

АНАЛИЗ КОНФОРМАЦИОННЫХ ОСОБЕННОСТЕЙ ФРАГМЕНТОВ УОТСОН-КРИКОВСКИХ ДУПЛЕКСОВ МЕТОДАМИ МОЛЕКУЛЯРНОЙ МЕХАНИКИ И КВАНТОВОЙ МЕХАНИКИ

© 2016 г. В.И. Полтев, В.М. Анисимов*, К. Санчес, А. Дерябина, Е. Гонсалес, Д. Гарсиа, Ф. Ривас, Н.А. Полтева**

Автономный Университет Пуэблы, Пуэбла, 72570, Мексика;

E-mail: poltev@fcfm.buap.mx

**Национальный центр суперкомпьютерных приложений, Университет штата Иллинойс, Урбана-Шампэйн, Иллинойс, 61820, США;*

***Институт теоретической и экспериментальной биофизики РАН, 142290, Пуцино Московской области*

Поступила в редакцию 30.11.15 г.

Общепринято, что важные характеристики уотсон-криковского дуплекса определяются молекулярным строением его субъединиц, но вопрос о том, какие свойства каждой из субъединиц определяют существенные характеристики уотсон-криковского дуплекса, остается не выясненным. Наши расчеты комплексов дезоксирибонуклеозидмонофосфатов с ионами натрия методом теории функционала плотности выявили ключевую роль конформационных свойств этих минимальных фрагментов одиночной цепи ДНК в возникновении уникальных черт уотсон-криковского дуплекса. Найдено, что направленность сахаро-фосфатного остова и предпочтительные области его торсионных углов вместе с различиями геометрии пуринов от пиримидинов определяют зависимость пространственной структуры уотсон-криковского дуплекса от нуклеотидной последовательности. Эти расчеты мы распространили теперь на минимальные фрагменты уотсон-криковского дуплекса, т.е. на комплементарные комплексы дезоксирибонуклеозидмонофосфатов с ионами натрия. Для поиска минимумов энергии VI-конформаций комплементарных комплексов дезоксирибонуклеозидмонофосфатов с разной нуклеотидной последовательностью использованы разные компьютерные программы и различные функционалы. Для двух комплексов расчеты выполнены также на уровне теории *ab initio* MP2/6-31++G**. Анализ торсионных углов остова, упаковки сахарного кольца и взаимного положения оснований оптимизированных структур показывает, что конформационные характеристики комплементарных комплексов дезоксирибонуклеозидмонофосфатов с ионами натрия остаются в пределах, характерных для семейства VI, и становятся ближе к соответствующим характеристикам в кристаллах уотсон-криковского дуплекса. Качественно, основные конформационные характеристики каждого из рассмотренных комплементарных комплексов дезоксирибонуклеозидмонофосфатов не зависят от используемых функционала и компьютерной программы, хотя численные значения некоторых конформационных параметров могут варьировать, оставаясь в пределах, характерных для соответствующего семейства. Многие популярные функционалы переоценивают расстояние между парами оснований, а MP2-расчеты и новые сложные функционалы приводят к структурам со слишком близкими межатомными контактами. Детальное изучение некоторых комплементарных комплексов дезоксирибонуклеозидмонофосфатов с ионами натрия указывает на существование нескольких минимумов энергии, соответствующих VI-конформациям, т.е. на сложность рельефа поверхности потенциальной энергии комплементарных комплексов дезоксирибонуклеозидмонофосфатов, что объясняет переменность конформационных параметров фрагментов дуплексов с одинаковой последовательностью. Популярные силовые поля молекулярной механики AMBER и CHARMM воспроизводят многие характеристики дезоксирибонуклеозидмонофосфатов и их комплементарных комплексов с ионами натрия, но не все детали зависимости конформации уотсон-криковского дуплекса от нуклеотидной последовательности.

Ключевые слова: ДНК, компьютерное моделирование, неэмпирические расчеты, теория функционала плотности, молекулярная механика.

Сокращения: СФО – сахаро-фосфатный остов, УКД – уотсон-криковский дуплекс, дДМФ – дезоксирибонуклеозидмонофосфат, кдДМФ – комплементарные комплексы дДМФ, ТФП – теория функционала плотности, ММ – молекулярная механика, ЯЭО – ядерный эффект Оверхаузера.

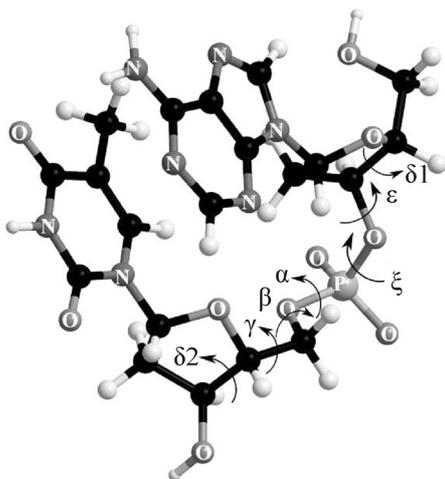


Рис. 1. Обозначения торсионных углов сахаро-фосфатного остова дДМФ.

Замечательным свойствам двойной спирали ДНК, позволяющим ей быть носителем генетической информации, посвящено много публикаций, начиная с 50-х годов прошлого века. В первых же статьях Уотсона и Крика [1,2] авторы открытия описывают биологическую целесообразность двойной спирали и рассматривают, как (пока, правда, только в принципе) решается вопрос о копировании генетического материала и ошибках такого копирования, достаточно редких, но необходимых для биологической эволюции, а значит, и для самой жизни. После этого открытия началось интенсивное изучение физических механизмов того, что Шредингер [3] назвал чудом – воспроизведения и самого существования живого. Первые же исследования молекулы ДНК и составляющих ее субъединиц (азотистые основания, дезоксирибоза, фосфатная группа) выявили удивительную приспособленность ее молекулярной и пространственной структуры к выполнению биологических функций. Но до настоящего времени нельзя считать решенным вопрос о том, какие свойства каких субъединиц определяют существенные для жизни характеристики ДНК. Прежде, чем излагать результаты наших компьютерных расчетов, мы суммируем ту достаточно известную информацию о строении ДНК, которая важна для анализа наших компьютерных расчетов разного уровня сложности.

