

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра биоинженерии

Грибкова Анна Кирилловна

ИССЛЕДОВАНИЕ ЭНЕРГЕТИЧЕСКОГО ПРОФИЛЯ ПРОЦЕССА
ОТВОРАЧИВАНИЯ ДНК ОТ ГИСТОНОВОГО КОРА НУКЛЕОСОМЫ

MODELING OF THE FREE ENERGY PROFILE OF DNA UNWRAPPING FROM
HISTONE OCTAMER

Выпускная квалификационная работа бакалавра

Научный руководитель:

канд. физ.-мат. наук, вед. науч. сотр.

А.К. Шайтан

Соруководитель:

аспирант Г.А. Армеев

Москва - 2017 г.

Список сокращений

МД - молекулярная динамика

ММ - молекулярная механика

н.п. - нуклеотидная пара

ФАСТ - хроматин-ремоделирующий фактор (facilitates chromatin transcription)

FEP - метод возмущения свободной энергии (free energy perturbation)

NAP 1 - шаперон гистонов (nucleosome assembly protein)

spFRET - метод Ферстеровского резонансного переноса энергии от одиночных молекул (single particle Förster resonance energy transfer)

Оглавление

1 Введение	3
2 Обзор литературы	4
2. 1 Нуклеосома	4
2.2 Нуклеосомные интермедиаты	5
2.3 Оценка энергии ДНК-гистоновых взаимодействий	6
2.3.1 Экспериментальные исследования	6
2.3.2 Теоретические подходы	8
2.4 Цель работы	8
2.5 Задачи:	8
3 Материалы и методы	10
3.1 MM/PBSA	10
3.2 Молекулярная динамика	16
3.3 3DNA	17
3.4 Методика работы	19
4 Результаты и обсуждение	22
5 Заключение	26
6 Выводы	27
7 Приложение	28
8 Список литературы	29

1 Введение

Геномная ДНК эукариотических организмов ассоциирована с белками, что приводит не только к компактизации генетического материала, но и к регуляции практически всех жизненно важных процессов посредством ограничения доступа белковым факторам. Взаимодействие ДНК с гистонами играет ключевую роль в различных процессах функционирования хроматина и, следовательно, функционировании живых организмов. Проблема организации хроматина, тем самым, является особо актуальной при рассмотрении социально значимых заболеваний и разработке подходов персонализированной медицины. Также известно, что многие белки, взаимодействующие с гистонами, являются потенциальными мишенями противораковых лекарственных средств.

Эксперименты по исследованию динамики хроматина обычно проводят на моонуклеосомах, однако полностью избежать случайного образования динуклеосом и других нуклеосомных интермедиатов не удастся. К тому же, многие биохимические и биофизические методы подразумевают использование экстремальных термических или химических условий, далеких от физиологических. Поэтому использование теоретических подходов для рассмотрения вопросов об энергии взаимодействия ДНК с гистоновым ядром считается весьма перспективным.

2 Обзор литературы

2. 1 Нуклеосома

Элементарной повторяющейся единицей хроматина является нуклеосомный кор, представляющий собой ~ 145 - 147 н.п. ДНК, обвивающих $1,65$ раз гистоновый октамер. Белковая часть кора имеет вид цилиндра радиусом 65 \AA и высотой 60 \AA и состоит из четырех димеров гистонов: два димера H2A-H2B и два H3-H4. Нуклеосомная ДНК содержит 14 сайтов связывания с гистонами, расположенных в малой бороздке. Преимущественно, эти контакты представлены водородными связями между гистонами и сахарофосфатным остовом нуклеиновой кислоты. Помимо водородных связей, взаимодействие ДНК и белка проявляется в гидрофобных взаимодействиях и солевых мостиках. Концевые части гистонов (N-концы) длиной 15-30 аминокислотных остатков называют гистоновыми хвостами, выступающими за пределы нуклеосомного кора. В них сосредоточены положительно заряженные аминокислоты, преимущественно лизин и аргинин. Несмотря на то, что гистоны являются высококонсервативными белками, внутри каждого семейства (кроме H4) существуют различные варианты, играющие важную роль в регуляции транскрипции, хромосомной сегрегации и репарации ДНК. Гистоны, а чаще всего именно гистоновые хвосты, подвержены посттрансляционным модификациям, в частности, ацетилированию и фосфорилированию, что уменьшает заряд белков и, соответственно, степень взаимодействия с ДНК, приводя к изменению структуры хроматина. Другие модификации влияют на узнавание и связывание ДНК с гистоновыми шаперонами и факторами ремоделирования хроматина.

В гистоновом октамере выделяют участок, называемый «acidic patch», который состоит из 8-ми отрицательно заряженных аминокислотных остатков, расположенных на гистонах H2A и H2B. Предполагается, что с ним

взаимодействует хвост гистона H4 соседней нуклеосомы при формировании хроматиновых фибрилл, а также некоторые белки (LANA, IL-33, RCC1, Sir3 and HMGN2). Нуклеосомный кор с линкерной ДНК длиной приблизительно 50 н.п. формирует нуклеосому. Образованию более плотной укладки ДНК (например, 30-ти нм фибриллы) способствует гистон H1, взаимодействующий с ДНК в местах ее входа и выхода из октамера гистонов.

Несмотря на плотную упаковку, для нуклеосом характерны высокочастотные тепловые движения и крупномасштабные конформационные переходы. Примером такой подвижности является частичное отделение концевых участков ДНК от гистоновой октамеры - “дыхание”, что говорит о меньшей энергии связывания этих участков ДНК. Таким образом, нуклеосома является стабильным, но в то же время динамичным комплексом, обеспечивающим помимо компактизации генетического материала еще и регуляцию экспрессии генов.

2.2 Нуклеосомные интермедиаты

Нуклеосомные интермедиаты, такие как гексасомы (шесть гистонов без димера H2A-H2B) и тетрасомы (четыре гистона) образуются при репликации, транскрипции и репарации ДНК. Было показано появление гексасом при прохождении через нуклеосому РНК-полимеразы II во время транскрипции, при взаимодействии с шапероном гистонов NAP 1, при восстановлении двухцепочечных разрывов (с участием фактора ремоделирования хроматина FACT) (Zlatanova et al., 2009).

Также было показано, что нуклеосомы склонны к частичному разворачиванию (Li et al. 2005). В работе (Koormans et al. 2007) наблюдалось обратимое отделение концевых участков нуклеосомной ДНК с временем жизни около 120 мс на иммобилизованных нуклеосомах; в работе (Gansen et al. 2009, а) наблюдалась зависимость скорости отворачивания ДНК от последовательности ДНК и ацетилирования гистонов; в работе (Gansen et al. 2009, б) при помощи

метода многопараметрической детекции флуоресценции было выявлено, что процесс разворачивания нуклеосом проходит с образованием нескольких конформационных подсостояний.

2.3 Оценка энергии ДНК-гистоновых взаимодействий

2.3.1 Экспериментальные исследования

Существует ряд экспериментальных методов для оценки энергии ДНК-гистоновых взаимодействий. Первые кристаллографические исследования рассматриваемых взаимодействий выявили следующее расположение наиболее сильных контактов: симметричные относительно диадной оси положения ± 5 н.п. и $\pm 36 - \pm 46$ н.п. (Brower-Toland et al., 2002). Большинство экспериментальных методов не позволяют оценить энергию взаимодействия нуклеосомной ДНК с октамером гистонов через измерение равновесных концентраций свободной ДНК и нуклеосом (Thåstro et al., 2004), однако существуют методы для прямого измерения профиля энергии взаимодействия отдельных участков ДНК с гистонами. —Одними из наиболее интересных примеров являются эксперименты по силовому растягиванию одиночных нуклеосом при помощи оптического пинцета (Mihardja et al., 2006).

