрическим центром квадруплекса показано сплошными линиями, а значения энергии отмечены пунктирными линиями соответствующего цвета. Граница центрального сайта связывания отмечена пунктирной розовой линией.

количество плоских волн описывающих волновую функцию напрямую зависит от размера ячейки.

В Таблице 3.7 представлены данные только для нижнего и центрального сайтов, положение катиона в верхнем сайте было слишком не стабильно для того, что бы найти средние значения энергии. Очевидно, что в ходе МД/КМ расчётов перемещение катиона калия из нижнего в центральный сайт связывания происходит с увеличением выигрыша в энергии взаимодействия. Надо отдельно отметить, что все рассчитанные значения энергий ниже нуля и находятся в диапазоне -330. . .-410 кДж/моль. Таким образом направление изменения



система	$\Delta E_{Interaction}$ , кДж/моль
$K^+ \bullet 15 - TBVA$ , нижний сайт	$-362.3 \pm 27.2$
$K^+ \bullet 15 - TBA$ , центральный сайт	$-391.2 \pm 21.8$

Таблица 3.6: Сведение значения энергии взаимодействия катиона калия с нижним и центральным сайтами, рассчитанные в этой работе на основе PW-DFT.

энергии рассчитанной таким образом подтверждает связывание в центральном сайте. С другой стороны значения энергий достаточно велики по отношению к экспериментальным данным и не могут предоставить реалистичного описания энергетики этого процесса. Подобного рода расчёты не учитывают ряд составляющих свободной энергии и очевидно, что система не находится в равновесном состоянии. Для сравнительного изучения известных позиций катионов в структурах Маратиаса и Болтона (PDBID:1c35,1c38) был применён тот же самый подход. Относительные значения энергий для индивидуальных положений катиона внутри структур сильно колеблются. Большинство значений выше нуля (Таблица 3.7), очевидно, что описанные Болтоном положения катионов не уравновешены для проведения подобных расчётов.

Надо отметить, что существует значимая разница между моделированием МД и МД/КМ в описании комплекса аптмера 15-ТВА с катионом. Основное направление схоже в обоих методах, но методы дают разную динамику процесса: в случае МД/КМ процесс перехода из нижнего и верхнего сайтов в центральный сайт связывания происходит в течении двух пикосекунд, в то время как, при классическом моделировании этот процесс занимает наносекунды. Разница возникает из сути методов: классическое моделирование не позволяет воспроизводить форму и поляризацию электронных облаков, и учитывает только достаточно далёкие электростатические взаимодействия катиона с аптамером. Простые потенциалы Ленорда-Джонса при описании Ван-дер-Ваальсовых вза-

PDBID		1c35	1c38
	ID	$\Delta E$ , кДж/моль	$\Delta E$ , кДж/моль
Нижний сайт связывания	101	689.9	100.0
	102	457.3	94.6
	103	1245.2	417.6
	104	769.4	632.2
	105	911.3	98.3
	106	469.0	149.8
	107	368.2	-158.6
	108	327.6	-165.7
	109	542.7	156.1
	110	883.7	174.1
Верхний сайт связывания	201	354.4	72.8
	202	350.6	115052.9
	203	603.3	-176.6
	204	233.5	570.7
	205	780.7	-181.2
	206	532.2	635.5
	207	1071.1	110.5
	208	1302.1	141692.8
	209	915.9	10.0
	210	236.0	107293.7

Таблица 3.7: Энергии взаимодействия аптмер : катион калия в комплексе на основе ЯМР данных с использованием PW-DFT подхода.

имодействий также не оптимальны, как уже не раз упоминалось ранее. МД/КМ потенциально может давать достаточно точное описание динамики, однако мы должны подчеркнуть, что полученные нами результаты могут существенно зависеть от приближений сделанных при моделировании квантовой системы, такие как: не полная релаксация системы и очень короткое время наблюдения. Дальнейшие исследования несомненно должны прояснить разницу между подходами и привести к полной сходимости подходов. Тем не менее предварительные МД/КМ расчёты, по крайней мере, примерно показывают различие классического описания от аккуратного описания всей электронной системы.

### 3.5.4 Изотермическая калориметрия

Для прояснения комплексообразования аптамера 15-ТВА и катиона калия, термодинамические параметры взаимодействия катионов калия и бария с 15-ТВА были изучены методом изотермической калориметрии. Микрокалориметрический подход позволяет полностью характеризовать взаимодействие субъектов в растворе. На Рисунке 3.46 приведены кривые титрования катионами калия и бария раствора 15-ТВА в буферном растворе. Очевидно, что барий имеет более высокую аффинность к аптамеру ( $1.1 \times 10^5 \text{ M}^{-1}$  против  $5.0 \times 10^3 \text{ M}^{-1}$ ) и практически такую же энтальпию связывания ( $-9.7 \text{ кКал/моль}$  против  $-8.2 \text{ кКал/моль}$ ) чем катион калия к аптамеру. Применение подгонки по модели 'один набор сайтов' показало, стехиометрия связывания обоих катионов с аптамером 15-ТВА близка к 1:1. Исследование конкуренции катионов калия и бария показало, что калий не способен связываться с аптамером если с ним уже связан катион бария (чёрная кривая на Рисунке 3.46).

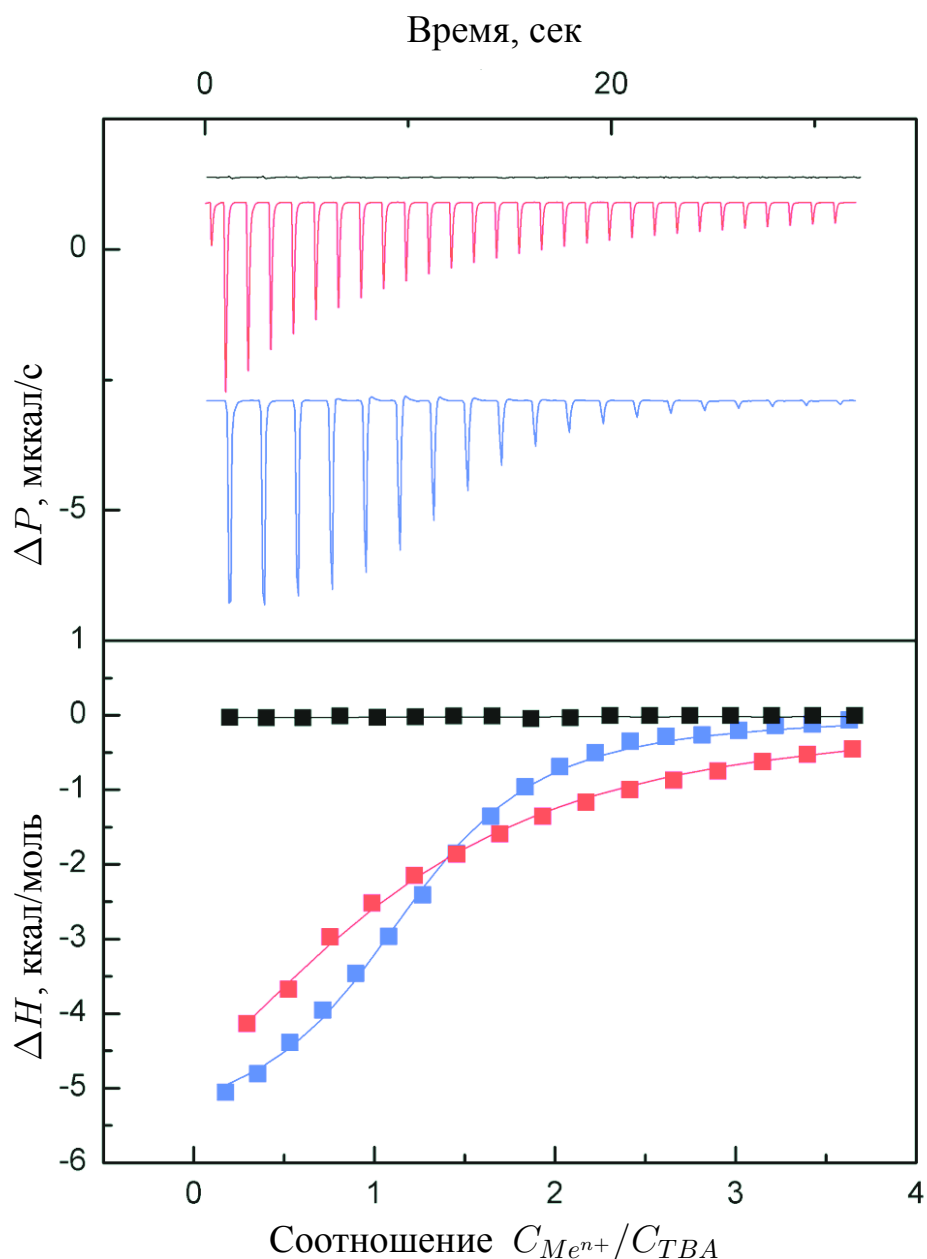


Рисунок 3.46: Кривые изотермического титрования связывания аптамера 15-ТВА с катионами калия и бария. А) Титрование аптамера катионами калия после образования комплекса с катионами бария (конкуренция), связывание не наблюдалось. В) связывание 15-ТВА с катионами калия. С) связывание аптамера с катионами бария

Результаты нашего моделирования, что основание остатка G8 влияет на эффективное блокирование катионов внутри центрального сайта и участвует в проникновении катиона во внутрь полости квадруплекса. Таким образом замена G8 меньшим по размеру основанием, например пиримидинового ряда, может облегчить обмен катионами между аптамером и раствором. Для проверки это-

аптамер	катион	$K_a, M^{-1}$	стехиометрия
125 мМ 15-ТВА	750 мМ КСl	$1.6 \times 10^4$	1:1
125 мМ 15-ТВА	750 мМ ВаСl <sub>2</sub>	$11 \times 10^4$	1:1
125 мМ 15-ТСТ	750 мМ КСl	Не обнаружено	-
250 мМ 15-ТСТ	833 мМ КСl	$0.07 \times 10^4$	1
250 мМ 15-ТСТ	833 мМ ВаСl <sub>2</sub>	$2.1 \times 10^4$	1:2

Таблица 3.8: Термодинамические параметры связывания аптамера 15-ТВА с ионами <sup>a</sup>

<sup>a</sup> Буфер: 140 мМ LiCl, 20 мМ Tris-HCl, pH=6.8

го предположения было исследовано взаимодействие мутантного 15-ТВА, где G8 был заменён на цитозин (15-ТСТ), с катионами. Результаты изотермической калориметрии всех исследованных систем приведены в Таблице 3.8. Олигонуклеотид 15-ТСТ связывается с катионами калия гораздо хуже, чем исходный аптамер 15-ТВА, и сохраняет уровень связывания с катионами бария.

### 3.5.5 Заключение об связывание катионов с аптамером 15-ТВА

Динамика структуры и стабильность гуаниновых квадруплексов плотно связаны со связывание катионов внутри квадруплексного стебля и иногда со связыванием катиона с полостями создаваемыми одно цепочечными фрагментами молекулы, такими как петли. Несмотря на ограничения [133], моделирование молекулярной динамики может предоставить интересную картину взаимодействия катиона с G-квадруплексной ДНК и дополнить экспериментальные данные атомистическими деталями процесса.

Ранее мы показали, что правильно собранная квадруплексная структура, а именно аптамер к тромбину 15-ТВА, в водном растворе может сохранять структуру достаточно долго для того, что бы захватить случайный катион из раствора [368]. Однако в той части работы мы не наблюдали образования комплексов со

стехиометрией 1:2 между аптамером и катионом о которой ранее сообщали Маратиас и Болтон [349]. Более того, положение катиона внутри квадруплексного участка ДНК отличалось от того, о котором сообщалось ранее в литературе [350; 369], где предполагалось существование двух сайтов связывания катиона калия в аптамере 15-ТВА. В соответствии с этими исследованиями, первый сайт располагался между петлей TGT и прилежащим к петле квартетом, а второй сайт был расположен между TT петель. Об положении катиона в центральном сайте связывания не сообщалось. В нашем исследовании мы исследуем эту систему методами моделирования МД, МД/КМ и изотермической калориметрии. Было рассчитано 30 траекторий молекулярной динамики. Наше исследование принесло несколько интересных наблюдений в картину взаимодействия G-ДНК с катионами.

### **Моделирование МД не обнаружило образование комплекса аптамера с катионом со стехиометрией 1:2.**

Наши результаты предполагают наличие 3-х потенциальных сайтов связывания катиона в структуре 15-ТВА. Один из них это известный сайт в центральной полости квадруплекса, в то время как два другие это те, которые обнаружили Маратиас и Болтон в месте контакта квадруплекса и петель. В моделировании только один центральный сайт оказался стабильным. Два других действуют как промежуточные ловушки при захвате катиона из раствора и не могут удержать катион надолго, после чего он перемещается во внутреннюю полость квадруплекса или возвращается в раствор, если центральный сайт уже занят жругим катионом. Мы считаем, что верхний сайт связывания является более важным, так как при связывании в нижнем сайте образование TT пары препят-

ствуется легкому прохождению катиона во внутреннюю полость. Наши наблюдения согласуются с известными экспериментальными данными [370; 371].

### **Атомистическое моделирование позволяет увидеть детали проникновения катионов во внутрь G-ДНК**

В нашей работе мы получили уникальную информацию об способах проникновения катиона во внутрь квадруплекса аптамера 15-ТВА. Существует по двое ворот с каждой стороны квадруплекса. Верхняя пара ворот формируется азотистым основанием остатка G8 из TGT петли (это и есть первые ворота в этой паре) и размером поры образуемой атомами кислорода Об гуанинов верхнего квартета. Нижняя пара ворот представлена неканонической TT парой на нижней стороне молекулы и, как и случае верхних ворот, размером поры между Об гуанинов нижнего квартета. Когда квадруплекс не связан с катионом все ворота открыты благодаря электростатическому отталкиванию карбонильных атомов кислорода оснований гуанинов. Атом азота N3 из азотистого основания остатка G8, расположенный над порой в квартете, тоже несёт отрицательный заряд, которые сдвигает основание остатка G8 в сторону от этой поры в центре квартета. Благодаря балансу гидрофобного эффекта и Ван-дер-Ваальсовых взаимодействий основание двигается только параллельно плоскости прилежащего квартета. Основание остатка T9 не имеет атомов с выраженным отрицательным зарядом рядом с центром квартета и его положение и динамика не сильно меняется с от присутствия катиона в центре квадруплекса (Рисунок 3.33).

Как только катион связался где либо с аптамером 15-ТВА его путь в центральную полость происходит достаточно быстро. При попадании катиона в центральный сайт связывания все перечисленные выше ворота стремятся закрыться. Размер поры между атомами кислорода в квартете уменьшается бла-

годаря электростатическому притягиванию к катиону. Это движение предотвращает выход катиона обратно в раствор. По той же причине основание остатка G8 смещается в сторону остатка T9 и TGT петля закрывает верхний выход из квадруплекса (Рисунки 3.33, 3.32, 3.40).

Очевидно, что петли не являются принципиально важными элементами для связывания катиона с G-ДНК. Однако они значительно модулируют весь процесс. Моделирование показывает TGT петля играет три независимые роли. Первое, она защищает структуру 15-TVA от разрушения структуры в отсутствие катиона, как мы наблюдали 40% вероятность того, что структура квадруплекса будет потеряна в ожидании захвата катиона, если удалить TGT петлю. В то же время TGT петля замедляет первоначальное связывание катиона заслоня вход в центральный сайт квадруплекса. Мы наблюдали как конструкция без этой петли захватывает катион быстрее, но и одновременно более подвержена структурной деградации. Как только катион оказывается внутри квадруплекса, петля помогает сохранить катион во внутренней полости, благодаря смещению основания G8. Суммируя выше сказанное можно заключить, что TGT петля оказывает стабилизирующую роль с структуре 15-TVA.

### **Корреляция движений катионов во время полного обмена ионами между G-ДНК и раствором**

Мы впервые в моделировании обнаружили случайный обмен катиона связанного с квадруплексом G-ДНК с катионом из раствора (Рисунок 3.42). Это случилось в одной из тех систем, где у аптамера 15-TVA были удалены TT петли. Выталкивание катиона происходит через нижнюю часть молекулы. Процесс выдавливания катиона начинается с приближения другого катиона с противоположной выходу стороны, т.е. со стороны TGT петли (верхний связывающий



сайт). Катион из раствора (промежуточное связанный в верхнем сайте) занимает центральный сайт связывания как только его покидает предыдущий катион. Благодаря этой корреляции между промежуточным связыванием и высвобождением катиона из центральной полости квадруплекса, центральный сайт не остаётся пустым. Это показывает, что вероятно для большинства G-ДНК катионный обмен происходит без потери стабильности всей структуры, по крайней мере это справедливо для двух квартетных квадруплексов. Факт корреляция между связыванием катиона из раствора и выталкиванием катиона из центральной полости может сильно стабилизировать как малые квадруплексы, так и промежуточные состояния при самосборке G-ДНК. Когда предполагается участие катиона из раствора в обмене, молекула с  $N$  последовательными квартетами может выпустить любой катион в центральной полости квадруплекса через промежуточные состояния с  $N-1$  связанными катионами (Рисунки 3.42, 3.47).

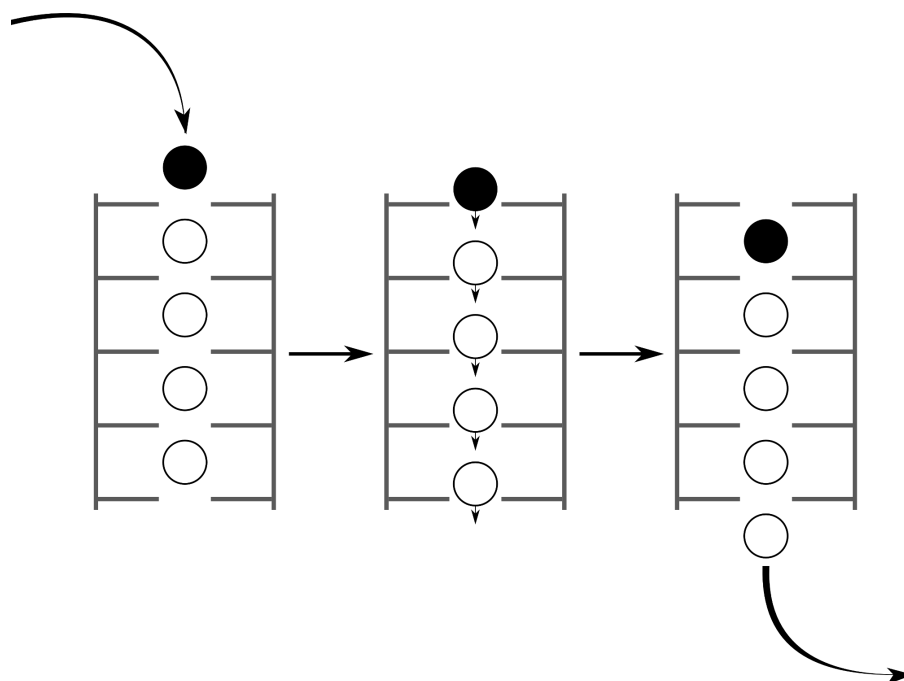


Рисунок 3.47: Обмен катионами в молекуле построенной из  $N$  последовательных квартетов. Приходящий катион отмечен чёрным, а связанные с квадруплексом, белым.

Этот эффект вероятно менее важен для протяжённых квадруплексов, состоящих, например, из четырёх квартетов. Такой квадруплекс чрезвычайно стабилен связывая всего два катиона во внутреннюю полость. Он легко может обменивать катион с раствором через последовательность переходных состояний имея два катиона внутри полости, разделяя события связывания катиона из раствора и выход катиона в раствор [117; 363].

### **Предварительные МД/КМ эксперименты**

Предварительное распознавание катионов верхним и нижним связывающими сайтам может быть описано в терминах электростатического взаимодействия. Однако, взаимодействия катиона с ДНК более сложны и только приблизительно описываются классическими силовыми полями используемыми в моделировании МД [347]. Таким образом, координационная составляющая взаимодействия катиона с аптамером не рассматривается. К сожалению полное квантово-химическое (КМ) описание такой небольшой молекулы как 15-ТВА, всё ещё за пределами современных вычислительных возможностей. Гибридные КМ/ММ моделирования стали распространёнными подходами в исследовании взаимодействий биополимеров с катионами, так как предложили разумный компромисс между скоростью и точностью расчётов [53; 347; 372]. Даже такие подходы весьма требовательны к вычислительной мощности, и современные работы сообщают о времени наблюдения в пикосекундной шкале. Этого к сожалению не достаточно для достижения равновесия системы. Однако, несколько пикосекунд предоставить грубое представление об происходящих процессах в системе.

Используя метод МД/КМ мы исследовали стабильность положения катиона калия во всех трёх сайтах связывания. Расчёты подтвердили результаты класси-

ческого моделирования указали на низкую вероятность стехиометрии 1:2 комплекса аптамера 15-ТВА с катионами калия. Один из катионов в комплексе 1:2 сместился в сторону раствора, но не покинул зону аптамера. Тем не менее мы считаем, что стехиометрия комплекса 1:1 и катион калия расположен в центре квадруплекса, между квартетами. В таком положении катион в среднем координируется атомами кислорода Об из оснований гуанина благодаря электростатическим и донор-акцепторным взаимодействиям. Для внутреннего контрольного эксперимента, мы выбрали катиона бария. Катион бария имеет близкий к стронцию размер и заряд, что приводит к практически идентичным спектрам кругового дихроизма комплексов аптамера 15-ТВА с катионами бария и стронция [373]. Стехиометрия комплекса между стронцием и аптамером 1:1 и весьма вероятно, что катион находится во внутренней полости между квартетами. Расчёты были проведены с катионом бария исключительно из практических аспектов работы со стронцием, стронций склонен к накапливанию в тканях, а барий не имеет такой особенности. В экспериментах по калориметрии мы использовали барий и поэтому в расчётах был использован тоже барий.

В МД/КМ расчётах мы показали, что находясь в центральном сайте связывания положение катиона бария очень стабильно. Катион координируют все восемь атомов кислорода Об из гуаниновых оснований остатков квадруплекса. В итоге не смотря на на все ограничения накладываемые в МД/КМ моделировании, эти расчёты показывают, что особенности электронной структуры не меняют основного заключения из классического моделирования о том, что существует стабильный комплекс аптамера с двумя катионами калия.

## **Калориметрические измерения**

Наши эксперименты по титрованию комплекса методом изотермической калориметрии указывают, что стехиометрия комплекса аптамера 15-ТВА как с катионом бария так и калия близка к соотношению 1:1. Более того, исследование конкуренции показало, что катионы занимают один и тот же связывающий сайт. Эти результаты полностью подтверждают наши расчёты.

## **Вариативность пути связывания участвует в динамике и образовании структуры квадруплексных ДНК**

Кроме механистического представления о путях попадания катиона в внутреннюю полость квадруплекса, наши результаты проливают свет на вопрос о том, что первично, образование структуры квадруплекса и захват катиона из раствора или сборка структуры вокруг катиона. Правильно собранная структура аптамера 15-ТВА без катиона способна сохранять свою структуру в течении сотен наносекунд, что очевидно, достаточно для того, чтобы встретить в растворе катион захватить его в растворе с ионной силой порядка  $\approx 0.1-0.2$  М. Экспериментально было показано, что 15-ТВА может организовать мономолекулярный G-квадруплекс в растворе при низких температурах при отсутствии возможности быть стабилизированным катионом [374]. Вероятно, что при комнатной температуре существует динамическое равновесие между собранной и расплавленной структурами, появление катионов подходящего радиуса смещает равновесие в сторону собранной структуры. Таким образом, катион не является инициатором сборки, а стабилизирует собранную и частично собранные структуры. Хотя такой процесс за пределами нашего времени наблюдения за системой, мы также полагаем, что сам процесс сборки происходит через набор

разнообразных и вероятно независимых путей. Это также в согласуется с более ранними и исследованиями по моделированию вероятных промежуточных состояний при формировании параллельного квадруплекса из четырёх тяжёлых  $d(G_4)_4$  [375]. Разные микросостояния (и вероятности их существования) могут по-разному реагировать на изменение как последовательности петель так и изменение внешних условий.

Сложность структурных перестроек в квадруплексных ДНК может быть проиллюстрирована захватом катиона аптмером 15-ТВА показанного в этой работе. Захват катиона молекулой аптамера это достаточно простое событие в структурной динамике квадруплексных ДНК включая процесс их самоорганизации. Уже собранные квадруплексные ДНК могут обмениваться катионами с раствором. В наших исследованиях включая как нативную структуру аптмера так и её модификации, у которых были удалены некоторые петли, мы обнаружили 23 события проникновения катиона в центральный сайт связывания из раствора. Моделирование МД ясно показывает, что кинетика между двумя состояниями, собранной 15-ТВА с катионом и аптамера без катиона, достаточно сложна. Мы наблюдали разнообразные пути проникновения катиона во внутрь квадруплекса со значительной разницей между катионами калия и натрия (Рисунки 3.33, 3.34, 3.38, 3.40). События, типа которых мы наблюдали, могут играть важную роль в процессах структурной самоорганизации, обмен катионами и взаимодействия с другими молекулами. Наши расчёты показывают, что даже такой простой процесс как захват катиона это достаточно сложный и вариативный процесс. Выбор между разными путями может строго зависеть от типа катиона. Мы полагаем, что в целом, все процессы связанные с самоорганизацией квадруплексных ДНК сложны и обладают вариативной природой индивидуальной для каждой последовательности. Практически невозможно получить такую

информацию на основе экспериментальных данных, так как это потребовало эффективной детекции сигнала в микросекундном диапазоне с атомистическим разрешением для ДНК и катионов. И так, не смотря на очевидные ограничения, длительные моделирования МД в будущем смогут принести важное понимание в картину переходов между как микросостояниями, так и макросостояниями, которые можно обнаружить экспериментально, в процессе организации и перестроении структуры квадруплексных ДНК.

## **3.6 Ключевые факторы, влияющие на конформационную динамику 15-ТВА, и их применение**

### **3.6.1 Динамика 15-ТВА и функциональная активность**

В настоящее время ДНК/РНК-аптамеры рассматриваются как потенциально перспективный класс лекарственных средств, обладающих высокой специфичностью и идеальными фармакокинетическими свойствами. Принципиальное отличие аптамеров от олигонуклеотидов, действие которых основано только на последовательности, – это формирование функционально стабильной третичной структуры. Именно стабильность структуры, скорее всего, есть тот фактор, улучшая который, мы можем увеличить эффективность действия аптамера.

Ранее методами калометрии и кругового дихроизма показано, что температура плавления структуры 15-ТВА близка к температуре человеческого тела [376]. Для повышения эффективности действия аптамера необходимо увеличить его структурную стабильность. При изучении структурно-динамических свойств аптамера 15-ТВА нами выявлено возможные пути повышения его структурной стабильности. Самый простой путь подразумевает повышение стабильности квадруплекса путём формирования стабильного комплекса аптамера с катионом. Однако тромбин взаимодействует с ТТ-петлями, подвижность которых связана с динамикой самого квадруплекса. По-видимому, сила связывания катиона с квадруплексом модулирует эффективность взаимодействия аптамера с тромбином. Этот путь легко проверить путём оценки тромбин-ингибирующей активности аптамера в присутствии катиона бария. Нами выявлено, что константа аффинности для комплекса аптамера с барием на порядок выше, чем при связывании катиона калия, а энтальпии комплексообразования для обоих

катионов сравнимы. Присутствие катиона бария значительно повышает температуру плавления квадруплекса, энергия сольватации катиона бария составляет -1305 кДж/моль, а для катиона калия -302 кДж/моль. И это согласуется с результатами моделирования: по сравнению с катионом калия барий практически неподвижен. В экспериментах, проведенных совместно с Е.В. Завьяловой, обнаружено, что в присутствии катиона бария антитромбиновая активность аптамера гораздо ниже, чем в присутствии калия (Рисунок 3.48).