1. Полинуклеотидная цепь одновременно содержит как практически жесткие и плоские субъединицы (основания), так и гибкие фрагменты (сахаро-фосфатный остов, СФО), что позволяет как хранить информацию (образовывать двойную спираль), так и считывать и копировать ее. Конформационная подвижность остова обеспечивает расплетание цепочек и раз-

ные способы образования внутри- и межцепочечных комплексов.

2. Сахаро-фосфатный остов состоит из одинаковых повторяющихся фрагментов, а вся информация содержится только в последовательности оснований. Остов имеет направление – две последовательные фосфатные группы связаны с атомами кольца дезоксирибозы в разных положениях. Две полинуклеотидные цепочки уотсон-криковского дуплекса (УКД) антипараллельны, последовательность оснований одной цепи определяется последовательностью другой.

3. Комплементарная пара УКД содержит основания разной химической структуры – пурины и пиримидины, различающиеся формой и размерами вдоль большой оси пары. Взаимное положение оснований соседних нуклеотидных пар, а значит, и вся пространственная структура УКД, зависит от последовательности оснований.

Расчеты минимальных фрагментов одиночной цепи ДНК (комплексов дезоксидинуклеозидмонофосфатов с ионами натрия, дДМФ) и минимальных фрагментов двойной спирали (комплементарных комплексов дДМФ, кдДМФ) позволили нам прийти к выводам о существенной роли гомогенного, с точки зрения химической структуры, СФО в зависимости пространственной структуры УКД от последовательности оснований. Эта зависимость уникальна для УКД. Другие двойные спирали не обеспечивают столь важную для биологической роли зависимость пространственной структуры от последовательности оснований.

МЕТОДИКА РАСЧЕТОВ

Рассматриваются минимальные фрагменты одиночной цепи ДНК (нейтрализованные ионом Na^+ дДМФ) и двойной спирали (комплексы кдДМФ). В таком фрагменте одиночной цепи реализуются все существенные для конформационного состояния цепи взаимодействия соседних субъединиц (между основаниями, оснований со своим и соседним сахаром, между соседними сахарами, обоих оснований и сахаров с фосфатной группой, а также каждой из субъединиц с ионом Na^+). В кдДМФ к этим взаимодействиям добавляются взаимодействия между комплементарными нуклеозидами. Молекулярная структура одного из дДМФ и обозначения торсионных углов СФО представлены на рис. 1.

Начальная конфигурация каждого из фрагментов строилась по координатам атомов соответствующих фрагментов экспериментально определенных структур комплементарных дуплексов, приведенных в банке данных нуклеи-

новых кислот [4] (атомы водорода и ионы натрия были помещены нами в наиболее вероятные для соответствующих молекул положения). Для некоторых дДМФ были построены и использованы в расчетах две или три начальные структуры, основываясь на данных для разных дуплексов. Если в банке данных не оказывалось какой-либо из последовательностей для данного семейства конформаций, эта последовательность строилась по координатам атомов другой последовательности, помещая нужное основание в положение того основания, для которого экспериментальные данные имеются.

Расчет и поиск минимумов энергии систем проводились методом теории функционала плотности (ТФП). Этот неэмпирический квантово-механический метод [5], который все шире применяется к разнообразным молекулярным системам, можно рассматривать как разумный компромисс между точностью получаемых результатов и затратами на их проведение. Детали методики расчетов и обоснование применения конкретных методов и компьютерных программ описаны в наших предыдущих работах [6–11]. Здесь мы только упомянем использованные компьютерные программы и некоторые функционалы. В первых наших расчетах дДМФ [6,7] были использована компьютерная программа DGAUSS [12] и функционал PW91 [13]. В последующих расчетах, кроме этой программы, была использована программа ADF (Amsterdam Density Functional) [14]. Эти программы позволили, за счет дополнительных обоснованных приближений, достаточно быстро получить результаты для всех возможных дДМФ, содержащихся в нативной ДНК в разных конфигурациях, провести оптимизацию из разных начальных точек, а затем начать более строгие расчеты по программе Gaussian [15] как для дДМФ, так и для более сложных фрагментов двойной спирали. По этой же программе были рассчитаны два дДМФ и два кдДМФ методом *ab initio* MP2/6-31++G**, а также упомянутые в последнем разделе «Результатов и обсуждения» расчеты взаимодействия отдельных оснований. Для сравнения результатов квантово-механических расчетов с расчетами методом молекулярной механики (ММ) проведены расчеты всех кдДМФ семейства VI и некоторых кдДМФ семейства AI с использованием силового поля [16] и программного обеспечения [17] AMBER, а также силового поля CHARMM [18] по программе NAMD [19]. Спиральные параметры рассчитаны по программе X3DNA [20].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Расчеты минимальных фрагментов одиночной цепи ДНК. Свидетельства важной роли молекулярной структуры и взаимодействий саха-

ро-фосфатного остова. Уже первые наши расчеты дДМФ [6,7], использовавшие конформации семейства VI в качестве исходных структур, изначально для некоторых нуклеозидных последовательностей, позволили сделать ряд интересных и, в некоторой степени, неожиданных выводов, которые были затем [8–11] подтверждены для всех последовательностей четырех семейств УКД. Резюмируя эти выводы здесь, мы учитываем и последующие наши результаты. Оказалось, что для дДМФ всех 16-ти возможных нуклеозидных последовательностей существуют локальные минимумы энергии, соответствующие AI, AII, VI и VII семействам дуплексов с уотсон-криковскими парами. Для некоторых последовательностей семейств AII и VII экспериментальных данных нет (может быть, пока нет), но минимумы энергии для этих дДМФ мы тоже получили. Оптимизированные конформации дДМФ имеют торсионные углы сахара-фосфатного остова и упаковку колец дезоксирибозы, характерные для соответствующего семейства конформаций. В таблице приведены средние значения углов СФО, рассчитанные в работе [21] по результатам анализа банка данных нуклеиновых кислот [4], вместе с усредненными по разным нуклеозидным последовательностям значениями этих углов в оптимизированных структурах дДМФ, описанных в наших предыдущих работах [7–11]. Сами численные значения торсионных углов имеют только приблизительный смысл; новые экспериментальные данные, появившиеся после публикации работы [21], могут сместить значения в некоторых строках на несколько градусов, а доступный сейчас более точный поиск минимумов энергии может изменить на несколько градусов некоторые из торсионных углов ранее рассчитанных структур. Но такие уточнения не могут изменить вывода о том, что экспериментально наблюдаемые В- и А-конформации УКД соответствуют локальным минимумам энергии дДМФ и что существование этих конформаций уже предопределено в минимальных фрагментах одиночной цепи. Дополнительные подтверждения этому соображению были затем [11] получены нами при расчетах минимальных фрагментов СФО ДНК (т.е. для дДМФ, у которых оба основания заменены на атомы водорода). Эти фрагменты имеют локальные минимумы энергии, соответствующие AI-, AII-, VI- и VII-конформациям УКД. Значения торсионных углов в этих минимумах тоже приведены в таблице.

