

На правах рукописи



Афанасьева Арина Сергеевна

**МОЛЕКУЛЯРНОЕ МОДЕЛИРОВАНИЕ ДИНАМИКИ И
МЕХАНИЗМОВ ДЕЙСТВИЯ АТФаз СЕМЕЙСТВА T1P49**

03.01.02 – биофизика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата физико-математических наук

Санкт-Петербург – 2015

Работа выполнена на кафедре биофизики ФГАОУ ВО Санкт-Петербургского политехнического университета Петра Великого. г. Санкт-Петербург и в НИЦ «Курчатовский институт» ФГБУ «Петербургский институт ядерной физики им. Б.П. Константинова»

Научный руководитель: **Петухов Михаил Геннадьевич**
доктор физико-математических наук,
доцент, ведущий научный сотрудник отделения молекулярной
и радиационной биофизики ПИЯФ

Официальные оппоненты: **Маслов Владимир Григорьевич**
д.ф.-м.н., профессор кафедры «Оптической физики и
современного естествознания» ФГАОУ ВО
«Санкт-Петербургский национальный исследовательский
университет информационных технологий, механики и
оптики»

Щёголев Борис Фёдорович
к.х.н., старший научный сотрудник лаборатории
нейрогенетики ФГБУН «Институт физиологии им. И.П.
Павлова» Российской академии наук

Ведущая организация: ФГБУН «Институт цитологии» Российской академии наук

Защита состоится 29 декабря 2015 г. в 14-00 часов на заседании диссертационного совета Д 212.229.25 при Федеральном государственном автономном образовательном учреждении высшего образования «Санкт-Петербургский политехнический университет Петра Великого» (ФГАОУ ВО «СПбПУ»), по адресу: 195251, Санкт-Петербург, ул. Политехническая, д. 29, корпус 2, ауд. 265

С диссертацией можно ознакомиться в фундаментальной библиотеке ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого».

Автореферат разослан «__» _____ 2015 г.

Ученый секретарь Диссертационного совета,
доктор биологических наук:



Линькова Наталья Сергеевна

Общая характеристика работы

Актуальность. АТФазы играют ключевую роль в трансформации химической энергии в важнейших биологических процессах живой клетки. Активные центры этих белков устроены таким образом, чтобы осуществлять связывание и гидролиз молекулы АТФ с образованием АДФ и свободного фосфат-иона. В ходе реакции дефосфорилирования выделяется энергия, которая реализуется ферментами для осуществления химических реакций, протекание которых без этой дополнительной энергии было бы невозможным, а также для механической работы. Множество различных типов АТФаз присутствует в живой клетке. Наиболее широкий класс представляют собой АТФазы Уолкер-типа, которые характеризуются наличием консервативных участков структуры Уолкер-А (или P-loop, фосфат связывающая петля) и Уолкер-В мотивы. Именно эти структурные участки осуществляют связывание и гидролиз АТФ молекул.

Белки семейства TP49, относящиеся к суперсемейству ДНК-зависимых АТФаз AAA+, являются неотъемлемыми для жизнедеятельности эукариотических клеток и архей. Как часть комплексов ремоделирования хроматина TP60, SWR1 и INO80 они играют важную роль в большинстве процессов жизнедеятельности клеток, таких как транскрипция, репарация ДНК, митоз и апоптоз. Недавно было показано, что эти белки играют ключевую роль в канцерогенезе, а селективные ингибиторы АТФазной активности этих белков являются перспективными лекарственными средствами в борьбе с несколькими видами рака человека.

Механизм гидролиза АТФ в белках семейства AAA+ предполагает протекание нуклеофильной атаки активированных молекул воды на γ -фосфат АТФ и образование пяти-координированного переходного состояния. Отрицательный заряд, накапливающийся вблизи γ -фосфата, компенсируется зарядом иона Mg^{2+} и зарядами доноров водородных связей из окружающих аминокислотных остатков боковых цепей активного центра (АЦ) белка. Следовательно, АЦ таких АТФаз должны включать остатки, способные к нуклеофильной активации воды, и электрофильные группы для стабилизации отрицательно заряженного переходного состояния, а также группы, осуществляющие координацию иона Mg^{2+} , который является необходимым кофактором реакции. АЦ АТФаз, как правило, содержат один или несколько «сенсоров», которые могут участвовать в реакции с АТФ, однако основная их функция заключается в получении информации о наличии или отсутствии γ -фосфата и в передаче сигнала к удаленным АЦ белка путем некоторых конформационных перестроек.

Исследование механизмов действия белков семейства TP49 приведет к более глубокому пониманию их функций в клетке, а также их роли в составе макромолекулярных комплексов, и откроет новые перспективы в лечении заболеваний, ассоциированных с нарушениями работы

этих белков.

Цели работы:

- выяснение механизмов предполагаемой геликазной и ДНК-зависимой АТФазной активности белков семейства TIR49 с помощью теоретических методов молекулярного моделирования и молекулярной динамики в периодическом водном боксе.
- разработка метода расчета позиций структурной воды в АЦ АТФаз путем исследования мотивов гидратации АТФ.

Для достижения этих целей были поставлены следующие **задачи**:

1. Исследование механизма гидролиза АТФ в белках семейства TIR49 методом молекулярной динамики
2. Анализ гидратации АЦ белков на примере репрезентативного набора кристаллических структур АТФаз человека.
3. Выявление характерных мотивов гидратации лигандов в АЦ белков из репрезентативного набора.
4. Разработка метода расчета возможных положений структурной воды в АЦ белка и сравнение полученных результатов с кристаллографическими данными (Protein Data Bank).
5. Сравнительное исследование эффективности докинга (структурно-ориентированного метода поиска низкоэнергетических конформаций лигандов в АЦ белков) в зависимости от включения молекул структурной воды, определенной разработанным методом.

Объектом исследования этой работы являются белки TIR49, относящиеся к семейству ДНК-зависимых АТФаз AAA+ и являющиеся неотъемлемыми для жизнедеятельности эукариотических клеток и архей. Кроме того для исследования характерных мотивов гидратации АТФ-связывающих центров использовался репрезентативный набор кристаллических структур АТФаз человека высокого разрешения. В качестве **методов** исследования использованы различные методы компьютерного моделирования: метод молекулярной динамики (МД) в периодическом водном боксе, минимизация энергии методом Монте-Карло, метод молекулярного докинга, а также методы собственной разработки с использованием их различных комбинаций.