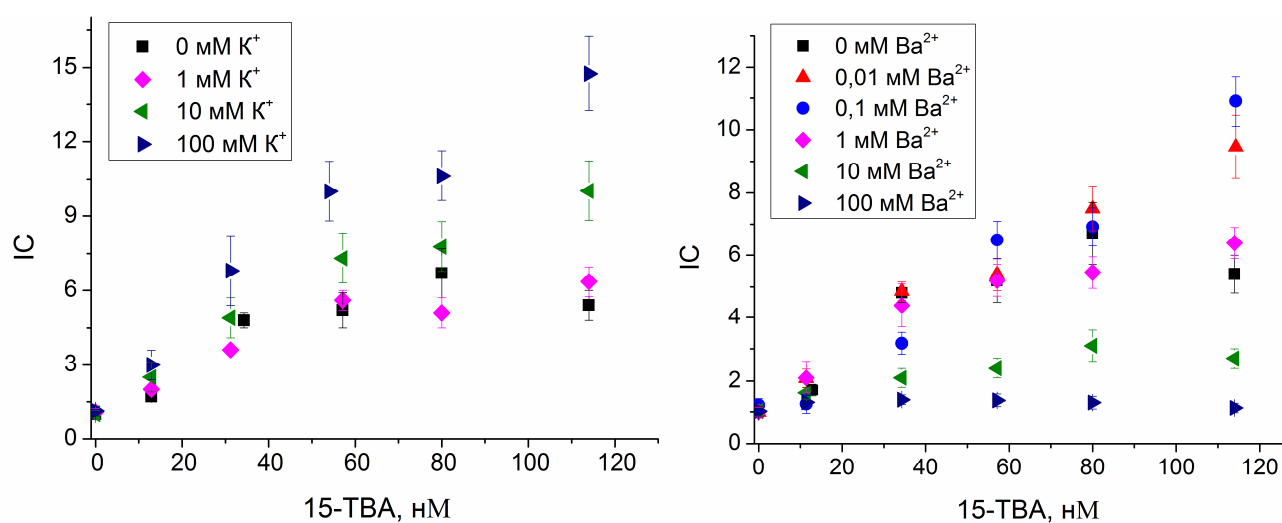


Рисунок 3.48: Сравнение ингибирующей активности 15-ТВА в присутствии катионов калия и бария.

Более того, в своей диссертационной работе Завьялова Е.В. утверждает, что ингибирующая активность аптамера 15-ТВА обратно пропорциональна степени стабилизации структуры катионами. То есть если можно составить ряд катионов по стабилизирующему влиянию на квадруплекса как:  $\text{Na}^+ < \text{K}^+ < \text{Ba}^{2+}$ , то влияние этих катионов на ингибирующую активность имеет обратную зависимость:  $\text{Na}^+ > \text{K}^+ > \text{Ba}^{2+}$ . Отсюда следует, что в ходе взаимодействия с тромбином квадруплекс или ГТ-петли не должны терять свою подвижность.

Для изучения атомистических причин этого явления мы использовали моделирование диссоциации комплекса 15-ТВА с тромбином методом зонтичного



интегрирования (“umbrella sampling”). Предполагается, что диссоциация комплекса тромбин–15-ТВА – это процесс линейного смещения центров масс тромбина и аптамера, тогда координатой реакции будет выступать расстояние между соответствующими центрами масс. В этом моделировании нас интересует изменение структуры аптамера в ходе реакции диссоциации с белком. В ходе сканирования координат реакции с шагом  $0.5\text{ \AA}$  квадруплексная часть аптамера претерпевает значительные изменения. В районе приблизительно  $35\text{ \AA}$  для расстояния между центрами масс аптамера и тромбина мы наблюдаем уменьшение числа водородных связей как между тромбином и аптамером, так и между гауниновыми основаниями квадруплекса (Рисунок 3.49).

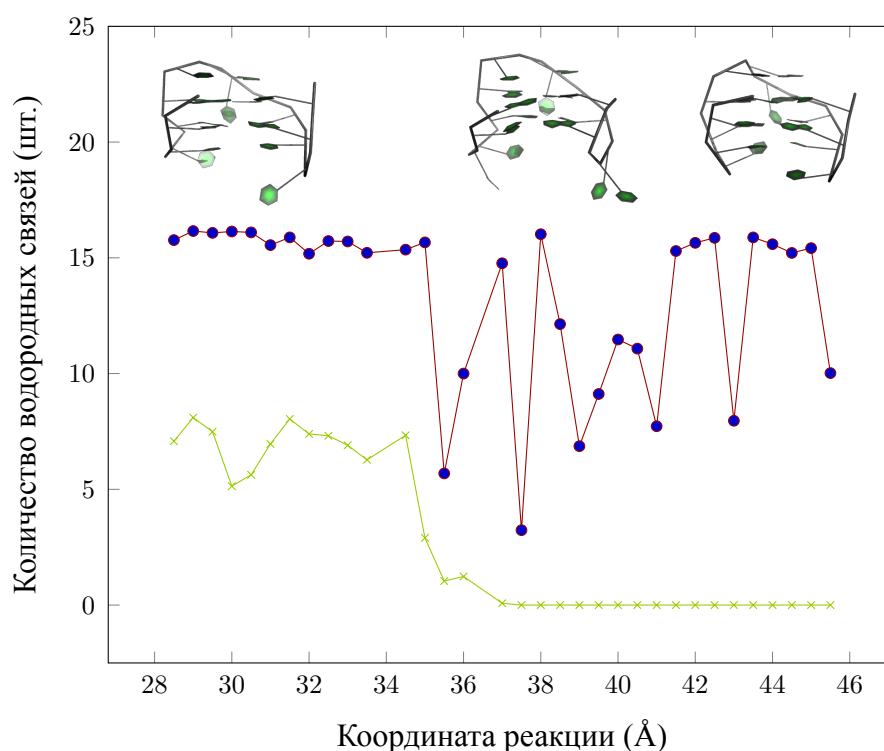


Рисунок 3.49: Изменение числа водородных связей в ходе моделирования диссоциации комплекса тромбин–15-ТВА. Над графиками указаны структуры аптамера, соответствующие участкам траектории. Синими символами отмечены водородные связи в квадруплексе, зелёными – в комплексе 15ТВА-тромбин.

Итак, квадруплекс теряет часть внутренних водородных связей при формировании конформации, которая способна связаться с тромбином. Чем силь-

нее связан катион внутри квадруплекса, тем меньше вероятность формирования такого переходного состояния. Стабилизация квадруплексной структуры не приводит к повышению ингибирующей активности аптамера. Необходимо сохранять конформационную подвижность аптамера, близкую к подвижности 15-ТВА, но при этом уменьшить представленность конформаций аналогичных полностью денатурированному состоянию.

### **3.6.2 Исследование структурного влияния дуплексной части ДНК на G-квадруплекс**

Известно, что просоединение предположительно дуплексного участка к концам молекулы 15-ТВА приводит к повышению активности такой конструкции как ингибитора тромбина [377]. С целью установить трехмерную структуру такого гибридного аптамера (в дальнейшем мы будем называть его 31-TGT) мы провели ЯМР-исследование. На основе соотнесения пиков спектров двумерных NOESY, DQF-COSY и HSQC найдено 247 дистанционных ограничений между атомами 31-TGT. С использованием этих ограничений построена модель структуры (Рисунок 3.50).

Как и в случае с 15-ТВА, аптамер 31-TGT содержит квадруплексный фрагмент с идентичным 15-ТВА состоянием ТТ-петель. Однако конформация TGT-петли претерпела существенные изменения (Рисунок 3.50). Так, G16 (соответствует G8 из 15-ТВА) из TGT-петли образует стэкинг-взаимодействия не с кватетом, а с остатком A8 из дополнительной части. Заметим, что в 31-TGT есть дуплексный участок, но он не вовлечен во взаимодействия с квадруплексом. Таким образом, в аптамере 31-TGT действительно реализован эффект отсутствия стабилизирующего влияния дуплексной части на квадруплекс. К сожалению,

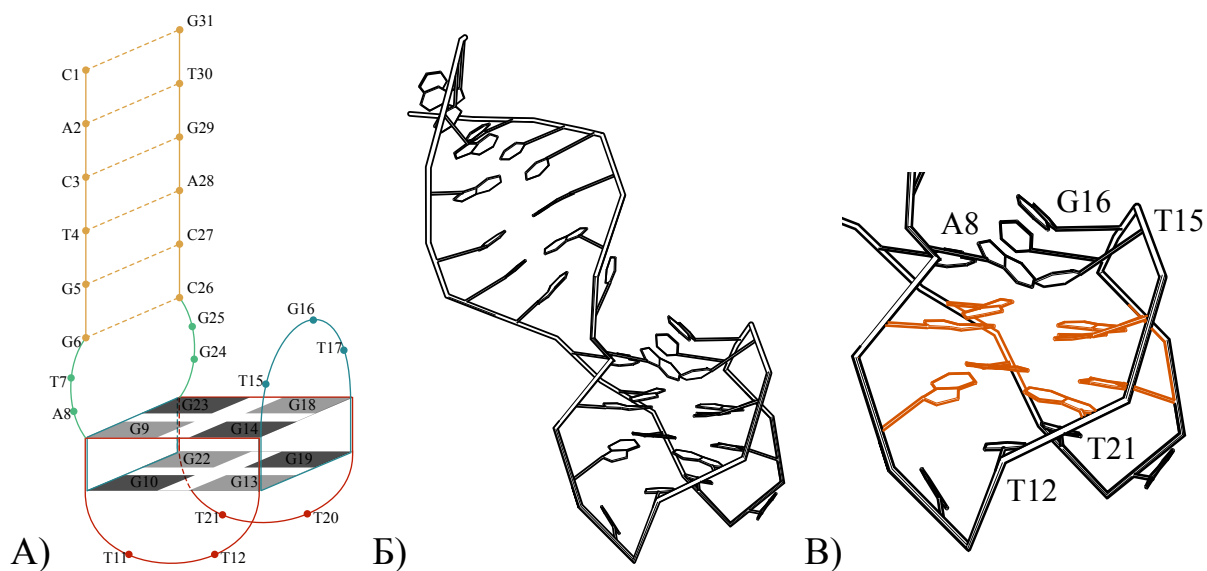


Рисунок 3.50: Структура олигонуклеотида 31-TGT. А) Вторичная структура 31-TGT. Красным отмечена квадруплексная часть, оранжевым – дуплексная часть. Б) Модель третичной структуры 31-TGT по данным ЯМР. В) Модель структуры квадруплексной части 31-TGT, оранжевым цветом отмечены G-квартеты

аптамер 31-TGT не отличается высокой специфичностью – он связывается и с протромбином, а это сильно ограничивает его терапевтический потенциал. Дело в том, что концентрация протромбина в крови намного выше концентрации тромбина, а значит и эффективные концентрации будут высокими.

### 3.6.3 Разработка терапевтического аптамера к тромбину

Мы предприняли попытку разработать аптамер на основе 15-ТВА с дополнительным структурным модулем, увеличивающим аффинность связывания аптамера с тромбином и при этом избежать неспецифических взаимодействий с другими белками крови. При разработке новой конструкции учитывали, во-первых, найденную корреляцию активности 15-ТВА с подвижностью его ТТ-петель, а во-вторых, наблюдение, что при захвате катиона из раствора происходит временная частичная дисбалансировка структуры квадруплекса с изменением подвижности концевых гуанинов 15-ТВА. При этом 5'-конец 15-ТВА под-

вержен заметным структурным перестройкам, а 3'-концевой нуклеотид не проявляет значимой подвижности. Так, введение модификаций на 5'-конец больше влияет на функциональную активность аптамера, чем при модификациях 3'-конца. Это напрямую связано с динамикой структуры всей молекулы.

Одним из предложенных нами решений было следующее: сделать двухквадруплексный вариант аптамера, т.е. две последовательности 15-ТВА, соединённые коротким нуклеотидным мостиком. У такой молекулы (в дальнейшем RA-36) на 5'-конце будет квадруплекс, подвижность которого не отличается от 15-ТВА, а квадруплекс на 3'-конце будет более подвижен, так как с его 5'-концом связан первый квадруплекс.

Сейчас методом кругового дихроизма выявлено наличие квадруплексной структуры в растворе RA-36. Данные о тромбин-ингибирующей активности RA-36 подтверждают его эффективность. Параметрически RA-36 обладает свойствами потенциально пригодного для терапии препарата и в настоящее время находится на стадии формальных доклинических испытаний. Таким образом, моделирование позволяет понять структурные особенности НК и на основе полученных данных проводить рациональный дизайн новых биологически активных соединений.

### **3.6.4 Разработка аптасенсора к тромбину**

Понимание структурной динамики аптамера 15-ТВА мы применили для создания сенсоров, которые можно использовать для определения концентрации тромбина. Надо признать, что на текущий момент в клинике практически не представлены методы прямого определения тромбина в крови, хотя эта информация очень важна для понимания статуса гомеостаза пациента.

Сенсор разрабатывали как элетропроводник на основе углеродных нанотрубок. Ранее показано, что одноцепочечная ДНК способна сорбироваться на поверхности углеродных нанотубок [378], образуя сеть стэкинг-взаимодействий. ДНК обладает высокой плотностью отрицательного заряда и это может влиять на проводящие свойства углеродной нанотрубки. Предполагалось, что аптамер при взаимодействии с нанотрубкой будет частично денатурировать и сорбироваться на поверхности, а в присутствие тромбина равновесие будет смещаться в сторону собранной квадруплексной структуры. Эти переходы будут сопровождаться значимым изменением электростатического окружения нанотрубки и её проводимости. Методами МД нами было изучено изменение структуры аптамера 15-ТВА в зависимости от способа и условий конъюгации с углеродной нанотрубкой.

При конъюгации аптамера с нанотрубкой рассматривали несколько линкеров, которые доступны как имидиты для встраивания в цепь ДНК (Рисунок 3.51).

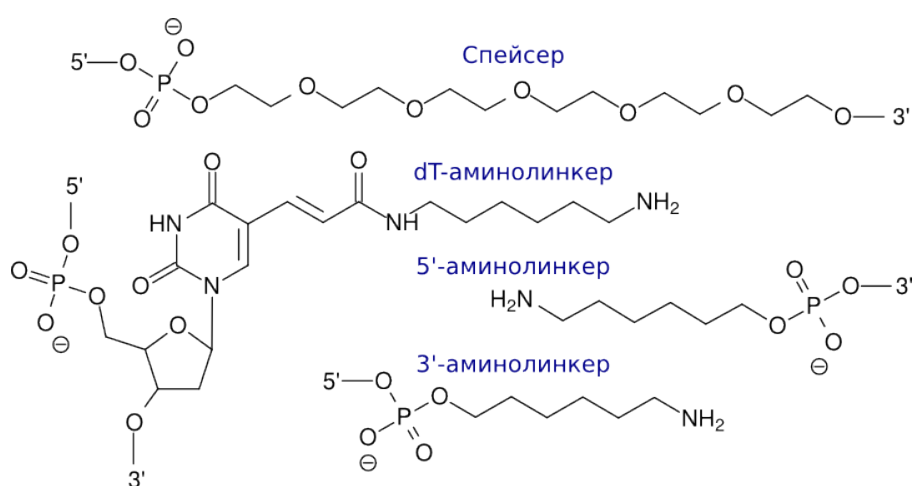


Рисунок 3.51: Аминолинкеры и спейсер для модификации 15-ТВА.

Модифицированные аптамеры содержали аминолинкеры на одном из концов молекулы либо на нуклеотиде Т7, а также спейсер который соединялся с аминогруппой (Рисунок 3.51). Такая “посадка” линкеров позволяет охватить

наиболее общие варианты связывания 15-ТВА с нанотрубкой: через 3'-конец, 5'-конец или внутренний нуклеотид. Исследование спейсеров разной длины нацелено на то, чтобы подобрать оптимально подходящее для функционирования расстояние между нанотрубкой и аптамером.

Наряду с ковалентным присоединением аптамера к нанотрубке мы рассматривали вариант нековалентной сорбции конъюгатов аптамеров с полисорбатом (Tween-20). Сорбция происходила благодаря гидрофобному эффекту между полисорбатной частью конъюгата и нанотрубкой. В выбранном силовом поле не содержится описания для спейсера, линкеров и Tween-20. С помощью программ R.E.D. и Gamess US [379; 380] были рассчитаны частичные заряды на атомах для соответствующих конъюгатов. Так как точно неизвестно, насколько плотно могут молекулы Tween-20 покрывать нанотрубки и как это можно регулировать, то в расчётной части были реализованы два крайних случая. В первом – на каждую молекулу аптамера приходится одна молекула Tween-20. Во втором варианте молекул Tween-20 столько, что они покрывают нанотрубки монослоем (30 штук).

Итак, было исследовано 30 систем, которые различаются:

- модификацией аптамера (5 групп)
  - 5'-аминолинкер
  - 5'-аминолинкер + спейсер
  - 3'-аминолинкер
  - 3'-аминолинкер + спейсер
  - dT-аминолинкер
- ионами (2 группы)

- противоион тетраметиламмоний
- противоион натрий + физиологическая концентрация NaCl;
- способом присоединения к нанотрубке, или функционализацией (3 группы)
  - ковалентная
  - гидрофобное взаимодействие (через Tween-20 с высокой концентрацией чистого Tween 20)
  - гидрофобное взаимодействие (через Tween 20 с низкой концентрацией чистого Tween 20)

Использование тетраметиламмония вызвано необходимостью иметь внутренний контроль для систем, где катион не проникает в центральную полость квадруплекса.

Модифицированные таким образом системы исследовали методом МД. Время наблюдения каждой системы в моделировании МД составило 1 мкс. Выявлено, что конформация аптамера в каждой системе по-своему уникальна. Так, в одних системах сохранялся G-квадруплекс, в других его не было, а в большинстве систем нельзя невозможно было сделать однозначные выводы.

Так как простого способа оценки стабильности квадруплекса пока нет, для формального анализа конформации мы использовали несколько критериев.

Во-первых, G-квадруплекс 15-ТВА состоит из двух квартетов, наличие которых является необходимым и достаточным условием существования квадруплексной конформации. Мы можем исследовать каждый квартет отдельно. Число водородных связей между азотистыми основаниями квартета колеблет-

ся от 0 до 8. С помощью этого критерия можно косвенно оценивать расстояние между азотистыми основаниями.

Во-вторых, характерная структурная особенность квартета заключается в том, что он плоский, то есть все атомы азотистых оснований лежат в одной плоскости. Оценить, насколько любая конкретная конформация гуанинов квартета отличается от планарной, можно с помощью системы расчетов.

В-третьих, Следующий критерий также происходит из оценки планарности плоскости квартета. Для его расчёта необходимо построить 2 плоскости: через атомы O6 гуанинов и через атомы C1' дезоксирибоз. Двугранный угол между полученными плоскостями и будет необходимой величиной планарности. В плоском квартете квадруплекса этот угол будет равен 0 градусов. Угол будет принимать минимальные значения, если азотистые основание симметрично выгнуты. Таким образом, в случае отклонения квартета от плоской конфигурации этот критерий покажет, насколько азотистые основания выходят из плоскости.

В-четвертых, мы оценивали взаимодействие аптамера с нанотрубкой. Известно, что поверхность нанотрубки может быть вовлечена в стэкинг-взаимодействия с ароматическими группами, что в случае с одноцепочечными НК приводит к их наматыванию вокруг нанотрубки. По этой причине нас интересовало, взаимодействует ли аптамер с нанотрубкой и в какова поверхность контакта. Измерить это можно, посчитав расстояние от центра масс квартетов до продольной оси нанотрубки. Радиус нанотрубки составляет 11 Å. Если один квартет аптамера в G-квадруплексной конформации вступает в стэкинг-взаимодействие с нанотрубкой, то рассчитанное расстояние будет больше радиуса нанотрубки примерно на два ван-дер-ваальсовых радиуса. При этом второй квартет будет находиться от поверхности нанотрубки на расстоянии примерно в два раза большем, чем первый. В случае же наматывания аптамера на нано-



трубку расстояние от центра масс атомов квартета может уменьшиться вплоть до 0 Å.

Для каждой системы проведён анализ всех вышеописанных параметров. Трёхмерные графики содержат точки для каждого момента траектории с шагом в 200 пс в пространстве описанных критериев. Для возможности оценки эволюции системы во времени каждая точка окрашена цветом в соответствии с приведённой шкалой – от чёрного до жёлтого. На двумерном графике показано расстояние от центра квартетов до оси нанотрубки и радиус нанотрубки.

Рассмотрим детально примеры стабильной и нестабильной структур. Стабильным оказался аптамер с 5'-аминолинкером и спейсером, в соли NaCl, с гидрофобной функционализацией при высокой концентрации Tween-20 (Рисунок 3.52).

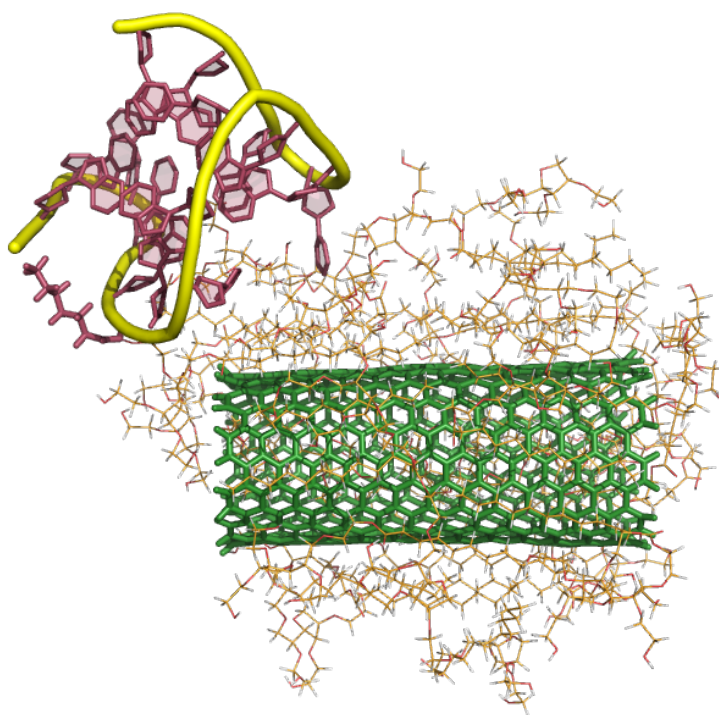


Рисунок 3.52: Пример системы аптамер-нанотрубка со стабильным G-квадруплексом: аптамер с 5'-аминолинкером, гидрофобная функционализация через Tween 20 при высокой концентрации Tween-20 в ячейке с раствором NaCl.

Распределение точек для обоих квартетов очень похоже (Рисунок 3.53). Высокая плотность точек в определённой области указывает на стабильную конформацию. Независимо от времени, значения планарности и двугранного угла близки к нулю, а количество водородных связей лежит в основном в пределах между 6 и 10, что говорит в пользу того, что это G-квадруплексная конформация. Хотя мы считали, что число водородных связей в аптамере не превышает 8, здесь обнаружены структуры с 9-ю и 10-ю водородными связями – и, действительно, по формальным критериям водород при атоме N1 может участвовать в двух водородных связях одновременно. Также можно заметить, что в течение всего времени моделирования МД квартеты находятся на значительном расстоянии от нанотрубки – не менее 15 Å от её поверхности. Такого расстояния недостаточно для прямого контакта. Очевидно, этому препятствует ‘шуба’, образованная из молекул Tween-20. Однако со временем квартеты постепенно приближаются к нанотрубке. Возможно, при более длительном наблюдении аптамер достигнет поверхности трубки.

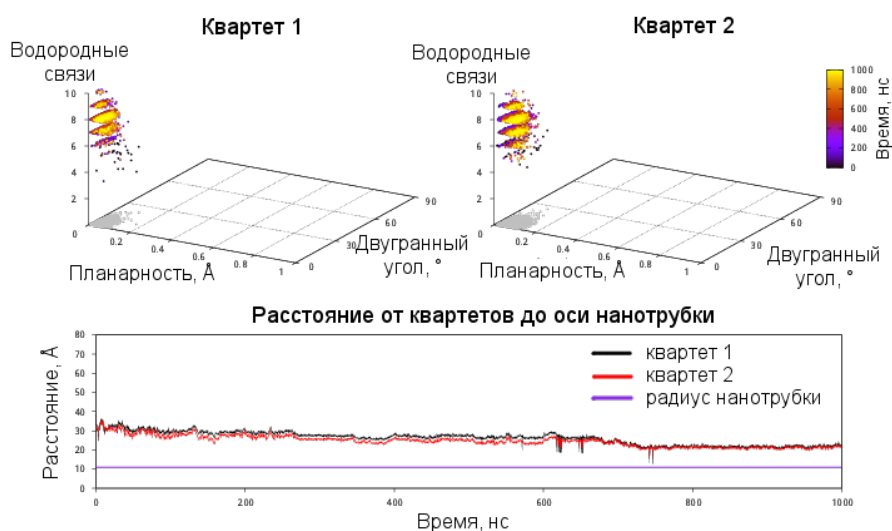


Рисунок 3.53: Пример графиков оценки конформации для системы аптамер-нанотрубка со стабильным G-квадруплексом: аптамер с 5'-амиолинкером, гидрофобная функционализация через Tween-20 при высокой концентрации Tween-20 в ячейке с раствором NaCl.

Теперь обратимся к системе, в которой аптамер потерял исходную конформацию и накрутился на нанотрубку. Это система с 3'-аминолинкером и спейсером на аптамере, с тетраметиламмонием в качестве противоиона, гидрофобной функционализацией без чистого Tween 20 (Рисунок 3.54).

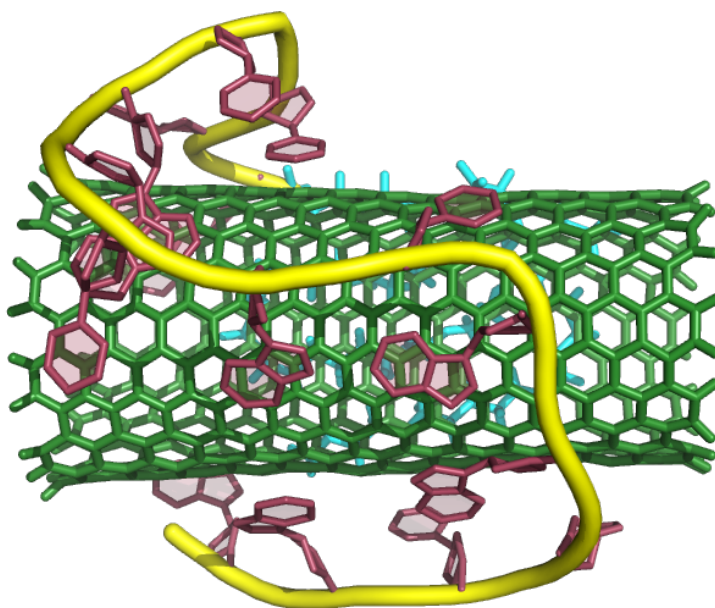


Рисунок 3.54: Пример системы аптамер-нанотрубка с обернутым вокруг нанотрубки аптамером: аптамер с 3'-аминолинкером и спейсером, гидрофобная функционализация через Tween 20 при отсутствии Tween-20 в ячейке с противоионом ТМА.

Распределение точек квартетов совершенно различны (Рисунок 3.55). Для второго квартета заметно скопление точек в области квадруплексной конформации в начальный момент времени (не более 100 нс), а после область точек становится более размытой. В ходе моделирования оба квартета не имеют квадруплексной конформации: водородных связей немного, а планарность и двугранный угол, наоборот, принимают высокие значения. На третьем графике показано, что центры масс квартетов лежат от оси трубки на расстоянии меньшем, чем её радиус, то есть аптамер намотан на нанотрубку.

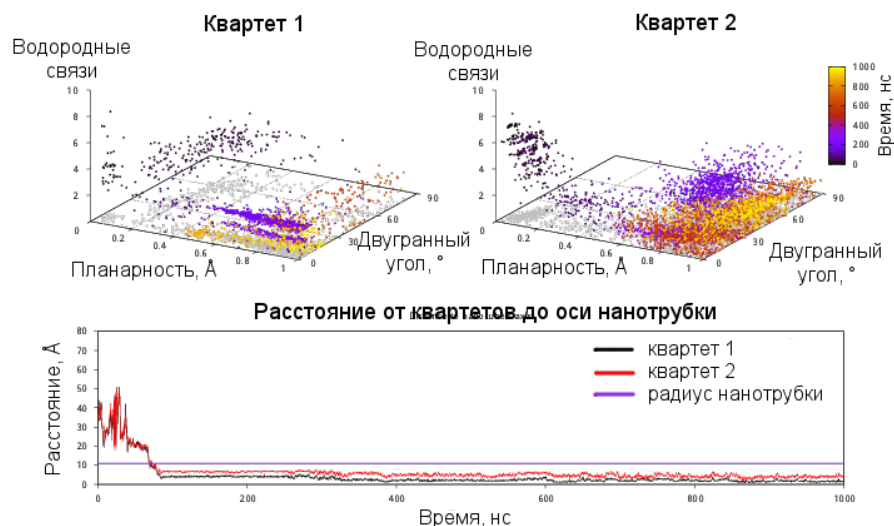


Рисунок 3.55: Пример графиков оценки конформации для системы аптамер-нанотрубка с намотанным на нанотрубку аптамером: аптамер с 3'-аминолинкером и спейсером, гидрофобной функционализацией через Tween 20 в отсутствие чистого Tween-20 в ячейке с противоионом TMA.