Экспериментально наблюдаемые ступенчатые удлинения нуклеиновой кислоты при растягивании описываются трехстадийной моделью (Brower-Toland et al., 2002). Согласно этой модели, в начале воздействия внешний виток ДНК может быть обратимо отделен, дальнейшее удлинение требует преодоления энергетического барьера и приводит к необратимому скручиванию внутреннего витка ДНК. На последней стадии октамер полностью отделяется от ДНК. Тем не менее, в рассматриваемой модели считается, что гистоновый кор - довольно стабильная структура, которая сохраняется на всем протяжении процесса растяжения ДНК. Однако исследования, выполненные методом Ферстеровского резонансного переноса энергии от одиночных

молекул (spFRET), продемонстрировали мобильность гистонового кора, способного разбираться с потерей димера H2A-H2B и даже тетрамера (H3-H4)₂ (Zhang et al., 2016). Помимо прочего, в экспериментах, совмещающих в себе spFRET и оптический пинцет, было показано, что гибкость ДНК зависит от нуклеотидной последовательности (Zhang et al., 2016). Началу отворачивания нуклеиновой кислоты мешает ее взаимодействие с N-концом (H3) и C-концом (H2A), в то время как N-концы (H2A, H2B, H4) препятствуют процессу отворачивания на поздних стадиях. Так, в исследованиях методом молекулярной динамики, контакты ДНК-гистоны (в случае нуклеосом с гистоновыми хвостами) оказались прочнее, чем межгистоновые (Ettig et al., 2011).

В проделанной нами работе для сравнения теоретически предсказанных профилей энергии использовались экспериментальные данные (Forties et al., 2011). Эти экспериментальные данные были получены следующим образом: к диэлектрической бусине, находящейся в оптической ловушке (Ashkin A. et al., 1989), прикладывалась постоянная механическая сила (28 pN). Одна цепь ДНК была связана с бусиной, а другой конец закреплен на поверхности. Схема эксперимента и полученный профиль изображены на рисунке 1.

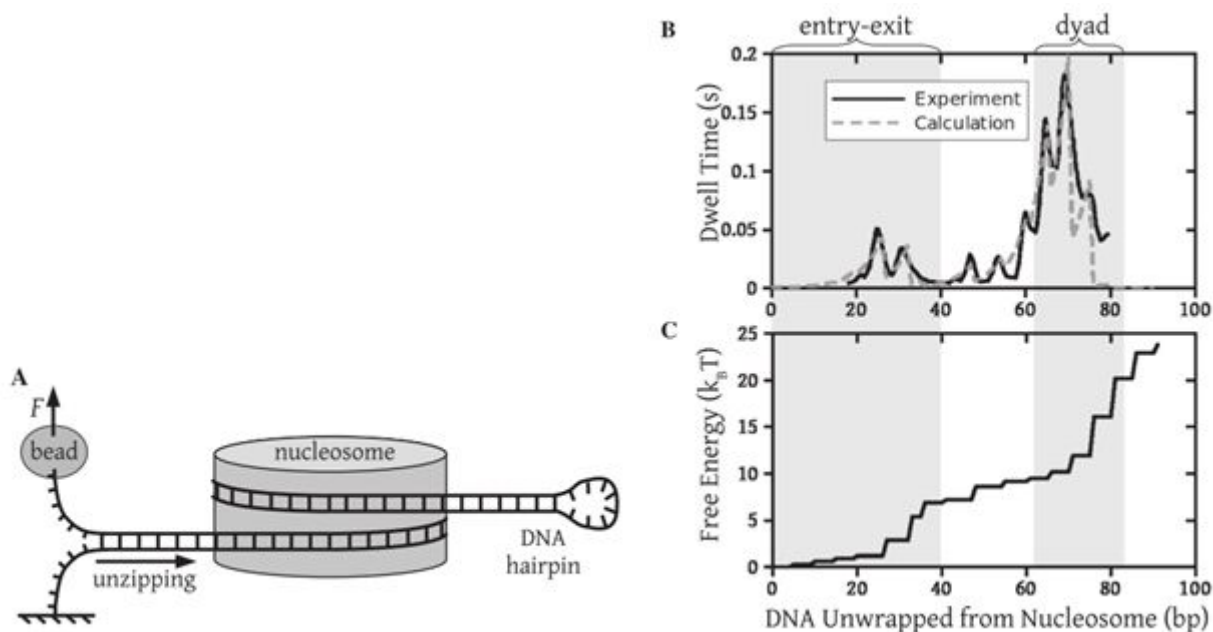


Рис. 1. Экспериментальная установка для растягивания нуклеосомы, В – зависимость времени задержки (с) от количества отвернутых н.п., С – зависимость свободной энергии связывания (кТ) от количества отвернутых н.п. Источник: Forties et al., 2011.

2.3.2 Теоретические подходы

К проблеме растяжения нуклеосом можно обратиться и со стороны компьютерных методов. Эксперименты по растягиванию нуклеосом проводили методом молекулярной динамики путем приложения внешних сил к линкерным участкам ДНК (Ettig et al., 2011). В результате было выявлено 5 стадий процесса отворачивания ДНК и два основных энергетических барьера в областях 70 и 45 н.п. относительно диадной оси.

В литературе описан ряд методов для теоретической оценки энергии взаимодействия ДНК-гистоны. Однако они не лишены ряда приближений и допущений. В принципе, проводить оценку энергетики взаимодействий можно как с помощью термодинамического интегрирования, так и эмпирических оценочных функций. Но для термодинамического интегрирования потребуется произвести очень большой объем вычислений, а для использования оценочных функций необходимо обучение на известном наборе экспериментальных данных.

2.4 Цель работы

Таким образом, нами была поставлена следующая цель работы: построение и исследование энергетического профиля процесса отворачивания ДНК от гистонового ядра гексасомы и нуклеосомы.

2.5 Задачи:

1. Создать уравновешенные модели гексасомы и нуклеосомы с отвернутой ДНК от 5-ти до 40-а нуклеотидных пар (н.п.) с шагом в 5 н.п. Для

гексасомы сторона отворота R, для нуклеосомы - R и L (R и L условные названия противоположных сторон отворота ДНК от гистонового октамера).

2. Получить устойчивые на всем времени симуляции модельные системы гексасомы и нуклеосомы в воде.
3. Построить энергетический профиль процесса отворачивания ДНК от гистонового ядра гексасомы и нуклеосомы с разрешением в 5 н.п., сравнить полученные результаты с экспериментальными данными, полученными на одиночных нуклеосомах (Forties et al., 2011) и исследовать соответствующие профили.

3 Материалы и методы

3.1 MM/PBSA

Использование компьютерного моделирования позволяет исследовать микроскопические параметры, измерение которых в реальном эксперименте невозможно. Также с помощью моделирования можно улучшать интерпретацию экспериментальных данных на молекулярном уровне.

В общем случае, прочность связывания комплекса лиганд - рецептор (в данной работе, лиганд - ДНК (DNA), рецептор - гистоны (H)) определяется свободной энергией связывания, ΔG_{bind} .