По результатам расчетов все оптимизированные конформации дДМФ характеризуются взаимным положением колец оснований, близким к параллельному, угол между плоскостями колец не превышает 25 градусов (обычно этот угол больше для А-, чем для В-конформаций).

Средние значения торсионных углов сахаро-фосфатного остова четырех семейств уотсон-криковских дуплексов по результатам анализа банка данных нуклеиновых кислот [21] и по результатам расчетов методами ТФП для дДМФ с разными последовательностями, а также для фрагментов сахаро-фосфатного остова

	δ_1	ϵ	ζ	α	β	γ	δ_2
AI [21]	82	206	285	295	173	54	82
AI дДМФ	80	210	290	290	171	58	81
AI СФО	82	210	288	292	176	57	80
AII [21]	85	197	289	146	192	183	85
AII дДМФ	81	204	284	154	201	180	84
AII СФО	83	201	291	155	191	188	80
VI [21]	133	182	263	299	179	48	133
VI дДМФ	139	179	264	296	186	55	136
VI СФО	136	183	260	299	180	56	132
VII [21]	143	251	168	293	143	46	143
VII дДМФ	131	227	175	300	134	46	127
VII СФО	130	242	171	299	134	50	135

В разных рассмотренных структурах перекрытие колец варьирует от существенного до практически нулевого, но отклонения плоскостей колец от параллельности оказываются не зависящими от степени их перекрытия. Этот результат свидетельствует о том, что почти параллельное расположение пар оснований в УКД не является следствием стекинга оснований, т.е. взаимодействия между основаниями, расположенными вдоль цепи ДНК (в таком случае степень параллельности или отклонения от параллельности должна бы зависеть от степени перекрытия колец), а возникает вследствие предпочтительных значений углов СФО в локальных минимумах его энергии.

Еще одним существенным результатом расчетов дДМФ является обнаружение зависимости взаимного положения почти параллельных оснований от их последовательности. В последовательностях Pur-Pur и Pur-Pur наблюдается существенное перекрытие колец оснований, а в последовательностях Pur-Pur и Pur-Pur только незначительное или вообще нулевое перекрытие. Рис. 2 демонстрирует существенное перекрытие оснований в последовательности dArdT, а рис. 3 – практически нулевое перекрытие колец в последовательности dTrdA (расчеты для семейства конформаций AI). Две последовательности тех же оснований имеют существенно разное их взаимное положение, а следовательно, и конформацию фрагмента и его конформационные возможности. Эта зависимость, впервые замеченная для В-семейства конформаций олигонуклеотидных дуплексов [22], исследованных к тому времени методом рентгеноструктурного анализа, неоднократно обсуждалась в связи с гибкостью и изломами

двойной спирали, необходимыми для упаковки ДНК в клетке. Наши результаты обобщают это наблюдение на все уотсон-криковские (и только уотсон-криковские) дуплексы и указывают на существенную роль СФО в существовании такой зависимости.

Могут ли уотсон-криковские дуплексы иметь другую зависимость пространственной структуры от нуклеотидной последовательности? Роль структуры оснований в этой зависимости. Получив описанные выше результаты, мы задались вопросом о том, не являются ли они случайными или предопределенными начальными структурами. Найденные нами структуры дДМФ соответствуют локальным минимумам их энергии. Начальные структуры для оптимизации были выбраны по координатам олигонуклеотидных дуплексов, депонированных в банке данных [4]. Начиная оптимизацию с конформаций, например, VI-семейства, имеющих отмеченную еще Дикерсоном [22] зависимость пространственной структуры от первичной, мы получаем и оптимизированные структуры этого же семейства с той же зависимостью. Необходимо было рассмотреть и гипотетическую возможность того, что, может быть, существуют и другие конформации УКД, которые пока не приведены в банке данных или, еще проще, не попали в нашу выборку начальных структур. Для некоторых последовательностей мы строили две или три начальных конформации, но структур для некоторых последовательностей VI-семейства существуют десятки, все их или большинство из них нельзя использовать для неэмпирических расчетов. Мы попробовали искусственно построить начальную структуру дДМФ, имеющую «неправильное» для такой

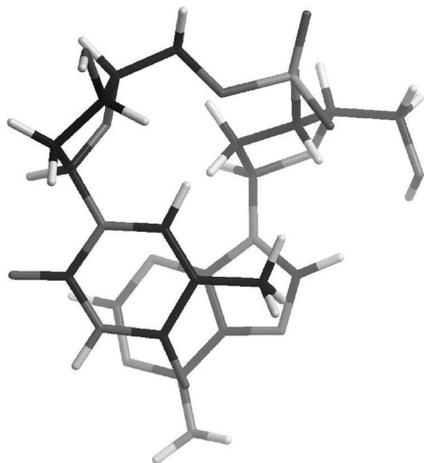


Рис. 2. Оптимизированная AI конформация dApdT.

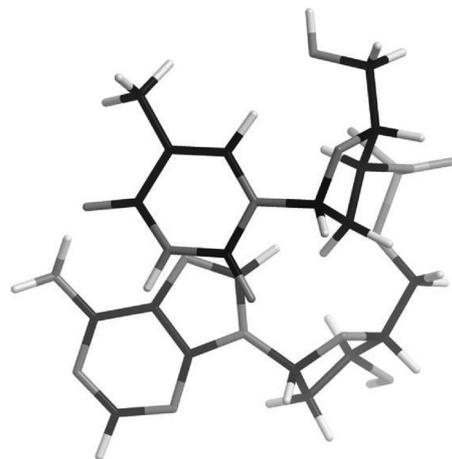


Рис. 3. Оптимизированная AI конформация dTrpdA.