### Сравнение стабильности 15-ТВА-содержащих систем

На основании анализа все системы были поделены на пять групп: с одним стабильным кватетом (✓), с двумя стабильными кватетами (✓✓), с полностью развёрнутыми кватетами (✗), с развёрнутыми кватетами и обёрнутой вокруг нанотрубки ДНК (✗●), а также группа с частично развёрнутыми кватетами (↔) (Таблица 3.9). Стоит отметить, что длины траектории даже в 1 мкс, возможно, является недостаточно для того, чтоб проследить кинетически стабильные и термодинамически выгодные конформации аптамера в составе конъюгата с нанотрубкой. Поэтому частично расплетённые конформации, вероятно, будут полностью развёрнуты при более длительном наблюдении. При гидрофобной функционализации легко заметить, что в присутствии катиона  $\text{Na}^+$  в основном наблюдаются стабильные конформации, а контрольные системы с тетраметиламмонием подтверждают, что без катиона металла аптамер теряет квадруплексную структуру. При изучении третьего графика по каждой системе

	Нековалентно				Ковалентно	
	Мало Tween 20		Много Tween 20		Na <sup>+</sup>	ТМА
	Na <sup>+</sup>	ТМА	Na <sup>+</sup>	ТМА		
3'-ал	✓✓	✗●	✓✓	✗	↔	✗
3'-ал+сп	✓	✗●	✓✓	✗	↔	↔
5'-ал	↔	✗	✓✓	✗	✓✓	✗
5'-ал+сп	↔	✓	✓✓	✗	✗	↔
dT-ал	✓	✓	✗	✓	✗	✗

Таблица 3.9: Анализ стабильности квадруплексной структуры 15-ТВА в конъюгатах с углеродными нанотрубками. Представлены 5 групп: с одним стабильным кватером (✓), с двумя стабильными кватертами (✓✓), с полностью развёрнутыми кватертами (✗), с развёрнутыми кватертами и обёрнутой вокруг нанотрубки ДНК (✗●), а также группа с частично развёрнутыми кватертами (↔).

также было замечено, что при высокой концентрации Tween 20, довольно быстро образуется монослой на поверхности нанотрубки, не позволяя ей взаимодействовать с аптамером напрямую. Можно заключить, что расплетание аптамера происходит как при воздействии нанотрубки, так и при экранировании её монослоем Tween 20. В системах с ковалентной функционализацией наличие Na<sup>+</sup> недостаточно для поддержания стабильного G-квадруплекса и аптамер разворачивается во всех системах, за исключением двух случаев. Во всех системах с Na<sup>+</sup> аптамер, модифицированный с 3'-конца, показал себя более устойчивым при воздействии с нанотрубкой.

### Предполагаемый механизм работы аптосенсора к тромбину

Как было показано, G-квадруплексная конформация может как сохраняться, так и разрушаться, и это зависит от способа модификации аптамера. Так, введение модификации по 5'-концу аптамера приводит к потере структуры квадруплекса при конъюгации с нанотрубкой. Такая же модификация на 3'-конце аптамера не приводит к столь значимым последствиям.

Основываясь на этих результатах, мы в сотрудничестве с группой Бобринецкого И.И. из национального исследовательского университета “МИЭТ” провели сборку тромбинового аптасенсора. Использование карбоксил-функционализированных нанотрубок позволило создать серию сенсоров с идентичными характеристиками. В экспериментах *in vitro* показано, что сенсоры достаточно чувствительны и специфичны к тромбину. При сравнении сенсоров с разными модификациями (на 5'- и 3'-концах) выявлено, что только модификация по 5'-концу позволяет детектировать тромбин.

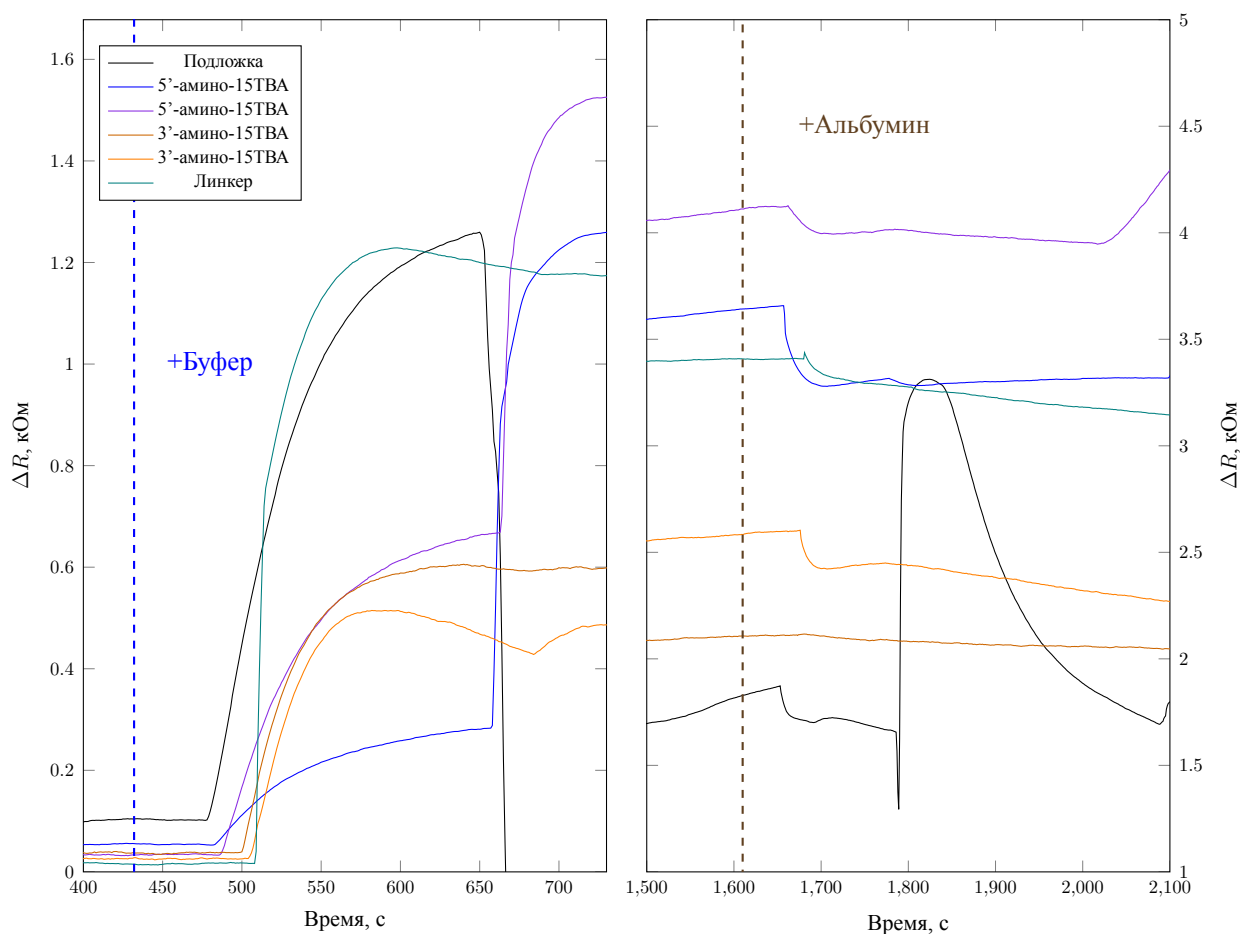


Рисунок 3.56: Ответ аптосенсоров на последовательное добавление буферного раствора, тромбина и альбумина.

На основании как экспериментальных (Рисунок 3.56), так и расчетных данных, можно предложить механизм работы подобных сенсоров. Проводимость нанотрубки зависит от того, участвует ли 15-ТВА в многочисленных стэкинг-

взаимодействиях с поверхностью нанотрубки. В отсутствие тромбина аптамер 15-ТВА не имеет выраженной структуры и распластан на поверхности нанотрубки, но при возможности образовать комплекс с тромбином конформационное равновесие смещается в сторону компактной структуры, которая практически не имеет контактов с нанотрубкой. Именно столь значимое конформационное изменение, вероятнее всего, и приводит к изменению проводимости трубки. Предполагаемый механизм представлен на Рисунке 3.57.

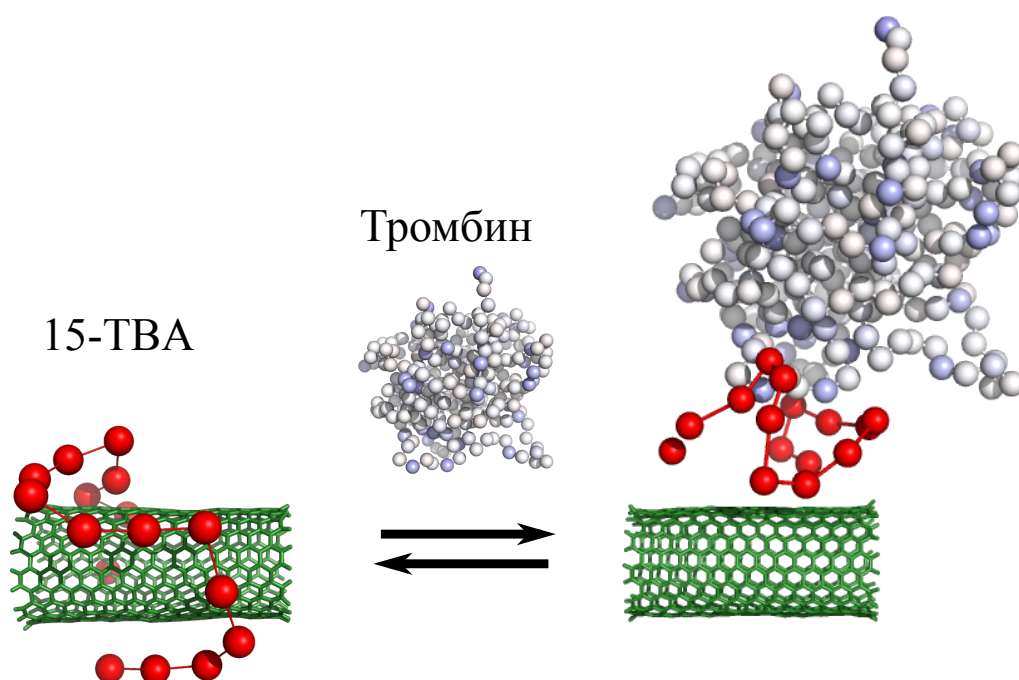


Рисунок 3.57: Изображение механизма работы аптасенсора на основе 5'-модифицированного варианта аптамера к тромбину. Красные сферы - это отрицательно заряженные фосфатные группы. Зеленым окрашена углеродная нанотрубка. Белым и голубым отмечен тромбин.

## **3.7 Температурная зависимость пути самоорганизации ДНК шпильки d(GCGCAGC)**

### **3.7.1 Метод обмена репликами для изучения фазового пространства биополимеров**

Как мы уже показали, компьютерное моделирование, такое как МД, может быть использовано для изучения сложных молекулярно-биологических систем. Для классического моделирования с использованием силовых полей существует два источника ошибок: первый – это неточность при составлении силового поля, второй – недостаточное сканирование фазового пространства. По отношению к моделированию МД главный фактор ошибок – это недостаточное время наблюдения. на основании полученных нами результатов можно говорить, что для такой небольшой молекулы как 15-ТВА время наблюдения должно находиться в микросекундном диапазоне. Неудивительно, что исследователи часто используют простые системы для проверки силовых полей и методологий сканирования фазового пространства. Надо отметить, что важен не только размер системы, но способность биополимера быстро самоорганизоваться в нативную структуру. Именно такие объекты выбираются для одного из самых сложных применений моделирования МД – моделирования процесса самоорганизации структуры биополимера линейной конформации. Такого рода моделирование должно быть очень продолжительным, время наблюдения за системой зависит от её природы и может достигать секунд, что в миллион раз больше, чем траектории, описанные в этой работе. Важно, чтобы силовые поля правильно описывали не только конформации, известные из экспериментальных данных, но и те состояния, которые являются нестабильными. Затраты на подобные экспери-



менты компенсируются значимостью получаемых результатов: возможно получение детальной информации о кинетике переходных состояний и взаимосвязи этапов самосборки. Надо признать, что моделирование МД не единственный путь получения подобной информации. Моделирование методом Монте-Карло тоже может дать реалистичный путь механизма самосборки, но кинетическая информация при этом может быть утеряна. Хотя другие методы предлагают эффективные алгоритмы расчетов, только поломоатомная молекулярная динамика позволяет получать детальную информацию о переходах между конформациями, что и является ключевым моментом в описании процесса самосборки. Традиционные экспериментальные методы ЯМР и РСА идентифицируют доминирующую в растворе или единственную в кристалле конформацию, хотя за выполнение биологической функции может отвечать одна или несколько относительно менее представленных конформаций. Такие состояния, а также кинетические барьеры между ними являются ключевыми для понимания функционирования молекулы. Именно эти данные мы искали в траекториях моделирования самосборки выбранной ДНК-шпильки.

Для решения этой задачи был использован подход, основанный на применении цепей Маркова с дискретным временем [381]. В этом подходе предполагается, что конформационные переходы внутри микросостояний происходят значительно быстрее, чем между микросостояниями. Поэтому, относящиеся к одному микросостоянию конформации должны быть структурно достаточно близкими. Вероятно, к одному микросостоянию должно относиться достаточно большое число конформаций.

Нами проведено исследование по сравнению конформационного ландшафта короткой ДНК-шпильки при разных температурах. Моделирование самосборки шпилечной структуры проводили при 27°C и 0°C. Одна из основных

проблем, возникающих при моделировании самосборки НК, равно как и других биологических макромолекул, – большое число локальных минимумов энергии. Существует ряд методов, направленных на решение этой задачи. Так, например, для изучения самосборки биологических макромолекул группа Pande [382] использует большое число коротких траекторий молекулярной динамики. Данный подход исходит из предположения о том, что самосборка – это очень редкий, но очень быстрый процесс. Ряд статей этой группы посвящен самосборке НК. Так, в одной из работ [383] обсуждается образование структуры 5'-GGGC[GCAA]GCCU шпильки в неявно заданном растворителе, а в статье [citermid15681648](#) сравниваются пути самосборки той же шпильки в явно и неявно заданных растворителях. Другим подходом к решению проблемы большого числа локальных минимумов является метод обмена репликами (Replica Exchange Molecular Dynamics, REMD). Существует лишь несколько работ, посвященных REMD ДНК [384—387]. Суть этого метода состоит в создании копий, или реплик, системы с их последующим моделированием при различных температурах. В ходе моделирования каждая копия системы совершает случайные блуждания по пространству температур в результате обмена состояниями между соседними по температурам репликами через равные промежутки времени. Вероятность обмена между соседними репликами выражается формулой:

$$P(1 \leftrightarrow 2) = \min \left( 1, \exp \left[ \left( \frac{1}{k_b T_1} - \frac{1}{k_b T_2} \right) (U_1 - U_2) \right] \right),$$

где  $k_b$  — константа Больцмана,  $T_1$  и  $T_2$  — температуры, а  $U_1$  и  $U_2$  — мгновенные потенциальные энергии реплик 1 и 2 соответственно. Иначе говоря, если две соседние реплики (реплика 1 с температурой  $T_1$  и реплика 2 с температурой  $T_2$ ) оказались близкими по энергии, то может произойти переход, в результате

которого конформация, содержащаяся в реплике 1, перейдет в реплику 2 и будет далее моделироваться при температуре  $T_2$ , а конформация, содержащаяся в реплике 2, перейдет в реплику 1 и будет далее моделироваться при температуре  $T_1$ .

На данном шаге попытки обмена совершаются не между всеми возможными соседними парами реплик: есть деление на “четные” и “нечетные” попытки обмена, осуществляемые по очереди. Для системы из 5 реплик с индексами температур 0 – 4 в ходе “нечетных” попыток обмена будут пробоваться пары 0 – 1, 2 – 3 на 1м, 3м и т. д. шагах, а в ходе “четных” — пары 1 – 2 и 3 – 4 на 2м, 4м и т. д. шагах (Рисунок 3.58).

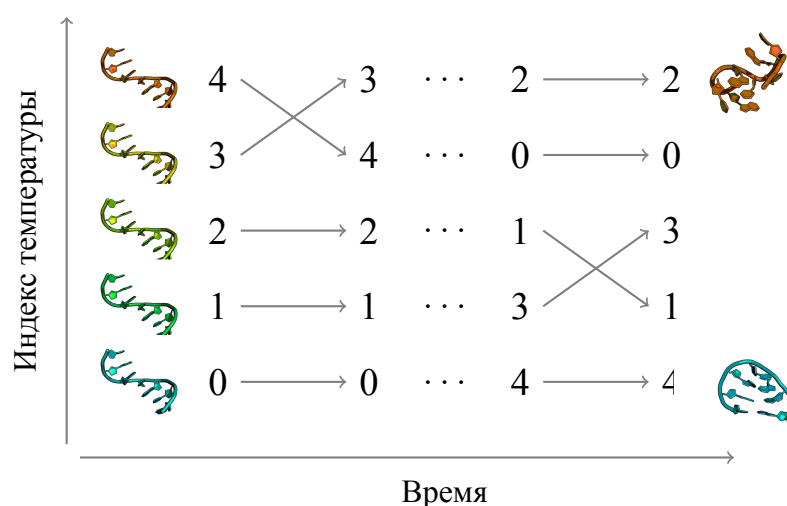


Рисунок 3.58: Молекулярное моделирование по методу обмена репликами.

Диапазон температур выбирается таким образом, чтобы нижняя температура соответствовала изучаемой, а при верхней легко осуществлялись переходы через энергетический барьер. В идеальном случае при верхней температуре у системы должен быть единственный глобальный минимум. Промежуточные температуры распределены экспоненциально и подбираются из такого расчета, чтобы частота обмена между репликами была примерно одинаковой и составляла около 0.15–0.45. Таким образом, в ходе моделирования молекулярной

динамики при высоких температурах происходят конформационные переходы, которые при более низких температурах являются очень редкими событиями, то есть осуществляется переход между соседними потенциальными “ямами”. При низких температурах система спускается в локальный минимум, находя наиболее энергетически выгодную конформацию. Все это делает REMD довольно эффективным методом для моделирования самосборки биологических макромолекул.

Есть ряд статей, посвященных моделированию самосборки НК методом обмена репликами или его вариациями. Так, в работе [386] проводили самосборку дуплекса ДНК 5'-GGCC х2 в неявно заданном растворителе. В работе [384] обсуждается самосборка ДНК шпильки 5'-GCGCAGC, а в работе [385], помимо моделирования самосборки, также сравнивается стабильность шпилек 5'-GCGCAGC и 5'-CGCGACG. Ряд статей [387—389] моделирует поведение ДНК в присутствии углеродной нанотрубки методом REMD.

Существует большое число вариаций метода обмена репликами, но лишь часть из них применима к НК. Так, в одной из работ [390] сравнивается традиционный REMD и вариация, называемая (BP)-RexMD, в приложении к моделированию поведения аденинового динуклеотида и ДНК-дуплекса без одного основания. В отличие от традиционного метода обмена репликами, в (BP)-RexMD реплики отличаются между собой не температурами, а коэффициентами перед потенциалами, описывающими взаимодействия в системе.

Метод SREMD – серийный метод обмена репликами – представляет собой REMD, адаптированный для распределенных вычислений на большом числе персональных компьютеров, как в проекте Folding@home [391]. Pande с соавт. моделируют самосборку 5'-GGGCGCAAGCCU РНК-шпильки и исследуют полученные микросостояния.

### 3.7.2 Шпилечные структуры

Шпильки представляют собой элементы вторичной структуры НК, состоящие из одноцепочечного участка (петли) и стебля двухцепочечного участка с уотсон-криковскими парами, ограничивающего петлю. Шпильки играют важную роль во многих биологических процессах. Образование подобных структур характерно для большинства мажорных клеточных РНК и важно для выполнения их биологических функций. Шпильки, обнаруживаемые в точках начала репликации [392; 393] и центромерах [394; 395]), участвуют в регуляции репликации. Экспрессия генов может регулироваться с участием шпилек на уровне транскрипции или трансляции [396; 397], где стабилизация или дестабилизация шпилек – в зависимости от различных условий (при изменении температуры, взаимодействии с низкомолекулярными лигандами, белками или РНК) – может вызывать преждевременную терминацию транскрипции или влиять на инициацию трансляции. Считается, что образование шпилек приводит к экспансии тринуклеотидных повторов, что является причиной таких заболеваний человека, как хорея Хантингтона, синдром ломкой X-хромосомы и миотоническая дистрофия [398]. Шпильки также являются типичным мотивом рибозимов – молекул РНК, способных к катализу [399]. Особой стабильностью отличаются шпильки, имеющие последовательность петли GNA (Рисунок 3.59). Возможной причиной такой стабильности является образование дополнительной неканонической пары между первым (G) и третьим (A) нуклеотидами петли [400].

Особой стабильностью отличаются шпильки, имеющие последовательность петли GNA (Рисунок 3.59). Возможной причиной такой стабильности является образование дополнительной не канонической пары между первым (G) и третьим (A) нуклеотидами петли [400].

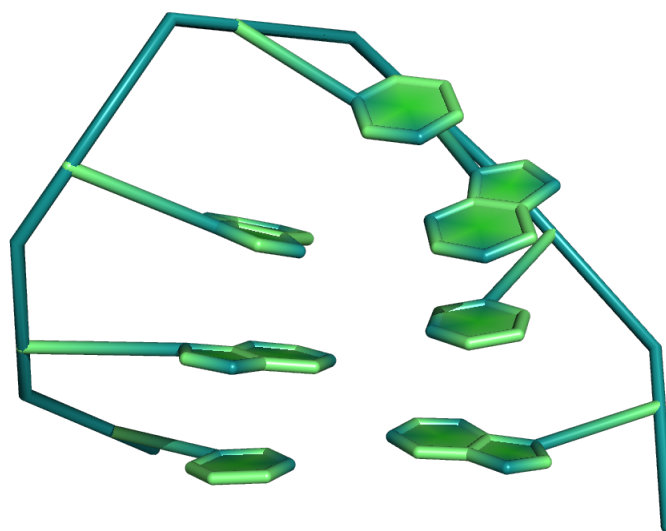


Рисунок 3.59: Пространственная структура 7-нуклеотидной шпильки, содержащей GCA-петлю

### 3.7.3 Анализ конформаций

Методы ЯМР и РСА позволяют, в основном, доминирующую в растворе или единственную в кристалле конформацию молекулы, однако, за выполнение биологической функции может отвечать одна или несколько относительно менее представленных конформаций. Такие состояния, а также кинетические барьеры между ними можно определить из траекторий моделирования самосборки биологических молекул.

Для решения этой задачи был разработан подход, основанный на применении цепей Маркова с дискретным временем [381]. На основе множества всех полученных в ходе моделирования конформаций строится множество элементарных микросостояний. Каждое микросостояние – это группа конформаций, выделенная на основе структурного критерия, – в данном случае среднеквадратичного отклонения по координатам тяжелых атомов (всех, кроме атомов во-

дорода). Предполагается, что конформационные переходы внутри микросостояний происходят значительно быстрее, чем между микросостояниями. Таким образом, относящиеся к одному микросостоянию конформации должны быть структурно достаточно близкими. Также к одному микросостоянию должно относиться достаточно большое число конформаций. Для определения микросостояний на основе траекторий моделирования могут быть использованы различные методы кластеризации.

Нами использован гибридный k-centers k-medoid алгоритм кластеризации, подробно описанный в работе [401]. Только структурный критерий, однако, не подходит для определения вероятностей переходов между двумя микросостояниями: два геометрически близких микросостояния могут быть разделены высоким кинетическим барьером. При достаточно большом наборе анализируемых данных кинетические параметры переходов между микросостояниями могут быть определены по траекториям моделирования как частота переходов между микросостояниями в ходе моделирования. При моделировании по методу обмена репликами подобный анализ позволяет сделать некоторые заключения о связи между микросостояниями, но не описывает точные кинетические переходы, так как в ходе моделирования каждая реплика совершает случайные блуждания по пространству температур.

### **3.7.4 Моделирование самосборки структуры олигонуклеотида 5'-GCCGAGC-3'**

Для моделирования самосборки шпильки было проведено 2 эксперимента: в температурном интервале от 300 и от 270 К. Длина моделирования составила 40 нс, а суммарная длительность траекторий по всем репликам в случае моде-

лирования от 300 К —1.52 мкс (38 реплик), а в случае моделирования от 273 К —1.68 мкс (42 реплики). Одним из индикаторов качества моделирования методом обмена репликами является то, что в ходе моделирования все реплики совершили случайные блуждания по всему пространству температур. Обратное может говорить или о слишком большой разнице между соседними температурами для данной системы, или же о недостаточном времени моделирования. На Рисунке 3.60 представлены графики температур, при которых моделировались две реплики: первая начинала с самой низкой, а вторая с самой высокой температуры. Рисунок 3.60А относится к моделированию от 300 до 450 К, где реплика 0 начинает моделирование с температурой 300 К, а реплика 37 —450 К. Рисунок 3.60Б относится к моделированию от 273 до 450 К, где реплика 0 начинает моделирование с температурой 273 К, а реплика 41 —450 К.

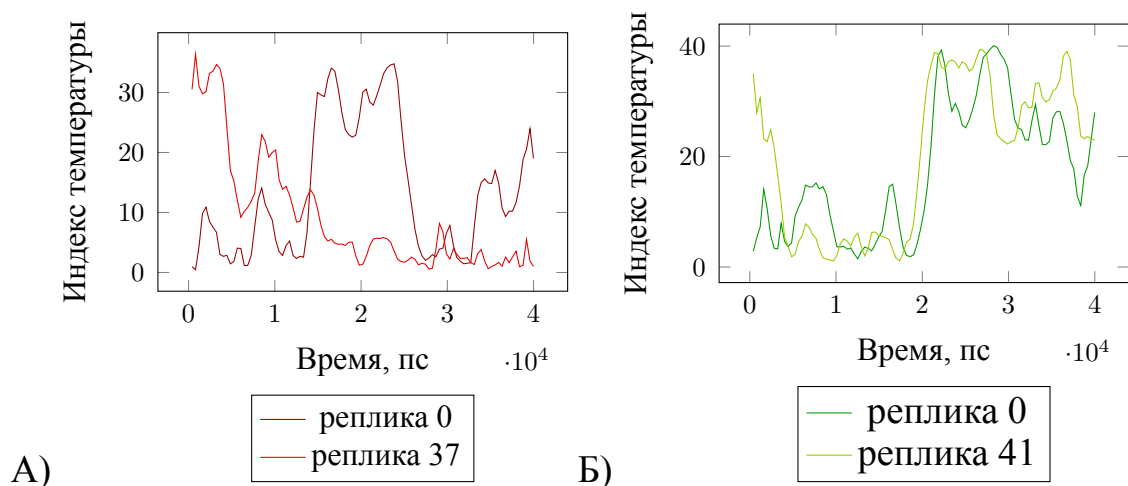


Рисунок 3.60: Случайные блуждания репрезентативных реплик по пространству температур. А) Нижняя температура 300К Б) Нижняя температура 273К.

По графикам видно, что в обоих случаях репрезентативные реплики пересекают все пространство температур.

Как же определить то, что происходит во время моделирования? Одним из традиционных способов контроля самосборки является подсчет минимального



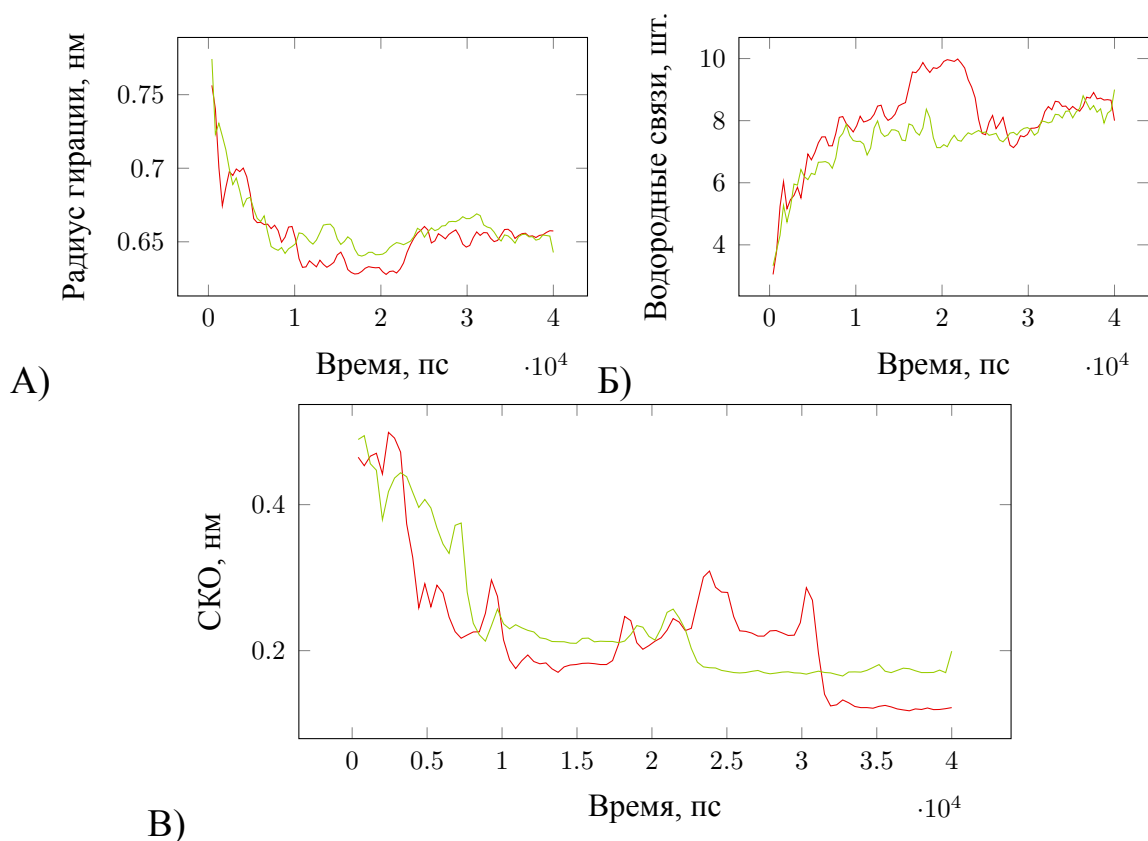


Рисунок 3.61: Моделирование самосборки от 273 К —зеленый, от 300 К —красный. А) Изменение во времени радиуса гирации олигонуклеотида. Б) Количество внутримолекулярных водородных связей во времени. В) Изменение среднеквадратичного отклонения координат тяжёлых атомов олигонуклеотида от экспериментально определённой структуры.