Выбор метода вычисления ΔG_{bind} обусловлен компромиссом между точностью и эффективностью. Например, в разработке лекарств широко применяется метод докинга и скоринга, однако он плохо распознает лиганды, различающиеся меньше, чем на порядок по величине сродства (< 6 кДж/моль в ΔG_{bind}). С другой стороны, есть методы расчета свободной энергии на основе полноатомных потенциалов (например, free energy perturbation - FEP), выводимые с использованием статистической механики. Эти более точные методы опираются на метод молекулярной динамики или метод Монте-Карло, и, следовательно, требуют большей статистической выборки конформаций и являются вычислительно сложными.

Между этими двумя группами методов, существуют промежуточные, использующие только конечные состояния, т.е. комплексы и, возможно, свободные рецепторы и лиганды, откуда и следует их название – методы конечных состояний. Они менее ресурсоемкие, чем FEP, и более точные, чем оценочные функции.

Далее будет рассмотрен один из популярных методов конечных состояний для вычисления ΔG_{bind} - MM/PBSA (Molecular

Mechanics/Poisson-Boltzmann Surface Area) (Kollman et al., 2000). В нем энергия связывания комплекса лиганд-рецептор, под которым в данной работе подразумевается ДНК с гистонами, определяется следующим образом:

$$\Delta G_{bind} = \langle G_{H-DNA} \rangle - \langle G_H \rangle - \langle G_{DNA} \rangle \quad (1)$$

где $H-DNA$ - комплекс, H - гистоны, DNA - ДНК, G_H и G_{DNA} - свободная энергия изолированных гистонов и ДНК в растворителе, угловые скобки - усреднение по кадрам МД.

Данный метод на сегодняшний день применяется не только для вычисления свободной энергии связывания лиганда в биологических макромолекулах, но и для дизайна белков, исследований белок-белковых взаимодействий, изучения конформационной стабильности. Кроме того, его используют для воспроизведения и рационализации экспериментальных данных, для улучшений результатов виртуальных скрининга и докинга. Преимущества метода в его модульной природе, также он не требует расчетов на обучающем наборе данных. Однако метод содержит в себе и несколько приближений, например, отсутствие оценки конформационной энтропии и детального учета взаимодействий с молекулами воды в сайтах связывания. Существует множество вариантов метода и их производительность зависит от исследуемой системы. Большинство попыток улучшить метод, например, использование более точных квантово-механических расчетов, поляризуемых силовых полей или усовершенствованной сольватации не приводило к существенному улучшению точности результатов.

Свободная энергия состояния (H , DNA , $H-DNA$) из уравнения (1) состоит из:

$$G = E_{bnd} + E_{el} + E_{VDW} + G_{pol} + G_{np} - TS \quad (2)$$

где: E_{bnd} - энергия связей, линейных и двугранных углов, E_{el} - электростатических взаимодействий, E_{VDW} - Ван-дер-Ваальсовых

взаимодействий; G_{pol} и G_{np} - полярный и неполярный вклады в свободную энергию сольватации, T - температура, S - энтропия.

Энергия электростатических взаимодействий (E_{el}) – вычисляется с использованием закона Кулона с учетом зарядов атомов, используемых в ММ силовых полях. Следовательно, результат зависит от величины зарядов, используемых для рецептора и лиганда. Было показано (Wang et al., 2006; Hou et al., 2011), что вариация диэлектрической постоянной влияет на точность: высокий заряд в месте связывания требует более высокого значения ϵ , нежели гидрофобные сайты. Считается оптимальным использовать для полярных и несущих заряд сайтов связывания $\epsilon=2-4$. Также можно улучшать значение ϵ с помощью поляризуемых потенциалов, мультипольных приближений или квантового-механических вычислений (Soderhjelm et al., 2010; Jiao et al., 2009).

Полярный вклад в энергию сольватации (G_{pol}) обычно рассчитывается путем численного решения уравнений Пуассона-Больцмана и представляет собой электростатическое взаимодействие между растворенным веществом и растворителем. Полярный фактор сольватации часто упраздняет E_{el} , поэтому сейчас большинство расчётов ведётся в неявном растворителе Борна, что, зачастую, быстрее и повышает точность, однако и здесь наблюдается зависимость от исследуемой системы.

Помимо вышеперечисленного, для воспроизведения экспериментальных данных необходимо учесть силы притяжения и отталкивания между растворителем и растворенным веществом, появляющиеся в результате сольватации. Эти силы входят в неполярную компоненту энергии сольватации G_{np} . Она вычисляется на основе модели неявного растворителя SASA (Solvent Accessible Surface Area) (Kollman et al., 2000). Данная модель базируется на том, что энергия взаимодействия белка с растворителем является суммой энергий взаимодействий с растворителем всех атомов белка, а энергия взаимодействия каждого атома считается пропорциональной площади поверхности атома,

контактирующей с растворителем. Таким образом, в модели используется допущение, что G_{np} линейно пропорциональна поверхности, доступной растворителю.

Последнее слагаемое из уравнения (2) – энтропия, умноженная на температуру, оценивается с помощью анализа нормальных мод колебательных частот. Gohlke и Case отмечают, что коррекция благодаря вращательной и поступательной энтропиям ($3 RT$) отсутствовала в первоначальной формулировке и что поступательная энтропия должна быть рассчитана с 1M стандартным состоянием, для сопоставления с экспериментальными значениями.

Энтропия S в оригинальном методе (Kollman et al., 2000; Srinivasan et al., 1998) рассчитывается путем удаления молекул воды и аминокислотных остатков на расстояние $> 8 \text{ \AA}$ от лиганда и последующей минимизации оставшегося участка с использованием зависящей от расстояния диэлектрической функции. Затем вычисляются колебательные частоты, которые используются для получения энтропии путем анализа нормальных мод с использованием приближения гармонического осциллятора и идеального газа. Энтропийный фактор дает наибольшую статистическую погрешность. Также известно, что энтропия нормальных мод, используемая в MM/PBSA, не включает важного вклада связывания лиганда в истинную энтропию. В частности, рассматривается только локальная неподвижность фактического связывания, без информации о всевозможных иных конформациях рецептора и лиганда. В принципе, можно было бы исследовать различные конформации рецептора и лиганда с помощью молекулярной динамики и рассчитать энтропию на основе полученных данных, однако фактически это невозможно (Genheden et al., 2012). Вычисления энтропии обычно самые ресурсозатратные в MM/PBSA, поэтому расчет производят лишь для некоторого количества мгновенных снимков (Kollman et al., 2000; Srinivasan et al., 1998; Gohlke et al.,

2004). Было показано, что вклад энтропии может быть опущен и в большинстве исследований он игнорируется (Foloppe et al., 2006; Wang et al., 2006; Homeyer et al., 2012), однако ясно, что учет энтропии необходим из-за потери поступательной и вращательной степеней свободы при связывании лиганда.

Обобщая, энтропия нормальных мод трудоемка в вычислении, особенно с учетом данных о конформационной энтропии. Альтернативные методы не дают статистически удовлетворительных результатов, поэтому вклад энтропии часто опускается.

В зависимости от того, из скольких различных моделирований берутся средние значения в уравнение (1), различают методы: 3A-MM/PBSA (three-average), в котором выполняется 3 различных моделирования для комплекса (PL), свободного лиганда (L) и не связанного рецептора (P), тогда:

$$\Delta G_{bind} = \langle G_{PL} \rangle_{PL} - \langle G_P \rangle_P - \langle G_L \rangle_L \quad (3)$$

и 1A-MM/PBSA (one-average), где моделируется комплекс и создаются средние значения по ансамблю для рецептора и лиганда путем простого удаления соответствующих атомов:

$$\Delta G_{bind} = \langle G_{PL} - G_P - G_L \rangle_{PL} \quad (4)$$

С одной стороны, 1A-MM/PBSA требует меньше моделирований, повышает точность и приводит к устранению E_{bnd} в уравнении (2). Но с другой стороны, он игнорирует изменения в структурах лиганда и рецептора, которые могут быть важными факторами для связывания (Swanson et al., 2004). Выбор метода 3A-MM/PBSA или 1A-MM/PBSA зависит от исследуемой системы и типа сольватации (Genheden et al., 2012). Было отмечено, что на практике, 1A-MM/PBSA дает более точный результат по сравнению с 3A-MM/PBSA (Spackova et al., 2003; Lepsik et al. 2004).