Научная новизна работы.

В работе впервые была получена регуляризованная структура гетерогексамерного комплекса TP49a/b с коротким фрагментом дн-ДНК внутри центрального канала с помощью молекулярного докинга, выполненного в программном пакете ICM-Pro. Конформационная стабильность комплекса была исследована путем расчета МД траектории исследуемой молекулярной системы (~350000 атомов, включая атомы воды) в водном окружении на протяжении 50 наносек. В ходе МД наблюдались значительные структурные изменения ДНК в результате взаимодействия с положительно заряженными аминокислотными остатками белка (Arg⁺ и Lys⁺). Это взаимодействие привело к локальному раскручиванию центральной части спирали ДНК, находящейся внутри канала, образованного гексамерным комплексом белков со стороны большой бороздки ДНК. Важным и неоднократно воспроизводимым результатом оказалось то, что присутствие АТФ в составе комплекса оказывает значительное влияние на структурные изменения в ДНК, приводя к разрыву некоторых комплементарных связей между основаниями ДНК, особенно в случае днДНК с высоким АТ-содержанием, что может играть значительную роль в механизмах геликазной активности этих белков.

В ходе работы был впервые разработан новый метод исследования динамики водного окружения в АЦ АТФаз, который проливает свет на роль динамики воды в активных центрах белков TP49, на их каталитическую активность и, таким образом, дает более глубокое представление о механизмах их действия. Кроме того, этот подход может быть весьма полезным для исследования механизмов каталитической активности не только ДНК зависимых АТФаз, но и ферментов других классов.

Разработан новый метод поиска всех возможных конформаций тесно-связанной (структурной) воды в АЦ АТФаз, основанный на расчетах параметров водородных связей воды с белком, а также расчетах свободной энергии молекулы воды в белковом комплексе методом Монте-Карло. Эффективность метода исследована на репрезентативном наборе кристаллических структур АТФаз человека. Продемонстрировано повышение эффективности молекулярного докинга при учете структурной воды в АЦ, полученной с помощью разработанного подхода.

Теоретическая и практическая значимость В данной работе предложен молекулярный механизм функциональной активности белков семейства TP49, который проясняет роль этих белков в составе хроматин-ремоделирующих комплексов в клетке. Разработанный в ходе работы метод исследования динамики водного окружения в АЦ АТФаз может быть использован для исследования механизмов каталитической активности не только ДНК-зависимых АТФаз, но и других ферментов.

Разработанный метод поиска структурной воды в АЦ значительно улучшает эффективность методов докинга лигандов в активных центрах различных белков и поэтому является

практически важным для конструирования высоко-специфичных лигандов белков методами компьютерного моделирования.

Положения, выносимые на защиту.

1. ДНК-зависимая АТФазная активность гексамерных комплексов белков ТР49 обусловлена ассоциативным механизмом реакции гидролиза АТФ.
2. Локальное раскручивание спирали ДНК является результатом взаимодействия с положительно заряженными белковыми петлями в центральном канале кольцевого гексамерного комплекса исследуемых белков.
3. Присутствие АТФ в АЦ гексамера ТР49 в процессе моделирования МД влияет на динамику и конечную структуру днДНК, приводя к разрыву части комплементарных связей в АТ-богатых последовательностях ДНК.
4. Присутствие ДНК в комплексе приводит к значительному увеличению как заселенности атакующей позиции в активном центре гексамерного комплекса ТР49 молекулами литической воды, так и вероятности формирования водородных связей с такими молекулами воды, и, таким образом, может увеличивать скорость гидролиза АТФ.
5. Учет структурной воды в активном центре белка, полученной с помощью разработанной утилиты AquaBridge на базе программного пакета молекулярного моделирования ICM-Pro, приводит к улучшению точности результатов молекулярного докинга.

Апробация работы и публикации. По материалам диссертации опубликованы 14 работ (в том числе 3 статьи в рецензируемых международных и отечественных научных журналах [1, 2, 3]).

Основные результаты работы были представлены на 38м, 39м и 40м международном конгрессе Федерации Европейского Биохимического Общества, FEBS (Санкт-Петербург, Россия, июнь 2013; Париж, Франция, сентябрь 2014; Берлин, Германия, июль 2015) [4, 5, 6, 7, 8, 9], XXI и XXII Российском национальном конгрессе «Человек и лекарство» (Москва, Россия, апрель 2014 и апрель 2015) [10, 11], III Всероссийском Конгрессе Молодых Ученых, ВКМУ ИТМО (Санкт-Петербург, Россия, апрель 2014) [12], 20м международном симпозиуме «EuroQSAR, Понимание Химико-Биологических Взаимодействий» (Санкт-Петербург, Россия, Сентябрь 2014) [13], 2й международной конференции Pontin&Reptin (Оэйраш, Португалия, октябрь 2014), международном научно-практическом семинаре «Ресурсы для компьютерной разработки лекарств» (Хинкстон, Англия, ноябрь 2014), а также международной конференции

молодых ученых «Экспериментальная и теоретическая биофизика'12» (Пушино, Россия, октябрь 2012) [14].

Личный вклад автора. состоял в проведении компьютерного моделирования исследуемых молекулярных систем, в обработке, анализе и интерпретации полученных результатов, а также в подготовке докладов и публикаций.

Структура и объем диссертации. Диссертационная работа состоит из введения, трех глав, выводов, списка цитируемой литературы (105 наименований). Работа изложена на 126 страницах, включает 33 рисунка, 9 таблиц и 1 приложение.

Содержание работы

Во введении обоснована актуальность диссертационной работы, сформулирована цель и аргументирована научная новизна исследования, показана практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В первой главе описаны основные известные биологические функции белков семейства TIP49, результаты биохимических исследований, проведенные различными исследовательскими группами, структурные особенности, а также имеющиеся на данный момент результаты исследований белков TIP49 методами численного моделирования. Можно отметить следующие основные положения.

Гены белков TIP49a и TIP49b являются неотъемлемыми для жизнедеятельности клеток дрожжей *Saccharomyces cerevisiae*, *Drosophila melanogaster* и *Caenorhabditis elegans*. Высокая степень эволюционной консервативности этих генов от дрожжей до человека свидетельствует, о том, что они вовлечены в важные клеточные функции. Эти белки склонны к образованию гетеро-комплексов (TIP49a/b), в которых они действуют совместно, участвуя в различных клеточных процессах, однако они также могут функционировать независимо друг от друга.