радиуса гирации по всем репликам. Радиус гирации рассчитывается согласно формуле

$$R_g = \left( \frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i} \right)^{\frac{1}{2}},$$

где  $m_i$  —масса  $i$ -ого атома, а  $r_i$  —позиция  $i$ -ого атома относительно центра масс. При анализе биологических макромолекул радиус гирации указывает на то, насколько компактна молекула, иначе говоря, находится ли она в свернутой или развернутой конформации. Снижение минимального радиуса гирации по всем репликам соответствует появлению в ходе моделирования все более и более компактных конформаций, что наблюдается в ходе обоих экспериментов (Рисунок 3.61А). Другим способом является изучение динамики накопле-

ния водородных связей внутри биополимера в ходе моделирования. Для определения того, что водородная связь существует, используется геометрический критерий: расстояние между донором и акцептором водорода не должно превышать  $3.5 \text{ \AA}$ , а угол водород-донор-акцептор не должен быть больше  $30^\circ$ . На Рисунке 3.61Б видно, что в ходе обоих REMD появляются структуры со все большим и большим количеством водородных связей, то есть структуры со все более низкой энергией. Этот процесс соответствует продвижению системы в сторону глобального минимума.

Процесс самосборки можно косвенно наблюдать по изменению минимального среднеквадратичного отклонения полученных в ходе моделирования conformаций от референсной структуры собранной шпильки (1ZHU с изменениями) по всем репликам (Рисунок 3.61В). В обоих случаях минимальное среднеквадратичное отклонение выходит на плато, однако в случае моделирования от  $300 \text{ K}$  (красная кривая) выход на плато происходит в районе  $30 \text{ ns}$ , а уровень опускается ниже  $1 \text{ \AA}$ . В случае же моделирования от  $273 \text{ K}$  (зеленая кривая) выход на плато происходит в районе  $25 \text{ ns}$ , а уровень не опускается ниже  $2 \text{ \AA}$ .

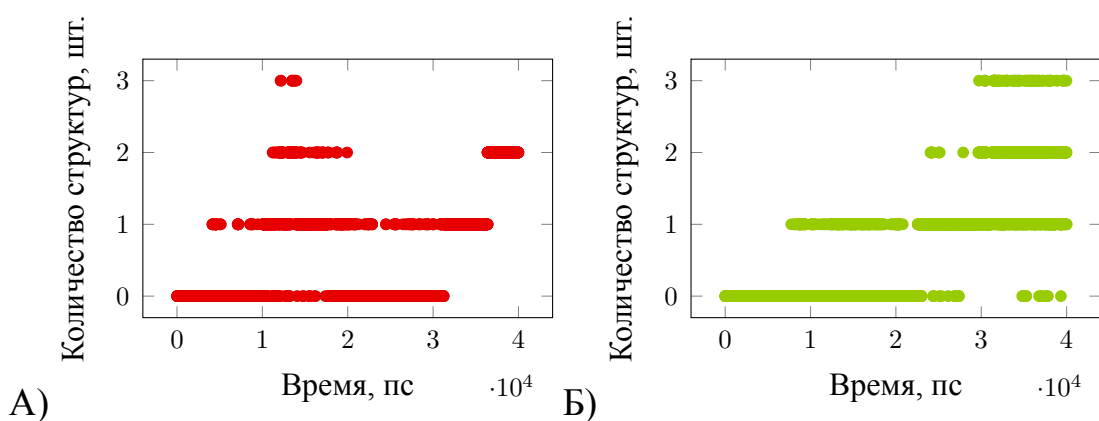


Рисунок 3.62: Количество реплик в ходе моделирования, среднеквадратичное отклонение которых от нативной структуры не превышает 2 ангстрем. А) Нижняя температура  $300 \text{ K}$  Б) Нижняя температура  $273 \text{ K}$

Но появляются ли такие структуры лишь однажды в ходе моделирования или же происходит их накопление? Рисунок 3.62 показывает число реплик, среднеквадратичное отклонение которых от референсной структуры в данных момент не превышает 2 Å. Как в случае моделирования от 300 К (Рисунок 3.62А), так и в случае моделирования от 273 К (Рисунок 3.62Б) видно накопление конформаций, структурно близких свернутой, вплоть до 3 единомоментно.

Может показаться, что 3 реплики из 38 или 42 — это довольно мало для наиболее представленного состояния. Однако, нельзя забывать, что рассматриваемые траектории совершали случайные блуждания по пространству температур, и состояния, характерные для высоких температур, могут быть очень редкими при биологически значимых температурах и практически не быть представленными в растворе. Также, вероятно, при продолжении моделирования может наблюдаться дальнейшее накопление свернутых конформаций.

Поскольку в ходе обоих экспериментов минимальное среднеквадратичное отклонение по всем репликам за последние 10 нс меняется слабо, этот отрезок времени был выбран для анализа наиболее представленных состояний.

В случае моделирования от 300 К наиболее представленным микросостоянием оказалась правильно свернутая шпилька (Рисунок 3.63). Среднеквадратичное отклонение составило 1.33 Å, а доля среди всех микросостояний за последние 10 нс моделирования — 3.4%.

В случае моделирования от 273 К наиболее представленными микросостояниями оказались неправильно свернутые шпильки (Рисунок 3.64). В каждой из наиболее представленных структур одно из оснований инвертировано относительно N-гликозидной связи: в структуре 3.64А перевернут седьмой цитозин, в структуре 3.64Б — третий гуанин, а в структуре 3.64В — первый гуанин. Пред-

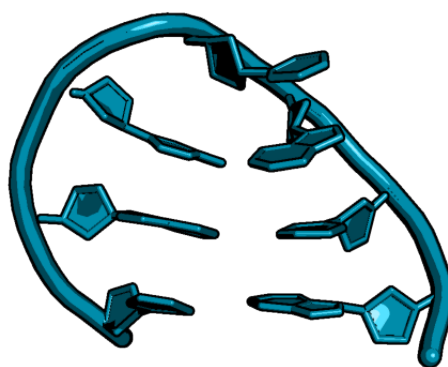


Рисунок 3.63: Наиболее представленное состояние в моделировании от 300 К  
 представленность данных структур в течение последних 10 нс моделирования составляет 2.9 – 3.6 %, а среднеквадратичное отклонение —1.75 – 2.15 Å.

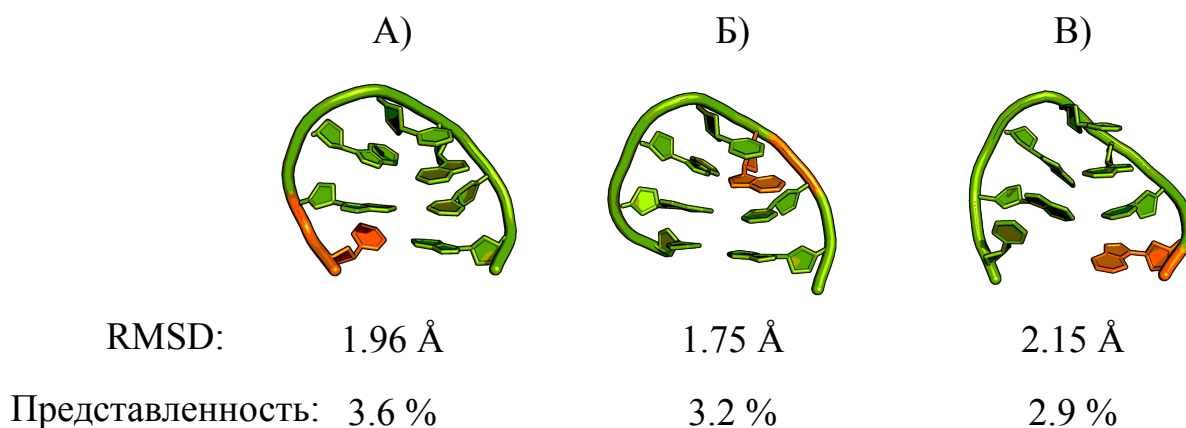


Рисунок 3.64: Наиболее представленные состояния в симуляции от 237 К.  
 Основания с неправильной ориентацией выделены оранжевым.

Для каждого моделирования были посчитаны наиболее вероятные пути (с максимальным потоком) конформационных переходов из стартового в наиболее представленные свернутые состояния (Рисунок 3.65 —моделирование от 300 К, Рисунок 3.66 —от 273 К). В обоих случаях пути сборки образуют сеть с большим числом компонентов, указывая на сложность механизмов самосборки, что соответствует современным представлениям об образовании структуры биополимеров , проходящем через большое количество переходных состояний[402].

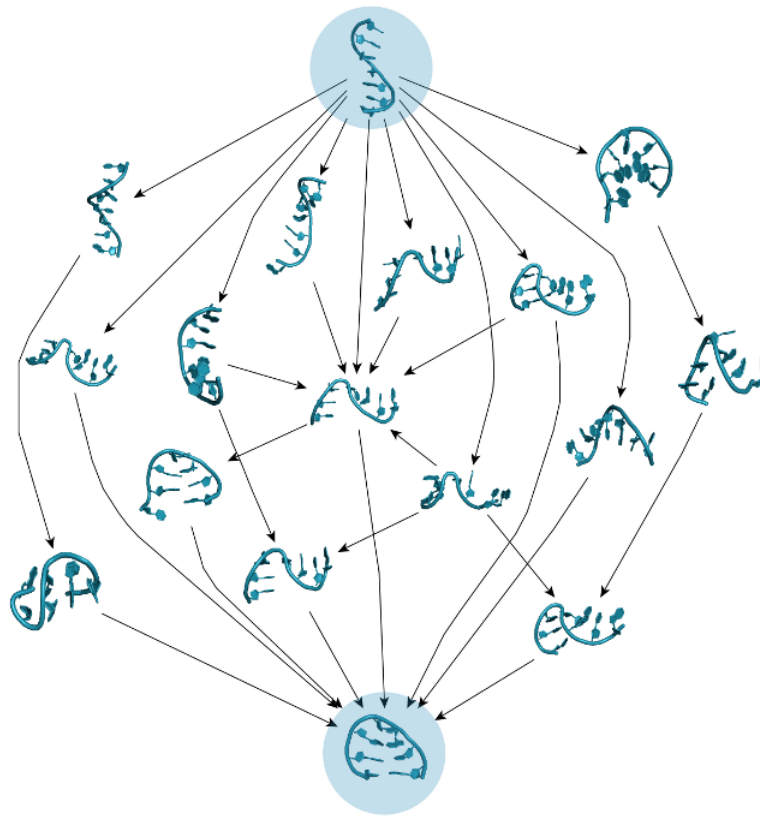


Рисунок 3.65: Моделирование от 300 К. 15 путей с наибольшим потоком, по которым возможны переходы из стартового состояния в наиболее представленные.

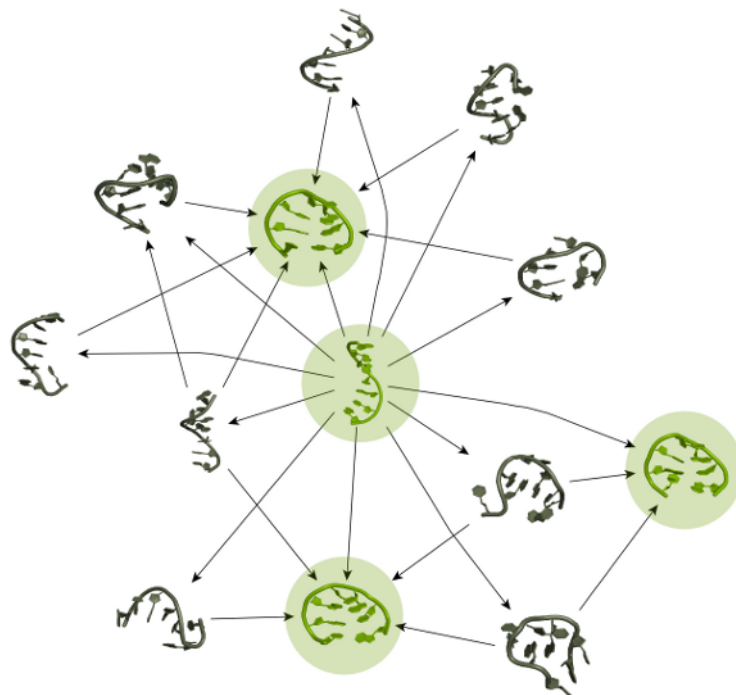


Рисунок 3.66: Моделирование от 273 К. 15 путей с наибольшим потоком, по которым возможны переходы из стартового состояния в наиболее представленные.

В ходе работы проведено 2 эксперимента по моделированию самосборки ДНК-шпильки с последовательностью 5'-GCGCAGC-3' по методу обмена репликами в температурных интервалах, различающихся нижней температурой. Возможность моделирования самосборки шпильки 5'-GCGCAGC-3' методом REMD продемонстрирована ранее [384; 385], однако пути сборки не анализировали. Полученные данные о сложности образовании структуры шпильки 5'-GCGCAGC и наличии большого числа переходных состояний хорошо согласуются как с экспериментальными данными [402], так и с результатом по моделированию самосборки РНК-шпильки похожего размера (8 нуклеотидов) [390]. Снижение нижней температуры моделирования привело к стабилизации группы структур с инвертированными парами оснований (Рисунок 3.64).

## **Глава 4**

# **Материалы и методы**

### **4.1 Материалы и методы**

#### **4.1.1 Алгоритм построения крупно-зернистой модели РНК**

##### **Первый этап сборки структуры**

Начальным процессом является анализ входного файла с описанием вторичной структуры РНК. На этом шаге проверяется принадлежность входных данных более крупной структуре, что может выражаться в начале нумерации не с первого нуклеотида и/или наличии вторичных взаимодействий вне заданной структуры. В результате работы модуля происходит замена входного файла на аналогичный, с учетом внесенных изменений. Далее происходит разбор вторичной структуры. Условием исполнения всех последующих операций является длина РНК не более 400 нуклеотидов. Это условие необходимо, для сохранения достоверности моделирования, так как точность построения модели уменьшается при увеличении длины молекулы, что связано с недостатками алгоритмов оптимизации геометрии. Для больших структур сложнее добиться достижения минимума энергии, а это иногда может приводить к заведомо



неправильной упаковке молекулы. Также слишком длинные структуры требуют значительного времени расчета. На основании данных разбора структуры производится поиск псевдоузлов и генерируется дополнительный файл с исключенными псевдоузлами. На основании этого файла производится построение схематической вторичной структуры, после чего эти данные преобразуются в форматы координат и топологии для пакета Gromacs. При этом генерируется файл топологии для вторичной структуры. Следующим шагом происходит минимизация энергии структуры и получается файл координат вторичной структуры без псевдоузлов. В том случае, если входной файл не содержал псевдоузлов, то эта структура является конечной. При наличии псевдоузлов модуль Top2st.pm генерирует файл топологии вторичной структуры с псевдоузлами, и происходит еще одна минимизация энергии, что приводит к формированию псевдоузлов во вторичной структуре.

На втором шаге генерируется файл топологии “промежуточной” структуры с заданными пользователем значениями “gap” и включением/выключением “автоматического поиска близких третичных взаимодействий”. На основании файла координат вторичной структуры придается левое закручивание спиральным участкам. Результатом его работы является новый файл координат. Затем производится минимизация энергии для получения промежуточной третичной структуры.

## **Второй этап построения модели**

На данном этапе пользователь, при использовании веб-сервера, может наложить дополнительные ограничения на моделируемую структуру путем непосредственного задания третичных контактов, указать места принудительного стекинга или использовать сравнительное моделирование. Вводимые пользо-

вателем данные записываются в соответствующие файлы. Далее проводится оптимизация геометрии на основании новой топологии и координат “промежуточной” структуры, результатом является модель третичной структуры РНК с заданными геометрическими свойствами.

### **Состав пакета программ для моделирования трехмерной структуры малых РНК**

Пакет программ состоит из 11 модулей, 1 файла базы данных и 4 скриптов. Два управляющих скрипта отвечают за две стадии процесса сборки структуры и используют указанные модули. Два командных скрипта - исполняются сервером, загружают управляющие скрипты и генерируют веб-страницы. Файл базы данных содержит информацию о геометрии различных видов тройных взаимодействий, он используется если задан их автоматический поиск в структуре.

#### **Модули**

- **Stpp.pm**: осуществляет первичную обработку входных файлов вторичной структуры. Позволяет пользователю в качестве входных данных указывать не только полные структуры, но и часть крупных молекул – осуществляет сдвиг нумерации входных структур к единице, удаляет все контакты вне введенной структуры. В результате работы заменяет входной СТ файл новым с учетом указанных изменений.
- **Readct.pm**: осуществляет анализ СТ файла, заданного в качестве аргумента, также генерирует дополнительные файлы визуализации на основании данного файла, в том случае, если эта функция включена. На основании СТ файла формирует ряд массивов. Данные множества экспортируются для использования другими модулями.

- **Findpk.pm**: осуществляет разбор СТ файла, заданного в качестве аргумента. Производит поиск псевдоузлов H-типа в структуре, генерирует дополнительный СТ файла, в котором исключаются спирали второй ветви псевдоузла путем преобразования в две петли.
- **Readss.pm**: осуществляет преобразование файла координат вторичной структуры в формате SS в файл координат в формате GRO (файл пакета программ GROMACS).
- **Top2st.pm**: создает файл топологии вторичной структуры основываясь на множествах, генерируемых модулями Readct.pm и Findpk.pm. Структура двухцепочечных участков описывается дистанционными ограничениями характерными для A-формы РНК (расстояние между нуклеотидами, длинны связей пары, углы). Между всеми спиралями, каждой из ветвей псевдоузла, задаются дистантные ограничения для принудительное формирование стекинга, в независимости от длинны петель между ними.
- **Top23st.pm**: создает файл топологии “промежуточной” трёхмерной структуры основываясь на множествах генерируемых модулями Readct.pm и Findpk.pm. Структура спиральных участков описывается дистантными ограничениями характерными для A-формы РНК (расстояние между нуклеотидами, длинны связей пары, углы закрутки спиралей и т.д.). Дистантные ограничения A-формы были созданы на основе статистической обработки структур молекул тРНК, 16S, 5S, 23S РНК из PDB банка, а также структуры 1QCU (классический РНК дуплекс). В качестве одного из аргументов функция модуля имеет “максимальное расстояние сохранения стекинга” (gap) – такое максимальное расстояние между соседними спиралями по одной из цепи, при котором спирали сохраняют

стекинг, при этом расстояние по второй цепи может быть любым. “Gap” может принимать значение 0 или 1. На основании заданного значения “gap” производится автоматический поиск стекинга спиралей и создание для него дистантных ограничений. В том случае, если для данной спирали под условия формирования стекинга попадает более одного варианта второй спирали, то стекинг не образуется, впоследствии пользователь может реализовать стекинг вручную для выбранной пары спиралей. В качестве другого аргумента функция имеет “включение автоматического поиска близких третичных взаимодействий”. Если эта функция включена, то в структуре происходит поиск близких третичных взаимодействий (тройных и прочих) и создание для них дистантных ограничений. Для петель из 3-5 нуклеотидов, которые являются достаточно структурно инвариантными, задается геометрия, аналогичная структурам из PDB банка. Для “промежуточной” структуры дистантные ограничения на петли не накладываются.

- **Gro23st.pm**: на основании файла координат оптимизированной вторичной структуры и множеств, генерируемых модулями Readct.pm и Findpk.pm, создает файл координат “промежуточной” структуры. Результатом работы модуля является левое закручивание спиралей, которое обеспечивает правильный ход цепи в процессе последующей минимизации энергии. Закручивание петель между ветвями псевдоузла не происходит, что необходимо для предотвращения спутывания петель псевдоузлов в процессе минимизации энергии.
- **Readtst.pm**: осуществляет разбор файла третичных контактов (tertiary structure contacts – TST) и преобразование данных в соответствующее

множество, используемое модулем Top3st.pm. Файл TST генерируется веб-сервером и описывает дополнительные третичные взаимодействия – пользователь задает номера пар взаимодействующих нуклеотидов и расстояние между ними в относительных единицах (расстояние в нанометрах деленное на 1.806 – расстояние между фосфорами пары нуклеотидов).

- **Readstk.pm**: осуществляет разбор файла “принудительного” (ручного) стекинга и преобразование данных в соответствующее множество, используемое модулем Top3st.pm, файл генерируется веб-сервером и описывает “ручной” стекинг выбранных спиралей - для этого пользователь задает пары номеров конечных нуклеотидов спиралей, которые будут образовывать стекинг.
- **Nomol.pm**: поддерживает применение гомологического моделирования – пользователь задает номер ID из PDB банка гомолога, название цепи (латинская буква), выравнивание образца и модели. На основании заданных параметров производится загрузка в оперативную память указанного файла из PDB банка и построение дополнительной (включенной) топологии модели на основе выравнивания и геометрии описываемой данным файлом молекулы.
- **Top3st.pm**: создает файл топологии третичной структуры используя множества генерируемые модулями Readct.pm, Findpk.pm, Readtst.pm, Readstk.pm и Nomol.pm. Дистантные ограничения, накладываемые на третичную структуру в обязательном порядке, полностью аналогичны ограничениям, создаваемым модулем Top23st.pm, причем включение автоматического поиска ближних взаимодействий, а также параметр “gap” имеют всегда то же значение, что и при построении “промежуточной

структуры’’. Дополнительно вводятся ограничение на петли, для которых задаются барьеры внутреннего вращения нуклеотидов в форме неправильных диэдралов и углов. Дополнительные дистантные ограничения задаются на основании исходящих данных модулей Readtst.pm, Readstk.pm и Homol.pm.

#### **4.1.2 Полноатомное моделирование молекулярной динамики рибосомного туннеля**

##### **Построение систем для полноатомному моделированию рибосомного туннеля**

Структура большой субъединицы рибосомы *Escherichia coli* была взята из структуры её комплекса с эритромицином, полученной методом рентгеноструктурного анализа (PDBID:3OFR). Структура комплекса тилозина и его производных с большой субъединицей рибосомы *Escherichia coli* была построена путем наложения молекулы тилозина на молекулу эритромицина. При наложении остаток микаминозы тилозина совмещался с остатком дезозамина эритромицина, а остальные участки тилозина перемещались согласованно с микаминозой. При этом лактонное кольцо тилозина заняло приблизительно ту же область пространства, что и лактонное кольцо эритромицина. Из полученной структуры был выделен фрагмент, полностью включающий в себя рибосомный туннель и находящийся в нем тилозин и имеющий в своей основе куб с ребром длиной 7 нм. Во фрагмент отбирались все остатки, у которых хотя бы один атом попадал в заданную кубическую область. Именно структура этого фрагмента служила основой для моделирования структур комплексов большой субъединицы рибосомы *E. coli* с исследуемыми производными, которое производилось наложением

ем лактонных колец производных на лактонное кольцо тилозина. Трехмерные структуры исследуемых производных получались оптимизацией геометрии в пакетах GAMESS и GAUSSIAN 09 по частям: структура производных 4 разделялась на блоки (остаток сахара, аминокислоты, тилонолид...), на функциональные группы, через которые блоки связывались между собой, устанавливались “заглушки” — метоксигруппы (связь блоков между собой имеет эфирную или сложноэфирную природу, так что такая замена уместна). Частичные заряды рассчитывались с помощью пакета R.E.D.-III Tools в сочетании с пакетом GAUSSIAN 09. При этом исходные координаты для тилозина были взяты из структуры его комплекса с 50S субъединицей рибосомы *Haloarcula marismortui*, полученной методом рентгеноструктурного анализа (PDBID:1K9M).

### **Моделирование молекулярной динамики производных тилозина в рибосомном тоннеле**

Для моделирования молекулярной динамики и анализа полученных траекторий использовался пакет GROMACS версии 4.5.4 с силовым полем parm99sb. Моделирование проводили в явно заданном растворителе при температуре 300 К с периодом привязки 0,1 пс и термостатом масштабирования скоростей и изотропном давлении 1 бар с периодом привязки 5 пс и баростатом Берендсена. Электростатические взаимодействия описывались методом сети частиц Эвальда с Использовалась кубическая ячейка, заполняемая молекулами воды модели TIP3P так, что моделируемый фрагмент рибосомы был покрыт 0,5 нм раствора с каждой стороны. Отрицательный заряд системы был компенсирован добавочными ионами магния и натрия, причем ионы магния добавлялись так, чтобы образовывались “магниевые мостики”, а ионы натрия добавлялись затем по

остаточному принципу так, чтобы они располагались вблизи от отрицательно заряженных групп. Все вычисления были проведены с шагом интегрирования 2 фс, запись состояний моделируемой системы происходила каждые 30 пс, длина каждой траектории составила 300 нс. Расчеты производились на суперкомпьютере “Ломоносов” НИВЦ МГУ.

### **4.1.3 Влияние топологии петель на геометрию квадруплексов**

#### **Создание выборки квартетов**

Список структур, содержащих G-квадруплексы, был составлен на основе списка базы данных PDB. Все найденные структуры были разделены на 8 групп согласно геометрии пространственной организации квадруплекса. На основе языка Perl и модулей Vector::Real и Statistics::Descriptive была разработана программа, которая определяет наличие квартетов в структуре, их расположение и измеряет геометрические параметры. Квартет определяется следующим образом: для гуанина должен существовать сосед, контактирующий атомом N7 с атомом O6 начального гуанина. Выбранная комбинация признается квартетом, если четвертый гуанин взаимодействует с первым и все гетероциклические основания находятся в одной плоскости с максимально допустимым удалением атомов от плоскости, составляющим 2Å. Следующим квартетом квадруплекса признавался квартет, располагающийся ближе всего к исходному и для которого расстояние между атомами C1' до ближайших нуклеотидов было не более 10Å (таким образом исключались варианты, когда тетраду составляли гуанины из разных квартетов). Структуры, принадлежащие разным ЯМР-моделям, обрабатывались как независимые квадруплексы.



## **Гистограмма распределения значений углов поворота квадруплекса и отклонений квартета от плоскости**

Для каждой из восьми полученных групп значения параметров объединялись, и строилась гистограмма распределения. Для углов поворота выбран диапазон от 0 до 600; отклонения от плоскости анализировались в диапазоне от 0 до 2Å; диапазоны углов и расстояний разбивались на 15 интервалов. Данные анализировали с помощью программы Gnuplot (<http://www.gnuplot.info>).

## **Расчёт угла поворота квартетов относительно друг друга в квадруплексе**

Для определения углов закрутки квадруплекса измерялся угол между двумя векторами. Первый вектор соединял атомы C1' двух соседних нуклеотидов в квартете, второй вектор соединял атомы C1' в соответствующих нуклеотидах в соседнем квартете (Рисунок 4.1).

## **Отклонения квартета от плоскости**

Для измерения степени нарушения симметричности и планарности отдельного квартета предложен оригинальный параметр - расстояние между центрами масс двух четырёхгранников. Первый четырёхгранник образован четырьмя атомами N9 гуанинов квартета, а второй четырёхгранник образован четырьмя атомами O6 тех же гуанинов (Рисунок 4.1). В случае если квартет симметричный и все гуанины образуют водородные связи друг с другом, то этим параметром фиксируется планарность квартета. В случае разрыва водородных связей симметрия квартета нарушается и параметр фиксирует степень искажения квартета. Этот подход также фиксирует оба описанных выше случая одновременно.