последовательности взаимное положение оснований, т.е. небольшими шагами смещать одно основание относительно другого, оптимизируя на каждом шаге геометрию при закрепленных (сравнительно «мягко», позволяя небольшие смещения) друг относительно друга основаниях. Такое построение нетрудно провести с помощью молекулярной механики, но наши попытки не привели к успеху [10]. По мере изменения взаимного положения оснований в сторону, характерного для другой последовательности (например, для конформации dTrpdA при попытке смещать Thy или Ade в сторону большего перекрытия гетероциклических колец), конформация становилась все менее выгодной энергетически и при этом теряла многие черты, характерные для спиральной структуры.

На следующем этапе выяснения роли отдельных компонентов в зависимости пространственной структуры ДНК от первичной мы заменяли основания (например, пурин на пиримидин, а пиримидин на пурин) в оптимизированной конформации при сохранении положения всех атомов СФО и атомов N1 и N9 оснований. Плоскости колец оснований оставались при таких заменах неизменными [11]. Оказалось, что простая, «механическая» замена приводит к взаимному положению оснований дДМФ, характерному для новой последовательности. Например, заменяем Ade на Thy, а Thy на Ade в оптимизированной конформации dApdT с существенным перекрытием колец оснований, и получаем конформацию dTrpdA, имеющую характерное для такой последовательности взаимное положение практически не перекрывающихся оснований [11]. Оптимизация этой конформации приводит к структуре, не отличимой от полученной ранее при оптимизации этого дДМФ соответствующего семейства. Эти результаты позволяют сделать вывод,

что зависимость пространственной структуры УКД от нуклеотидной последовательности, в частности существенное различие во взаимном положении оснований в последовательностях, имеющих пурин на 5'-конце, от положения в последовательностях, имеющих на этом конце пиримидин, имеет две причины – (1) различия в структуре и форме молекул пурина и пиримидина, в частности, их размеров вдоль большой оси уотсон-криковской пары и (2) различия в положении атомов N9/N1 двух оснований, вызванных «направленностью» СФО и областями его торсионных углов, соответствующих А- и В-семействам конформаций.

Описанные выше закономерности связи пространственной структуры минимальных фрагментов полинуклеотидной цепи со структурой ее субъединиц справедливы только для УКД, но не для других двойных спиралей. Например, конформации дДМФ с последовательностями dGpdC и dCpdG с гуанозином в *син*-конформации, соответствующие левоспиральной Z-форме, далеки от локальных минимумов энергии. После длительной оптимизации эти конформации теряют черты, характерные для Z-формы и вообще спиральной структуры [9]. Для конформаций дДМФ, соответствующих дуплексу с хугстиновскими парами Ade:Thy и *син*-конформацией аденозина, нами были найдены минимумы энергии с конформационными параметрами, близкими к соответствующим параметрам дуплекса в кристалле (структура d(ATATAT) из работы [23]). Оптимизированные конформации dApdT и dTrpdA сохраняют существенное перекрытие колец оснований, характерное для дуплекса в кристалле, т.е. демонстрируют отсутствие зависимости взаимного положения оснований от их последовательности [10,11]. Для последовательности dApdA, отсутствующей в дуплексе d(ATATAT) и по-

строенной нами заменой в последовательностях dApdT и dTpdA тимина на аденин в *син*-конформации, также наблюдается существенное перекрывание колец оснований [11]. Таким образом, существенное перекрывание колец оснований наблюдается в конформациях дДМФ с последовательностями Pur–Pur, Pur–Pur и Pur–Pur, совместимых с образованием хугстиновского дуплекса. В таком дуплексе только в конформации dTpdT перекрывание оснований незначительно, поскольку она идентична конформации для этой последовательности в УКД. Эти результаты позволяют предсказать, что для хугстиновских дуплексов с любой последовательностью аденинов и тимinov (когда они будут экспериментально исследованы) не будет наблюдаться та зависимость взаимного положения оснований от их последовательности, которая характерна для УКД [11]. Две соседние пары оснований в таких дуплексах всегда будут иметь существенное перекрывание оснований по крайней мере в одной из цепей, тогда как соответствующие УКД имеют незначительное перекрывание оснований в обеих цепях для последовательностей Pur–Pur, т.е. хугстиновские дуплексы, в противоположность УКД, не имеют последовательностей, благоприятных для изломов спирали. Таким образом, мы получаем дополнительное свидетельство в пользу того, что хугстиновские пары не могут быть использованы для построения двойных спиралей, выполняющих все функции генетического материала, хотя это не исключает их возможную биологическую роль (см., например, [24]). Конечно, другие характеристики хугстиновских пар (например, возможность их образования для пар Gua:Cyt только при положительном зарядном цитозине, невозможность образования трех водородных связей в паре Gua:Cyt для обеспечения существенно разной стабильности двух пар) уже исключают их использование в построении генетического материала вместо уотсон-криковских пар. Однако, даже если бы природа выбрала для этого другие пурины и пиримидины, исключающие такие ограничения, хугстиновское спаривание оснований не могло бы обеспечить функционирование механизма наследственности.

Еще один тип двойных спиралей, а именно спиралей с параллельными цепями, также должен быть исключен из возможных кандидатов на формирование вещества наследственности. В банке данных нуклеиновых кислот имеются структуры параллельного дуплекса, содержащего пары Ade:Thy в обратной уотсон-криковской конфигурации [25], т.е. пары, в которых водородная связь образуется между аминогруппой аденина и O2 (а не O4) тимина, а вторая связь – такая же, как в уотсон-криковской паре, т.е. N1(Ade)...H–N3(Thy). Эта конформация экспе-

риментально исследовалась методом ядерного магнитного резонанса (ядерный эффект Оверхаузера, ЯЭО), а координаты атомов получены используя молекулярную механику (силовое поле AMBER) с наложением ограничений, следующих из результатов ЯЭО. Авторы предлагают 20 моделей дуплекса. В этих моделях последовательности dApdT:dTpdA характеризуются существенным перекрыванием в цепи dApdT и незначительным в цепи dTpdA (как в УКД), а последовательности dApdA:dTpdT имеют в цепи dTpdT даже большее перекрывание оснований, чем в цепи dApdA (в противоположность УКД). Такие структуры тоже не обеспечивают возможности изломов спирали. Детальный анализ этих структур требует анализа кдДМФ, т.е. минимальных фрагментов двойной спирали, а применение молекулярной механики к построению атомных моделей дуплексов (в частности, силового поля AMBER, как это делается в работе [25]) имеет свои ограничения, о которых мы будем говорить ниже.