В уравнениях (3), (4) для оценки средних значений по ансамблю используют модели явных растворителей. Все молекулы растворителя затем удаляются из мгновенных снимков, потому что для вычисления энергии растворения используются неявные PBSA модели растворителей. Однако, это приводит к непоследовательности в методе, потому что во время моделирования и вычисления энергии используется не одна и та же энергетическая функция. Существующую проблему можно обойти путем моделирования в неявном растворителе, однако в таком случае не определена точность неявной модели и иногда случается диссоциация субъединиц лиганда или белка (Weis et al., 2006).

Было показано, что для повышения точности расчетов следует использовать набор коротких независимых расчетов, по сравнению с одной продолжительной (Sadiq et al., 2010; Adler et al., 2013; Wright et al., 2014). Для получения независимых моделирований используют различные начальные скорости. Изменчивость может быть также получена за счет использования произвольного выбора при настройке MM/PBSA вычислений, например сольватации рецептора или выбора альтернативной конформации в исходной кристаллической структуре.

В методе LIE (linear interaction energy) (Aqvist et al., 1994; Hansson et al., 1998) моделируются только комплекс и свободный лиганд, а неполярный член добавляется, согласно уравнению:

$$\Delta G = \alpha(\langle E_{ele}^{L-S} \rangle_{PL} - \langle E_{ele}^{L-S} \rangle_L) + \beta(\langle E_{vdW}^{L-S} \rangle_{PL} - \langle E_{vdW}^{L-S} \rangle_L) \quad (5)$$

Где E_{ele}^{L-S} - электростатическое взаимодействие между лигандом и окружением (рецептором или водой), E_{vdW}^{L-S} - соответствующее Ван-дер-Ваальсово взаимодействие, α и β – эмпирические параметры. Было отмечено, что по сравнению с MM/PBSA, LIE в 2-7 раза эффективнее благодаря трудоемким вычислениями энтропии (Genheden et al., 2011).

Таким образом, рассмотренный метод ММ/PBSA для предсказания эффективности связывания лиганда и рецептора, включает в себя 6 энергетических компонент, каждая из которых может быть отдельно рассмотрена и улучшена. Точность метода выше, чем в докинге и скоринге, но хуже, чем в AP методах и сопоставима с другими методами конечных состояний (LIE, LRA). В методе используются следующие термодинамические приближения: игнорирование структурных изменений в рецепторе и лиганде при связывании, отсутствие данных о конформационной энтропии, о количестве и энтропии молекул воды в сайте связывания до и после связывания. Кроме того, не дают результатов попытки улучшить подход с помощью более точных методов (квантово-механические расчеты, поляризуемые силовые поля и т.д.). С другой стороны, требуется провести множество независимых моделирований для получения пригодной точности, что уменьшает преимущества ММ/PBSA по сравнению с AP методами. Как правило, метод показывает слишком большие энергии, тем самым давая хороший коэффициент детерминации (r^2), но неудовлетворительное среднее абсолютное отклонение (MAD), часто худшее, чем нулевая гипотеза сходной аффинности для всех лигандов.

Однако задачей данной работы является построение теоретического профиля энергии отворачивания ДНК от нуклеосомы, что не требует точного установления абсолютной энергии связывания.

3.2 Молекулярная динамика

Наработать траектории для последующего усреднения можно с помощью молекулярной динамики. В классическом методе молекулярной динамики (МД) реальному молекулярному объекту сопоставляют математический образ системы взаимодействующих материальных точек, движение которых описывается классическими уравнениями Ньютона. В качестве материальной точки обычно рассматривается атом. Задание потенциальной энергии

модельной молекулярной системы определяет взаимодействие между частицами. Формулы и таблицы параметров, необходимые для вычисления потенциальной энергии молекулярной системы по координатам атомов, называются силовым полем. В силовом поле Amber (Weiner et al., 1981) полная потенциальная энергия представлена следующим образом:

$$U = U_{\text{вал.св.}} + U_{\text{вал.угл.}} + U_{\text{торс.угл.}} + U_{\text{плоских гр.}} + U_{vdW} + U_{qq} \quad (6)$$

Идея метода заключается в численном решении классических уравнений Ньютона, то есть макроскопические свойства системы находят из известных микроскопических характеристик.

Классическая МД отличается от экспериментальных исследований тем, что в ней зафиксированы энергия и объем, а не температура и давление. В терминах статистической механики МД дает канонический ансамбль (NVT) и, если использовать постоянное давление, то изотермический-изобарический ансамбль (NPT).

Появление краевых эффектов во время моделирования можно избежать благодаря использованию периодических граничных условий. Для ускорения расчетов обычно вводят радиусы обрезания для электростатических и Ван-дер-Ваальсовых взаимодействий.

3.3 3DNA

Для приготовления начальных конфигураций ДНК можно использовать программное обеспечение 3DNA (Xiang-Jun et al., 2003), которое также используется для анализа, восстановления и визуализации трехмерных структур, содержащих нуклеиновые кислоты. Сначала все возможные пары оснований как канонические, так и не канонические, идентифицируются, затем каждая пара характеризуется шестью степенями свободы (рис. 2, слева). Они включают в себя три вращательные и три поступательных координаты, и описывают позицию и ориентацию каждого из двух оснований. Но для более

полного описания системы вводится еще шесть параметров (рис. 2, справа), описывающих взаимное расположение двух пар оснований, то есть стэкинг-геометрию. Таким образом, геометрию ДНК можно полностью охарактеризовать набором из $12N - 6$ параметров (где N - число нуклеотидных пар). Помимо вышесказанного, программа позволяет восстанавливать атомную модель из набора соответствующих параметров. Для получения В-формы ДНК заданной последовательности используется утилита fiber.

