

## НЕКОТОРЫЕ ПРОБЛЕМЫ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ НЕВАЛЕНТНЫХ ВЗАИМОДЕЙСТВИЙ ДНК

© 2019 г. В. Полтев, А. Дерябина, В. Домингес, К. Санчес,  
Э. Гонсалес, Н.А. Полтева\*

*Автономный Университет Пуэблы, Пуэбла, 72570, Мексика*

*\*Институт теоретической и экспериментальной биофизики РАН,  
142290, Пушкино Московской области, ул. Институтская, 3*

*E-mail: poltev@scfm.buap.mx*

Поступила в редакцию 07.12.18 г.

После доработки 25.01.19 г.

Принята к публикации 29.01.19 г.

Проанализированы некоторые сложности применения разных методов компьютерного моделирования к исследованию закономерностей формирования пространственной структуры ДНК. Учитывая важность вклада азотистых оснований, проведены расчеты энергии межмолекулярных взаимодействий в разных взаимных положениях молекул для всех десяти попарных комбинаций их метилпроизводных (1-метилпиримидинов и 9-метилпуринов). С использованием разных силовых полей молекулярной механики и квантовомеханических методов *ab initio* и теории функционала плотности найдены локальные минимумы энергии, отвечающие разным взаимным положениям молекул. Подробно рассматриваются минимумы энергии, отвечающие трем типам взаимного положения молекул, соответствующим: 1) образованию двух N–H...O и/или N–H...N водородных связей; 2) почти параллельному расположению и перекрыванию колец (стекингу) оснований и 3) T-образному (близкому к перпендикулярному) положению колец оснований с образованием H-связей. Многие из этих минимумов были получены с использованием методов разной степени сложности, но от метода расчета зависит, какой минимум энергии окажется наиболее глубоким (глобальным) для каждой из комбинаций оснований, а также относительные глубины минимумов. Анализ результатов этих расчетов, расчетов простых фрагментов ДНК, а также обширных литературных данных приводит к заключению, что до сих пор нет метода, который количественно описывал бы все экспериментальные данные о невалентных взаимодействиях ДНК. Сравнение значений энергии и геометрических характеристик комплексов, полученных разными методами, позволяет выявить их возможности и недостатки. Вместе с тем адекватное описание невалентных взаимодействий нуклеиновых кислот требует дополнительных исследований их простых фрагментов и комплексного применения разных методов.

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

**DOI:** 10.1134/S0006302919020042

Правильно оценить энергию взаимодействия между основаниями и другими субъединицами нуклеиновых кислот в разных их взаимных положениях важно для понимания принципов формирования пространственных структур и конформационных возможностей этих биополимеров, а следовательно, и физических механизмов генетических процессов. Уже более

полувекa разные группы исследователей предлагают для этого самые разные методы классической и квантовой физики, от первого, сверхупрощенного описания молекул оснований как точечных молекулярных диполей [1] до недавних, самых строгих из возможных на данный момент неэмпирических методов квантовой механики (КМ) (см., например, работы [2,3]). Количественное рассмотрение взаимодействия между основаниями ДНК было начато в 1962 г. [1] еще до первых попыток применения к таким исследованиям только зарождавшегося тогда метода молекулярной механики (ММ) и методов КМ разного уровня строгости. Хотя

Сокращения: КМ – квантовая механика, ММ – молекулярная механика, ТФП – теория функционала плотности, УКД – уотсон-криковские дуплексы, дДМФ – дезоксирибонуклеозидфосфат, кдДМФ – комплементарные дезоксирибонуклеозидфосфаты, BSSE – суперпозиционная ошибка базисного набора (basis set superposition error).

вскоре после этой публикации было показано, что расчеты в приближении молекулярных диполей не дают не только количественно, но и качественно правильных значений энергии взаимодействия между основаниями, эта статья оказала влияния на многие последующие как расчетно-теоретические, так и экспериментальные исследования взаимодействий компонентов нуклеиновых кислот. Современная компьютерная техника и программное обеспечение позволяют оценить энергию попарных комплексов оснований в разных взаимных положениях и найти ее минимумы как с использованием различных силовых полей ММ, так и достаточно строгими методами КМ. Однако систематический поиск этих минимумов наиболее строгими из существующих КМ-методов до сих пор не доступен. Поэтому все имеющиеся оценки энергии как методами ММ, так и КМ стоит рассматривать как приближенные. Эта работа посвящена расчетам и поиску минимумов энергии взаимодействия между основаниями во всех десяти возможных попарных комбинациях оснований ДНК и систематическому сравнению энергетических и структурных характеристик, полученных разными методами ММ и КМ.