Расчеты минимальных фрагментов двойной спирали ДНК. Основные результаты и закономерности, полученные при рассмотрении дДМФ, были подтверждены расчетами минимальных фрагментов двойной спирали, кдДМФ. Одна из оптимизированных конформаций кдДМФ, соответствующая VI-семейству, приведена на рис. 4. Рассчитанные структуры кдДМФ, как правило, несколько ближе к структурам олигонуклеотидных дуплексов в кристаллах. Различия в конформационных характеристиках кдДМФ от дДМФ не изменяют основных выводов о зависимости пространственной структуры фрагментов от нуклеотидной последовательности и роли структурных компонентов в этой зависимости. Эти различия касаются небольших (в пределах соответствующего семейства) изменений некоторых торсионных углов, планарности аминогруппы, образования водородных связей между компонентами дДМФ. Аминогруппы гуанина, аденина и цитозина всех дДМФ, оптимизированных методом ТФП, получаются существенно не плоскими, как и при всех расчетах отдельных оснований неэмпирическими корреляционными методами квантовой механики. Иногда они образуют водородные связи с другим основанием или с дезоксирибозой. В комплементарных комплексах дДМФ такие связи не возникают, аминогруппы, поскольку они образуют водородные связи с атомами кислорода комплементарного основания, становятся практически плоскими, как в кристаллах оснований, нуклеозидов, нуклеотидов. Для некоторых кдДМФ VI-семейства конформаций мы провели оптимизацию геометрии, используя разные функционалы, разные компьютерные программы и разные стартовые

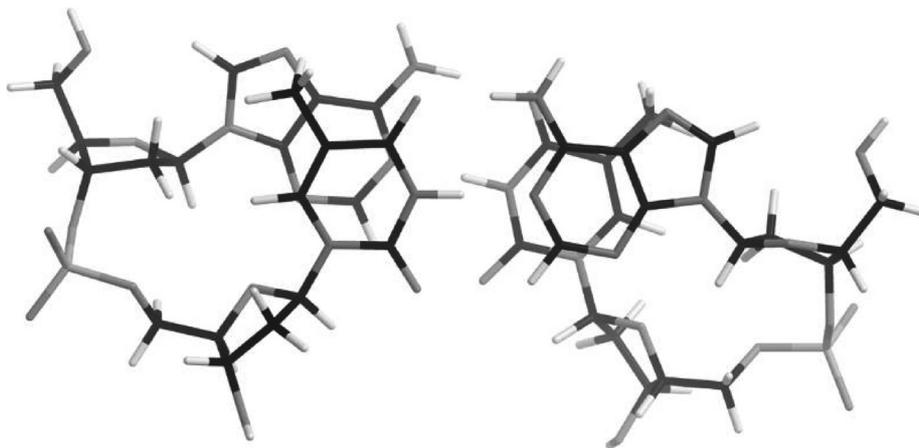


Рис. 4. Локальный минимум энергии фрагмента dApdT:dApdT, конформация семейства VI.

конформации, как построенные по координатам разных структур банка данных нуклеиновых кислот, так и полученные предварительной оптимизацией методом ММ. Хотя численные значения торсионных углов оптимизированных структур могут немного отличаться, конформации остаются в пределах VI-семейства, и закономерности, относящиеся к зависимости взаимного положения оснований, сохраняются. Детальное изучение дуплекса dApdT:dApdT с использованием программы Gaussian [15] показало, что имеется несколько структур, принадлежащих VI-семейству и отвечающих локальным минимумам энергии. Для каждого из них был проведен анализ частот, чтобы подтвердить истинность минимума. Эти минимумы разделены небольшими энергетическими барьерами, отличаются по энергии до 1,5 ккал/мол, имеют торсионные углы, отличающиеся от 1–2 до 20 градусов. Параметры двойной спирали структур, отвечающих разным минимумам, также могут несколько отличаться. Эти результаты, которые, можно ожидать, справедливы и для других кДМФ, демонстрируют сложность рельефа потенциальной энергии минимальных фрагментов УКД. Эта сложность является одной из причин существования конформаций фрагментов олигонуклеотидных дуплексов одного и того же семейства с одной и той же последовательностью оснований, но с разными конформационными параметрами.

Влияние метода расчета на основные характеристики оптимизированных структур дДМФ и кДМФ. Все результаты, описанные выше, хорошо согласуются с экспериментальными данными (структурами кристаллов олигонуклеотидов) и инвариантны по отношению к деталям метода ТФП (функционал, базисный набор функций) и используемому программному обеспечению. Но имеется одно заметное несогласие с экспериментальными структурами, которое

следует специально обсудить и рассмотреть его влияние на выводы, а также возможные пути его устранения. Это несогласие, отмеченное еще в наших первых публикации по результатам расчетов дДМФ методом ТФП [7,9], связано с увеличенным, по сравнению с экспериментальными данными, расстоянием между основаниями. Этого можно было и ожидать. Используемый нами в той работе функционал PW91PW91, как и многие другие популярные функционалы (например, функционал PBE [26], использованный нами для сравнения в расчетах некоторых систем), не требующие больших компьютерных ресурсов, не в полной мере учитывает дисперсионные взаимодействия, т.е. стеккинг оснований не может быть рассчитан правильно. Этот факт явился тогда [7–10] дополнительным указанием на существенную роль СФО в формировании двойной спирали ДНК и в зависимости ее пространственной структуры от последовательности оснований. Рассмотрение кДМФ, т.е. учет взаимодействий между комплементарными цепями, изменяет расстояние между основаниями незначительно. Можно было бы ожидать, что более строгие расчеты рассмотренных систем, использующие новые функционалы, разрабатываемые для учета дисперсионных взаимодействий и требующие существенно больших компьютерных ресурсов, а также расчеты методом *ab initio* с учетом электронной корреляции помогут получить структуры, более близкие к экспериментальным данным, т.е. с меньшими расстоянием между соседними вдоль цепи основаниями или парами. Однако результаты таких расчетов во многом не оправдали наши ожидания. Использование функционалов PWB6K [27] и M06-2X [28] действительно сокращало расстояние между плоскостями оснований, при небольших (по сравнению со структурами, рассчитанными с функционалом PW91PW91 [13] или PBE [26]) изме-