TIP49a и TIP49b были найдены в составе группы хроматин-ремоделирующих комплексов, расщепляющих АТФ, таких как p400, TIP60, INO80 и других. Кроме того, белки семейства TIP49 регулируют транскрипцию не только путем ассоциации с хроматин-ремоделирующими комплексами, но также посредством взаимодействия с различными факторами транскрипции и РНК-полимеразой II холоферментным комплексом. TIP49a и TIP49b также участвуют в сборке snoRNP и ее последующем созревании в нуклеоплазме.

На основе приведенных данных можно предположить, что TIP49a и TIP49b могут иметь различающиеся индивидуальные функции в процессе митоза, которые отличаются от их роли в транскрипции и ремоделировании хроматина. Несколько групп показали наличие

АТФазной активности в экспериментах *in vitro* как для ТР49а, так и для ТР49б белков человека, что соответствует принадлежности этих белков семейству AAA+ АТФаз. Наблюдаемые индивидуальные показатели АТФазной активности были менее 0,1 (моль гидрализованной АТФ / моль мономера в минуту), для смешанного двойного гексамера ТР49а/б было показано значительное увеличение АТФазной активности в экспериментах *in vitro*, в связи с чем, было предположено, что такая форма белка является каталитически активной в клетке. Наличие АТФазной активности ТР49а/ТР49б для всех аналогов, найденных в различных организмах, не подлежит сомнению; однако, геликазная активность этих белков до сих пор является спорным вопросом.

Взятые вместе, ТР49а и ТР49б участвуют во многих жизненно важных клеточных процессах. Однако механизм их функционирования до сих пор неясен.

На основании анализа первой главы были сформулированы цели и задачи исследования.

Во **второй главе** описаны результаты исследования динамики и механизмов взаимодействия гетерогексамерных кольцевых комплексов белков ТР49а/б с дн-ДНК, в эту главу входит описание используемого метода моделирования молекулярной динамики с применением метода GPU PMEMD.

Для исследования динамики взаимодействия ДНК с белками ТР49а/б были получены полноатомные регуляризованные пространственные структуры комплексов ТР49а/б с фрагментами днДНК разных вариантов нуклеотидного состава в центральном канале гексамерного кольца на основе кристаллических структур белков семейства ТР49 (коды PDB: 2C9O, 3UK6 и 2XSZ) и структуры человеческого ТР49б с помощью моделирования по гомологии (рис. 1).

Далее проводилось моделирование МД полученных систем в периодическом водном боксе. Расчеты молекулярной динамики проводились с использованием программного пакета AMBER 12 в силовом поле AMBER ff99bsc0, дополненном параметрами для расчетов аминокислот и ДНК, использовалась модели молекулы воды типа ТР3Р. При расчете задавалась постоянная температура 37 °С (за исключением моделирования систем mkТР49 — аналог из термофильной археи моделировался при 75 °С). Длительность расчета составляла 50нс, временной шаг - 2фс.

В ходе МД наблюдались значительные структурные изменения ДНК в результате взаимодействия с положительно заряженными аминокислотными остатками белка (Arg⁺ и Lys⁺) в составе белковых петель, расположенных внутри центрального канала гексамерного комплекса. Это взаимодействие привело к частичному раскручиванию спирали центральной части фрагмента ДНК со стороны большой бороздки.

Присутствие АТФ в комплексе сильно влияет на динамику связанного в центральном канале

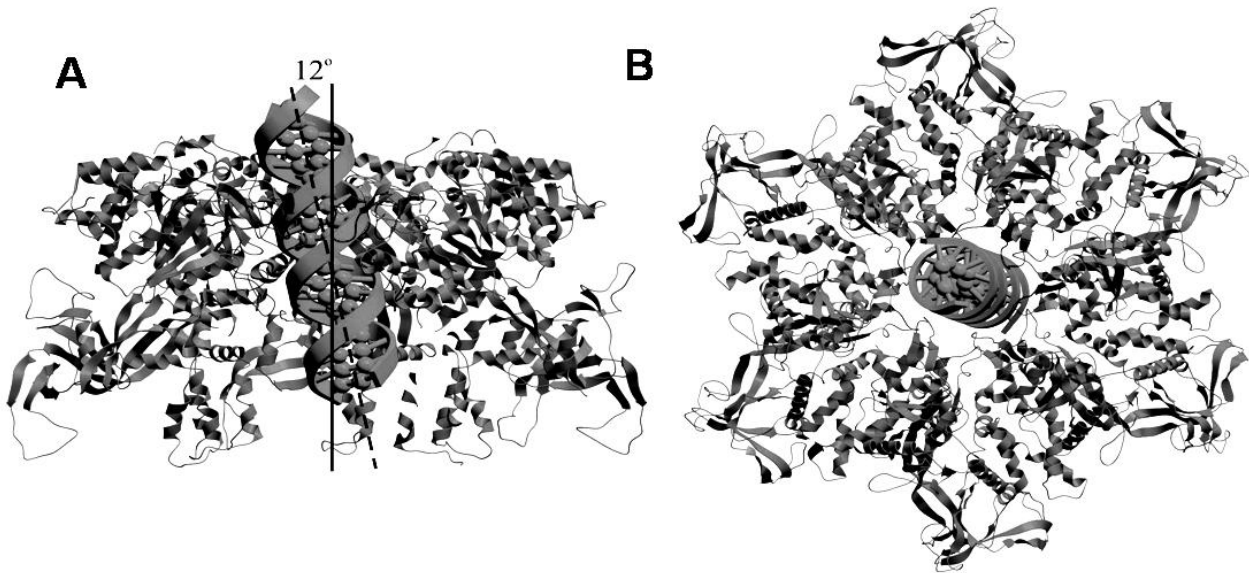


Рис. 1. Структура комплекса гексамера TIR49 с фрагментом днДНК в В-форме в центральном канале А) вид сбоку в продольном разрезе, В) вид сверху

смешанного гетерогексамера TIR49a/b днДНК (рис. 2). В частности, для днДНК со 100%-м АТ составом (повторяющийся 'АТАТ' мотив) наблюдается разрыв части водородных связей между основаниями, ведущий к потере спаренного состояния на ограниченном участке ДНК (7-10-я пары оснований), при этом в комплексах без АТФ в этом случае наблюдается только раскручивание спирали ДНК с сохранением сопряженного состояния ДНК оснований (рис. 2).