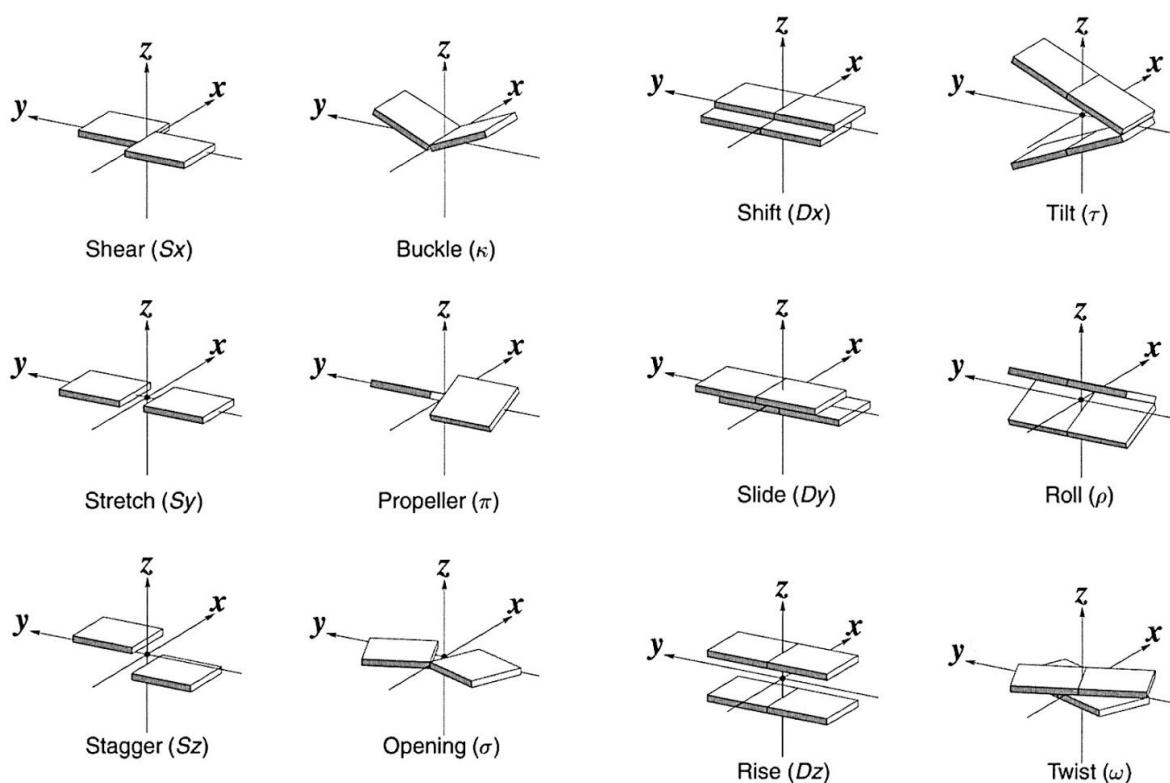


Рис. 2. Графическое определение степеней свободы, используемых для описания геометрии комплементарных и не комплементарных нуклеотидных пар и шагов между ними. Источник: Xiang-Jun Lu et al., 2003.

3.4 Методика работы

Для проведения исследования на основе кристаллической структуры нуклеосомы из базы данных PDB с индексом 1kx5 (Davey et al., 2002) были созданы модельные системы в полноатомном приближении: от нуклеосом удалялись подвижные гистоновые хвосты (Banks et al., 2003) и отворачивалась

ДНК (от 5-ти до 40-ка н.п. с шагом в 5 н.п.) с разных сторон, как показано на рис. 3.

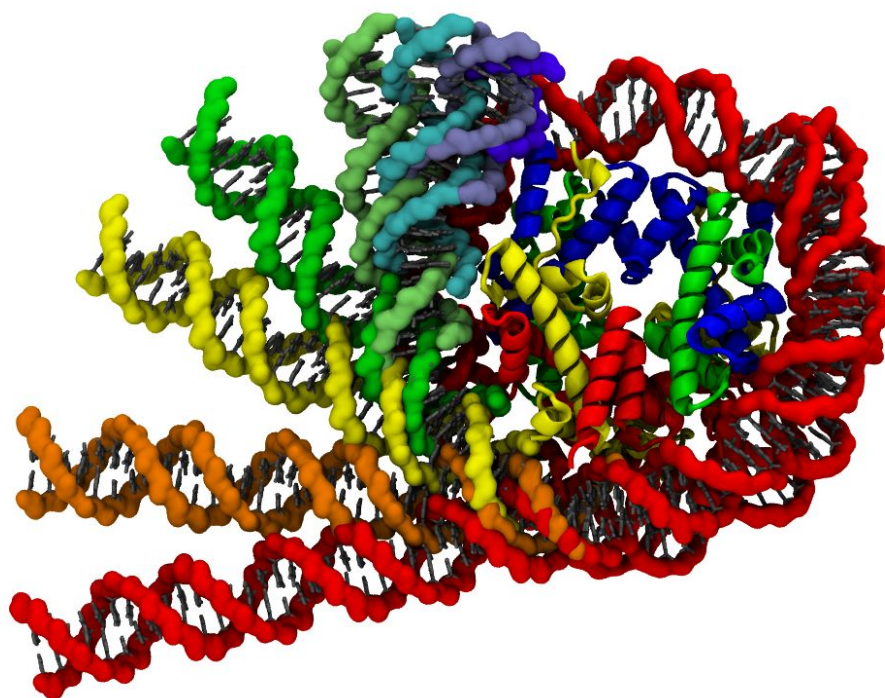


Рис. 3. Восемь совмещенных нуклеосом с отвернутыми от 5-ти до 40-ка н.п. ДНК (сторона отворота R).

Далее системы помещались в прямоугольные расчетные ячейки, с учетом того, чтобы наименьшее расстояние от нуклеосомы до одной из граней нуклеосомы было не менее 1 нм. Свободное пространство ячейки заполнялось водой (модель TIP3P, которая является трехточечной, жесткой и неполяризуемой) и ионами Na^+ Cl^- с концентрацией 150 мМ. Для предотвращения краевых эффектов были использованы периодические граничные условия.

Чтобы привести системы к термодинамическому равновесию была проведена минимизация методом наискорейшего спуска. Дальнейшее уравнивание системы производилось также, как в работе (Армеев и др., 2016). Для термостатирования расчетной ячейки применялся

модифицированный термостат Берендсена со стохастической прибавкой (Bussi et al., 2007), причем вода и молекулы термостатировались отдельно. В качестве баростата был использован изотропный баростат Паринелло-Рамана (Nose et al., 1983). Далее проводилась наработка траекторий длительностью от 1 до 5 нс методом молекулярной динамики. Стабильность систем была проверена методом расчета наименьшего среднеквадратичного отклонения структуры по траектории (RMSD), типичный график представлен на рис. 4.

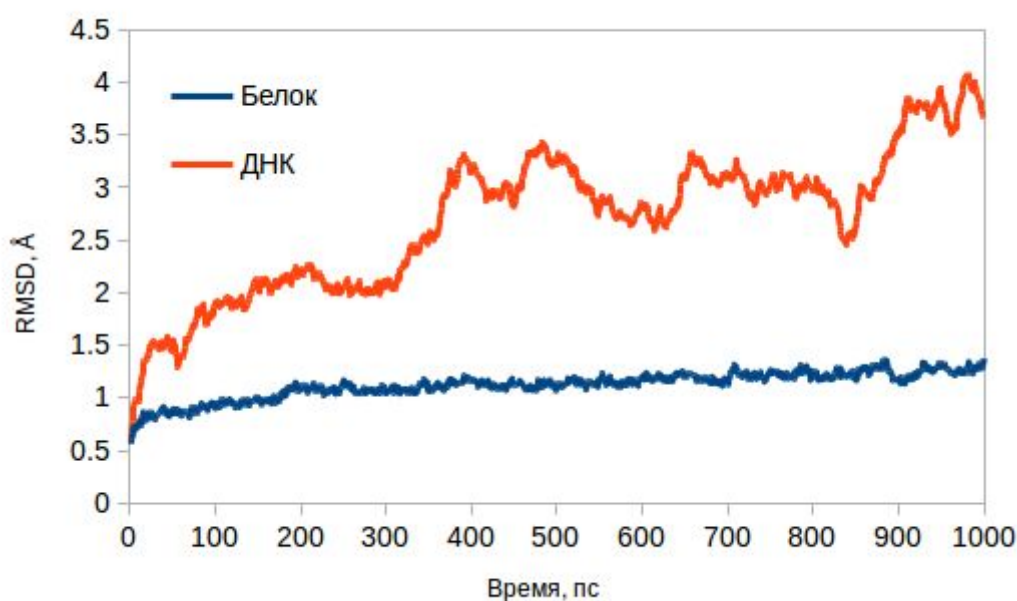


Рис. 4. Типичный график зависимости RMSD от времени моделирования.