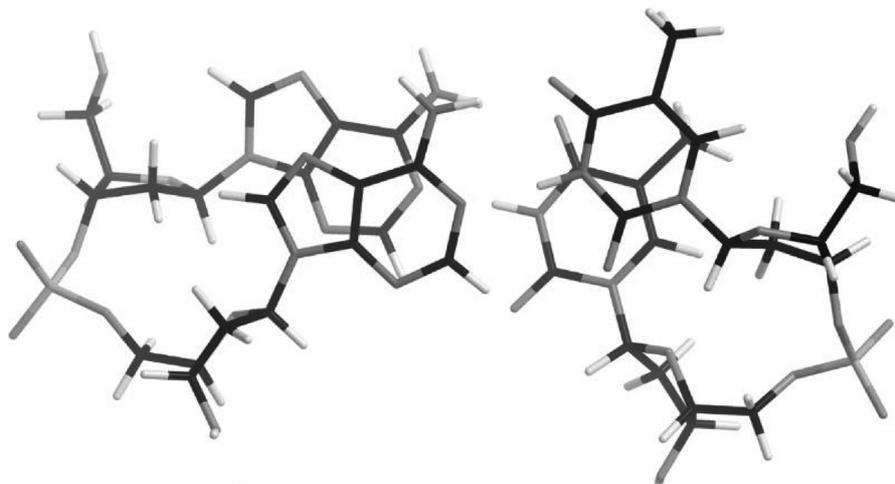


Рис. 5. Наиболее глубокий VI-минимум энергии dApdA:dTpdT, полученный в расчетах с использованием силового поля и программы AMBER.

нениях торсионных углов, конформации дезоксирибозы и взаимного положения оснований, характерного для соответствующей последовательности; при этом время счета увеличивалось в несколько раз. Детальный анализ таких структур выявил сокращенные (до 3,0 Å между атомами колец оснований) межатомные контакты, не встречающиеся в кристаллах не только субъединиц ДНК, но и в других органических молекулах. Структуры, оптимизированные на уровне теории *ab initio* MP2/6-31++G**, т.е. в результате расчетов, учитывающих корреляцию электронов и достаточно широкий базисный набор волновых функций, а поэтому возможных только для единичных систем такой сложности, страдают аналогичным недостатком (отдельные С...С-, С...О- или С...N-расстояния оказываются сокращенными до 2,9 Å). Эти результаты позволяют сделать обобщение, что современные неэмпирические квантово-механические расчеты не позволяют построить физически обоснованную модель минимальных фрагментов ДНК независимо от имеющихся реальных компьютерных ресурсов. Основные конформационные особенности УКД (зависимость пространственной структуры от нуклеотидной последовательности и вклад в ее формирование отдельных субъединиц) были выявлены с помощью расчетов дДМФ методом ТФП с использованием компьютерных программ и функционалов, не требующих больших компьютерных ресурсов, и подтверждены затем упомянутыми здесь более строгими расчетами. Можно надеяться, что дальнейшее совершенствование методики (в частности, разработка новых функционалов, более правильно воспроизводящих дисперсионные взаимодействия) позволит получить физически более приемлемые структуры фрагментов ДНК. Однако стоит заметить, что

такие расчеты уже не могут считаться в строгом смысле «неэмпирическими», так как функционалы содержат подбираемые параметры (например, авторы функционала PWB6K [27] указывают, что он является шестипараметрическим), а во-вторых, они вряд ли выявят новые общие закономерности. Как мы показали ранее, основные закономерности не изменяются по мере уточнения неэмпирических методов.

Применение молекулярной механики к изучению конформационных особенностей уотсон-криковских дуплексов. Возможности усовершенствования метода молекулярной механики с учетом конформационных особенностей уотсон-криковских дуплексов. Применение неэмпирических методов квантовой механики, необходимое для выяснения конформационных особенностей разных нуклеотидных последовательностей и вклада отдельных структурных компонентов в их формирование, ограничено расчетами минимальных фрагментов ДНК. Систематические расчеты этими методами более сложных фрагментов практически невозможны, тогда как метод ММ позволяет рассматривать системы, состоящие из сотен нуклеотидов вместе с окружающей средой, включающей молекулы воды, противоионы и другие молекулы. Результаты, получаемые методом ММ, зависят от применяемого силового поля, т.е. от того, какие и насколько тщательно экспериментальные и квантово-механические результаты и соображения использованы при выборе математических выражений и подбираемых параметров поля. В последние годы наметилась тенденция на усложнение математических формул и повышение роли квантово-механических результатов в выработке силовых полей ММ. Несмотря на появление физически более обоснованных, но более сложных, «поляризуемых» силовых полей,

наиболее часто употребляемыми остаются простые, аддитивные механические модели молекул, возможности усовершенствования которых нам кажутся далеко не исчерпанными [29].

В связи с этим возникает вопрос о том, в какой мере популярные силовые поля ММ воспроизводят те конформационные особенности минимальных фрагментов, которые естественным образом объясняются с помощью расчетов неэмпирическими методами. С этим вопросом связан другой – не могут ли быть общие закономерности пространственной структуры УКД быть использованы для проверки, а может быть, и улучшения метода ММ в целом и отдельных силовых полей в частности?