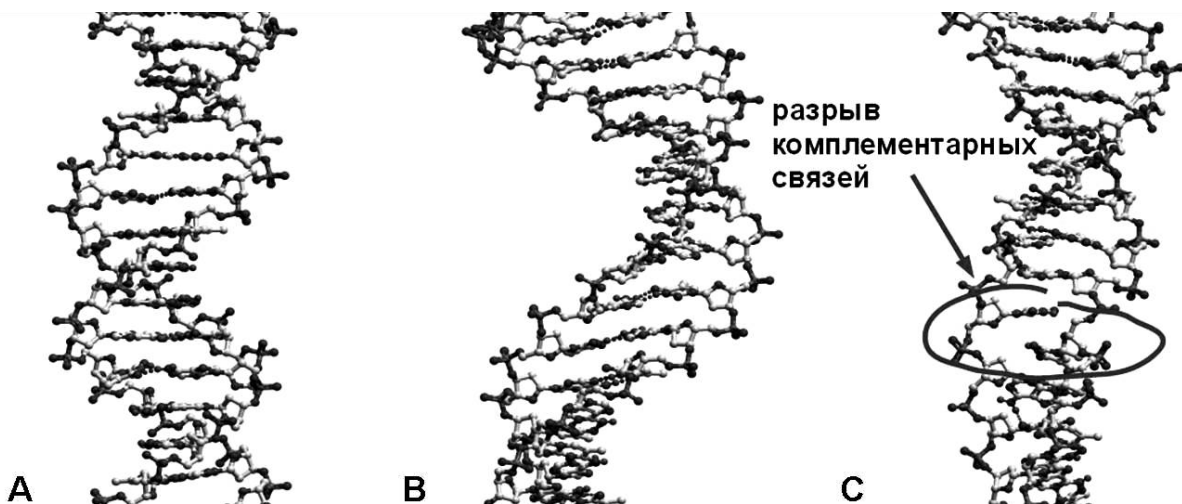


Рис. 2. Отличия конечной структуры ДНК после 50 нс молекулярной динамики в присутствии или отсутствии АТФ в АТФ-связывающих центрах белкового комплекса, А) начальная структура ДНК, В) после МД без АТФ, С) после МД в присутствии АТФ

Можно предположить, что такие структурные изменения являются результатом изменения электростатического потенциала в присутствии АТФ в комплексе, дополнительный отрицательный заряд которого приводит к смещению электростатического потенциала на

протон-донорных группах ДНК оснований в сторону более отрицательных значений. Подобный эффект влияния присутствия АТФ в комплексе на разрыв комплиментарных взаимодействий в двунитевой ДНК наблюдается нами впервые и может объяснять АТФ-зависимое расщепление спирали ДНК с высоким АТ-содержанием.

В третьей главе приведены результаты *исследования динамики активной воды в АТФ-связывающих центрах гексамеров TIP49a/b*, описана методика исследования динамики активной воды в АТФ-связывающих белках семейства TIP49, приведены исследования механизмов гидролиза АТФ в белках *mkTIP49* (аналог TIP49, содержащийся в термофильной архее *Methanopyrus kandleri* (mkTIP49; UniProtKB ID: Q8TZC3_МЕТКА)), а также *исследование динамики литической воды в АТФ-связывающих центрах гексамеров человеческого TIP49a/b в комплексе с днДНК различного состава*.

В данной главе описан разработанный новый метод анализа динамики активированных молекул воды в окрестности их целевых групп. Активность белка дикого типа и его мутантных форм считалась пропорциональной доступности АТФ для атакующей молекулы воды в правильной конформации. Доступность рассчитывали как накапливаемую вероятность присутствия воды в минимальном объеме, расположенном в центральной части передней полусферы γ -фосфатной группы АТФ. Этот объем, вмещающий только одну молекулу воды в каждый момент времени, был определен как сфера с радиусом 1.4\AA , и был помещен на расстоянии 2.5\AA от γ -фосфатной группы (рис. 3).

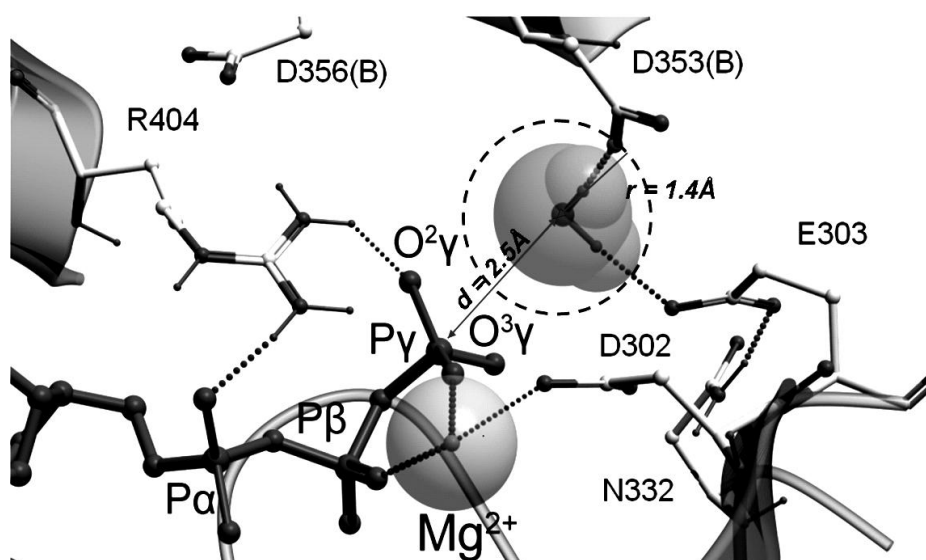


Рис. 3. Структура формирования активированного комплекса с молекулой воды в атакующей позиции напротив γ -фосфата АТФ

Вероятности присутствия воды P_w рассчитывались как доля времени МД, когда молекула воды присутствует в этом объеме. Поскольку гидролиз АТФ требует активации молекулы

воды на соседних группах белка, способных акцептировать протон воды, мы также вычисляли вероятность образования водородных связей P_{Whb} (1) между литической молекулой воды и несколькими отрицательно заряженными остатками белка (E297, D346 и D349), т.е. вероятность присутствия воды в правильной активированной конформации.

$$P_W = \frac{\sum_i n_W \cdot \Delta t}{t_i}, \quad P_{Whb} = \frac{\sum_i n_{hb} \cdot \Delta t}{t_i}. \quad (1)$$

Расчеты проводились с учетом всех известных стерических ограничений для образования соответствующих водородных связей. Все расстояния и углы между атомами были рассчитаны по координатам атомов на каждом временном шаге МД с использованием выходного файла траектории МД.