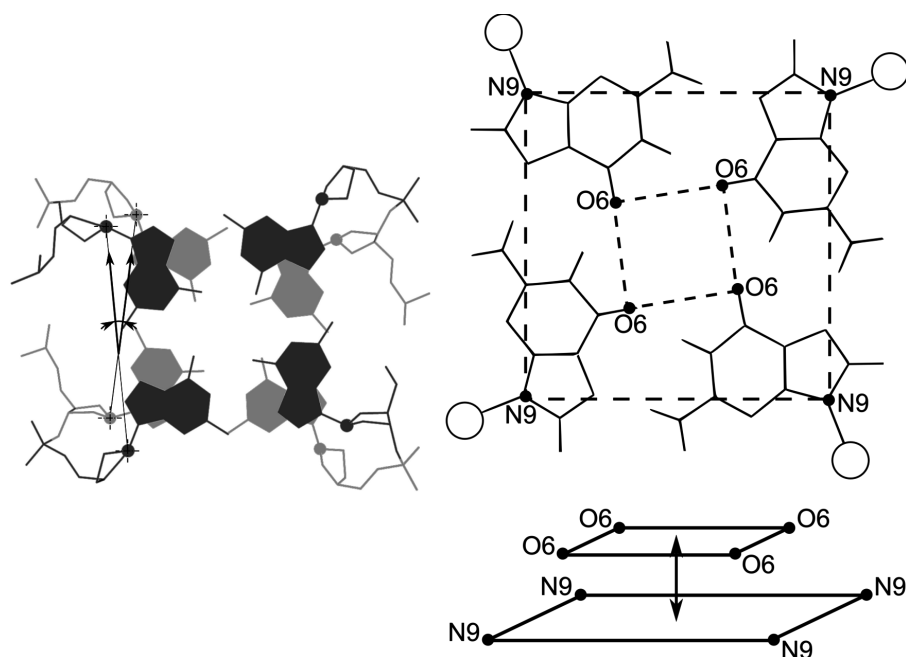


Рисунок 4.1: Определение измеряемых параметров. Слева: определение углов закрутки квадруплекса. Кругами показаны атомы C1'. Справа: схема взаимного расположения двух четырёхгранников: внешнего, образованного атомами N9, и внутреннего, образованного атомами O6. Если планарность квартета нарушается, то внутренний четырёхгранник выходит из плоскости квартета. Расстояние между центрами масс внутреннего и внешнего квартетов служит численным параметром отклонения от плоскости.

Для каждой из восьми полученных групп значения параметров объединялись и строилась гистограмма распределения. Для углов закрутки выбран диапазон от 0 до 600; отклонения от плоскости анализировались в диапазоне от 0 до 2Å; диапазоны углов и расстояний разбивались на 15 интервалов. Данные анализировали с помощью программы Gnuplot.

#### 4.1.4 Методы моделирования молекулярной динамики систем с аптамером 15-ТВА

Для всех систем моделирование и анализ траекторий молекулярной динамики проводилось в программном пакете GROMACS 4.0 [403]. В этой части работы мы использовали два силовых поля. AMBER-99 $\phi$ , которое известно как

parm99, которое было модифицировано для уточнения параметров угла  $\phi$  некоторых аминокислот Сориным и коллегами [404] и адаптировано для применения в программном пакете GROMACS [405]. Второе использованное нами силовое поле это parmbsc0, которое было адаптированное нами для GROMACS, в основе которого лежит модификация parm99sb [406]. Суть наших модификаций сводится к оптимизации описания торсионных углов сахарно-фосфатного остова. Всё вычислительные эксперименты проводились в явно заданном растворителе. Для поддержания температуры соответствующей 300K использовался алгоритм “V-rescale” [407], который основан на алгоритме скалирования скоростей Берендсена с дополнительным стохастическим членом, который позволяет точно воспроизводить канонический ансамбль. Для сохранения канонического ансамбля биополимеры и вода с ионами термостатировались отдельно. Это связано с тем, что вода и ионы не имеют вибрационных степеней свободы, а только поступательное и вращательное движение. Для моделирования в NPT ансамбле использовался баростат Берендсена [408], применяемый каждые 5 пс. Для описания удалённых электростатических взаимодействий использовалось модифицированное суммирование Эвальда, где вычислялось взаимодействие частица-решётка, а не частица-частица [409]. Исследуемые молекулы и комплексы помещали в триклинную ячейку с отступом 15 ангстрем от крайнего атома, которую наполняли молекулами воды, которые описывались четырёх точечной моделью TIP4P [410]. Избыток отрицательного заряда в системе нейтрализовали заменяя случайные молекулы воды на катионы натрия. Также во всех расчётах использовали ионную силу свойственную крови, т.е. 0.14 NaCl. Для достижения этого значения необходимое количество молекул воды, также, заменялось на катионы натрия и анионы хлора, при это расстояние между ионами не должно было превышать 6Å. В качестве интегратора, алгоритма который

предсказывает позиции атомов на основе действующих сил, использовался алгоритм leap-frog с шагом 3 фемтосекунды. Каждые 6 пикосекунд конфигурация системы сохранялась в файл.

Необходимо отметить, что большинство G-квадруплексов содержащих молекул стабилизируются катионом калия. В нашем исследовании есть системы где калий не располагался специально, в остальных случаях катион калия помещали в геометрический центр фигуры, образованной атомами O6 гуаниновых оснований остатков квадруплекса. Мы использовали стандартные параметры для поля parm99 описание катиона калия, радиус Ленарда-Джонса  $2.658\text{\AA}$  и глубина потенциальной ямы  $0.00137\text{ кДж/моль}$ . Катион натрия описывался радиусом равным  $1.868\text{\AA}$  и глубина потенциальной ямы функции составляла  $0.01589\text{ кДж/моль}$ .

Все расчёты были проведены на суперкомпьютере “Чебышев”, доступ на который был предоставлен вычислительным центром МГУ. Для достижения оптимальной производительности нами была оценена производительность расчёта системы с аптамером (примерно 23000 атомов) в зависимости от количества используемых вычислительных узлов.

Анализ файлов траекторий проводился с помощью программ представленных в пакете GROMACS или собственно разработанных. При анализе молекулярной динамики водородная связь считалась существующей если её можно было обнаружить в половине кадров траектории и геометрия положения её участников удовлетворяла условию: расстояние между донором и акцептором не более  $3.5\text{\AA}$  и угол акцептор-донор-водород не должен превышать 30 градусов.

Все расчётные эксперименты и описания систем приведены в Таблице 4.1

объект <sup>a</sup>	Силовое поле	Длина траектории, нс	количество атомов					
			ДНК	Белок	Вода	Na <sup>+</sup>	K <sup>+</sup>	Cl <sup>-</sup>
G-стебель	parmbsc0	700	260	0	12516	3	1	0
G-стебель без ионов	parmbsc0	10	260	0	12516	4	0	0
РСА 15-ТВА	parm99	900	488	0	22516	13	1	0
РСА 15-ТВА	parmbsc0	900	488	0	19240	13	1	0
ЯМР 15-ТВА	parm99	900	488	0	22524	13	1	0
ЯМР 15-ТВА	parmbsc0	900	488	0	22532	13	1	0
ЯМР 15-ТВА с Na <sup>+</sup> в центре	parmbsc0	900	488	0	22532	14	0	0
ЯМР 15-ТВА без катиона в центре	parmbsc0	900	488	0	22528	14	0	0
<b>Конформации с модифицированной TGT петлёй</b>								
TG(+T) РСА 15-ТВА	parmbsc0	900	488	0	18688	13	1	0
TG(-T) ЯМР 15-ТВА	parmbsc0	900	488	0	18312	13	1	0
TG(-T) ЯМР <sup>p</sup> 15-ТВА <sup>b</sup>	parmbsc0	900	488	0	18968	13	1	0
T(-GT) ЯМР <sup>p</sup> 15-ТВА <sup>b</sup>	parmbsc0	900	488	0	20532	13	1	0
<b>Комплексы аптамер-тромбин</b>								
РСА аптамер-тромбин комплекс	parmbsc0 + parm99SB	600	488	4657	104964	49	1	45
ЯМР аптамер-тромбин комплекс	parmbsc0 + parm99SB	600	488	4658	105104	48	1	45
1:2 РСА аптамер-тромбин комплекс	parmbsc0 + parm99SB	600	488	9314	155404	71	1	78
1:2 ЯМР аптамер-тромбин комплекс	parmbsc0 + parm99SB	600	488	9316	155732	71	1	78

Таблица 4.1: Параметры моделирования молекулярной динамики

<sup>a</sup> Один катион калия помещён в центральную полость квадруплекса, если не указано обратное. <sup>b</sup> ЯМР<sup>p</sup> Как стартовая структура использовалась конфигурация полученная после моделирования молекулярной динамики ЯМР структуры. Последние 100 нс из 900 нс траектории были использованы что бы построить среднюю структуру, которая и являлась объектом манипуляций для получения конформаций TG(-T) или T(-GT).

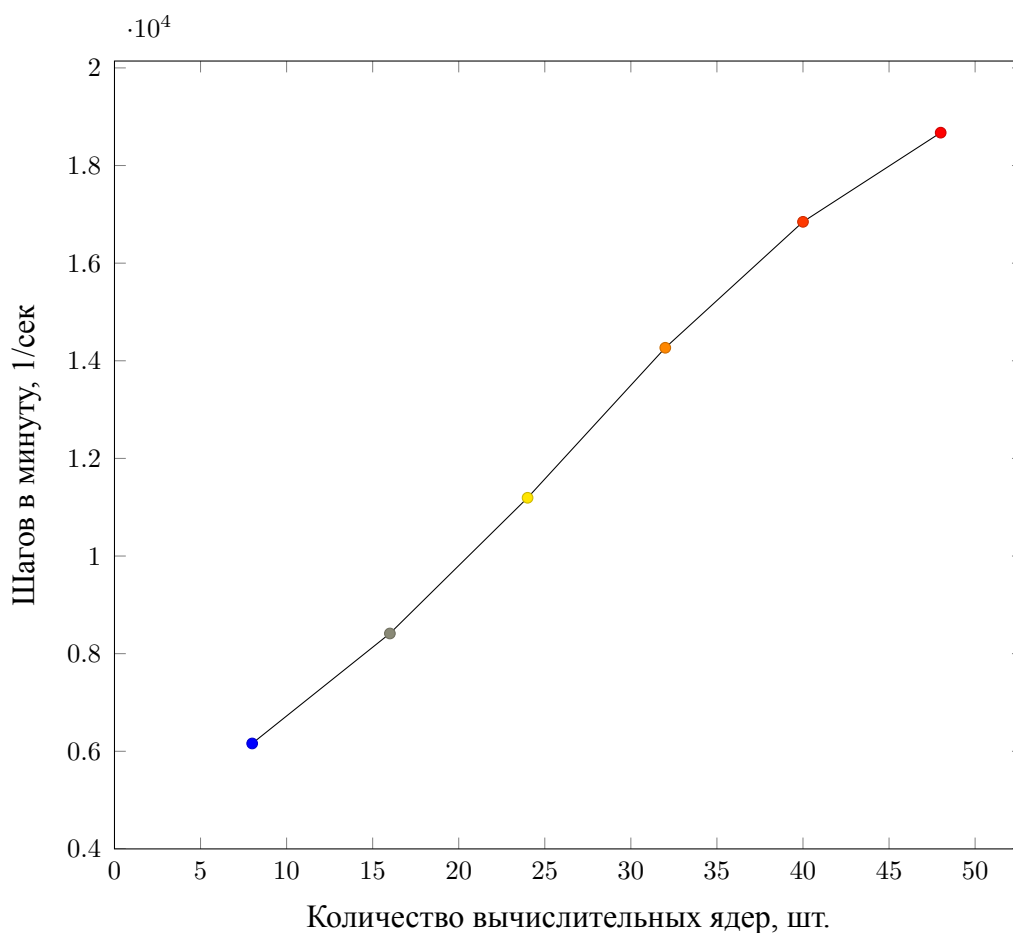


Рисунок 4.2: Эффективность масштабирования вычислений для выбранной аппаратно-программной платформы Gromacs-“Чебышев” для системы содержащей 23000 частиц. Описание конфигурации суперкомпьютера можно найти по адресу <http://www.top500.org/system/9240>.

#### 4.1.5 Гибридное молекулярно механическое / квантово-механическое моделирование

Стартовые конформации для 15-ТВА брали из PDB, ID 1c35 для комплексов с  $K^+$  и 1rde для комплекса с  $Ba^{2+}$ . Было сконструировано 5 систем. Верхнюю позицию катиона (системы 1, 4) определяли как среднее арифметическое от координат  $K^+$ , располагающихся над верхним G-квartetом (остатки G1, G6, G10, G15) в структуре 1c35. Центральную позицию (система 2) определяли как среднее арифметическое от координат атомов O6 гуанинов квадруплекса. Нижнюю позицию катиона (системы 3, 4) определяли как среднее арифмети-

ческое от координат  $K^+$ , располагающихся под нижним G-квартетом (остатки G2, G5, G11, G14) в структуре 1c35. Системы 1-4 были смоделированы с катионами калия, система 2 была также смоделирована с катионом бария. Молекулярно-механическую часть описывали с помощью параметров силового поля parmbsc0. Квантово-механическая часть системы была описана в терминах теории функционала плотности с использованием псевдопотенциалов и базиса плоских волн (PW-DFT), с применением спин-поляризованного формализма и функционала PW91. Взаимодействия между валентными электронами и ионным остовом описывали сверхмягкими псевдопотенциалами Вандербильта. Квантово-механическая часть состояла из азотистых оснований 15-ТВА (за исключением Т7, не принимающего участия во взаимодействиях с катионом) и катионов калия или бария. Разбиение систем на квантово-механическую и молекулярно-механическую части осуществляли по N-гликозидной связи с введением атомов водорода для насыщения разорванных связей. В конечном итоге квантово-механическая система состояла из 178 атомов (179 в случае системы 4). Взаимодействия Ван-дер-Ваальса, плохо описываемые теорией функционала плотности, корректировали аналитическим потенциалом Гримми (161). Каждая из моделируемых систем была наполнена водой модели tip4p, заряд компенсировали добавлением ионов натрия. Вода и ионы натрия были уравновешены вокруг комплекса ДНК: $K^+$  ( $Ba^{2+}$ ) проведением МД в течение 100 пикосекунд с закреплением позиций комплекса. Подготовленные системы были подвержены комбинированному квантово-механическому и молекулярно-механическому моделированию (КМ/ММ) с помощью пакета программ GROMACS/CPMD с траекторией длиной 1,5 пикосекунд, с шагом 0,12 фемтосекунды ( 5 а.е.) и электронной массой 500 под контролем термостата Ноза-Хувера при температуре 310К. Базис плоских волн, описывающий

поведение валентных электронов, был ограничен пределом 30 Ry. Использовали кубическую квантово-механическую подъячейку с длиной стороны 40 Ry, что составляло в итоге примерно 90000 плоских волн на волновую функцию.

#### 4.1.6 Изотреническая калориметрия

Термодинамические параметры связывания катионов калия и бария с олигонуклеотидами d(GGTTGGTGTGGTTGG) и d(GGTTGGTCTGGTTGG) оценивали с помощью прибора iTC200, как описано ранее [411]. Эксперименты проводили при температуре 25°C в буфере, содержащем 20 mM Tris, pH 6.8, в присутствии 140 mM LiCl. Двухмикролитровые аликвоты растворов KCl или BaCl<sub>2</sub> вносили в ячейку объёмом 0.2 мл, содержащую раствор олигонуклеотидов до достижения концентрации насыщения. Концентрация ДНК составляла 50–250 мкМ, концентрация катионов в инжекторе от 1 до 10 mM. При введении соли в ячейку измеряли выделяемую теплоту и полученные значения вычитали из значений, полученных для реакции комплексообразования, получая величину эффективной теплоты реакции связывания. Полученные кривые аппроксимировали с помощью процедуры нелинейного регрессивного анализа и рассчитывали константу связывания и стехиометрию реакции. Соответствующее изменение энтропии определяли из классических уравнений термодинамики. Все измерения для каждого олигонуклеотида повторяли не менее трёх раз и рассчитывали среднее значение.



## 4.1.7 Исследование структуры аптамера 31-TGT методами ЯМР

Аптамеры были получены твердофазным амидитным синтезом и очищены ООО “ЕвроГен”. Все ЯМР спектры были измерены на спектрометре Bruker AVANCE 600 MHz. ЯМР спектры регистрировались в 20 mM KCl и 10 mM фосфатного буфера и при pH 6. Концентрация ДНК аптамера в каждой пробе составляла 2 mM, объем пробы 0,35 мл. Температура образца с аптамером при NOESY D<sub>2</sub>O, HSQC и COSY экспериментах составляла 25°C. При NOESY H<sub>2</sub>O эксперименте температура образца с аптамером составляла 10°C. Перед снятием спектров проводился отжиг ДНК. Температура раствора с аптамером была поднята до 70°C с последующим постепенным охлаждением до 4°C. Вся работа по расшифровке спектров производится в специальном программном обеспечении по отнесению и интеграции спектров Sparky [412]. Данные о химических сдвигах необходимые для отнесения спектров были получены с сайта Biological Magnetic Resonance Data Bank [413]. Преобразование интенсивности кросс-пиков в ограничение расстояний между протонами производилось в программном обеспечении NMRConstraints, входящий в пакет программ NMRTABLE. Начальная топология по последовательности 31-TVA и развернутая структура была построена, используя пакет программ Crystallography and NMR Systems (CNS) [414]. Минимизация энергии была произведена с помощью GROMACS (с применением силового поля parmbsc0). Силовые константы 800 (кДж/мол)/Å<sup>2</sup> были использованы для экспериментально полученных дистанционных ограничений Уотсон-Криковских водородных связей. Радиус обрезания для нековалентных взаимодействий был установлен в 20Å. Растворитель задан в неявном виде по методу GBSA.

Финальная модель была создана моделированием отжига в течении 2нс от 1000К до 100К при заданных дистанционных ограничениях на основе ЯМР анализа. Все 247 ограничений расстояний, полученные из ЯМР спектров, были использованы при моделировании структуры аптамера.

## **4.1.8 Моделирование конъюгатов 15-ТВА и нанотрубок**

### **Моделирование углеродной нанотрубки**

В связи с наличием проводимости CNT и пи-электронной системы возникает вопрос об их поляризуемости. Квантово-механические расчёты показали анизотропность этого явления. В направлении продольной оси поляризуемость в несколько раз превышает таковую в перпендикулярном направлении. Так как нанотрубки и ДНК вступают в стэкинг-взаимодействия при участии пи-электронных облаков, при проведении молекулярно-динамических расчётов систем, включающих ДНК и CNT, важно знать, насколько точно используемые методы воспроизводят указанные взаимодействия. Силовые поля молекулярно-механических методов можно разделить на две группы: с фиксированным зарядом и поляризуемые. Последние представляют больший простор для параметризации, однако первые менее ресурсоёмкие. В работе [415] было продемонстрировано, что различные силовые поля с фиксированным зарядом достаточно точно воспроизводят расстояния между взаимодействующими частями молекул и силу взаимодействия, хотя положение и ориентация ароматической группы относительно поверхности CNT искажены. Так как нас интересует сила взаимодействия азотистых оснований с нанотрубкой, а не конкретная ориентация, то для неё подойдёт то же силовое поле с фиксированным зарядом parmbsc0. Пространственная структура CNT представляет из себя свёрнутую

в трубку поверхность гексагональной решётки с атомами углерода в узлах. Её можно создать, например, с помощью такой программы, как buildCstruct. Так как мы не учитываем электронные эффекты в наших расчётах, конформация нанотрубки не имеет значения и мы выбрали стандартную открытую armchair-нанотрубку с конфигурацией (10, 10) и длиной около 30Å. Ввиду регулярности структуры описание топологии CNT легко поддаётся автоматизации средствами GROMACS. Частичные заряды на атомах нанотрубки выбраны нулевыми, так как нанотрубка имеет множество групп симметрии и практически все атомы неотличимы друг от друга.

### **Подготовка систем к расчёту молекулярной динамики**

Моделирования молекулярной динамики выполнены на суперкомпьютерах МГУ им. Ломоносова с помощью пакета программ GROMACS. Для каждого из пяти модифицированных аптамеров созданы системы с ковалентной сшивкой аптамера и карбоксилированной нанотрубки, а также со сшивкой аптамера и Tween 20 в расчёте на гидрофобное взаимодействие этой молекулы с нанотрубкой. Ковалентное взаимодействие накладывает более жёсткие ограничения на взаимное расположение аптамера и нанотрубки. В рамках подготовки систем к динамике молекулы подвергнуты оптимизации. После этого добавлен явно заданный растворитель вода по модели TIP4P/Ew [416]. Молекула ДНК содержит фосфаты и имеет отрицательный заряд, который нужно нейтрализовать. Стоит отметить, что в силовых полях parm ион  $K^+$  параметризован со слишком большим радиусом для канала в G-ДНК. Это иногда приводит к его выталкиванию из структуры, что негативно сказывается на стабилизации квадруплекса. Кроме того, в моделированиях  $K^+$  с  $Cl^-$  происходит кристаллизация в соль KCl. Ион  $Na^+$  имеет меньший радиус и подобных артефактов в моделировании не возникает,

поэтому в данной работе мы использовали катион  $\text{Na}^+$ . Было создано два типа систем: с противоионами в виде катионов  $\text{Na}^+$ , которые заменяют  $\text{K}^+$  в качестве иона, стабилизирующего G-квадруплекс, в физиологической концентрации соли  $\text{NaCl}$ ; а так же с заведомо крупным катионом тетраметиламмония, который, вероятно, не оказывает значительного влияния на структуру аптамера.

## Глава 5

# Выводы

1. Крупнозернистое моделирование длинных молекул РНК применимо для супрамакромолекулярных комплексов при наличии структурных ограничений, не связанных с собственной структурой РНК. Расположение тмРНК длиной 363 нуклеотида определяется позиционированием псевдоузла рК1 и белка SmpB.
2. Локальное представление функциональных районов супрамакромолекулярных комплексов позволяет достоверно моделировать полноатомные модели. Эффективность взаимодействия тилозина и его производных с 23S рРНК в рибосоме определяется сетью водородных связей.
3. Моделирование без ограничений для коротких молекул ДНК позволяет определить конформационное пространство всех топологических элементов структуры. Из трёх внутренних петель минимального 15-членного G-квадруплекса (15-ТВА) только одна тринуклеотидная петля стабилизирует структуру.

4. Точное моделирование с учётом электронной плотности указывает на сложное поведение катионов металлов при взаимодействии с ДНК. Эффективное хелатирование катионов в центре минимального G-квадруплекса (15-ТВА) определяется уменьшением вероятности диссоциации комплекса за счёт динамики латеральных петель.
5. Раскрытие основ формирования структуры и динамики G-квадруплексных ДНК позволяет целенаправленно оптимизировать их узнающие свойства. Изученный в работе структурно-функциональный профиль аптамера 15-ТВА использован для разработки антитромботического препарата и наноаптасенсоров.

# Список литературы

1. *Berendsen H.* Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics. — Cambridge University Press, 2007. — (EngineeringPro collection).
2. *Glattli A., Daura X., Gunsteren W. F. van* Derivation of an improved simple point charge model for liquid water: SPC/A and SPC/L // The Journal of Chemical Physics. — 2002. — Vol. 116, no. 22. — Pp. 9811–9828.
3. *Guillot B.* A reappraisal of what we have learnt during three decades of computer simulations on water // Journal of Molecular Liquids. — 2002. — Vol. 101, no. 13. — Pp. 219–260.
4. *Gunsteren W. van, Berendsen H.* Algorithms for macromolecular dynamics and constraint dynamics // Molecular Physics. — 1977. — Vol. 34, no. 5. — Pp. 1311–1327. — eprint: <http://www.tandfonline.com/doi/pdf/10.1080/00268977700102571>.
5. *Gunsteren W. F. van* [et al.] Biomolecular Simulation: The GROMOS96 manual and userguide. — Zrich, Switzerland : Hochschulverlag AG an der ETH Zrich, 1996.

6. *Ryckaert J.-P., Ciccotti G., Berendsen H. J.* Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes // *Journal of Computational Physics*. — 1977. — Vol. 23, no. 3. — Pp. 327–341.
7. *Hess B.* [et al.] LINCS: A linear constraint solver for molecular simulations // *Journal of Computational Chemistry*. — 1997. — Vol. 18, no. 12. — Pp. 1463–1472.
8. *Muller M., Katsov K., Schick M.* Biological and synthetic membranes: What can be learned from a coarse-grained description? // *Physics Reports*. — 2006. — Vol. 434, no. 56. — Pp. 113–176.
9. *Marrink S.* [et al.] The MARTINI force field: coarse grained model for biomolecular simulations. // *J Phys Chem B*. — 2007. — July. — Vol. 111, no. 27. — Pp. 7812–24.
10. *Marrink S. J.* [et al.] The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations // *The Journal of Physical Chemistry B*. — 2007. — Vol. 111, no. 27. — Pp. 7812–7824. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/jp071097f>.
11. *Johnson M., Head-Gordon T., Louis A.* Representability problems for coarse-grained water potentials. // *J Chem Phys*. — 2007. — Apr. — Vol. 126, no. 14. — P. 144509.
12. *Riniker S., Gunsteren W. van* A simple, efficient polarizable coarse-grained water model for molecular dynamics simulations. // *J Chem Phys*. — 2011. — Feb. — Vol. 134, no. 8. — P. 084110.
13. *Olson W., Zhurkin V.* Working the kinks out of nucleosomal DNA. // *Curr Opin Struct Biol*. — 2011. — June. — Vol. 21, no. 3. — Pp. 348–57.



14. *Wing R.* [et al.] Crystal structure analysis of a complete turn of B-DNA. // *Nature*. — 1980. — Oct. — Vol. 287, no. 5784. — Pp. 755–8.
15. *Burge S.* [et al.] Quadruplex DNA: sequence, topology and structure // *Nucleic Acids Res.* — 2006. — Vol. 34, no. 19. — Pp. 5402–5415.
16. *Mathews D., Turner D.* Prediction of RNA secondary structure by free energy minimization. // *Curr Opin Struct Biol.* — 2006. — June. — Vol. 16, no. 3. — Pp. 270–8.
17. *Jaeger L., Verzemnieks E., Geary C.* The UA handle: a versatile submotif in stable RNA architectures. // *Nucleic Acids Res.* — 2009. — Jan. — Vol. 37, no. 1. — Pp. 215–30.
18. *Leontis N., Stombaugh J., Westhof E.* The non-Watson-Crick base pairs and their associated isostericity matrices. // *Nucleic Acids Res.* — 2002. — Aug. — Vol. 30, no. 16. — Pp. 3497–531.
19. *Lescoute A.* [et al.] Recurrent structural RNA motifs, Isostericity Matrices and sequence alignments. // *Nucleic Acids Res.* — 2005. — Oct. — Vol. 33, no. 8. — Pp. 2395–409.
20. *Sponer J.* [et al.] Quantum chemical studies of nucleic acids: can we construct a bridge to the RNA structural biology and bioinformatics communities? // *J Phys Chem B.* — 2010. — Dec. — Vol. 114, no. 48. — Pp. 15723–41.
21. *Zirbel C.* [et al.] Classification and energetics of the base-phosphate interactions in RNA. // *Nucleic Acids Res.* — 2009. — Aug. — Vol. 37, no. 15. — Pp. 4898–918.
22. *Noller H.* RNA structure: reading the ribosome. // *Science.* — 2005. — Sept. — Vol. 309, no. 5740. — Pp. 1508–14.

23. *Dickerson R.* DNA bending: the prevalence of kinkiness and the virtues of normality. // *Nucleic Acids Res.* — 1998. — Apr. — Vol. 26, no. 8. — Pp. 1906–26.
24. *El Hassan M., Calladine C.* Two distinct modes of protein-induced bending in DNA. // *J Mol Biol.* — 1998. — Sept. — Vol. 282, no. 2. — Pp. 331–43.
25. *Olson W.* [et al.] DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. // *Proc Natl Acad Sci U S A.* — 1998. — Sept. — Vol. 95, no. 19. — Pp. 11163–8.
26. *Srinivasan A., Olson W.* Nucleic acid model building: the multiple backbone solutions associated with a given base morphology. // *J Biomol Struct Dyn.* — 1987. — June. — Vol. 4, no. 6. — Pp. 895–938.
27. *Suzuki M.* [et al.] Use of a 3D structure data base for understanding sequence-dependent conformational aspects of DNA. // *J Mol Biol.* — 1997. — Dec. — Vol. 274, no. 3. — Pp. 421–35.
28. *Olson W.* [et al.] A standard reference frame for the description of nucleic acid base-pair geometry. // *J Mol Biol.* — 2001. — Oct. — Vol. 313, no. 1. — Pp. 229–37.
29. *Lu X., Olson W.* 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. // *Nucleic Acids Res.* — 2003. — Sept. — Vol. 31, no. 17. — Pp. 5108–21.
30. *Sponer J., Leszczynski J., Hobza P.* Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. // *Biopolymers.* — 2002. — Aug. — Vol. 61, no. 1. — Pp. 3–31.