Все моделирование проводилось с использованием силового поля AMBER99SB-BSC0 (Guy et al., 2012), интегрирование проходило методом leap-frog с шагом 2 фс. Электростатические взаимодействия учитывались при помощи метода PME (Essmann et al., 1995) с шагом решетки 1 Å, радиусом обрезания взаимодействий в прямом пространстве 1 нм и кубической интерполяцией. Для учета ван-дер-ваальсовых взаимодействий использовался радиус отсечки в 1 нм с плавным выравниванием потенциала на границе. Расчеты проводились в программе Gromacs (Pronk et al., 2013). Для визуализации моделей применялась программа VMD (Humphrey et al., 1996). Используемые параметры представлены в таблице 1 в приложении.

4 Результаты и обсуждение

Рассчитанные при помощи метода ММ/PBSA энергетические профили процесса отделения ДНК от гистонового ядра нуклеосомы, представленные на рис. 5, по своей форме хорошо согласуются с экспериментальными данными, полученными на одночных нуклеосомах (Forties et al., 2011)

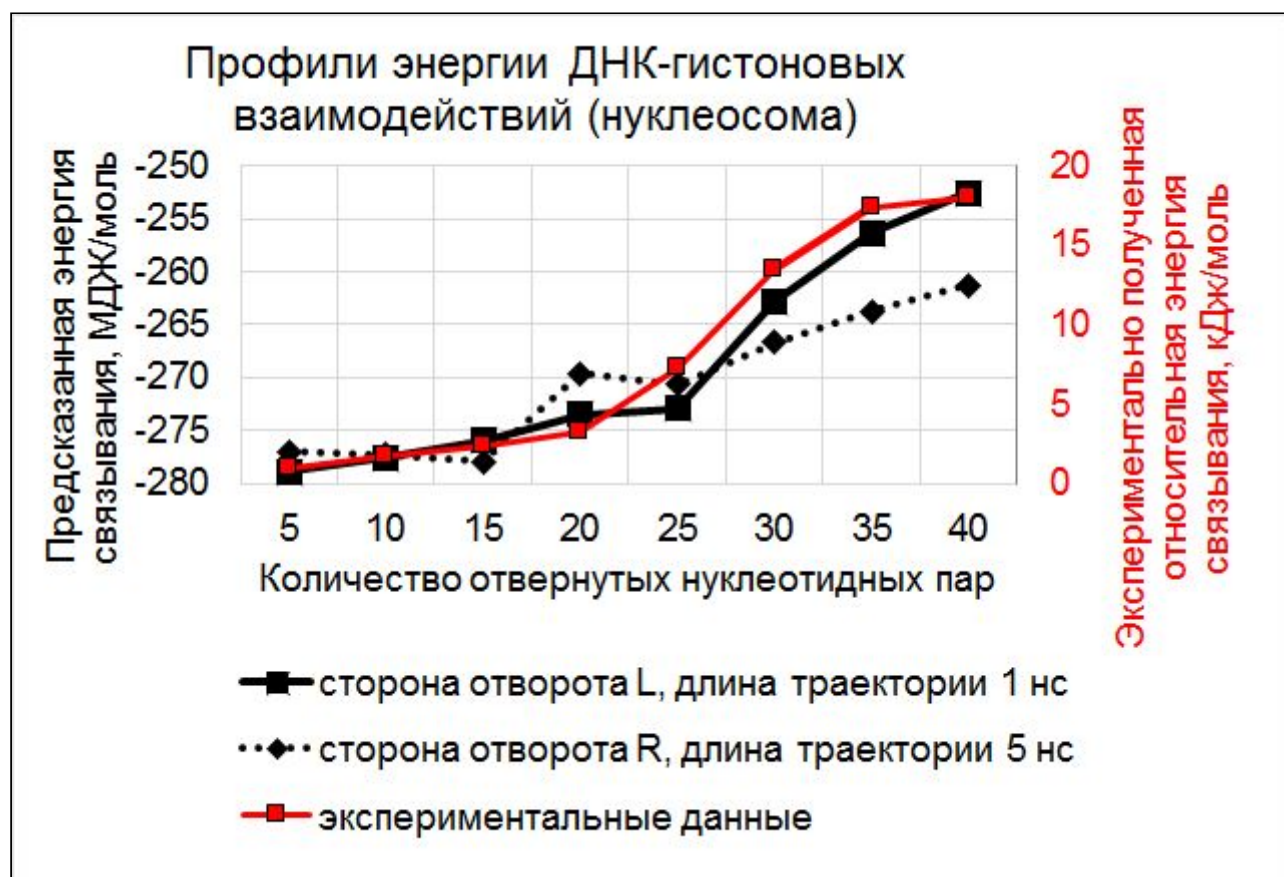


Рис. 5. Предсказанные и экспериментально измеренные профили энергии ДНК-гистоновых взаимодействий (R и L - отворачиваемые концы ДНК).

Разница в абсолютных значениях свободной энергии связывания объясняется тем, что используемый метод ММ/PBSA, в силу достаточно серьезных приближений, дает физически некорректные значения энергии. Однако в данной работе интерес представляет вид теоретического профиля энергии отворачивания ДНК от нуклеосомы, для построения которого достаточно относительных значений.

Данное замечание отражено и в полученных вкладах различных энергетических компонент в энергию связывания. Основной вклад за электростатическими взаимодействиями, далее следует полярная компонента энергии сольватации, и практически незаметные по сравнению с остальными вклады Ван-дер-Ваальсовых взаимодействий и энергии гидратации поверхности, доступной растворителю. При расчете энергии взаимодействия ДНК с гистонами по начальным неуравновешенным структурам был показан сильно завышенный вклад Ван-дер-Ваальсовых взаимодействий, что связано со значительными стерическими перекрытиями в моделях и подтверждает необходимость использования МД траекторий. Однако с увеличением длины траектории (от 1 до 5 нс) точность метода с учетом погрешности измерений не увеличивалась на сопоставимо значимые значения. Из чего можно сделать вывод, что для соблюдения баланса между затрачиваемым на расчеты временем и точностью полученных результатов можно ограничиться траекторией длиной в 1 нс.

Помимо расчета профиля свободной энергии для собранной нуклеосомы был также произведен аналогичный расчет для гексасомы - нуклеосомы без димера гистонов H2A-H2B (отворачивание производили с противоположной удаленным гистонам стороны), профили изображены на рис. 6 (экспериментальные данные, взятые для сравнения, получены на одиночных нуклеосомах).