Исследование *механизмов гидролиза АТФ в белках mkTIP49* проводилось путем анализа влияния аминокислотных замен основных участников этого процесса в активном сайте белка. Белок mkTIP49 имеет достаточно высокую степень гомологии по последовательности с человеческими белками семейства TIP49a/b (45% и 42% соответственно) и одновременно позволяет достаточно простое выделение, очистку и биохимическое исследование этих белков, которое проводилось в сотрудничестве с лабораторией Финна Вернера в University College London (Лондон, Великобритания).

Было исследовано влияние мутаций в структурном мотиве Уолкер-Б на АТФазную активность белков mkTIP49 в гексамерной форме, путем наблюдения за динамикой активной воды в АТФ-связывающих центрах. С этой целью были построены полноатомные модели гексамеров mkTIP49 дикого типа, а также мутационных форм с аминокислотными заменами в мотиве Уолкер-Б D296N, и E297Q в комплексе с АТФ и Mg^{2+} и проанализирована их конформационная подвижность путем моделирования МД.

Анализ результатов МД моделирования гексамера D296N показал значительное расширение каталитических центров, возникающее на шаге уравнивания системы – подготовительном этапе, предшествующем продуктивной динамике. Анализ динамики гексамера E297Q также выявил явную дезорганизацию АЦ mkTIP49. Введение этой мутации приводит к сдвигу отрицательно заряженных остатков D346 и D349 от фосфатной цепи АТФ и к формированию водородных связей с мутированным Q297 остатком. Полученные результаты, таким образом, позволяют продемонстрировать, на атомном уровне, последствия удаления отрицательного заряда внутри Уокер-Б мотива для общей конфигурации каталитического центра mkTIP49 и подчеркнуть важность ключевых остатков, участвующих в процессе активации воды напротив γ -фосфата АТФ.

Для изучения влияния мотива сенсор I на динамику литической воды и гидролиз АТФ в белках mkTIP49, было исследовано влияние аминокислотной замены N326Q. Эта замена приводит к появлению полярной аминогруппы мутированного остатка 326 ближе к АТФ по сравнению с белком дикого типа и может привести к формированию донорной водородной связи с кислородом воды. Такое взаимодействие противоположно формированию акцепторной водородной связи водородов воды с отрицательно заряженными остатками, которое необходимо для активации воды, таким образом, происходит переориентация молекулы воды в нелитическое положение. (рис. 4) Полученные экспериментальные данные подтверждают, что АТФазная активность очищенного гексамера mkTIP49 с мутацией N326Q была практически не детектируема по сравнению с АТФазной активностью дикого типа, подтверждая результаты МД моделирования и динамики литической воды в АЦ mkTIP49. Эта мутантная форма белка TIP49a/b может использоваться для защиты АТФ в АЦ от спонтанного гидролиза.