31. *Pérez A.* [et al.] Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. // *Biophys J.* — 2007. — June. — Vol. 92, no. 11. — Pp. 3817–29.
32. *Fürtig B.* [et al.] NMR spectroscopy of RNA. // *Chembiochem.* — 2003. — Oct. — Vol. 4, no. 10. — Pp. 936–62.
33. *Sponer J., Kypr J.* Relationships among rise, cup, roll and stagger in DNA suggested by empirical potential studies of base stacking. // *J Biomol Struct Dyn.* — 1993. — Aug. — Vol. 11, no. 1. — Pp. 27–41.
34. *Neidle S.* Principles of Nucleic Acid Structure. — Elsevier Science, 2010.
35. *Altona C., Sundaralingam M.* Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. // *J Am Chem Soc.* — 1972. — Nov. — Vol. 94, no. 23. — Pp. 8205–12.
36. *Privé G., Yanagi K., Dickerson R.* Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G. // *J Mol Biol.* — 1991. — Jan. — Vol. 217, no. 1. — Pp. 177–99.
37. *Svozil D.* [et al.] DNA conformations and their sequence preferences. // *Nucleic Acids Res.* — 2008. — June. — Vol. 36, no. 11. — Pp. 3690–706.
38. *Schneider B., Neidle S., Berman H.* Conformations of the sugar-phosphate backbone in helical DNA crystal structures. // *Biopolymers.* — 1997. — July. — Vol. 42, no. 1. — Pp. 113–24.
39. *Várnai P.* [et al.] Alpha/gamma transitions in the B-DNA backbone. // *Nucleic Acids Res.* — 2002. — Dec. — Vol. 30, no. 24. — Pp. 5398–406.

40. *Krepl M.* [et al.] Reference simulations of noncanonical nucleic acids with different variants of the AMBER force field: quadruplex DNA, quadruplex RNA and Z-DNA. // *J Chem Theory Comput.* — 2012. — July. — Vol. 8, no. 7. — Pp. 2506–2520.
41. *Jones S.* [et al.] Protein-DNA interactions: A structural analysis. // *J Mol Biol.* — 1999. — Apr. — Vol. 287, no. 5. — Pp. 877–96.
42. *Lu X., Shakked Z., Olson W.* A-form conformational motifs in ligand-bound DNA structures. // *J Mol Biol.* — 2000. — July. — Vol. 300, no. 4. — Pp. 819–40.
43. *Matthews B.* Protein-DNA interaction. No code for recognition. // *Nature.* — 1988. — Sept. — Vol. 335, no. 6188. — Pp. 294–5.
44. *Pabo C., Nekludova L.* Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? // *J Mol Biol.* — 2000. — Aug. — Vol. 301, no. 3. — Pp. 597–624.
45. *Rohs R.* [et al.] Nuance in the double-helix and its role in protein-DNA recognition. // *Curr Opin Struct Biol.* — 2009. — Apr. — Vol. 19, no. 2. — Pp. 171–7.
46. *Richardson J.* [et al.] RNA backbone: consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). // *RNA.* — 2008. — Mar. — Vol. 14, no. 3. — Pp. 465–81.
47. *Murray L.* [et al.] RNA backbone is rotameric. // *Proc Natl Acad Sci U S A.* — 2003. — Nov. — Vol. 100, no. 24. — Pp. 13904–9.
48. *Schneider B., Morávek Z., Berman H.* RNA conformational classes. // *Nucleic Acids Res.* — 2004. — June. — Vol. 32, no. 5. — Pp. 1666–77.

49. *Duarte C., Pyle A.* Stepping through an RNA structure: A novel approach to conformational analysis. // *J Mol Biol.* — 1998. — Dec. — Vol. 284, no. 5. — Pp. 1465–78.
50. *Hershkovitz E.* [et al.] Automated identification of RNA conformational motifs: theory and application to the HMLSU 23S rRNA. // *Nucleic Acids Res.* — 2003. — Nov. — Vol. 31, no. 21. — Pp. 6249–57.
51. *Le Faucheur X.* [et al.] Nonparametric clustering for studying RNA conformations. // *IEEE/ACM Trans Comput Biol Bioinform.* — 2013. — June. — Vol. 8, no. 6. — Pp. 1604–19.
52. *Hershkovitz E.* [et al.] Statistical analysis of RNA backbone. // *IEEE/ACM Trans Comput Biol Bioinform.* — 2006. — Nov. — Vol. 3, no. 1. — Pp. 33–46.
53. *Banás P.* [et al.] Theoretical studies of RNA catalysis: hybrid QM/MM methods and their comparison with MD and QM. // *Methods.* — 2009. — Oct. — Vol. 49, no. 2. — Pp. 202–16.
54. *Lavery R.* [et al.] A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. // *Nucleic Acids Res.* — 2010. — Jan. — Vol. 38, no. 1. — Pp. 299–313.
55. *Vokacova Z.* [et al.] Structure and dynamics of the ApA, ApC, CpA, and CpC RNA dinucleoside monophosphates resolved with NMR scalar spin-spin couplings. // *J Phys Chem B.* — 2009. — Jan. — Vol. 113, no. 4. — Pp. 1182–91.
56. *Beveridge D.* [et al.] Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and

- results on d(CpG) steps. // *Biophys J.* — 2004. — Dec. — Vol. 87, no. 6. — Pp. 3799–813.
57. *Mlídaek A.* [et al.] Conformational Energies of DNA Sugar–Phosphate Backbone: Reference QM Calculations and a Comparison with Density Functional Theory and Molecular Mechanics // *Journal of Chemical Theory and Computation.* — 2010. — Vol. 6, no. 12. — Pp. 3817–3835. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct1004593>.
58. *Sychrovský V.* [et al.] Calculation of structural behavior of indirect NMR spin-spin couplings in the backbone of nucleic acids. // *J Phys Chem B.* — 2006. — Nov. — Vol. 110, no. 45. — Pp. 22894–902.
59. *Krasovska M.* [et al.] Cations and hydration in catalytic RNA: molecular dynamics of the hepatitis delta virus ribozyme. // *Biophys J.* — 2006. — July. — Vol. 91, no. 2. — Pp. 626–38.
60. *Denning E., MacKerell A.* Intrinsic contribution of the 2'-hydroxyl to RNA conformational heterogeneity. // *J Am Chem Soc.* — 2012. — Feb. — Vol. 134, no. 5. — Pp. 2800–6.
61. *Cheatham T., Cieplak P., Kollman P.* A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. // *J Biomol Struct Dyn.* — 1999. — Feb. — Vol. 16, no. 4. — Pp. 845–62.
62. *Hart K.* [et al.] Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. // *J Chem Theory Comput.* — 2012. — Jan. — Vol. 8, no. 1. — Pp. 348–362.
63. *Mlýnský V.* [et al.] Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H<sup>+</sup> forms are most consistent with crystal

- structures of hairpin ribozyme. // *J Phys Chem B*. — 2010. — May. — Vol. 114, no. 19. — Pp. 6642–52.
64. *Zgarbova M.* [et al.] Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. // *J Chem Theory Comput*. — 2011. — Sept. — Vol. 7, no. 9. — Pp. 2886–2902.
65. *Jurecka P.* [et al.] Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. // *J Comput Chem*. — 2007. — Jan. — Vol. 28, no. 2. — Pp. 555–69.
66. *Zgarbovi M.* [et al.] A Novel Approach for Deriving Force Field Torsion Angle Parameters Accounting for Conformation-Dependent Solvation Effects // *Journal of Chemical Theory and Computation*. — 2012. — Vol. 8, no. 9. — Pp. 3232–3242. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct3001987>.
67. *Wolfe-Simon F.* [et al.] A bacterium that can grow by using arsenic instead of phosphorus. // *Science*. — 2011. — June. — Vol. 332, no. 6034. — Pp. 1163–6.
68. *Mlidek A.* [et al.] On the Geometry and Electronic Structure of the As-DNA Backbone // *The Journal of Physical Chemistry Letters*. — 2011. — Vol. 2, no. 5. — Pp. 389–392. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/jz200015n>.
69. *Denning E., Mackerell A.* Impact of arsenic/phosphorus substitution on the intrinsic conformational properties of the phosphodiester backbone of DNA

- investigated using ab initio quantum mechanical calculations. // *J Am Chem Soc.* — 2011. — Apr. — Vol. 133, no. 15. — Pp. 5770–2.
70. *Long J. W., Ray W. J.* Kinetics and thermodynamics of the formation of glucose arsenate. Reaction of glucose arsenate with phosphoglucomutase // *Biochemistry.* — 1973. — Vol. 12, no. 20. — Pp. 3932–3937. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/bi00744a023>.
71. *Baer C. D., Edwards J. O., Rieger P. H.* Kinetics of the hydrolysis of arsenate(V) triesters // *Inorganic Chemistry.* — 1981. — Vol. 20, no. 3. — Pp. 905–907. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ic50217a052>.
72. *Baer C. D.* [et al.] Kinetics of an associative ligand-exchange process: alcohol exchange with arsenate(V) triesters // *Journal of the American Chemical Society.* — 1980. — Vol. 102, no. 18. — Pp. 5793–5798. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja00538a016>.
73. *Reaves M.* [et al.] Absence of detectable arsenate in DNA from arsenate-grown GFAJ-1 cells. // *Science.* — 2012. — July. — Vol. 337, no. 6093. — Pp. 470–3.
74. *Micklefield J.* Backbone modification of nucleic acids: synthesis, structure and therapeutic applications. // *Curr Med Chem.* — 2001. — Aug. — Vol. 8, no. 10. — Pp. 1157–79.
75. *Joyce G.* Evolution. Toward an alternative biology. // *Science.* — 2012. — Apr. — Vol. 336, no. 6079. — Pp. 307–8.
76. *Pinheiro V.* [et al.] Synthetic genetic polymers capable of heredity and evolution. // *Science.* — 2012. — Apr. — Vol. 336, no. 6079. — Pp. 341–4.



77. *Schoning K.* [et al.] Chemical etiology of nucleic acid structure: the alpha-threofuranosyl-(3'→2') oligonucleotide system. // *Science*. — 2000. — Nov. — Vol. 290, no. 5495. — Pp. 1347–51.
78. *Zhang L., Peritz A., Meggers E.* A simple glycol nucleic acid. // *J Am Chem Soc*. — 2005. — Mar. — Vol. 127, no. 12. — Pp. 4174–5.
79. *Lescrinier E.* [et al.] Solution structure of a HNA-RNA hybrid. // *Chem Biol*. — 2000. — Sept. — Vol. 7, no. 9. — Pp. 719–31.
80. *Allart B.* [et al.] D-Altritol Nucleic Acids (ANA): Hybridisation Properties, Stability, and Initial Structural Analysis // *Chemistry – A European Journal*. — 1999. — Vol. 5, no. 8. — Pp. 2424–2431.
81. *Nielsen P.* [et al.] Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. // *Science*. — 1991. — Dec. — Vol. 254, no. 5037. — Pp. 1497–500.
82. *Betts L.* [et al.] A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex. // *Science*. — 1995. — Dec. — Vol. 270, no. 5243. — Pp. 1838–41.
83. *Ovaere M.* [et al.] How does hydroxyl introduction influence the double helical structure: the stabilization of an altritol nucleic acid:ribonucleic acid duplex. // *Nucleic Acids Res*. — 2012. — Aug. — Vol. 40, no. 15. — Pp. 7573–83.
84. *Topham C., Smith J.* Orientation preferences of backbone secondary amide functional groups in peptide nucleic acid complexes: quantum chemical calculations reveal an intrinsic preference of cationic D-amino acid-based chiral PNA analogues for the P-form. // *Biophys J*. — 2007. — Feb. — Vol. 92, no. 3. — Pp. 769–86.

85. *Soliva R.* [et al.] Molecular Dynamics Simulations of PNADNA and PNARNA Duplexes in Aqueous Solution // Journal of the American Chemical Society. — 2000. — Vol. 122, no. 25. — Pp. 5997–6008. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja000259h>.
86. *Sen S., Nilsson L.* Molecular Dynamics of Duplex Systems Involving PNA: Structural and Dynamical Consequences of the Nucleic Acid Backbone // Journal of the American Chemical Society. — 1998. — Vol. 120, no. 4. — Pp. 619–631. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja972234x>.
87. *Shields G. C., Laughton C. A., Orozco M.* Molecular Dynamics Simulation of a PNADNAPNA Triple Helix in Aqueous Solution // Journal of the American Chemical Society. — 1998. — Vol. 120, no. 24. — Pp. 5895–5904. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja9723444>.
88. *Denning E.* [et al.] Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. // J Comput Chem. — 2011. — July. — Vol. 32, no. 9. — Pp. 1929–43.
89. *Case D.* [et al.] The Amber biomolecular simulation programs. // J Comput Chem. — 2005. — Dec. — Vol. 26, no. 16. — Pp. 1668–88.
90. *Cornell W. D.* [et al.] A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules // Journal of the American Chemical Society. — 1995. — Vol. 117, no. 19. — Pp. 5179–5197. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja00124a002>.
91. *Bešševó I.* [et al.] Simulations of A-RNA duplexes. The effect of sequence, solute force field, water model, and salt concentration. // J Phys Chem B. — 2012. — Aug. — Vol. 116, no. 33. — Pp. 9899–916.

92. *Ode H.* [et al.] Force field parameters for rotation around chi torsion axis in nucleic acids. // *J Comput Chem.* — 2008. — Nov. — Vol. 29, no. 15. — Pp. 2531–42.
93. *Yildirim I.* [et al.] Reparameterization of RNA chi Torsion Parameters for the AMBER Force Field and Comparison to NMR Spectra for Cytidine and Uridine. // *J Chem Theory Comput.* — 2010. — May. — Vol. 6, no. 5. — Pp. 1520–1531.
94. *Yildirim I.* [et al.] Revision of AMBER Torsional Parameters for RNA Improves Free Energy Predictions for Tetramer Duplexes with GC and iGiC Base Pairs. // *J Chem Theory Comput.* — 2012. — Jan. — Vol. 8, no. 1. — Pp. 172–181.
95. *Ditzler M.* [et al.] Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in. // *Acc Chem Res.* — 2010. — Jan. — Vol. 43, no. 1. — Pp. 40–7.
96. *Baniš P.* [et al.] Can We Accurately Describe the Structure of Adenine Tracts in B-DNA? Reference Quantum-Chemical Computations Reveal Overstabilization of Stacking by Molecular Mechanics // *Journal of Chemical Theory and Computation.* — 2012. — Vol. 8, no. 7. — Pp. 2448–2460. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct3001238>.
97. *Miller M.* [et al.] Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3' sequence of d(AG<sub>3</sub>(TTAG<sub>3</sub>)<sub>3</sub>). // *J Am Chem Soc.* — 2010. — Dec. — Vol. 132, no. 48. — Pp. 17105–7.

98. *Zheng K.* [et al.] Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. // *Nucleic Acids Res.* — 2010. — Jan. — Vol. 38, no. 1. — Pp. 327–38.
99. *Fujimoto T.* [et al.] The effects of molecular crowding on the structure and stability of g-quadruplexes with an abasic site. // *J Nucleic Acids.* — 2011. — June. — Vol. 2011. — P. 857149.
100. *Webba da Silva M.* Geometric formalism for DNA quadruplex folding. // *Chemistry.* — 2007. — June. — Vol. 13, no. 35. — Pp. 9738–45.
101. *Crnugelj M., Sket P., Plavec J.* Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations. // *J Am Chem Soc.* — 2003. — July. — Vol. 125, no. 26. — Pp. 7866–71.
102. *Strahan G., Keniry M., Shafer R.* NMR structure refinement and dynamics of the K<sup>+</sup>-[d(G3T4G3)]<sub>2</sub> quadruplex via particle mesh Ewald molecular dynamics simulations. // *Biophys J.* — 1998. — Aug. — Vol. 75, no. 2. — Pp. 968–81.
103. *Smith F., Feigon J.* Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. // *Nature.* — 1992. — Mar. — Vol. 356, no. 6365. — Pp. 164–8.
104. *Haider S., Parkinson G., Neidle S.* Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. // *J Mol Biol.* — 2002. — July. — Vol. 320, no. 2. — Pp. 189–200.
105. *Stefl R.* [et al.] Molecular dynamics of DNA quadruplex molecules containing inosine, 6-thioguanine and 6-thiopurine. // *Biophys J.* — 2001. — Jan. — Vol. 80, no. 1. — Pp. 455–68.

106. *Spacková N.* [et al.] Theoretical study of the guanine → 6-thioguanine substitution in duplexes, triplexes, and tetraplexes. // *J Am Chem Soc.* — 2004. — Nov. — Vol. 126, no. 44. — Pp. 14642–50.
107. *Lee M.* [et al.] Large-scale conformational dynamics of the HIV-1 integrase core domain and its catalytic loop mutants. // *Biophys J.* — 2005. — May. — Vol. 88, no. 5. — Pp. 3133–46.
108. *Koller A.* [et al.] Aromatic N versus aromatic F: bioisosterism discovered in RNA base pairing interactions leads to a novel class of universal base analogs. // *Nucleic Acids Res.* — 2010. — May. — Vol. 38, no. 9. — Pp. 3133–46.
109. *Pronk S.* [et al.] Copernicus: A new paradigm for parallel adaptive molecular dynamics // *High Performance Computing, Networking, Storage and Analysis (SC), 2011 International Conference for.* — Nov. 2011. — Pp. 1–10.
110. *Lindorff-Larsen K.* [et al.] How fast-folding proteins fold. // *Science.* — 2011. — Oct. — Vol. 334, no. 6055. — Pp. 517–20.
111. *Hünenberger P., McCammon J.* Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study. // *Biophys Chem.* — 1999. — Apr. — Vol. 78, 1-2. — Pp. 69–88.
112. *Klein D.* [et al.] The kink-turn: a new RNA secondary structure motif. // *EMBO J.* — 2001. — Aug. — Vol. 20, no. 15. — Pp. 4214–21.
113. *Rázga F.* [et al.] Hinge-like motions in RNA kink-turns: the role of the second a-minor motif and nominally unpaired bases. // *Biophys J.* — 2005. — May. — Vol. 88, no. 5. — Pp. 3466–85.

114. *Schroeder K.* [et al.] A structural database for k-turn motifs in RNA. // RNA. — 2010. — Aug. — Vol. 16, no. 8. — Pp. 1463–8.
115. *Shankar N.* [et al.] The NMR structure of an internal loop from 23S ribosomal RNA differs from its structure in crystals of 50s ribosomal subunits. // Biochemistry. — 2006. — Oct. — Vol. 45, no. 39. — Pp. 11776–89.
116. *Reblove K.* [et al.] An RNA molecular switch: Intrinsic flexibility of 23S rRNA Helices 40 and 68 5'-UAA/5'-GAN internal loops studied by molecular dynamics methods. // J Chem Theory Comput. — 2010. — Jan. — Vol. 2010, no. 6. — Pp. 910–929.
117. *Stefl R.* [et al.] Formation pathways of a guanine-quadruplex DNA revealed by molecular dynamics and thermodynamic analysis of the substates. // Biophys J. — 2003. — Sept. — Vol. 85, no. 3. — Pp. 1787–804.
118. *Dickerson R.* [et al.] The effect of crystal packing on oligonucleotide double helix structure. // J Biomol Struct Dyn. — 1987. — Dec. — Vol. 5, no. 3. — Pp. 557–79.
119. *Jain S., Sundaralingam M.* Effect of crystal packing environment on conformation of the DNA duplex. Molecular structure of the A-DNA octamer d(G-T-G-T-A-C-A-C) in two crystal forms. // J Biol Chem. — 1989. — Aug. — Vol. 264, no. 22. — Pp. 12780–4.
120. *Shakke Z.* [et al.] The conformation of the DNA double helix in the crystal is dependent on its environment. // Nature. — 1989. — Nov. — Vol. 342, no. 6248. — Pp. 456–60.
121. *Dieckmann T.* [et al.] Solution structure of an ATP-binding RNA aptamer reveals a novel fold. // RNA. — 1996. — July. — Vol. 2, no. 7. — Pp. 628–40.

122. *Ramos A., Varani G.* Structure of the acceptor stem of Escherichia coli tRNA Ala: role of the G3.U70 base pair in synthetase recognition. // *Nucleic Acids Res.* — 1997. — June. — Vol. 25, no. 11. — Pp. 2083–90.
123. *Gueron M., Kochoyan M., Leroy J.* A single mode of DNA base-pair opening drives imino proton exchange. // *Nature.* — 1987. — Apr. — Vol. 328, no. 6125. — Pp. 89–92.
124. *Furse K., Corcelli S.* Dynamical signature of abasic damage in DNA. // *J Am Chem Soc.* — 2011. — Feb. — Vol. 133, no. 4. — Pp. 720–3.
125. *Dornberger U., Leijon M., Fritzsche H.* High base pair opening rates in tracts of GC base pairs. // *J Biol Chem.* — 1999. — Mar. — Vol. 274, no. 11. — Pp. 6957–62.
126. *Warmlander S.* [et al.] The influence of the thymine C5 methyl group on spontaneous base pair breathing in DNA. // *J Biol Chem.* — 2002. — Aug. — Vol. 277, no. 32. — Pp. 28491–7.
127. *Lane A.* [et al.] Stability and kinetics of G-quadruplex structures. // *Nucleic Acids Res.* — 2008. — Oct. — Vol. 36, no. 17. — Pp. 5482–515.
128. *Tidor B.* [et al.] Dynamics of DNA oligomers. // *J Biomol Struct Dyn.* — 1983. — Oct. — Vol. 1, no. 1. — Pp. 231–52.
129. *Seibel G., Singh U., Kollman P.* A molecular dynamics simulation of double-helical B-DNA including counterions and water. // *Proc Natl Acad Sci U S A.* — 1985. — Oct. — Vol. 82, no. 19. — Pp. 6537–40.
130. *Chen A., Draper D., Pappu R.* Molecular simulation studies of monovalent counterion-mediated interactions in a model RNA kissing loop. // *J Mol Biol.* — 2009. — July. — Vol. 390, no. 4. — Pp. 805–19.

131. *Auffinger P., Hashem Y.* Nucleic acid solvation: from outside to insight. // *Curr Opin Struct Biol.* — 2007. — June. — Vol. 17, no. 3. — Pp. 325–33.
132. *Ke A.* [et al.] Structural roles of monovalent cations in the HDV ribozyme. // *Structure.* — 2007. — Mar. — Vol. 15, no. 3. — Pp. 281–7.
133. *Sponer J., Spacková N.* Molecular dynamics simulations and their application to four-stranded DNA. // *Methods.* — 2007. — Dec. — Vol. 43, no. 4. — Pp. 278–90.
134. *Joung I., Cheatham T.* Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. // *J Phys Chem B.* — 2008. — July. — Vol. 112, no. 30. — Pp. 9020–41.
135. *Balasubramanian S., Neidle S.* G-quadruplex nucleic acids as therapeutic targets. // *Curr Opin Chem Biol.* — 2009. — June. — Vol. 13, no. 3. — Pp. 345–53.
136. *Neidle S.* The structures of quadruplex nucleic acids and their drug complexes. // *Curr Opin Struct Biol.* — 2009. — June. — Vol. 19, no. 3. — Pp. 239–50.
137. *Alcaro S.* The impact of the G-quadruplex conformation in the development of novel therapeutic and diagnostic agents. // *Curr Pharm Des.* — 2012Sen. — Vol. 18, no. 14. — Pp. 1865–6.
138. *Bhattacharya S.* [et al.] Symmetrical bisbenzimidazoles with benzenediyl spacer: the role of the shape of the ligand on the stabilization and structural alterations in telomeric G-quadruplex DNA and telomerase inhibition. // *Bioconjug Chem.* — 2010. — July. — Vol. 21, no. 7. — Pp. 1148–59.



139. *Haider S., Autiero I., Neidle S.* Surface area accessibility and the preferred topology of telomeric DNA quadruplex-ligand complexes. // *Biochimie.* — 2011. — Aug. — Vol. 93, no. 8. — Pp. 1275–9.
140. *Sparapani S.* [et al.] Rational design of acridine-based ligands with selectivity for human telomeric quadruplexes. // *J Am Chem Soc.* — 2010. — Sept. — Vol. 132, no. 35. — Pp. 12263–72.
141. *Ma D.* [et al.] Molecular modeling of drug-DNA interactions: virtual screening to structure-based design. // *Biochimie.* — 2011. — Aug. — Vol. 93, no. 8. — Pp. 1252–66.
142. *Trotta R.* [et al.] A more detailed picture of the interactions between virtual screening-derived hits and the DNA G-quadruplex: NMR, molecular modelling and ITC studies. // *Biochimie.* — 2011. — Aug. — Vol. 93, no. 8. — Pp. 1280–7.
143. *Haider S., Neidle S.* Molecular modeling and simulation of G-quadruplexes and quadruplex-ligand complexes. // *Methods Mol Biol.* — 2010. — Vol. 608, None. — Pp. 17–37.
144. *Han H.* [et al.] Selective interactions of cationic porphyrins with G-quadruplex structures. // *J Am Chem Soc.* — 2001. — Sept. — Vol. 123, no. 37. — Pp. 8902–13.
145. *Arora A.* [et al.] Binding of berberine to human telomeric quadruplex - spectroscopic, calorimetric and molecular modeling studies. // *FEBS J.* — 2008. — Aug. — Vol. 275, no. 15. — Pp. 3971–83.
146. *Li M.* [et al.] Molecular dynamics studies of the 3D structure and planar ligand binding of a quadruplex dimer. // *J Mol Model.* — 2011. — Mar. — Vol. 17, no. 3. — Pp. 515–26.

147. *Collie G.* [et al.] A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. // *Nucleic Acids Res.* — 2010. — Sept. — Vol. 38, no. 16. — Pp. 5569–80.
148. *Hou J.* [et al.] New insights into the structures of ligand-quadruplex complexes from molecular dynamics simulations. // *J Phys Chem B.* — 2010. — Nov. — Vol. 114, no. 46. — Pp. 15301–10.
149. *Li M.* [et al.] The 3D structures of G-quadruplexes of HIV-1 integrase inhibitors: molecular dynamics simulations in aqueous solution and in the gas phase. // *J Mol Model.* — 2010. — Apr. — Vol. 16, no. 4. — Pp. 645–57.
150. *Li M., Luo Q., Li Z.* Molecular dynamics study on the interactions of porphyrin with two antiparallel human telomeric quadruplexes. // *J Phys Chem B.* — 2010. — May. — Vol. 114, no. 18. — Pp. 6216–24.
151. *Agrawal S., Ojha R., Maiti S.* Energetics of the human Tel-22 quadruplex-telomestatin interaction: a molecular dynamics study. // *J Phys Chem B.* — 2008. — June. — Vol. 112, no. 22. — Pp. 6828–36.
152. *Garner T.* [et al.] Selectivity of small molecule ligands for parallel and antiparallel DNA G-quadruplex structures. // *Org Biomol Chem.* — 2009. — Oct. — Vol. 7, no. 20. — Pp. 4194–200.
153. *Aixiao L.* [et al.] Molecular modeling study of binding site selectivity of TQMP to G-quadruplex DNA. // *Eur J Med Chem.* — 2010. — Mar. — Vol. 45, no. 3. — Pp. 983–91.
154. *Bazzicalupi C.* [et al.] Modeling and biological investigations of an unusual behavior of novel synthesized acridine-based polyamine ligands in the binding of double helix and G-quadruplex DNA. // *ChemMedChem.* — 2010. — Dec. — Vol. 5, no. 12. — Pp. 1995–2005.

155. *Petraccone L.* [et al.] The triazatruxene derivative azatrux binds to the parallel form of the human telomeric G-quadruplex under molecular crowding conditions: biophysical and molecular modeling studies. // *Biochimie*. — 2011. — Aug. — Vol. 93, no. 8. — Pp. 1318–27.
156. *Rao L.* [et al.] Interactions of a platinum-modified perylene derivative with the human telomeric G-quadruplex. // *J Phys Chem B*. — 2011. — Nov. — Vol. 115, no. 46. — Pp. 13701–12.
157. *Read M.* [et al.] Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: structure-activity relationships. // *J Med Chem*. — 1999. — Nov. — Vol. 42, no. 22. — Pp. 4538–46.
158. *Murat P., Singh Y., Defrancq E.* Methods for investigating G-quadruplex DNA/ligand interactions. // *Chem Soc Rev*. — 2011. — Nov. — Vol. 40, no. 11. — Pp. 5293–307.
159. *Xu Y., Komiyama M.* Structure, function and targeting of human telomere RNA. // *Methods*. — 2012. — May. — Vol. 57, no. 1. — Pp. 100–5.
160. *Li J.* [et al.] Identification of nonplanar small molecule for G-quadruplex grooves: molecular docking and molecular dynamic study. // *Bioorg Med Chem Lett*. — 2011. — Dec. — Vol. 21, no. 23. — Pp. 6969–72.
161. *Ibrahim M.* Molecular mechanical study of halogen bonding in drug discovery. // *J Comput Chem*. — 2011. — Sept. — Vol. 32, no. 12. — Pp. 2564–74.
162. *Rendine S.* [et al.] Halogen bonding in ligand-receptor systems in the framework of classical force fields. // *Phys Chem Chem Phys*. — 2011. — Nov. — Vol. 13, no. 43. — Pp. 19508–16.