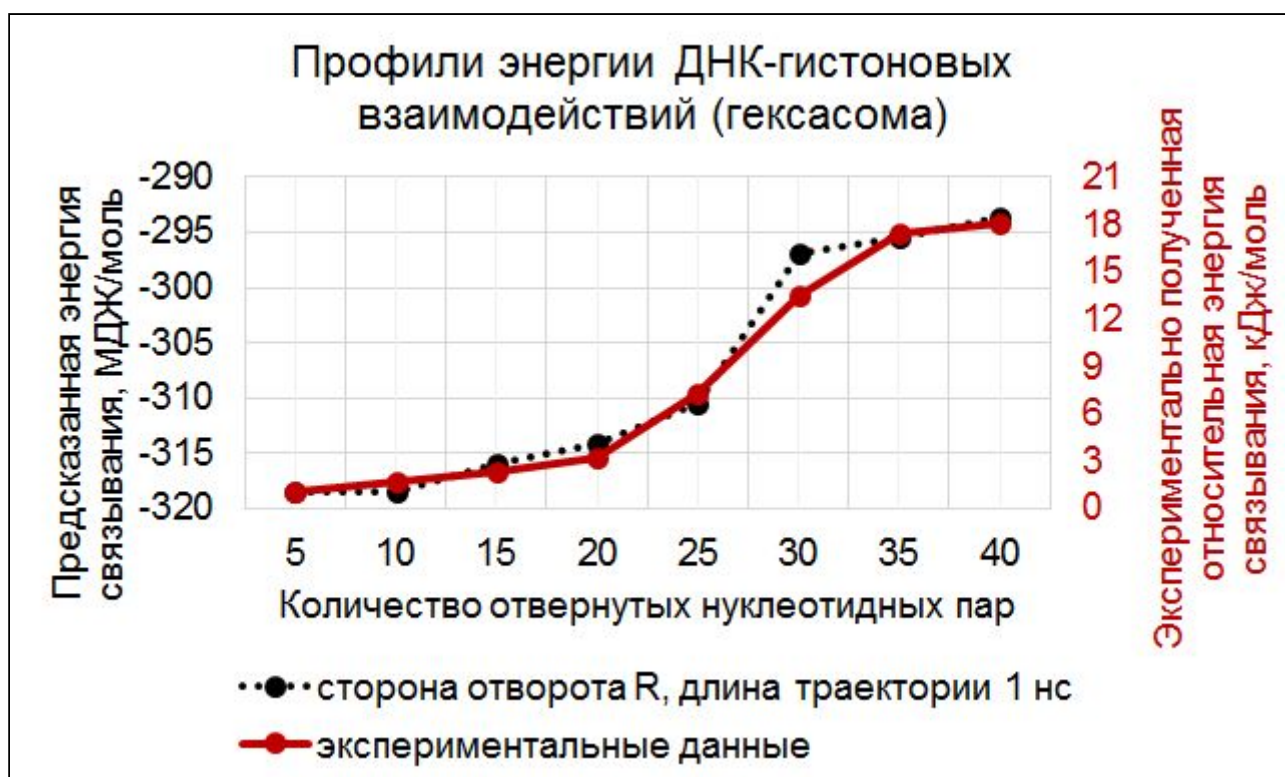


Рис. 6. Предсказанный и экспериментально измеренный профиль энергии ДНК-гистоновых взаимодействий.

Интересно, что кривая, полученная для этого расчета лучше соответствует экспериментальным значениям. Можно предположить, что в эксперименте (Forties et al., 2011) участвовали не только нуклеосомы, но и гексасомы, ведь гистоны H2A-H2B способны к спонтанной обратимой диссоциации. Данное предположение верно лишь в том случае, если погрешность метода невелика. Действительно, дисперсия рассчитанных по траекториям величин свободной энергии мала относительно абсолютных величин, но она, по всей видимости, и не может быть напрямую использована для оценки погрешности метода.

Оценку погрешности подхода можно произвести иначе. Нуклеосома - псевдосимметричная структура относительно диадной оси, в данной работе использовалась структура 1KX5 с палиндромной (относительно диадной оси) ДНК. Таким образом, профили свободной энергии при отворачивании ДНК с разных сторон должны совпадать, однако соответствующие вычисленные профили отличаются друг от друга (рис. 5). Разница между профилями может

служить мерой погрешности метода. Различия в энергии для разных сторон ДНК можно объяснить тем, что расчета коротких траекторий молекулярной динамики недостаточно для полного исследования конфигурационного пространства вблизи начальной структуры. Помимо этого, гистоновые хвосты были удалены не полностью (не до первых элементов вторичной структуры), а согласно обычным позициям их удаления в экспериментах. Оставшиеся короткие участки гистоновых хвостов конформационно лабильны и приводят к “зашумлению” рассчитанного профиля свободной энергии на коротких временах расчета.

5 Заключение

В результате данной работы были созданы молекулярные модели нуклеосом и гексасом в воде с отвернутыми участками ДНК (от 5 до 40 нуклеотидных пар с шагом 5 н.п., с двух сторон для нуклеосомы (R и L) и с одной стороны (R) для гексасомы). Конформация данных моделей была уравновешена, были рассчитаны траектории движения атомов (длиной от 1 до 5 нс). На основе полученных траекторий были рассчитаны энергии взаимодействия ДНК и гистонов методом ММ/ПБСА. Расчеты проводились последовательно для различных степеней отворачивания, что позволило получить энергетический профиль процесса отворачивания ДНК от октамера. Предсказанная форма профиля согласуется с экспериментальными данными, полученными на одночных нуклеосомах (Forties et al., 2011). Были определены (с точностью до констант) вклады энергетических компонент в энергию связывания и показаны границы применимости метода ММ/ПБСА для исследования нуклеосомных систем.

Таким образом, мы показали, что разработанная методика позволяет исследовать профили энергии связывания ДНК с гистонами в различных нуклеосомных интермедиатах, образуемых при откручивании ДНК от гистоновой ядра. Данная методика может быть использована для изучения влияния посттрансляционных модификаций гистонов и гистоновых вариантов на энергетику нуклеосом, что важно для понимания механизмов регуляции транскрипции в хроматине

6 Выводы

1. Были проведены расчеты 8-ми гексасомных систем и 16-ти нуклеосомных с отвернутыми от 5-ти до 40-ка н.п. ДНК (с шагом в 5 н.п.) методами МД и ММ/PBSA. Суммарное время симуляции 80 нс. Получены соответствующие профили энергии взаимодействия ДНК с гистонами.
2. Теоретически предсказанный профиль энергии связывания ДНК с гистонами согласуется с экспериментальными данными (Forties et al., 2011) в местах входа и выхода нуклеиновой кислоты в нуклеосому.
3. Было показано, что разработанная методика позволяет исследовать профили энергии связывания ДНК с гистонами в различных нуклеосомных интермедиатах, образуемых при откручивании ДНК от гистонового ядра.

7 Приложение

Таблица 1. Набор параметров, используемых на различных стадиях моделирования методом МД.

	Минимизация энергии	Термостатирование	Баростатирование	Свободная динамика
Интегратор	Steepest descent	Leap-Frog	Leap-Frog	Leap-Frog
Шаг интегрирования	-	2 фс	2 фс	2 фс
Число шагов	-	100 000	1000 000	2 500 000
Термостат	-	V-rescale Tref=300 K ttau=0.1 пс	V-rescale Tref=300 K ttau=0.1 пс	V-rescale Tref=300 K ttau=0.1 пс
Баростат	-	-	Parrinello-Rahman Pref=1 бар ttau=2 пс	Parrinello-Rahman Pref=1 бар ttau=2 пс
Обновление списка соседних атомов	1 шаг	5 шагов	5 шагов	5 шагов
Радиусы обрезания взаимодействий	Электростатические взаимодействия в прямом пространстве – 2 нм; Ван-дер-Ваальсовы взаимодействия -2 нм	Электростатические взаимодействия в прямом пространстве – 1 нм; Ван-дер-Ваальсовы взаимодействия -1 нм		