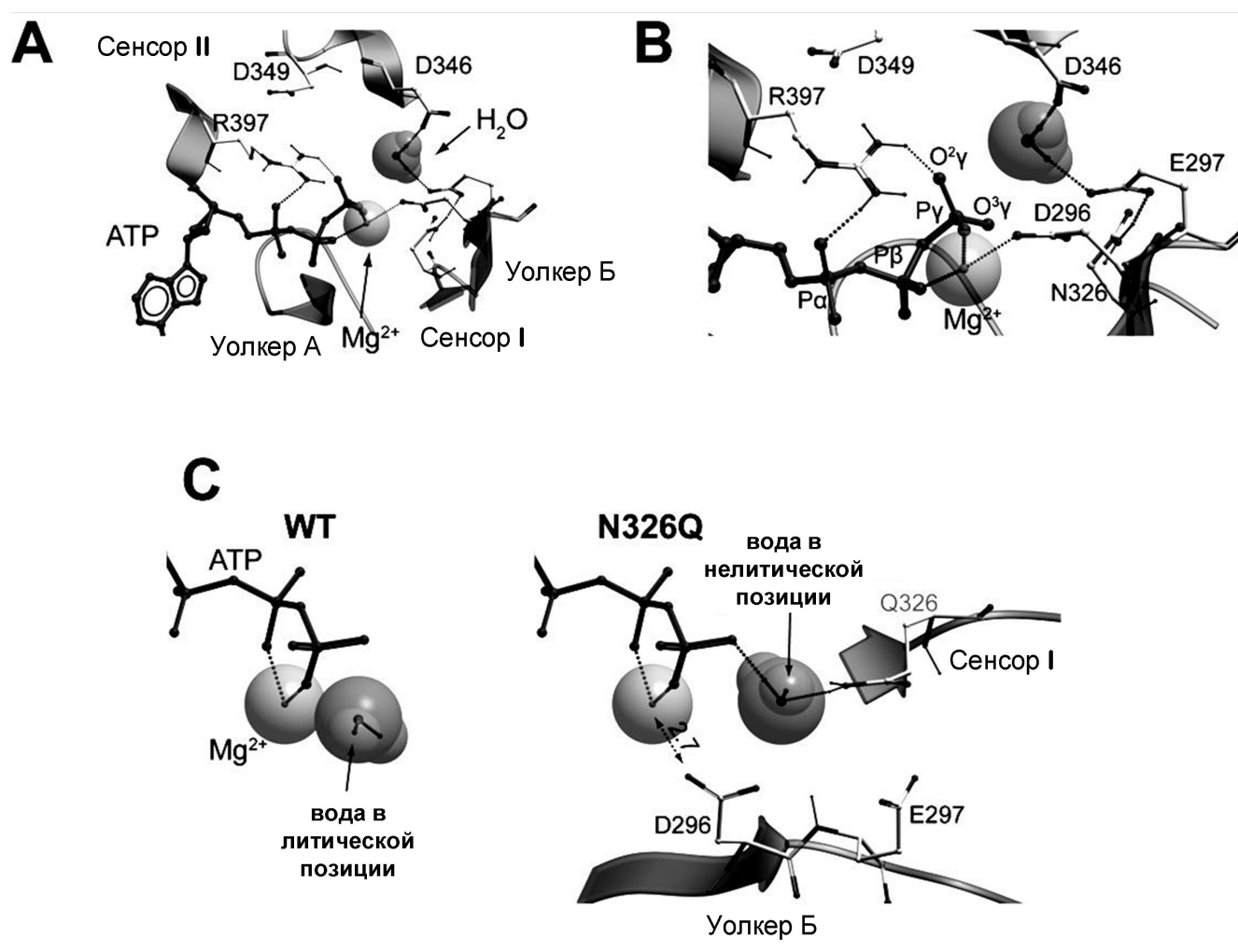


Рис. 4. Предложенная структура АЦ белка mkTIP49 и механизм гидролиза АТФ. Общий вид АЦ (А) и окружение γ -фосфатной группы АТФ (В), ион Mg^{2+} , а также атакующая молекула воды. На рис. (С) – положение атакующей молекулы воды в белках mkTIP49 дикого типа (слева) и в мутанте N326Q (справа)

Другим важным аспектом данного исследования было обнаружение дополнительного

структурного мотива в АТФ-связывающих сайтах гексамерной формы ТР49. Консервативный α -спиральный мотив GDLLDR участвует в формировании гомо- и гетерогексамеров ТР49. Этот структурный мотив, предоставляемый в АЦ соседним протомером, содержит остатки D346 и D349, которые совместно с E297 и D296 могут принимать участие в гидролизе АТФ, выступая в качестве акцепторов протона воды. Наличие этого структурного мотива в гексамерной форме ТР49 вероятно является причиной увеличения АТФазной активности гексамерных комплексов белка по сравнению с мономерной формой.

Описанные эффекты влияния точечных аминокислотных замен в АЦ ТР49, полученные на основе моделирования МД, согласуются с экспериментальными данными по АТФазной активности белков ТР49 дикого типа и мутантных форм.

Далее на основании данных о механизме гидролиза АТФ, полученных для mkТР49, была предпринята попытка исследования динамики литической воды в АЦ гексамеров человеческого ТР49a/b в комплексе с днднк различного состава.

Водное окружение АТФ-связывающего центра является ключевой характеристикой, которая может определять скорость гидролиза АТФ, при этом молекула воды должна располагаться в атакующей позиции в правильной конформации относительно γ -фосфатной группы АТФ в АЦ белка для протекания нуклеофильной атаки. Образование данного комплекса с водой в атакующей позиции регулируется двумя характеристиками АЦ:

- заселенность атакующей позиции водой, или время нахождения воды в подходящей позиции для нуклеофильной атаки АТФ;
- правильное позиционирование молекулы воды в атакующей позиции за счет формирования водородных связей с отрицательно заряженными остатками в АЦ в непосредственной близости от γ -фосфата АТФ (аминокислотные остатки Asp и Glu в составе структурного мотива Уолкер-Б и два остатка Asp в составе GDLLDR структурного мотива из соседней субъединицы, подробно исследованных в нашей работе [3]).

Мы пришли к выводу, что нуклеотидный состав ДНК влияет на обе описанные выше характеристики. На графиках P_w и P_{whb} (рис. 5) видно, что наибольших значений исследуемые параметры достигают в комплексах с ДНК 100% GC состава ($P_w(GC) = 15,30 \pm 0,07\%$ и $P_{whb}(GC) = 1,091 \pm 0,015\%$ - рассчитаны по последним 10 нс времени МД), для комплекса с ДНК 100% AT состава: $P_w(AT) = 11,02 \pm 0,06\%$ и $P_{whb}(AT) = 0,545 \pm 0,018\%$, при этом для ДНК смешанного состава эти параметры принимают промежуточные значения.

Многие белковые комплексы в клетке, участвующие в процессинге ДНК, обладают свойством регуляции АТФазной активности в присутствии ДНК, однако механизм этой

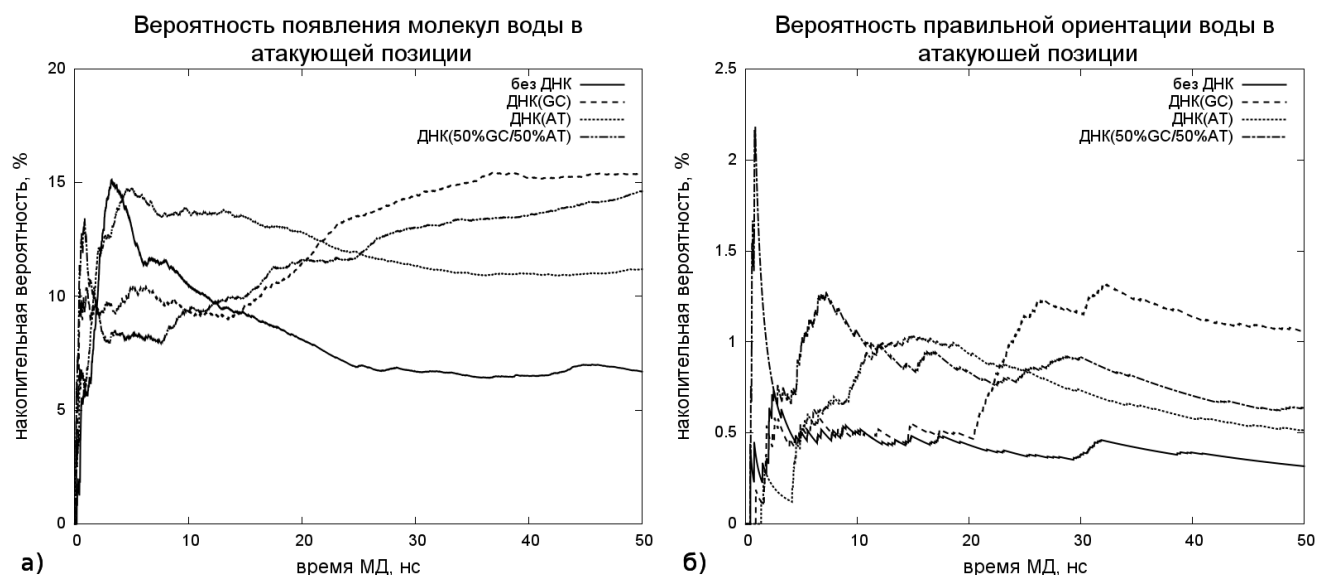


Рис. 5. а) заселенность атакующей позиции водой, б) вероятность формирования водородных связей литической воды с протон-акцепторными остатками в АТФ-связывающем центре в комплексах ТР49а/б с АТФ и дДНК различного состава: 100% GC, 100% AT, 50%AT + 50%GC, а также в комплексе без ДНК, усредненные по шести субъединицам гексамера

регуляции до сих пор остается неизвестным. В этой части работы был впервые предложен молекулярный механизм, объясняющий ДНК-зависимую регуляцию скорости гидролиза АТФ.