Мы начали рассмотрение этих вопросов с расчетов тех же простых модельных систем, что обсуждались в предыдущих разделах, с помощью наиболее популярных аддитивных силовых полей AMBER [16] и CHARMM [18]. Здесь мы кратко упомянем результаты оптимизации геометрии ВІ- и АІ-конформаций кдДМФ, проведенной с помощью программы и силового поля AMBER, и первые предположения о возможности улучшения метода ММ с учетом конформационных особенностей разных нуклеозидных последовательностей. Мы провели оптимизацию нескольких десятков кдДМФ, используя в качестве стартовых их конформации в структурах олигонуклеотидных дуплексов, а также конформации, оптимизированные нами ранее методом ТФП. Предварительные результаты позволяют заключить, что многие закономерности, проявляющиеся в неэмпирических расчетах, воспроизводятся в расчетах методом ММ с использованием упомянутых силовых полей. В частности, многие оптимизированные конформации минимальных фрагментов УКД сохраняют такие характерные черты соответствующих семейств конформаций, как области торсионных углов СФО, упаковка сахарного кольца, взаимное положение оснований. Имеет место существенное перекрытие колец оснований в кдДМФ с последовательностями P_{up}-P_{up} и незначительное или нулевое перекрытие для последовательностей P_{up}-P_{up}. При оптимизации разных конформаций кдДМФ с одинаковой последовательностью обнаруживается несколько минимумов энергии, соответствующих одному семейству конформаций. Детальное рассмотрение этих минимумов выявило (пока) одно существенное отличие в результатах, полученных методом ММ, от результатов неэмпирических расчетов. Из нескольких локальных минимумов энергии кдДМФ семейства ВІ с последовательностью dApdA:dTrpdT наиболее глубоким оказывается минимум, в котором существенно перекрываются кольца не только аденинов, но и тиминов (расчеты с использованием силового поля

AMBER). Минимумы с незначительным перекрытием тиминов тоже существуют, но имеют энергию на 1,5 ккал/моль выше, чем наиболее глубокий. Та же тенденция была выявлена при рассмотрении оптимизированных конформаций dGpdG:dCpdC. Мы предполагаем, что силовое поле AMBER (а также CHARMM, но результаты расчетов кдДМФ с этим силовым полем требуют более подробного обсуждения, которое мы пока откладываем) переоценивает стеккинг-взаимодействия оснований и может быть улучшено с учетом результатов расчетов минимальных фрагментов ДНК. Основанием для такого предположения являются не только приведенные выше результаты, но и расчеты различных минимумов энергии взаимодействия между отдельными основаниями, проводимые нами как для разных силовых полей ММ, так и корреляционным методом квантовой механики на уровне MP2/6-31G** (с учетом ошибки суперпозиции базисов). Эти результаты будут представлены в отдельных публикациях, здесь мы только упомянем выводы из них, важные для рассмотрения конформационных характеристик УКД. Все использованные методы показывают наличие трех групп локальных минимумов энергии, соответствующих разным взаимным положениям двух оснований: (1) пары с двумя (для Gua и Cyt – с тремя) водородными связями, многие из этих минимумов соответствуют положению оснований почти в одной плоскости; (2) расположенными друг над другом основаниями в почти параллельных плоскостях, стеккингу оснований; (3) расположению оснований в почти перпендикулярных плоскостях и образованию между ними одной водородной связи. Расчеты для этих трех групп минимумов были проведены нами ранее [30] с помощью нашего собственного силового поля, подобранного специально для расчетов оснований нуклеиновых кислот. Наиболее глубокими минимумами для всех 10 сочетаний оснований (а также их N1- или N9-метилпроизводных) являются минимумы первого типа, а другие минимумы в разной степени менее выгодны [30]. Эта закономерность (а также относительная выгодность разных минимумов энергии для всех последовательностей) воспроизводится в расчетах методом MP2/6-31G**, тогда как результаты расчетов с помощью силовых полей AMBER и CHARMM дают для ряда сочетаний оснований другую последовательность энергий. Мы планируем учесть результаты расчетов простых фрагментов ДНК разными методами в своих предложениях по усовершенствованию силовых полей ММ, которые будут базироваться как на упомянутых выше результатах, так и на расчетах и поиске локальных и глобальных минимумов энергии кристаллов компонентов нуклеиновых кислот.

Тщательный подбор параметров метода ММ особо важен при построении молекулярных моделей при ограниченном наборе экспериментальных данных, например, конформаций биополимеров в водных растворах. Проиллюстрируем это на примере одной из не уотсон-криковских двойных спиралей, уже упомянутых выше, параллельных дуплексов, содержащих пары Ade:Thu в обратной уотсон-криковской конфигурации. В работе [25] конформация содержащего такие пары дуплекса экспериментально исследовалась с помощью ядерного эффекта Оверхаузера, а координаты атомов 20 предложенных моделей были получены, используя силовое поле AMBER с наложением ограничений, следующих из результатов ЯЭО. Предложение нескольких довольно похожих по конформационным характеристикам моделей можно понять, учитывая сложность построения молекулярных моделей, удовлетворяющих одновременно критериям низкой внутримолекулярной энергии и экспериментальным результатам, отражающим взаимное положение атомов водорода. Все модели имеют торсионные углы СФО, близкие к соответствующим углам УКД семейства VI, а параметры спирали и стекинг оснований от них существенно отличаются. Минимальные фрагменты этих моделей, соответствующие последовательностям dArdT:dTrpA, имеют в цепи dArdT существенное перекрытие оснований, а в цепи dTrpA – практически нулевое, как и для УКД. Но фрагменты этих моделей, соответствующие последовательностям dArdA:dTrpT, имеют в цепи dTrpT даже большее перекрытие оснований, чем в цепи dArdA (в противоположность УКД). Используя в качестве стартовых конформаций кДДФ фрагменты ряда этих моделей, мы провели оптимизацию их геометрии методом ТФП (функционалы PW91 и PBE) и методом ММ (силовое поле AMBER). Оба метода позволили найти несколько локальных минимумов энергии. Взаимное положение оснований в оптимизированных методом ТФП структурах демонстрирует существенное перекрытие колец аденинов и незначительное перекрытие тиминных, подобно УКД. Расчеты методом ММ дают наиболее глубокий минимум энергии, соответствующий существенному перекрытию колец тиминных, которое характерно для построенных в работе [25] моделях. Это и не удивительно, поскольку модели строились с использованием того же силового поля AMBER. Применение «улучшенных» силовых полей должно помочь строить более обоснованные молекулярные модели в подобных случаях.

Данная статья подготовлена по материалам доклада на V Съезде биофизиков России (Ростов-на-Дону, 2015). Авторы посвящают эту публикацию памяти талантливого ученого

В.И. Данилова (1936–2014), инициатора серии наших работ по применению метода теории функционала плотности к расчетам минимальных фрагментов ДНК, одного из первых в мире исследователей, начавших успешно применять методы квантовой механики к изучению «главной молекулы жизни».

Работа выполнена при финансовой поддержке Автономного университета Пуэблы (Пуэбла, Мексика).