163. *Bayly C. I.* [et al.] A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model // *The Journal of Physical Chemistry*. — 1993. — Vol. 97, no. 40. — Pp. 10269–10280. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/j100142a004>.
164. *Dupradeau F.* [et al.] R.E.DD.B.: a database for RESP and ESP atomic charges, and force field libraries. // *Nucleic Acids Res.* — 2008. — Jan. — Vol. 36, Database issue. — Pp. D360–7.
165. *Wang J.* [et al.] Development and testing of a general amber force field. // *J Comput Chem.* — 2004. — July. — Vol. 25, no. 9. — Pp. 1157–74.
166. *Wang J.* [et al.] Automatic atom type and bond type perception in molecular mechanical calculations. // *J Mol Graph Model.* — 2006. — Oct. — Vol. 25, no. 2. — Pp. 247–60.
167. *Noy A.* [et al.] Theoretical study of large conformational transitions in DNA: the B $\leftrightarrow$ A conformational change in water and ethanol/water. // *Nucleic Acids Res.* — 2007. — Nov. — Vol. 35, no. 10. — Pp. 3330–8.
168. *Beck D., White G., Daggett V.* Exploring the energy landscape of protein folding using replica-exchange and conventional molecular dynamics simulations. // *J Struct Biol.* — 2007. — Mar. — Vol. 157, no. 3. — Pp. 514–23.
169. *Meyer M., Hocquet A., Sühnel J.* Interaction of sodium and potassium ions with sandwiched cytosine-, guanine-, thymine-, and uracil-base tetrads. // *J Comput Chem.* — 2005. — Mar. — Vol. 26, no. 4. — Pp. 352–64.
170. *Mourik T. van, Dingley A.* Characterization of the monovalent ion position and hydrogen-bond network in guanine quartets by DFT calculations of NMR parameters. // *Chemistry*. — 2005. — Oct. — Vol. 11, no. 20. — Pp. 6064–79.

171. *Fonseca Guerra C.* [et al.] Telomere structure and stability: covalency in hydrogen bonds, not resonance assistance, causes cooperativity in guanine quartets. // *Chemistry*. — 2011. — Nov. — Vol. 17, no. 45. — Pp. 12612–22.
172. *Bogdanov A.* [Some structural aspects of the peptidyltransferase reaction]. // *Mol Biol (Mosk)*. — 2003. — July. — Vol. 37, no. 3. — Pp. 511–4.
173. *Tu G.* [et al.] C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. // *J Biol Chem*. — 1995. — Apr. — Vol. 270, no. 16. — Pp. 9322–6.
174. *Keiler K., Waller P., Sauer R.* Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. // *Science*. — 1996. — Feb. — Vol. 271, no. 5251. — Pp. 990–3.
175. *Sundermeier T.* [et al.] A previously uncharacterized role for small protein B (SmpB) in transfer messenger RNA-mediated trans-translation. // *Proc Natl Acad Sci U S A*. — 2005. — Feb. — Vol. 102, no. 7. — Pp. 2316–21.
176. *Nonin-Lecomte S., Felden B., Dardel F.* NMR structure of the *Aquifex aeolicus* tmRNA pseudoknot PK1: new insights into the recoding event of the ribosomal trans-translation // *Nucleic Acids Res*. — 2006. — Vol. 34, no. 6. — Pp. 1847–1853.
177. *Gutmann S.* [et al.] Crystal structure of the transfer-RNA domain of transfer-messenger RNA in complex with SmpB // *Nature*. — 2003. — Aug. — Vol. 424, no. 6949. — Pp. 699–703.
178. *Valle M.* [et al.] Visualizing tmRNA entry into a stalled ribosome // *Science*. — 2003. — Apr. — Vol. 300, no. 5616. — Pp. 127–130.

179. *Kaur S.* [et al.] Cryo-EM visualization of transfer messenger RNA with two SmpBs in a stalled ribosome. // *Proc Natl Acad Sci U S A.* — 2006. — Oct. — Vol. 103, no. 44. — Pp. 16484–9.
180. *Moore S., Sauer R.* The tmRNA system for translational surveillance and ribosome rescue. // *Annu Rev Biochem.* — 2007. — Apr. — Vol. 76. — Pp. 101–24.
181. *Bugaeva E.* [et al.] Structural features of the tmRNA-ribosome interaction. // *RNA.* — 2009. — Dec. — Vol. 15, no. 12. — Pp. 2312–20.
182. *Berman H. M.* [et al.] The Protein Data Bank // *Nucleic Acids Res.* — 2000. — Jan. — Vol. 28, no. 1. — Pp. 235–242.
183. *Murthy V. L., Rose G. D.* RNABase: an annotated database of RNA structures // *Nucleic Acids Res.* — 2003. — Jan. — Vol. 31, no. 1. — Pp. 502–504.
184. *Sharma S., Ding F., Dokholyan N. V.* iFoldRNA: three-dimensional RNA structure prediction and folding // *Bioinformatics.* — 2008. — Sept. — Vol. 24, no. 17. — Pp. 1951–1952.
185. *Jonikas M. A.* [et al.] Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters // *RNA.* — 2009. — Feb. — Vol. 15, no. 2. — Pp. 189–199.
186. *Burks J.* [et al.] Comparative 3-D modeling of tmRNA // *BMC Mol. Biol.* — 2005. — Vol. 6. — P. 14.
187. *Massire C., Westhof E.* MANIP: an interactive tool for modelling RNA // *J. Mol. Graph. Model.* — 1998. — Vol. 16, 4-6. — Pp. 197–205.

188. *Tsai H.* [et al.] Molecular modeling of the three-dimensional structure of the bacterial RNase P holoenzyme. // *J Mol Biol.* — 2003. — Jan. — Vol. 325, no. 4. — Pp. 661–75.
189. *J. M. T., A. C. D.* Modeling Unusual Nucleic Acid Structures // *Molecular Modeling of Nucleic Acids.* — Chap. 25. Pp. 379–393. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/bk-1998-0682.ch024>.
190. *Parisien M., Major F.* The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. // *Nature.* — 2008. — Mar. — Vol. 452, no. 7183. — Pp. 51–5.
191. *Boyapati V.* [et al.] Basis for ligand discrimination between ON and OFF state riboswitch conformations: the case of the SAM-I riboswitch. // *RNA.* — 2012. — June. — Vol. 18, no. 6. — Pp. 1230–43.
192. *Gillet R.* [et al.] Scaffolding as an organizing principle in trans-translation. The roles of small protein B and ribosomal protein S1 // *J. Biol. Chem.* — 2007. — Mar. — Vol. 282, no. 9. — Pp. 6356–6363.
193. *Yusupova G.* [et al.] Structural basis for messenger RNA movement on the ribosome. // *Nature.* — 2006. — Nov. — Vol. 444, no. 7117. — Pp. 391–4.
194. *Harms J.* [et al.] High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. // *Cell.* — 2001. — Nov. — Vol. 107, no. 5. — Pp. 679–88.
195. *Shpanchenko O.* [et al.] Stepping transfer messenger RNA through the ribosome. // *J Biol Chem.* — 2005. — May. — Vol. 280, no. 18. — Pp. 18368–74.

196. *Bugaeva E.* [et al.] One SmpB molecule accompanies tmRNA during its passage through the ribosomes. // FEBS Lett. — 2008. — Apr. — Vol. 582, no. 10. — Pp. 1532–6.
197. *Hallier M.* [et al.] Pre-binding of small protein B to a stalled ribosome triggers trans-translation. // J Biol Chem. — 2004. — June. — Vol. 279, no. 25. — Pp. 25978–85.
198. *Ramrath D.* [et al.] The complex of tmRNA-SmpB and EF-G on translocating ribosomes. // Nature. — 2012. — May. — Vol. 485, no. 7399. — Pp. 526–9.
199. *Mankin A.* Nascent peptide in the "birth canal" of the ribosome. // Trends Biochem Sci. — 2006. — Jan. — Vol. 31, no. 1. — Pp. 11–3.
200. *Bogdanov A.* [et al.] Ribosomal tunnel and translation regulation. // Biochemistry (Mosc). — 2010. — Dec. — Vol. 75, no. 13. — Pp. 1501–16.
201. *Cruz-Vera L.* [et al.] Nascent polypeptide sequences that influence ribosome function. // Curr Opin Microbiol. — 2011. — Apr. — Vol. 14, no. 2. — Pp. 160–6.
202. *Ōmura S.* Macrolide Antibiotics: Chemistry, Biology, and Practice. — Academic Press, 2002.
203. *Wilson D.* On the specificity of antibiotics targeting the large ribosomal subunit. // Ann N Y Acad Sci. — 2011. — Dec. — Vol. 1241. — Pp. 1–16.
204. *Kannan K., Mankin A.* Macrolide antibiotics in the ribosome exit tunnel: species-specific binding and action. // Ann N Y Acad Sci. — 2011. — Dec. — Vol. 1241. — Pp. 33–47.



205. *Hansen J.* [et al.] The structures of four macrolide antibiotics bound to the large ribosomal subunit. // *Mol Cell*. — 2002. — July. — Vol. 10, no. 1. — Pp. 117–28.
206. *Kannan K., Vázquez-Laslop N., Mankin A.* Selective protein synthesis by ribosomes with a drug-obstructed exit tunnel. // *Cell*. — 2012. — Oct. — Vol. 151, no. 3. — Pp. 508–20.
207. *Sumbatyan N., Korshunova G., Bogdanov A.* Peptide derivatives of antibiotics tylosin and desmycosin, protein synthesis inhibitors. // *Biochemistry (Mosc)*. — 2003. — Oct. — Vol. 68, no. 10. — Pp. 1156–8.
208. *Korshunova G.* [et al.] [Peptide derivatives of tylosin-related macrolides]. // *Bioorg Khim*. — 2007. — May. — Vol. 33, no. 2. — Pp. 235–44.
209. *Starosta A.* [et al.] Interplay between the ribosomal tunnel, nascent chain, and macrolides influences drug inhibition. // *Chem Biol*. — 2010. — May. — Vol. 17, no. 5. — Pp. 504–14.
210. *Trabuco L.* [et al.] Recognition of the regulatory nascent chain TnaC by the ribosome. // *Structure*. — 2010. — May. — Vol. 18, no. 5. — Pp. 627–37.
211. *Petrone P.* [et al.] Side-chain recognition and gating in the ribosome exit tunnel. // *Proc Natl Acad Sci U S A*. — 2008. — Oct. — Vol. 105, no. 43. — Pp. 16549–54.
212. *Dunkle J.* [et al.] Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. // *Proc Natl Acad Sci U S A*. — 2010. — Oct. — Vol. 107, no. 40. — Pp. 17152–7.

213. *Kowalak J., Bruenger E., McCloskey J.* Posttranscriptional modification of the central loop of domain V in *Escherichia coli* 23 S ribosomal RNA. // *J Biol Chem.* — 1995. — July. — Vol. 270, no. 30. — Pp. 17758–64.
214. *Hornak V.* [et al.] Comparison of multiple Amber force fields and development of improved protein backbone parameters. // *Proteins.* — 2006. — Nov. — Vol. 65, no. 3. — Pp. 712–25.
215. *Petropoulos A.* [et al.] Stepwise binding of tylosin and erythromycin to *Escherichia coli* ribosomes, characterized by kinetic and footprinting analysis. // *J Biol Chem.* — 2008. — Feb. — Vol. 283, no. 8. — Pp. 4756–65.
216. *Karahalios P.* [et al.] On the mechanism of action of 9-O-arylalkyloxime derivatives of 6-O-mycaminosyltylonolide, a new class of 16-membered macrolide antibiotics. // *Mol Pharmacol.* — 2006. — Oct. — Vol. 70, no. 4. — Pp. 1271–80.
217. *Bulkley D.* [et al.] Revisiting the structures of several antibiotics bound to the bacterial ribosome. // *Proc Natl Acad Sci U S A.* — 2010. — Oct. — Vol. 107, no. 40. — Pp. 17158–63.
218. *Llano-Sotelo B.* [et al.] Binding and action of CEM-101, a new fluoroketolide antibiotic that inhibits protein synthesis. // *Antimicrob Agents Chemother.* — 2010. — Dec. — Vol. 54, no. 12. — Pp. 4961–70.
219. *Poehlsgaard J.* [et al.] Visualizing the 16-membered ring macrolides tildipirosin and tilmicosin bound to their ribosomal site. // *ACS Chem Biol.* — 2012. — Aug. — Vol. 7, no. 8. — Pp. 1351–5.
220. *GELLERT M., LIPSETT M. N., DAVIES D. R.* Helix formation by guanylic acid // *Proceedings of the National Academy of Sciences of the United States of America.* — 1962. — Dec. — Vol. 48. — Pp. 2013–2018.

221. *Arnott S., Chandrasekaran R., Marttila C. M.* Structures for polyinosinic acid and polyguanylic acid // *The Biochemical journal*. — 1974. — Aug. — Vol. 141, no. 2. — Pp. 537–543.
222. *Zimmerman S. B.* X-ray study by fiber diffraction methods of a self-aggregate of guanosine-5'-phosphate with the same helical parameters as poly(rG) // *Journal of molecular biology*. — 1976. — Sept. — Vol. 106, no. 3. — Pp. 663–672.
223. *Simonsson T.* G-quadruplex DNA structures—variations on a theme // *Biological chemistry*. — 2001. — Apr. — Vol. 382, no. 4. — Pp. 621–628.
224. *De Lange T.* Telomere-related genome instability in cancer // *Cold Spring Harbor symposia on quantitative biology*. — 2005. — Vol. 70. — Pp. 197–204.
225. *Wright W. E.* [et al.] Normal human chromosomes have long G-rich telomeric overhangs at one end // *Genes & development*. — 1997. — Nov. — Vol. 11, no. 21. — Pp. 2801–2809.
226. *Lei M., Podell E. R., Cech T. R.* Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection // *Nature structural & molecular biology*. — 2004. — Dec. — Vol. 11, no. 12. — Pp. 1223–1229.
227. *Sen D., Gilbert W.* Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis // *Nature*. — 1988. — July. — Vol. 334, no. 6180. — Pp. 364–366.
228. *Sundquist W. I., Klug A.* Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops // *Nature*. — 1989. — Dec. — Vol. 342, no. 6251. — Pp. 825–829.

229. *Gilbert D. E., Feigon J.* Multistranded DNA structures // Current opinion in structural biology. — 1999. — June. — Vol. 9, no. 3. — Pp. 305–314.
230. *Cogoi S., Xodo L. E.* G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription // Nucleic acids research. — 2006. — Vol. 34, no. 9. — Pp. 2536–2549.
231. *Rankin S.* [et al.] Putative DNA quadruplex formation within the human c-kit oncogene // Journal of the American Chemical Society. — 2005. — Aug. — Vol. 127, no. 30. — Pp. 10584–10589.
232. *Dexheimer T. S., Sun D., Hurley L. H.* Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter // Journal of the American Chemical Society. — 2006. — Apr. — Vol. 128, no. 16. — Pp. 5404–5415.
233. *Siddiqui-Jain A.* [et al.] Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription // Proceedings of the National Academy of Sciences of the United States of America. — 2002. — Sept. — Vol. 99, no. 18. — Pp. 11593–11598.
234. *Hurley L. H.* [et al.] Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element // Seminars in oncology. — 2006. — Aug. — Vol. 33, no. 4. — Pp. 498–512.
235. *Nimjee S. M., Rusconi C. P., Sullenger B. A.* Aptamers: an emerging class of therapeutics // Annual review of medicine. — 2005. — Vol. 56. — Pp. 555–583.
236. *Shamah S. M., Healy J. M., Cload S. T.* Complex target SELEX // Accounts of chemical research. — 2008. — Jan. — Vol. 41, no. 1. — Pp. 130–138.

237. *Bock L. C.* [et al.] Selection of single-stranded DNA molecules that bind and inhibit human thrombin // *Nature*. — 1992. — Feb. — Vol. 355, no. 6360. — Pp. 564–566.
238. *Jing N.* [et al.] Targeting signal transducer and activator of transcription 3 with G-quartet oligonucleotides: a potential novel therapy for head and neck cancer // *Molecular cancer therapeutics*. — 2006. — Feb. — Vol. 5, no. 2. — Pp. 279–286.
239. *McMicken H. W., Bates P. J., Chen Y.* Antiproliferative activity of G-quartet-containing oligonucleotides generated by a novel single-stranded DNA expression system // *Cancer gene therapy*. — 2003. — Dec. — Vol. 10, no. 12. — Pp. 867–869.
240. *Jayapal P.* [et al.] Structure-activity relationships of a caged thrombin binding DNA aptamer: insight gained from molecular dynamics simulation studies // *Journal of structural biology*. — 2009. — June. — Vol. 166, no. 3. — Pp. 241–250.
241. *Lim K. W.* [et al.] Structure of the human telomere in K<sup>+</sup> solution: a stable basket-type G-quadruplex with only two G-tetrad layers // *Journal of the American Chemical Society*. — 2009. — Apr. — Vol. 131, no. 12. — Pp. 4301–4309.
242. *Wang Y., Patel D. J.* Solution structure of the Tetrahymena telomeric repeat d(T2G4)<sub>4</sub> G-tetraplex // *Structure* (London, England: 1993). — 1994. — Dec. — Vol. 2, no. 12. — Pp. 1141–1156.
243. *Parkinson G. N., Lee M. P. H., Neidle S.* Crystal structure of parallel quadruplexes from human telomeric DNA // *Nature*. — 2002. — June. — Vol. 417, no. 6891. — Pp. 876–880.

244. *Haider S., Parkinson G. N., Neidle S.* Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex // *Journal of molecular biology*. — 2002. — July. — Vol. 320, no. 2. — Pp. 189–200.
245. *Horvath M. P., Schultz S. C.* DNA G-quartets in a 1.86 Å resolution structure of an *Oxytricha nova* telomeric protein-DNA complex // *Journal of molecular biology*. — 2001. — July. — Vol. 310, no. 2. — Pp. 367–377.
246. *Phillips K.* [et al.] The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution // *Journal of molecular biology*. — 1997. — Oct. — Vol. 273, no. 1. — Pp. 171–182.
247. *Hud N. V.* [et al.] The selectivity for K<sup>+</sup> versus Na<sup>+</sup> in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by <sup>1</sup>H NMR // *Biochemistry*. — 1996. — Dec. — Vol. 35, no. 48. — Pp. 15383–15390.
248. *Kankia B. I., Marky L. A.* Folding of the thrombin aptamer into a G-quadruplex with Sr(2<sup>+</sup>): stability, heat, and hydration // *Journal of the American Chemical Society*. — 2001. — Nov. — Vol. 123, no. 44. — Pp. 10799–10804.
249. *Deng J., Xiong Y., Sundaralingam M.* X-ray analysis of an RNA tetraplex (UGGGGU)<sub>4</sub> with divalent Sr(2<sup>+</sup>) ions at subatomic resolution (0.61 Å) // *Proceedings of the National Academy of Sciences of the United States of America*. — 2001. — Nov. — Vol. 98, no. 24. — Pp. 13665–13670.
250. *Hassan M. A. el, Calladine C. R.* The assessment of the geometry of dinucleotide steps in double-helical DNA; a new local calculation scheme // *Journal of molecular biology*. — 1995. — Sept. — Vol. 251, no. 5. — Pp. 648–664.

251. *Patel P. K., Koti A. S., Hosur R. V.* NMR studies on truncated sequences of human telomeric DNA: observation of a novel A-tetrad // *Nucleic acids research*. — 1999. — Oct. — Vol. 27, no. 19. — Pp. 3836–3843.
252. *Gavathiotis E., Searle M. S.* Structure of the parallel-stranded DNA quadruplex d(TTAGGGT)<sub>4</sub> containing the human telomeric repeat: evidence for A-tetrad formation from NMR and molecular dynamics simulations // *Organic & biomolecular chemistry*. — 2003. — May. — Vol. 1, no. 10. — Pp. 1650–1656.
253. *Gavathiotis E.* [et al.] Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d(TTAGGGT)<sub>4</sub> containing the human telomeric repeat // *Journal of molecular biology*. — 2003. — Nov. — Vol. 334, no. 1. — Pp. 25–36.
254. *Clark G. R.* [et al.] Structure of the first parallel DNA quadruplex-drug complex // *Journal of the American Chemical Society*. — 2003. — Apr. — Vol. 125, no. 14. — Pp. 4066–4067.
255. *Wang Y., Patel D. J.* Guanine residues in d(T<sub>2</sub>AG<sub>3</sub>) and d(T<sub>2</sub>G<sub>4</sub>) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution // *Biochemistry*. — 1992. — Sept. — Vol. 31, no. 35. — Pp. 8112–8119.
256. *Wang Y., Patel D. J.* Guanine residues in d(T<sub>2</sub>AG<sub>3</sub>) and d(T<sub>2</sub>G<sub>4</sub>) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution // *Biochemistry*. — 1992. — Sept. — Vol. 31, no. 35. — Pp. 8112–8119.

257. *Patel P. K., Hosur R. V.* NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by *Saccharomyces cerevisiae* telomere repeats // *Nucleic acids research*. — 1999. — June. — Vol. 27, no. 12. — Pp. 2457–2464.
258. *Patel P. K., Bhavesh N. S., Hosur R. V.* NMR observation of a novel C-tetrad in the structure of the SV40 repeat sequence GGGCGG // *Biochemical and biophysical research communications*. — 2000. — Apr. — Vol. 270, no. 3. — Pp. 967–971.
259. *Phan A. T.* [et al.] Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K<sup>+</sup> solution // *Nucleic acids research*. — 2007. — Vol. 35, no. 19. — Pp. 6517–6525.
260. *Wang Y., Patel D. J.* Solution structure of the *Tetrahymena* telomeric repeat d(T2G4)<sub>4</sub> G-tetraplex // *Structure* (London, England: 1993). — 1994. — Dec. — Vol. 2, no. 12. — Pp. 1141–1156.
261. *Luu K. N.* [et al.] Structure of the human telomere in K<sup>+</sup> solution: an intramolecular (3 + 1) G-quadruplex scaffold // *Journal of the American Chemical Society*. — 2006. — Aug. — Vol. 128, no. 30. — Pp. 9963–9970.
262. *Dai J.* [et al.] Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation // *Nucleic acids research*. — 2007. — Vol. 35, no. 7. — Pp. 2440–2450.
263. *Dai J.* [et al.] Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K<sup>+</sup> solution: insights into structure polymorphism of the human telomeric sequence // *Nucleic acids research*. — 2007. — Vol. 35, no. 15. — Pp. 4927–4940.



264. *Dai J.* [et al.] NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region // *Nucleic acids research*. — 2006. — Vol. 34, no. 18. — Pp. 5133–5144.
265. *Schultze P., Macaya R. F., Feigon J.* Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG) // *Journal of molecular biology*. — 1994. — Feb. — Vol. 235, no. 5. — Pp. 1532–1547.
266. *Marathias V. M., Bolton P. H.* Structures of the potassium-saturated, 2:1, and intermediate, 1:1, forms of a quadruplex DNA // *Nucleic acids research*. — 2000. — May. — Vol. 28, no. 9. — Pp. 1969–1977.
267. *Marathias V. M.* [et al.] Determination of the number and location of the manganese binding sites of DNA quadruplexes in solution by EPR and NMR in the presence and absence of thrombin // *Journal of molecular biology*. — 1996. — July. — Vol. 260, no. 3. — Pp. 378–394.
268. *Mao X., Marky L. A., Gmeiner W. H.* NMR structure of the thrombin-binding DNA aptamer stabilized by Sr<sup>2+</sup> // *Journal of biomolecular structure & dynamics*. — 2004. — Aug. — Vol. 22, no. 1. — Pp. 25–33.
269. *Lim K. W.* [et al.] Sequence variant (CTAGGG)<sub>n</sub> in the human telomere favors a G-quadruplex structure containing a G.C.G.C tetrad // *Nucleic acids research*. — 2009. — Oct. — Vol. 37, no. 18. — Pp. 6239–6248.
270. *Lim K. W.* [et al.] Structure of the human telomere in K<sup>+</sup> solution: a stable basket-type G-quadruplex with only two G-tetrad layers // *Journal of the American Chemical Society*. — 2009. — Apr. — Vol. 131, no. 12. — Pp. 4301–4309.
271. *Zhang Z.* [et al.] Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K<sup>+</sup> solution: insights into the

- interconversion of human telomeric G-quadruplex structures // *Nucleic acids research*. — 2010. — Jan. — Vol. 38, no. 3. — Pp. 1009–1021.
272. *Hu L.* [et al.] Giardia telomeric sequence d(TAGGG)<sub>4</sub> forms two intramolecular G-quadruplexes in K<sup>+</sup> solution: effect of loop length and sequence on the folding topology // *Journal of the American Chemical Society*. — 2009. — Nov. — Vol. 131, no. 46. — Pp. 16824–16831.
273. *Wang Y., Patel D. J.* Solution structure of the human telomeric repeat d[AG<sub>3</sub>(T<sub>2</sub>AG<sub>3</sub>)<sub>3</sub>] G-tetraplex // *Structure* (London, England: 1993). — 1993. — Dec. — Vol. 1, no. 4. — Pp. 263–282.
274. *Smith F. W., Feigon J.* Quadruplex structure of Oxytricha telomeric DNA oligonucleotides // *Nature*. — 1992. — Mar. — Vol. 356, no. 6365. — Pp. 164–168.
275. *Wang Y., Patel D. J.* Solution structure of the Oxytricha telomeric repeat d[G<sub>4</sub>(T<sub>4</sub>G<sub>4</sub>)<sub>3</sub>] G-tetraplex // *Journal of molecular biology*. — 1995. — Aug. — Vol. 251, no. 1. — Pp. 76–94.
276. *Parkinson G. N., Lee M. P. H., Neidle S.* Crystal structure of parallel quadruplexes from human telomeric DNA // *Nature*. — 2002. — June. — Vol. 417, no. 6891. — Pp. 876–880.
277. *Parkinson G. N., Cuenca F., Neidle S.* Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes // *Journal of molecular biology*. — 2008. — Sept. — Vol. 381, no. 5. — Pp. 1145–1156.
278. *Ambrus A.* [et al.] Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization // *Biochemistry*. — 2005. — Feb. — Vol. 44, no. 6. — Pp. 2048–2058.

279. *Phan A. T.* [et al.] Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter // *Nature chemical biology*. — 2005. — Aug. — Vol. 1, no. 3. — Pp. 167–173.
280. *Hsu S.-T. D.* [et al.] A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics // *Journal of the American Chemical Society*. — 2009. — Sept. — Vol. 131, no. 37. — Pp. 13399–13409.
281. *Kuryavyi V., Phan A. T., Patel D. J.* Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter // *Nucleic acids research*. — 2010. — Oct. — Vol. 38, no. 19. — Pp. 6757–6773.
282. *Matsugami A.* [et al.] An intramolecular quadruplex of (GGA)(4) triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction // *Journal of molecular biology*. — 2001. — Oct. — Vol. 313, no. 2. — Pp. 255–269.
283. *Phan A. T.* [et al.] An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase // *Proceedings of the National Academy of Sciences of the United States of America*. — 2005. — Jan. — Vol. 102, no. 3. — Pp. 634–639.
284. *Kettani A.* [et al.] Solution structure of a Na cation stabilized DNA quadruplex containing G-G-G-G and G-C-G-C tetrads formed by G-G-G-C repeats observed in adeno-associated viral DNA // *Journal of molecular biology*. — 1998. — Sept. — Vol. 282, no. 3. — Pp. 619–636.
285. *Bouaziz S., Kettani A., Patel D. J.* A K cation-induced conformational switch within a loop spanning segment of a DNA quadruplex containing G-G-G-C

- repeats // *Journal of molecular biology*. — 1998. — Sept. — Vol. 282, no. 3. — Pp. 637–652.
286. *Kettani A.* [et al.] A two-stranded template-based approach to G.(C-A) triad formation: designing novel structural elements into an existing DNA framework // *Journal of molecular biology*. — 2000. — Aug. — Vol. 301, no. 1. — Pp. 129–146.
287. *Smith F. W., Feigon J.* Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides // *Nature*. — 1992. — Mar. — Vol. 356, no. 6365. — Pp. 164–168.
288. *Smith F. W., Feigon J.* Strand orientation in the DNA quadruplex formed from the *Oxytricha* telomere repeat oligonucleotide d(G4T4G4) in solution // *Biochemistry*. — 1993. — Aug. — Vol. 32, no. 33. — Pp. 8682–8692.
289. *Haider S. M., Parkinson G. N., Neidle S.* Structure of a G-quadruplex-ligand complex // *Journal of molecular biology*. — 2003. — Feb. — Vol. 326, no. 1. — Pp. 117–125.
290. *Campbell N. H.* [et al.] Selectivity in ligand recognition of G-quadruplex loops // *Biochemistry*. — 2009. — Mar. — Vol. 48, no. 8. — Pp. 1675–1680.
291. *Gill M. L., Strobel S. A., Loria J. P.* 205Tl NMR methods for the characterization of monovalent cation binding to nucleic acids // *Journal of the American Chemical Society*. — 2005. — Nov. — Vol. 127, no. 47. — Pp. 16723–16732.
292. *Gill M. L., Strobel S. A., Loria J. P.* Crystallization and characterization of the thallium form of the *Oxytricha nova* G-quadruplex // *Nucleic acids research*. — 2006. — Vol. 34, no. 16. — Pp. 4506–4514.