В третьей главе продемонстрировано как структурное окружение в АЦ белков ТР49 оказывает координальное влияние на функционирование белка, в связи с этим следующим этапом исследования стало *изучение водного окружения АЦ различных АТФ-связывающих белков и влияния водного окружения на связывание лигандов в данных АЦ (четвертая глава).*

Некоторые молекулы воды в АЦ могут считаться тесно-связанными с белком, это определяется числом и силой водородных связей, такая вода считается структурной и должна учитываться как часть структуры белка в применении различных структурно-ориентированных методов разработки лекарств, в частности докинга. В этой главе приведено описание разработанного нового метода поиска всех возможных конформаций структурных молекул воды внутри активных центров белков на основе минимизации свободной энергии молекул воды методом Монте-Карло и дальнейшей оптимизации водородных связей, образующихся между такими молекулами воды и функциональными группами белка. Метод был реализован в виде утилиты AquaBridge, предназначенной для использования в коммерчески доступном программном пакете молекулярного моделирования ICM-Pro. Процедура работы модуля включает несколько шагов: 1) выделяется расчетная область вокруг лиганда; 2) генерируется молекула воды, которая последовательно перемещается внутри выделенной области; 3) в каждой позиции осуществляется поиск оптимальной конформации путем минимизации энергии молекулы воды методом Монте-Карло; 4) если обнаруживается низкоэнергетическая

конформация молекулы воды в данной позиции, проверяется возможность формирования водородных связей такой молекулы воды с белком либо лигандом. В случае формирования двух и более водородных связей с белком, молекула воды считается структурной. Критерий отбора по формирующимся водородным связям включает расчет базовых параметров, таких как длина водородной связи, значение донорного и акцепторного углов. Вероятность выбора той или иной конформации в методе Монте-Карло определяется следующим выражением:

$$W(j \rightarrow j') = \begin{cases} \exp\left(-\frac{(u_{j'} - u_j)}{kT}\right), & u_{j'} - u_j > 0 \\ 1. & u_{j'} - u_j < 0 \end{cases} \quad (2)$$

где j и j' предыдущая и текущая пространственные конформации белка соответственно, u_j и $u_{j'}$ — значения свободной энергии белка в предыдущей и в текущей конформации, T — температурный параметр, который может варьироваться в процессе минимизации энергии, с целью оптимизации числа итераций процедуры.

Работа утилиты AquaBridge, основанной на описанном методе поиска структурной воды в АЦ белков была протестирована на репрезентативном наборе кристаллических структур АТФ-связывающих белков. Данный репрезентативный набор включает 51 кристаллическую структуру АТФ-связывающих белков человека высокого разрешения ($< 3\text{\AA}$) с молекулой АТФ либо АДФ в АЦ (база данных кристаллических структур PDB — Protein Data Bank). На рис. 6 приведена гистограмма распределения расстояний между атомами кислорода структурной воды, полученной с помощью разработанной нами процедуры, и положениями кислородов воды в соответствующей кристаллической структуре белка из репрезентативного набора. Из данных гистограммы видно, что большая часть позиций структурной воды, найденных с помощью AquaBridge, практически совпадает или находится вблизи позиций воды, найденной в кристалле (77.8% найденных молекул воды находятся на расстоянии менее 1\AA от позиций кристаллографической воды), оставшаяся доля представляет собой позиции тесно-связанной воды в АЦ, которых нет в кристалле.

Эффективность работы процедуры была исследована различными методами, в том числе с помощью моделирования молекулярной динамики нескольких белковых систем из составленного репрезентативного набора, а также на примере докинга небольших лигандов в АЦ белков. Анализ эффективности докинга небольших органических молекул в АЦ исследуемых белков свидетельствует о значительном улучшении селективности алгоритма докинга и повышении точности предсказания конформации лигандов в случае учета структурной воды в АЦ белка.

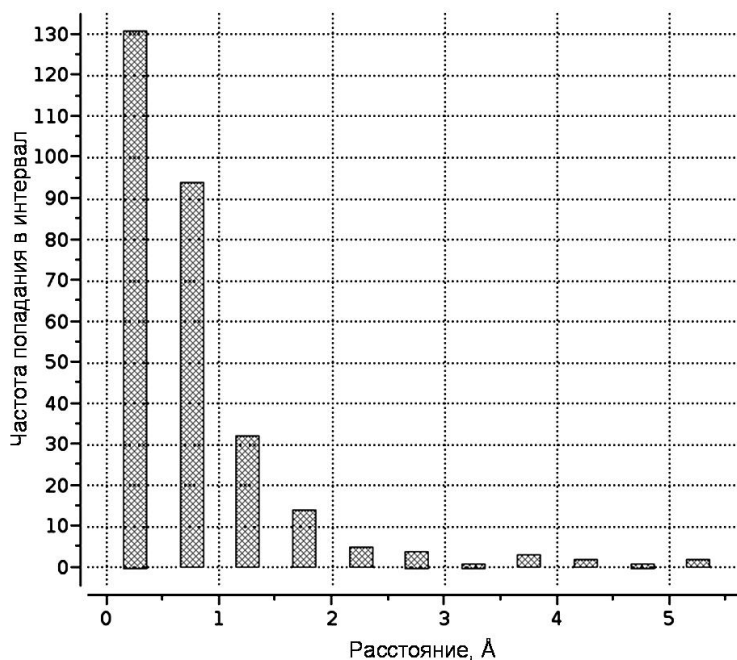


Рис. 6. Распределение конформаций структурной воды, найденной в АЦ белков из репрезентативного набора, по расстояниям до положений атомов кислорода кристаллографической воды

Разработанный метод поиска структурной воды в АЦ белков является полезным инструментом для многих задач молекулярного моделирования, структурно-ориентированных компьютерных методов разработки лекарственных средств и теоретического предсказания энергии связывания комплекса белок-лиганд.