СПИСОК ЛИТЕРАТУРЫ

1. J. D. Watson and F. H. C. Crick, *Nature* **171**, 737 (1953).
2. J. D. Watson and F. H. C. Crick, *Nature* **171**, 964 (1953).
3. E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell* (Cambridge: University Press, 1944). (Русский перевод: Э. Шрёдингер, *Что такое жизнь? Физический аспект живой клетки*, 3-е изд. (РХД, Ижевск, 2002).
4. H. M. Berman, W. K. Olson, D. L. Beveridge, et al., *Biophys. J.* **63**, 751 (1992).
5. В. Кон, *Успехи физ. наук* **172**, 336 (2002).
6. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *J. Biomol. Struct. Dyn.* **24**, 660 (2007).
7. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *J. Biomol. Struct. Dyn.* **25**, 563 (2008).
8. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *J. Molec. Structure: THEOCHEM* **912**, 53 (2009).
9. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *Int. J. Quant. Chem.* **110**, 2548 (2010).
10. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *Comput. Theor. Chem.* **975**, 69 (2011).
11. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *Biopolymers*, **101**, 640 (2014).
12. C. Sosa, J. Andzelm, B. C. Elkin, et al., *J. Phys. Chem.* **96**, 6630 (1992).
13. J. P. Perdew and Y. Wang, *Phys. Rev B* **45**, 13244 (1992).
14. G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, et al., *J. Comput. Chem.* **22**, 931 (2001).
15. M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., *Gaussian 09, Revision E.01* (Gaussian, Inc., Wallingford, CT, 2011).
16. W. D. Cornell, P. Cieplak, C. I. Bayly, et al., *J. Amer. Chem. Soc.* **117**, 5179 (1995).
17. D. A. Case, T. A. Darden, T. E. Cheatham III, et al., *AMBER 9* (University of California, San Francisco, 2006).
18. K. Vanommeslaeghe and A. D. MacKerell, Jr., *Biochim. Biophys. Acta: General Subjects* **1850**, 861 (2015).
19. J. C. Phillips, R. Braun, and W. Wang, *J. Comput. Chem.* **26**, 1781 (2005).
20. X.-J. Lu and W. K. Olson, *Nucl. Acids Res.* **31**, 5108 (2003).

21. D. Svozil, J. Kalina, M. Omelka, and B. Schneider, *Nucl. Acids Res.* **36**, 3690 (2008).
22. R. E. Dickerson, in *Structure, Motion, Interaction and Expression of Biological Macromolecules*, Ed. by R. H. Sarma and M. H. Sarma (Adenine Press, Albany, NY, 1998), Vol. 1, pp. 17–36.
23. N. G. A. Abrescia, C. Gonzalez, C. Gouyette, and J. A. Subirana, *Biochemistry* **43**, 4092 (2004).
24. M. D. Frank-Kamenetskii and S. M. Mirkin, *Annu. Rev. Biochem.* **64**, 65 (1995).
25. V. R. Parvathy, S. R. Bhaumik, K. V. R. Chary, et al., *Nucl. Acids Res.* **30**, 1500 (2002).
26. J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
27. Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A* **109**, 5656 (2005).
28. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215 (2007).
29. V. Poltev, In: *Handbook of Computational Chemistry*, Ed. by J. Leszczynski (Springer Science), pp. 259–291 (2012).
30. В. И. Полтев, А. С. Дерябина, Э. Гонсалес и Т. И. Грохлина, *Биофизика* **47**, 996 (2002).

Analysis of Conformational Features of Watson-Crick Duplex Fragments by Molecular Mechanics and Quantum Mechanics Methods

V.I. Poltev*, V.M. Anisimov**, C. Sanchez*, A. Deriabina*, E. Gonzalez*, D. Garcia*, F. Rivas*, and N.A. Polteva***

*Autonomous University of Puebla, Puebla, 72570 Mexico

**National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

***Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia

It is generally accepted that the important characteristic features of the Watson–Crick duplex originate from the molecular structure of its subunits. However, it still remains to elucidate what properties of each subunit are responsible for the significant characteristic features of the DNA structure. The computations of deoxydinucleoside monophosphates complexes with Na-ions using density functional theory revealed a pivotal role of DNA conformational properties of single-chain minimal fragments in the development of unique features of the Watson–Crick duplex. We found that directionality of the sugar-phosphate backbone and the preferable ranges of its torsion angles, combined with the difference between purines and pyrimidines in ring bases, define the dependence of three-dimensional structure of the Watson–Crick duplex on nucleotide base sequence. In this work, we extended these density functional theory computations to the minimal fragments of DNA duplex, complementary deoxydinucleoside monophosphates complexes with Na-ions. Using several computational methods and various functionals, we performed a search for energy minima of BI-conformation for complementary deoxydinucleoside monophosphates complexes with different nucleoside sequences. Two sequences are optimized using *ab initio* method at the MP2/6-31++G** level of theory. The analysis of torsion angles, sugar ring puckering and mutual base positions of optimized structures demonstrates that the conformational characteristic features of complementary deoxydinucleoside monophosphates complexes with Na-ions remain within BI ranges and become closer to the corresponding characteristic features of the Watson–Crick duplex crystals. Qualitatively, the main characteristic features of each studied complementary deoxydinucleoside monophosphates complex remain invariant when different computational methods are used, although the quantitative values of some conformational parameters could vary lying within the limits typical for the corresponding family. We observe that popular functionals in density functional theory calculations lead to the overestimated distances between base pairs, while MP2 computations and the newer complex functionals produce the structures that have too close atom–atom contacts. A detailed study of some complementary deoxydinucleoside monophosphate complexes with Na-ions highlights the existence of several energy minima corresponding to BI-conformations, in other words, the complexity of the relief pattern of the potential energy surface of complementary deoxydinucleoside monophosphate complexes. This accounts for variability of conformational parameters of duplex fragments with the same base sequence. Popular molecular mechanics force fields AMBER and CHARMM reproduce most of the conformational characteristics of deoxydinucleoside monophosphates and their complementary complexes with Na-ions but fail to reproduce some details of the dependence of the Watson–Crick duplex conformation on the nucleotide sequence.

Key words: DNA, computer modeling, non-empiric computations, density functional theory, molecular mechanics