293. *Balkwill G. D.* [et al.] Folding topology of a bimolecular DNA quadruplex containing a stable mini-hairpin motif within the diagonal loop // *Journal of molecular biology.* — 2009. — Feb. — Vol. 385, no. 5. — Pp. 1600–1615.
294. *Sket P., Crnugelj M., Plavec J.* d(G3T4G4) forms unusual dimeric G-quadruplex structure with the same general fold in the presence of K<sup>+</sup>, Na<sup>+</sup> or NH<sub>4</sub><sup>+</sup> ions // *Bioorganic & medicinal chemistry.* — 2004. — Nov. — Vol. 12, no. 22. — Pp. 5735–5744.
295. *Crnugelj M., Hud N. V., Plavec J.* The solution structure of d(G(4)T(4)G(3))(2): a bimolecular G-quadruplex with a novel fold // *Journal of molecular biology.* — 2002. — July. — Vol. 320, no. 5. — Pp. 911–924.
296. *Keniry M. A.* [et al.] Solution structure of the Na<sup>+</sup> form of the dimeric guanine quadruplex [d(G3T4G3)]<sub>2</sub> // *European journal of biochemistry / FEBS.* — 1995. — Oct. — Vol. 233, no. 2. — Pp. 631–643.
297. *Parkinson G. N., Ghosh R., Neidle S.* Structural basis for binding of porphyrin to human telomeres // *Biochemistry.* — 2007. — Mar. — Vol. 46, no. 9. — Pp. 2390–2397.
298. *Campbell N. H.* [et al.] Structural basis of DNA quadruplex recognition by an acridine drug // *Journal of the American Chemical Society.* — 2008. — May. — Vol. 130, no. 21. — Pp. 6722–6724.
299. *Kuryavyi V., Phan A. T., Patel D. J.* Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter // *Nucleic acids research.* — 2010. — Oct. — Vol. 38, no. 19. — Pp. 6757–6773.

300. *Webba da Silva M.* Association of DNA quadruplexes through G:C:G:C tetrads. Solution structure of d(GCGGTGGAT) // *Biochemistry*. — 2003. — Dec. — Vol. 42, no. 49. — Pp. 14356–14365.
301. *Kettani A.* [et al.] A dimeric DNA interface stabilized by stacked A.(G.G.G.G).A hexads and coordinated monovalent cations // *Journal of molecular biology*. — 2000. — Mar. — Vol. 297, no. 3. — Pp. 627–644.
302. *Webba da Silva M.* Experimental demonstration of T:(G:G:G:G):T hexad and T:A:A:T tetrad alignments within a DNA quadruplex stem // *Biochemistry*. — 2005. — Mar. — Vol. 44, no. 10. — Pp. 3754–3764.
303. *Tuerk C., Gold L.* Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase // *Science (New York, N.Y.)* — 1990. — Aug. — Vol. 249, no. 4968. — Pp. 505–510.
304. *Ikebukuro K.* [et al.] A novel method of screening thrombin-inhibiting DNA aptamers using an evolution-mimicking algorithm // *Nucleic Acids Research*. — 2005. — Jan. — Vol. 33, no. 12. — e108–e108.
305. *Schultze P., Macaya R. F., Feigon J.* Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG) // *Journal of molecular biology*. — 1994. — Feb. — Vol. 235, no. 5. — Pp. 1532–1547.
306. *Padmanabhan K.* [et al.] The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer // *The Journal of biological chemistry*. — 1993. — Aug. — Vol. 268, no. 24. — Pp. 17651–17654.
307. *Kelly J. A., Feigon J., Yeates T. O.* Reconciliation of the X-ray and NMR structures of the thrombin-binding aptamer d(GGTTGGTGTGGTTGG) // *Journal of molecular biology*. — 1996. — Mar. — Vol. 256, no. 3. — Pp. 417–422.

308. *Tasset D. M., Kubik M. F., Steiner W.* Oligonucleotide inhibitors of human thrombin that bind distinct epitopes // *Journal of molecular biology.* — 1997. — Oct. — Vol. 272, no. 5. — Pp. 688–698.
309. *Pagano B.* [et al.] Stability and binding properties of a modified thrombin binding aptamer // *Biophysical journal.* — 2008. — Jan. — Vol. 94, no. 2. — Pp. 562–569.
310. *Fadrní E.* [et al.] Molecular dynamics simulations of Guanine quadruplex loops: advances and force field limitations // *Biophysical journal.* — 2004. — July. — Vol. 87, no. 1. — Pp. 227–242.
311. *Sponer J., Spacková N.* Molecular dynamics simulations and their application to four-stranded DNA // *Methods (San Diego, Calif.)* — 2007. — Dec. — Vol. 43, no. 4. — Pp. 278–290.
312. *Fadrní E.* [et al.] Single Stranded Loops of Quadruplex DNA As Key Benchmark for Testing Nucleic Acids Force Fields // *Journal of Chemical Theory and Computation.* — 2009. — Sept. — Vol. 5, no. 9. — Pp. 2514–2530.
313. *Haider S., Neidle S.* Molecular modeling and simulation of G-quadruplexes and quadruplex-ligand complexes // *Methods in molecular biology (Clifton, N.J.)* — 2010. — Vol. 608. — Pp. 17–37.
314. *Hazel P., Parkinson G. N., Neidle S.* Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures // *Nucleic acids research.* — 2006. — Vol. 34, no. 7. — Pp. 2117–2127.
315. *Cavallari M.* [et al.] Stability and migration of metal ions in G4-wires by molecular dynamics simulations // *The journal of physical chemistry. B.* — 2006. — Dec. — Vol. 110, no. 51. — Pp. 26337–26348.

316. *Cheatham T E 3., Cieplak P., Kollman P. A.* A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat // *Journal of biomolecular structure & dynamics*. — 1999. — Feb. — Vol. 16, no. 4. — Pp. 845–862.
317. *Wang J., Cieplak P., Kollman P. A.* How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? // *Journal of Computational Chemistry*. — 2000. — Sept. — Vol. 21, no. 12. — Pp. 1049–1074.
318. *Pérez A.* [et al.] Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers // *Biophysical journal*. — 2007. — June. — Vol. 92, no. 11. — Pp. 3817–3829.
319. *Pérez A., Luque F. J., Orozco M.* Dynamics of B-DNA on the microsecond time scale // *Journal of the American Chemical Society*. — 2007. — Nov. — Vol. 129, no. 47. — Pp. 14739–14745.
320. *Byrd R. H.* [et al.] A Limited Memory Algorithm for Bound Constrained Optimization // *SIAM Journal on Scientific Computing*. — 1995. — Sept. — Vol. 16, no. 5. — Pp. 1190–1208.
321. *Ahmed H. U.* [et al.] The determination of protonation states in proteins // *Acta crystallographica. Section D, Biological crystallography*. — 2007. — Aug. — Vol. 63, Pt 8. — Pp. 906–922.
322. *Stefl R.* [et al.] Formation pathways of a guanine-quadruplex DNA revealed by molecular dynamics and thermodynamic analysis of the substates // *Biophysical journal*. — 2003. — Sept. — Vol. 85, no. 3. — Pp. 1787–1804.
323. *Russo Krauss I.* [et al.] High-resolution structures of two complexes between thrombin and thrombin-binding aptamer shed light on the role of cations in



- the aptamer inhibitory activity // *Nucleic acids research*. — 2012. — Sept. — Vol. 40, no. 16. — Pp. 8119–8128.
324. *Hud N. V.* [et al.] Binding sites and dynamics of ammonium ions in a telomere repeat DNA quadruplex // *Journal of molecular biology*. — 1999. — Jan. — Vol. 285, no. 1. — Pp. 233–243.
325. *Vairamani M., Gross M. L.* G-quadruplex formation of thrombin-binding aptamer detected by electrospray ionization mass spectrometry // *Journal of the American Chemical Society*. — 2003. — Jan. — Vol. 125, no. 1. — Pp. 42–43.
326. *Majhi P. R.* [et al.] Heat capacity changes associated with guanine quadruplex formation: an isothermal titration calorimetry study // *Biopolymers*. — 2008. — Apr. — Vol. 89, no. 4. — Pp. 302–309.
327. *Car, Parrinello* Unified approach for molecular dynamics and density-functional theory // *Physical review letters*. — 1985. — Nov. — Vol. 55, no. 22. — Pp. 2471–2474.
328. *Eichinger M.* [et al.] A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields // *The Journal of Chemical Physics*. — 1999. — Vol. 110, no. 21. — P. 10452.
329. *Smirnov I., Shafer R. H.* Effect of loop sequence and size on DNA aptamer stability // *Biochemistry*. — 2000. — Feb. — Vol. 39, no. 6. — Pp. 1462–1468.
330. Aptamers specific for biomolecules and methods of making.
331. *Baldrich E., O'Sullivan C. K.* Ability of thrombin to act as molecular chaperone, inducing formation of quadruplex structure of thrombin-binding ap-

- tamer // Analytical biochemistry. — 2005. — June. — Vol. 341, no. 1. — Pp. 194–197.
332. *Tsiang M.* [et al.] Functional mapping of the surface residues of human thrombin // The Journal of biological chemistry. — 1995. — July. — Vol. 270, no. 28. — Pp. 16854–16863.
333. *Hazel P.* [et al.] Loop-length-dependent folding of G-quadruplexes. // J Am Chem Soc. — 2004. — Dec. — Vol. 126, no. 50. — Pp. 16405–15.
334. *Smargiasso N.* [et al.] G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation. // J Am Chem Soc. — 2008. — Aug. — Vol. 130, no. 31. — Pp. 10208–16.
335. *Gaudin A.* [et al.] How long is too long? Effects of loop size on G-quadruplex stability. // Nucleic Acids Res. — 2010. — Nov. — Vol. 38, no. 21. — Pp. 7858–68.
336. *Risitano A., Fox K.* Influence of loop size on the stability of intramolecular DNA quadruplexes. // Nucleic Acids Res. — 2004None. — Vol. 32, no. 8. — Pp. 2598–606.
337. *Phillips K.* [et al.] The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution. // J Mol Biol. — 1997. — Oct. — Vol. 273, no. 1. — Pp. 171–82.
338. *Ida R., Wu G.* Direct NMR detection of alkali metal ions bound to G-quadruplex DNA. // J Am Chem Soc. — 2008. — Mar. — Vol. 130, no. 11. — Pp. 3590–602.

339. *Hazel P., Parkinson G., Neidle S.* Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures. // *Nucleic Acids Res.* — 2006None. — Vol. 34, no. 7. — Pp. 2117–27.
340. *Rueda M., Luque F., Orozco M.* G-quadruplexes can maintain their structure in the gas phase. // *J Am Chem Soc.* — 2006. — Mar. — Vol. 128, no. 11. — Pp. 3608–19.
341. *Li H., Cao E., Gisler T.* Force-induced unfolding of human telomeric G-quadruplex: a steered molecular dynamics simulation study. // *Biochem Biophys Res Commun.* — 2009. — Jan. — Vol. 379, no. 1. — Pp. 70–5.
342. *Pagano B.* [et al.] Stability and cations coordination of DNA and RNA 14-mer G-quadruplexes: a multiscale computational approach. // *J Phys Chem B.* — 2008. — Sept. — Vol. 112, no. 38. — Pp. 12115–23.
343. *Petraccone L.* [et al.] An integrated molecular dynamics (MD) and experimental study of higher order human telomeric quadruplexes. // *Biopolymers.* — 2010. — June. — Vol. 93, no. 6. — Pp. 533–48.
344. *Spacková N., Berger I., Sponer J.* Structural dynamics and cation interactions of DNA quadruplex molecules containing mixed guanine/cytosine quartets revealed by large-scale MD simulations. // *J Am Chem Soc.* — 2001. — Apr. — Vol. 123, no. 14. — Pp. 3295–307.
345. *Cavallari M.* [et al.] Stability and migration of metal ions in G4-wires by molecular dynamics simulations. // *J Phys Chem B.* — 2006. — Dec. — Vol. 110, no. 51. — Pp. 26337–48.
346. *Cavallari M., Garbesi A., Di Felice R.* Porphyrin intercalation in G4-DNA quadruplexes by molecular dynamics simulations. // *J Phys Chem B.* — 2009. — Oct. — Vol. 113, no. 40. — Pp. 13152–60.

347. *Fadrní E.* [et al.] Single Stranded Loops of Quadruplex DNA As Key Benchmark for Testing Nucleic Acids Force Fields // *Journal of Chemical Theory and Computation*. — 2009. — Vol. 5, no. 9. — Pp. 2514–2530. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct900200k>.
348. *Mao X., Marky L., Gmeiner W.* NMR structure of the thrombin-binding DNA aptamer stabilized by Sr<sup>2+</sup>. // *J Biomol Struct Dyn*. — 2004. — Aug. — Vol. 22, no. 1. — Pp. 25–33.
349. *Marathias V., Bolton P.* Determinants of DNA quadruplex structural type: sequence and potassium binding. // *Biochemistry*. — 1999. — Apr. — Vol. 38, no. 14. — Pp. 4355–64.
350. *Marathias V., Bolton P.* Structures of the potassium-saturated, 2:1, and intermediate, 1:1, forms of a quadruplex DNA. // *Nucleic Acids Res*. — 2000. — May. — Vol. 28, no. 9. — Pp. 1969–77.
351. *Trajkovski M., Sket P., Plavec J.* Cation localization and movement within DNA thrombin binding aptamer in solution. // *Org Biomol Chem*. — 2009. — Nov. — Vol. 7, no. 22. — Pp. 4677–84.
352. *Jing N.* [et al.] Mechanism of inhibition of HIV-1 integrase by G-tetrad-forming oligonucleotides in Vitro. // *J Biol Chem*. — 2000. — July. — Vol. 275, no. 28. — Pp. 21460–7.
353. *Schultze P., Macaya R., Feigon J.* Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG). // *J Mol Biol*. — 1994. — Feb. — Vol. 235, no. 5. — Pp. 1532–47.
354. *Cordomi A., Edholm O., Perez J. J.* Effect of Force Field Parameters on Sodium and Potassium Ion Binding to Dipalmitoyl Phosphatidylcholine Bilayers // *Journal of Chemical Theory and Computation*. — 2009. — Vol. 5,

- no. 8. — Pp. 2125–2134. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct9000763>.
355. *Van Der Spoel D.* [et al.] GROMACS: fast, flexible, and free. // *J Comput Chem.* — 2005. — Dec. — Vol. 26, no. 16. — Pp. 1701–18.
356. *Payne M. C.* [et al.] Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients // *Rev. Mod. Phys.* — 1992. — Oct. — Vol. 64, issue 4. — Pp. 1045–1097.
357. *Pant M., Rajagopal A.* Theory of inhomogeneous magnetic electron gas // *Solid State Communications.* — 1972. — Vol. 10, no. 12. — Pp. 1157–1160.
358. *Perdew J.* [et al.] Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. // *Phys Rev B Condens Matter.* — 1992. — Sept. — Vol. 46, no. 11. — Pp. 6671–6687.
359. *Vanderbilt D.* Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. // *Phys Rev B Condens Matter.* — 1990. — Apr. — Vol. 41, no. 11. — Pp. 7892–7895.
360. *Grimme S.* Accurate description of van der Waals complexes by density functional theory including empirical corrections. // *J Comput Chem.* — 2004. — Sept. — Vol. 25, no. 12. — Pp. 1463–73.
361. *Biswas P., Gogonea V.* A regularized and renormalized electrostatic coupling Hamiltonian for hybrid quantum-mechanical-molecular-mechanical calculations. // *J Chem Phys.* — 2005. — Oct. — Vol. 123, no. 16. — P. 164114.
362. *Hoover W.* Canonical dynamics: Equilibrium phase-space distributions. // *Phys Rev A.* — 1985. — Mar. — Vol. 31, no. 3. — Pp. 1695–1697.

363. Špačková N., Berger I., Šponer J. Nanosecond Molecular Dynamics Simulations of Parallel and Antiparallel Guanine Quadruplex DNA Molecules // Journal of the American Chemical Society. — 1999. — Vol. 121, no. 23. — Pp. 5519–5534. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja984449s>.
364. Fadrná E. [et al.] Molecular dynamics simulations of Guanine quadruplex loops: advances and force field limitations. // Biophys J. — 2004. — July. — Vol. 87, no. 1. — Pp. 227–42.
365. Zacharias M. Simulation of the structure and dynamics of nonhelical RNA motifs. // Curr Opin Struct Biol. — 2000. — June. — Vol. 10, no. 3. — Pp. 311–7.
366. Faralli C. [et al.] The solvation dynamics of Na<sup>+</sup> and K<sup>+</sup> ions in liquid methanol // Theoretical Chemistry Accounts. — 2007. — Vol. 118, no. 2. — Pp. 417–423.
367. Hancock R. D. [et al.] The structure of the 11-coordinate barium complex of the pendant-donor macrocycle 1,4,7,10-tetrakis(carbamoylmethyl)-1,4,7,10-tetraazacyclododecane: an analysis of the coordination numbers of barium(II) in its complexes // Inorganica Chimica Acta. — 2004. — Vol. 357, no. 3. — Pp. 723–727.
368. Reshetnikov R. [et al.] Structural Dynamics of Thrombin-Binding DNA Aptamer d(GGTTGGTGTGGTTGG) Quadruplex DNA Studied by Large-Scale Explicit Solvent Simulations // Journal of Chemical Theory and Computation. — 2010. — Vol. 6, no. 10. — Pp. 3003–3014. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ct100253m>.

369. *Hong E.* [et al.] Mass spectrometric studies of alkali metal ion binding on thrombin-binding aptamer DNA. // *J Am Soc Mass Spectrom.* — 2010. — July. — Vol. 21, no. 7. — Pp. 1245–55.
370. *Shim J., Tan Q., Gu L.* Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity. // *Nucleic Acids Res.* — 2009. — Feb. — Vol. 37, no. 3. — Pp. 972–82.
371. *Vairamani M., Gross M. L.* G-Quadruplex Formation of Thrombin-Binding Aptamer Detected by Electrospray Ionization Mass Spectrometry // *Journal of the American Chemical Society.* — 2003. — Vol. 125, no. 1. — Pp. 42–43. — eprint: <http://pubs.acs.org/doi/pdf/10.1021/ja0284299>.
372. *Turjanski A., Hummer G., Gutkind J.* How mitogen-activated protein kinases recognize and phosphorylate their targets: A QM/MM study. // *J Am Chem Soc.* — 2009. — May. — Vol. 131, no. 17. — Pp. 6141–8.
373. *Kankia B., Marky L.* Folding of the thrombin aptamer into a G-quadruplex with Sr(2+): stability, heat, and hydration. // *J Am Chem Soc.* — 2001. — Nov. — Vol. 123, no. 44. — Pp. 10799–804.
374. *Nagatoishi S., Tanaka Y., Tsumoto K.* Circular dichroism spectra demonstrate formation of the thrombin-binding DNA aptamer G-quadruplex under stabilizing-cation-deficient conditions. // *Biochem Biophys Res Commun.* — 2007. — Jan. — Vol. 352, no. 3. — Pp. 812–7.
375. *Stadlbauer P.* [et al.] Structural dynamics of possible late-stage intermediates in folding of quadruplex DNA studied by molecular simulations. // *Nucleic Acids Res.* — 2013. — Aug. — Vol. 41, no. 14. — Pp. 7128–43.

376. *Dolinnaya N.* [et al.] Coexistence of G-quadruplex and duplex domains within the secondary structure of 31-mer DNA thrombin-binding aptamer. // *J Biomol Struct Dyn.* — 2012. — June. — Vol. 30, no. 5. — Pp. 524–31.
377. *Ikebukuro K.* [et al.] Novel strategy for DNA aptamers inhibiting enzymatic activity using algorithm mimicking evolution. // *Nucleic Acids Res Suppl.* — 2003. — July. — No. 3. — Pp. 205–6.
378. *Yarotski D.* [et al.] Scanning tunneling microscopy of DNA-wrapped carbon nanotubes. // *Nano Lett.* — 2009. — Jan. — Vol. 9, no. 1. — Pp. 12–7.
379. *Alexeev Y.* [et al.] GAMESS as a free quantum-mechanical platform for drug research. // *Curr Top Med Chem.* — 2012None. — Vol. 12, no. 18. — Pp. 2013–33.
380. *Dupradeau F.* [et al.] The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. // *Phys Chem Chem Phys.* — 2010. — July. — Vol. 12, no. 28. — Pp. 7821–39.
381. *Weber J., Pande V.* Characterization and rapid sampling of protein folding Markov state model topologies. // *J Chem Theory Comput.* — 2011. — Oct. — Vol. 7, no. 10. — Pp. 3405–3411.
382. *Schwantes C., Pande V.* Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. // *J Chem Theory Comput.* — 2013. — Apr. — Vol. 9, no. 4. — Pp. 2000–2009.
383. *Sorin E. J.* [et al.] RNA simulations: probing hairpin unfolding and the dynamics of a GNRA tetraloop // *J. Mol. Biol.* — 2002. — Apr. — Vol. 317, no. 4. — Pp. 493–506.



384. *Kannan S., Zacharias M.* Folding of a DNA hairpin loop structure in explicit solvent using replica-exchange molecular dynamics simulations // *Biophys. J.* — 2007. — Nov. — Vol. 93, no. 9. — Pp. 3218–3228.
385. *Kannan S., Zacharias M.* Role of the closing base pair for d(GCA) hairpin stability: free energy analysis and folding simulations // *Nucleic Acids Res.* — 2011. — Oct. — Vol. 39, no. 19. — Pp. 8271–8280.
386. *Kannan S., Zacharias M.* Simulation of DNA double-strand dissociation and formation during replica-exchange molecular dynamics simulations // *Phys Chem Chem Phys.* — 2009. — Dec. — Vol. 11, no. 45. — Pp. 10589–10595.
387. *Roxbury D., Jagota A., Mittal J.* Structural characteristics of oligomeric DNA strands adsorbed onto single-walled carbon nanotubes // *J Phys Chem B.* — 2013. — Jan. — Vol. 117, no. 1. — Pp. 132–140.
388. *Roxbury D., Mittal J., Jagota A.* Molecular-basis of single-walled carbon nanotube recognition by single-stranded DNA // *Nano Lett.* — 2012. — Mar. — Vol. 12, no. 3. — Pp. 1464–1469.
389. *Roxbury D., Jagota A., Mittal J.* Sequence-specific self-stitching motif of short single-stranded DNA on a single-walled carbon nanotube // *J. Am. Chem. Soc.* — 2011. — Aug. — Vol. 133, no. 34. — Pp. 13545–13550.
390. *Curuksu J., Zacharias M.* Enhanced conformational sampling of nucleic acids by a new Hamiltonian replica exchange molecular dynamics approach // *J Chem Phys.* — 2009. — Mar. — Vol. 130, no. 10. — P. 104110.
391. *Hagen M.* [et al.] Serial replica exchange // *J Phys Chem B.* — 2007. — Feb. — Vol. 111, no. 6. — Pp. 1416–1423.

392. *Zhou L.* [et al.] Precise determination, cross-recognition, and functional analysis of the double-strand origins of the rolling-circle replication plasmids in haloarchaea // *J. Bacteriol.* — 2008. — Aug. — Vol. 190, no. 16. — Pp. 5710–5719.
393. *Ruiz-Maso J. A.* [et al.] Interactions between the RepB initiator protein of plasmid pMV158 and two distant DNA regions within the origin of replication // *Nucleic Acids Res.* — 2007. — Vol. 35, no. 4. — Pp. 1230–1244.
394. *Jaishree T. N., Wang A. H.* Human chromosomal centromere (AATGG)<sub>n</sub> sequence forms stable structures with unusual base pairs // *FEBS Lett.* — 1994. — June. — Vol. 347, no. 1. — Pp. 99–103.
395. *Chou S. H., Zhu L., Reid B. R.* On the relative ability of centromeric GNA triplets to form hairpins versus self-paired duplexes // *J. Mol. Biol.* — 1996. — June. — Vol. 259, no. 3. — Pp. 445–457.
396. *Naville M., Gautheret D.* Transcription attenuation in bacteria: theme and variations // *Brief Funct Genomic Proteomic.* — 2009. — Nov. — Vol. 8, no. 6. — Pp. 482–492.
397. *Smith A. M.* [et al.] Riboswitch RNAs: regulation of gene expression by direct monitoring of a physiological signal // *RNA Biol.* — 2010. — Vol. 7, no. 1. — Pp. 104–110.
398. *Mitas M.* Trinucleotide repeats associated with human disease // *Nucleic Acids Res.* — 1997. — June. — Vol. 25, no. 12. — Pp. 2245–2254.
399. *Talini G., Branciamore S., Gallori E.* Ribozymes: Flexible molecular devices at work // *Biochimie.* — 2011. — Nov. — Vol. 93, no. 11. — Pp. 1998–2005.

400. *Yoshizawa S.* [et al.] GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins // *Biochemistry*. — 1997. — Apr. — Vol. 36, no. 16. — Pp. 4761–4767.
401. *Beauchamp K. A.* [et al.] MSMBuild2: Modeling Conformational Dynamics at the Picosecond to Millisecond Scale // *J Chem Theory Comput*. — 2011. — Oct. — Vol. 7, no. 10. — Pp. 3412–3419.
402. *Ma H.* [et al.] DNA folding and melting observed in real time redefine the energy landscape // *Proc. Natl. Acad. Sci. U.S.A.* — 2007. — Jan. — Vol. 104, no. 3. — Pp. 712–716.
403. *Hess B.* [et al.] GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation // *Journal of Chemical Theory and Computation*. — 2008. — Mar. — Vol. 4, no. 3. — Pp. 435–447.
404. *Sorin E. J., Pande V. S.* Exploring the helix-coil transition via all-atom equilibrium ensemble simulations // *Biophysical journal*. — 2005. — Apr. — Vol. 88, no. 4. — Pp. 2472–2493.
405. *Van Der Spoel D.* [et al.] GROMACS: fast, flexible, and free // *Journal of computational chemistry*. — 2005. — Dec. — Vol. 26, no. 16. — Pp. 1701–1718.
406. *Hornak V.* [et al.] Comparison of multiple Amber force fields and development of improved protein backbone parameters // *Proteins*. — 2006. — Nov. — Vol. 65, no. 3. — Pp. 712–725.
407. *Bussi G., Donadio D., Parrinello M.* Canonical sampling through velocity rescaling // *The Journal of Chemical Physics*. — 2007. — Vol. 126, no. 1. — P. 014101.

408. *Berendsen H. J. C.* [et al.] Molecular dynamics with coupling to an external bath // *The Journal of Chemical Physics*. — 1984. — Vol. 81, no. 8. — P. 3684.
409. *Darden T., York D., Pedersen L.* Particle mesh Ewald: An N log(N) method for Ewald sums in large systems // *The Journal of Chemical Physics*. — 1993. — Vol. 98, no. 12. — P. 10089.
410. *Jorgensen W. L.* [et al.] Comparison of simple potential functions for simulating liquid water // *The Journal of Chemical Physics*. — 1983. — Vol. 79, no. 2. — P. 926.
411. *Tsvetkov F., Devred F., Makarov A.* [Thermodynamics of zinc binding to human S100A2]. // *Mol Biol (Mosk)*. — NoneNone. — Vol. 44, no. 5. — Pp. 938–42.
412. *Lee W.* [et al.] PINE-SPARKY: graphical interface for evaluating automated probabilistic peak assignments in protein NMR spectroscopy. // *Bioinformatics*. — 2009. — Aug. — Vol. 25, no. 16. — Pp. 2085–7.
413. *Markley J.* [et al.] BioMagResBank (BMRB) as a partner in the Worldwide Protein Data Bank (wwPDB): new policies affecting biomolecular NMR depositions. // *J Biomol NMR*. — 2008. — Mar. — Vol. 40, no. 3. — Pp. 153–5.
414. *Brnnger A. T.* [et al.] *Crystallography & NMR system*: A new software suite for macromolecular structure determination // *Acta Crystallographica Section D*. — 1998. — Sept. — Vol. 54, no. 5. — Pp. 905–921.
415. *Yang Z.* [et al.] Amino acid analogues bind to carbon nanotube via interactions: comparison of molecular mechanical and quantum mechanical calculations. // *J Chem Phys*. — 2012. — Jan. — Vol. 136, no. 2. — P. 025103.

416. *Vega C., Abascal J., Nezbeda I.* Vapor-liquid equilibria from the triple point up to the critical point for the new generation of TIP4P-like models: TIP4P/Ew, TIP4P/2005, and TIP4P/ice. // *J Chem Phys.* — 2006. — July. — Vol. 125, no. 3. — P. 34503.