Заключение

1. В работе впервые предложен молекулярный механизм, объясняющий эффект ДНК-зависимой АТФазной активности гетерогексамерных комплексов белков ТР49а/б.
2. С помощью моделирования МД показано, что взаимодействие фосфатной цепи ДНК с положительно заряженными аминокислотными остатками белка ТР49а/б (Arg^+ и Lys^+) приводит к частичному раскручиванию спирали центральной части фрагмента ДНК, находящегося внутри канала, образованного гексамерным кольцом белкового комплекса.
3. Показано, что присутствие АТФ в составе комплекса ТР49а/б/ДНК оказывает значительное влияние на структурные изменения в ДНК, приводя к разрыву некоторых комплементарных связей между основаниями ДНК, особенно в случае дДНК с высоким АТ-составом, что может играть значительную роль в механизмах геликазной активности этих белков.

4. Анализ динамики активной воды в АТФ-связывающих центрах гексамера TIR49 показал, что присутствие ДНК в комплексе приводит к значительному увеличению как заселенности водой атакующей позиции в АТФ-связывающих центрах белка, так и вероятности формирования водородных связей с такими молекулами воды, таким образом, приводя к увеличению скорости гидролиза АТФ.
5. Показано, что учет структурной воды в АЦ белков значительно улучшает результаты молекулярного докинга небольших органических соединений в АЦ белка, повышая селективность метода докинга, а также точность конформации лиганда в АЦ белка, полученной в результате докинга.

Благодарности

Выражаю благодарность своему научному руководителю Михаилу Геннадьевичу Петухову за предоставленную возможность работы над интересным проектом, а также за проявленное терпение и понимание. А также Григорьеву Михаилу Юрьевичу (CNRS, Тулуза) за стимулирующее обсуждение проекта и продуктивные идеи в ходе работы над проектом. Илатовскому Андрею, Рычкову Георгию и Якимову Александру за поддержку и оказание неоценимой помощи в моей исследовательской деятельности. Сотрудникам кафедры биофизики и сотрудникам лаборатории биофизики макромолекул ОМРБ ПИЯФ за поддержку и интерес к проекту. Также хочу выразить благодарность сотрудникам НИК «Нанобиотехнологии» и его руководителю Ходорковскому Михаилу Алексеевичу за понимание и поддержку моего исследовательского проекта.

Список публикаций по теме диссертации

1. Афанасьева А.С. Григорьев М.Ю. Петухов М.Г., Якимов А.П. Динамика и механизмы взаимодействия гетерогексамерных кольцевых комплексов белков TIP49a/b с двунитовой ДНК // Цитология. 2015. Т. 57, № 10. С. 671–678.
2. Afanasyeva A, Izmailov S, Grigoriev M, Petukhov M. AquaBridge, a novel method for systematic search of structural water molecules within the protein active sites // Journal of Computational Chemistry. 2015. Vol. 36, no. 26. P. 1973–1977.
3. Afanasyeva A, Hirtreiter A, Schreiber A *et al.* Lytic water dynamics reveal evolutionarily conserved mechanisms of ATP hydrolysis by TIP49 AAA+ ATPases // Structure. 2014. Vol. 22, no. 4. P. 549–559.
4. Afanasyeva A. Petukhov M., Grigoriev M. The structural basis of the TIP49a/b dodecamerization // FEBS Journal. Vol. 282. Berlin, Germany: 2015. P. 333.
5. Chervyakova D.B. Afanasyeva A. Khodorkovsky M.A. Isaev-Ivanov V.V. Petukhov M., Lebedev D.V. TIP49a protein forms active rod-like structures in solution // FEBS Journal. Vol. 282. Berlin, Germany: 2015. P. 335.
6. Afanasyeva A. Petukhov M., Grigoriev M. Mechanisms of DNA stretching in complexes with TIP49A/B proteins // The FEBS EMBO 2014 Conference.Late Abstracts. Paris, France: 2014. P. 51.
7. Afanasyeva A, Hirtreiter A, Schreiber A *et al.* Mechanism of ATP hydrolysis by the archeal TIP49 AAA+ protein // FEBS JOURNAL. Vol. 280. Saint-Petersburg, Russia: 2013. P. 156.
8. Livinskaya V, Afanasyeva A, Dolle C *et al.* Potential role of cytosolic 5'-nucleotidases in human NAD metabolism // FEBS JOURNAL. Vol. 280. Saint-Petersburg, Russia: 2013. P. 181.
9. Petukhov M, Ilatovskiy A, Artamonova T *et al.* Dynamics of the dsDNA/TIP49a hexameric complexes // FEBS JOURNAL. Vol. 280. Saint-Petersburg, Russia: 2013. P. 156.
10. Афанасьева А.С. Петухов М.Г., Григорьев М.Ю. Влияние ДНК на динамику литической воды в ДНК зависимых АТФазах семейства TIP49 // XXII Российский национальный конгресс "Человек и лекарство 6 - 10 апреля. Сборник трудов. Москва: ЗАО РИЦ "Человек и лекарство": 2015. С. 172.
11. Афанасьева А.С. Петухов М.Г. Новый метод поиска белков, способных связываться с одинаковыми лигандами // Труды XXI российского национального конгресса «Человек и лекарство». Москва, Россия: 2014. С. 200.
12. Афанасьева А.С. Измайлов С.А. Исследование молекулярных механизмов гидратации лигандов в АТФазах человека методами молекулярного моделирования и молекулярной

динамики // Сборник тезисов докладов III всероссийского конгресса молодых ученых, Выпуск 2. НИУ ИТМО, Санкт-Петербург, Россия: 2014. С. 222.

13. Afanasyeva A. Petukhov M., Izmailov S. The use of new molecular modelling method for investigation of structural water // 20th EuroQSAR, Understanding Chemical-Biological Interactions, St.-Petersburg, August 31 - September 4. Saint-Petersburg, Russia: 2014. P. 179.
14. Афанасьева А.С. Сабанцев А.В. Побегалов Г.Е. Илатовский А.В. Петухов М.Г., Руденко Е.Д. Исследование механизмов растяжения ДНК и комплексов ДНК-TP49 с помощью молекулярного моделирования и молекулярной динамики в периодическом водном боксе // Экспериментальная и теоретическая биофизика'12, Пущино. Сборник трудов. Пущино, Россия: 2012. С. 62.