8 Список литературы

1. Армеев Г., Шайтан К., Шайтан А. Релаксация структуры нуклеосомы при отворачивании днк: исследование методом молекулярной динамики // Вестник Московского университета. Серия 16. Биология. 2016. № 3. С. 34–37.
2. Adler M., Beroza P. Improved ligand binding energies derived from molecular dynamics: replicate sampling enhances the search of conformational space // J Chem Inf Model 2013; 53:2065-72.
3. Aqvist J., Medina C., Samuelsson J.E. A new method for predicting binding affinity in computer-aided drug design // Proteins Eng 1994; 7:385-91.
4. Ashkin A., Dziedzic J., Yamane T. Optical trapping and manipulation of single cells using infrared laser beams // Nature, vol. 330, issue 6150 (1989) pp. 769-71.
5. Banks D.D., Gloss L.M. Equilibrium folding of the core histones: the H3-H4 tetramer is less stable than the H2A-H2B dimer // Biochemistry. 2003. Vol. 42. N 22. P. 6827–6839.
6. Brower-Toland B., Smith C., Yeh R., Lis J., Peterson C., Wang M. Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA // Proceedings of the National Academy of Sciences of the United States of America. 2002 vol: 99 (4) pp: 1960-5.
7. Bussi G. [et al.] Canonical sampling through velocity rescaling // Journal of Chemical Physics. -2007. -1: Vol. 126.
8. Davey C.A., Sargent D.F., Luger K., Maeder A.W., Richmond T.J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution // J. Mol. Biol. Elsevier, 2002. Vol. 319. N 5. P. 1097–1113.
9. Essmann [et al.] A Smooth Particle Mesh Ewald Method // Journal of Chemical Physics. -1995. -19: Vol. 103. -p. 8577-8593.

10. Ettig R., Kepper N., Stehr R., Wedemann G., Rippe K. Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping // *Biophys. J.* 2011. Vol. 101. N 8. P. 1999–2008.
11. Foloppe N., Hubbard R. Towards predictive ligand design with free-energy based computational methods // *Curr Med Chem* 2006; 13:3583-608.
12. Forties R.A., North J.A., Javaid S., Tabbaa O.P., Fishel R., Poirier M.G., Bundschuh R. A quantitative model of nucleosome dynamics // *Nucleic Acids Res.* 2011. Vol. 39. N 19. P. 8306–8313.
13. Gansen A., Tóth K., Schwarz N., Langowski J. Structural variability of nucleosomes detected by single-pair Förster resonance energy transfer: histone acetylation, sequence variation, and salt effects // *J Phys Chem B.* 2009;113: 2604–2613.
14. Gansen A., Valeri A., Hauger F., Felekyan S., Kalinin S., Tóth K. et al. Nucleosome disassembly intermediates characterized by single-molecule FRET // *Proc Natl Acad Sci U S A.* 2009;106: 15308–15313.
15. Genheden S., Ryde U. Comparison of end-point continuum-solvation methods for the calculation of protein-ligand binding free energies // *Protein* 2012; 80:1326-42.
16. Genheden S., Ryde U. Comparison of the efficiency of the LIE and MM/GBSA methods to calculate ligand-binding energies // *J Chem Theory Comput* 2011; 7:3768-78.
17. Gohlke H., Case D.A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf // *J Comput Chem* 2004; 25:238-50.
18. Guy A.T., Piggot T.J., Khalid S. Single-stranded DNA within nanopores: conformational dynamics and implications for sequencing; a molecular dynamics simulation study // *Biophys. J.* 2012. Vol. 103. N 5. P. 1028–1036.

19. Hansson T., Marelius J., Aqvist J. Ligand binding affinity prediction by linear interaction energy methods // J Comput-Aided Mol Des 1998; 12:27-35.
20. Homeyer N, Gohlke H. Free energy calculations by the molecular mechanics Poisson-Boltzmann surface area method // Mol Inf 2012; 31:114-22.
21. Hou T., Wang J., Li Y. et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. The accuracy of binding free energy calculations based on molecular dynamics simulations // J Chem Inf Model 2011; 51:69-82.
22. Humphrey W., Dalke A., Schulten K. VMD: visual molecular dynamics // J. Mol. Graph. 1996. Vol. 14. N 1. P. 33–38, 27–28.
23. Jiao D., Zhang J., Duke R.E. et al. Trypsin-ligand binding free energies from explicit and implicit solvent simulations with polarizable potential // J Comput Chem 2009; 30:1701-11.
24. Kollman P.A., Massova I., Reyes C. et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models // Acc Chem Res 2000;33:889-97.
25. Koopmans W.J.A., Brehm A., Logie C., Schmidt T., van Noort J. Single-pair FRET microscopy reveals mononucleosome dynamics // J Fluoresc. 2007;17: 785–795.
26. Lepsik M., Kriz Z., Havlas Z. Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations // Proteins 2004;57:279-93.
27. Li G., Levitus M., Bustamante C., Widom J. Rapid spontaneous accessibility of nucleosomal DNA // Nat Struct Mol Biol. 2005;12: 46–53.
28. Mihaudja S., Spakowitz A., Zhang Y., Bustamante C.. Effect of force on mononucleosomal dynamics // Proc Natl Acad Sci USA, vol. 103, issue 43 (2006) pp. 15871-15876.
29. Nose S., Klein M. L. Constant Pressure Molecular-Dynamics for Molecular-Systems [Journal] // Molecular Physics. -1983. -5: Vol. 50. -p. 1055-1076.

30. Pronk S., Páll S., Schulz R., Larsson P., Bjelkmar P., Apostolov R., Shirts M.R., Smith J.C., Kasson P.M., van der Spoel D., Hess B., Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit // *Bioinformatics*. 2013. Vol. 29. N 7. P. 845–854.
31. Sadiq S.K., Wright D.W., Kenway O.A. et al. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases // *J Chem Inf Model* 2010;50:890-905.
32. Soderhjelm P., Kongsted J., Ryde U. Ligand affinities estimated by quantum chemical calculations // *J Chem Theory Comput* 2010;6:1726-37.
33. Spackova N., Cheatham T.E., Ryjacek F. et al. Molecular dynamics simulations and thermodynamics analysis of DNA-drug complexes. Minor groove binding between 4',6-diamidino-2- phenylidole and DNA duplexes in solution // *J Am Chem Soc* 2003;125:1759-69.
34. Srinivasan J., Cheatham T.E., Cieplak P. et al. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices // *J Am Chem Soc* 1998;120:9401-9.
35. Swanson J.M., Henchman R.H., McCammon J.A. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy // *Biophys J* 2004;86:67-74.
36. Thåstro A., Gottesfeld J. M., Luger K., Widom J. Histone-DNA Binding Free Energy Cannot Be Measured in Dilution-Driven Dissociation Experiments // *Biochemistry*. 2004 vol: 43 (3) pp: 736-741.
37. Wang J., Hou T., Xu X. Recent advances in free energy calculations with a combination of molecular mechanics and continuum models // *Curr Comput-Aided Drug Design* 2006;2:95-103.
38. Weiner P. K., Kollman P. A. // *J. Comput. Chem*. 1981. Vol. 2. № 3. P. 287–303.

39. Weis A., Katebzadeh K., Soderhjelm P. et al. Ligand affinities predicted with the MM/PBSA method: Dependence on the simulation method and the force field // J Med Chem 2006;49:6596-606.
40. Wright D.W., Hall B.A., Kenway O.A. et al. Computing clinically relevant binding free energies of HIV-1 protease inhibitors // J Chem Theory Comput 2014;10:1228-41.
41. Xiang-Jun L., Olson W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures // Nucleic Acids Res. 2003. 31(17), 5108-21.
42. Zhang B., Zheng W., Papoian G., Wolynes P. Exploring the Free Energy Landscape of Nucleosomes // Journal of the American Chemical Society, vol. 138, issue 26 (2016) pp. 8126-8133.
43. Zlatanova J., Bishop T., Victor J., Jackson V., van Holde K. The Nucleosome Family: Dynamic and Growing // Structure. 2009 vol: 17 (2) pp: 160-171.