#### ОБОСНОВАНИЕ РАБОТЫ И ПОСТАНОВКА ВОПРОСОВ

В течение последних двенадцати лет мы проводили с помощью квантовомеханического метода теории функционала плотности (ТФП) [4] расчеты минимальных фрагментов одиночной цепи уотсон-криковских дуплексов (УКД), т.е. комплексов дезоксирибонуклеозид монофосфатов с ионами натрия (дДМФ) для всех последовательностей оснований (см., например, работы [5–9]). Мы называем уотсон-криковскими дуплексы с комплементарными парами А:Т и G:C нуклеозидами в антиконформации, у которых обе антипараллельные сахаро-фосфатные цепи имеют торсионные углы в одной из областей, характерных для А- или В-семейств конформаций. Эти расчеты позволили показать, что многие общие закономерности формирования пространственной структуры УКД, включая зависимость пространственной структуры от нуклеотидной последовательности, определяются взаимодействиями субъединиц в дДМФ. Конформационные параметры В1-, В2-, А1- и А2-семейств УКД соответствуют локальным минимумам дДМФ и минимального фрагмента сахаро-фосфатного остова. Расчеты минимальных фрагментов В1- и А1-конформаций УКД, т.е. комплементарных дДМФ (кдДМФ), подтвердили эти выводы [8]. Однако эти выводы справедливы только для УКД, но не для других

конформаций полинуклеотидов, не для других двойных спиралей, таких как Z-форма, параллельные или хугстиновские дуплексы, или дуплексы с неправильными парами оснований [8,9]. Естественным продолжением работы в этом направлении является распространение таких расчетов на минимальные фрагменты других двойных спиралей, в том числе на дДМФ с торсионными углами в других областях (дуплексы с такими фрагментами есть в Банке данных нуклеиновых кислот [10]), на минимальные фрагменты дуплексов с уотсон-криковскими нуклеозидными парами, но с разными конформациями остова двух цепей, а также исследование более протяженных одноцепочечных и двуцепочечных полинуклеотидных фрагментов. Последняя задача в настоящий момент практически не решается КМ-методами и требует применения метода ММ. Многие конформации кдДМФ, оптимизированные разными силовыми полями ММ с использованием в качестве начальной структуры соответствующего фрагмента олигонуклеотидных дуплексов из Банка данных, воспроизводят закономерности формирования их пространственных структур, упомянутые выше. Однако уже первые наши расчеты [8] дДМФ и кдДМФ с помощью наиболее распространенных силовых полей АМБЕР [11] или CHARM [12] не позволили воспроизвести все те закономерности формирования пространственной структуры, которые были выявлены в расчетах методом ТФП и соответствуют экспериментальным структурам. В частности, не воспроизводится существенное отличие перекрывания колец оснований в последовательностях «пурин–пурин» и «пурин–пиримидин» от последовательностей «пиримидин–пиримидин» и «пиримидин–пурин». Эти результаты позволили нам высказать предположение о переоценке стекинг-взаимодействий данными силовыми полями [8], что мы и наблюдали в последующих, более систематических расчетах методом ММ. В ряде случаев оптимизированные конформации кдДМФ не соответствуют результатам КМ-расчетов и экспериментальным данным, а иногда и не могут быть фрагментами протяженной двойной спирали. Это указывает на необходимость дальнейшего уточнения описания взаимодействий между основаниями в моделях силовых полей ММ.

Непростым оказалось и распространение расчетов методом ТФП на фрагменты дуплексов с разными конформациями сахаро-фосфатного остова двух цепей. Используя ранее простые функционалы, не учитывающие в полной мере дисперсионные взаимодействия, а значит, и стекинг оснований, часто приводят к конформациям с большим углом

между парами оснований, не соответствующим экспериментальным данным, не совместимым с протяженной двойной спиралью. Применение других, более сложных функционалов, в том числе специально разработанных для учета стекинг-взаимодействий, а также «гибридного» метода ТФП с учетом дисперсионных взаимодействий в виде дополнительного слагаемого энергии, часто приводит к структурам со слишком короткими межатомными расстояниями между некоторыми атомами соседних по цепи оснований. Для многих случаев такие расчеты с применением метода ТФП дают структуры, похожие на результаты оптимизации методом ММ, но только с еще одним недостатком – нереалистично короткими межатомными расстояниями. Мы уже отмечали в прежних работах [7–9], что использование наиболее строгих из возможных сейчас методов *ab initio*, учитывающих корреляцию электронов и использующих довольно широкий набор базисных волновых функций, также приводит к конформациям дДМФ и кдДМФ с недопустимо короткими расстояниями между некоторыми атомами оснований.

Все эти трудности заставляют нас вернуться к систематическому рассмотрению одной из наиболее простых систем, относящихся к нуклеиновым кислотам – попарным комплексам, состоящим из двух оснований в разных взаимных положениях.

## МЕТОДИКА РАСЧЕТОВ

**Рассматриваемые системы.** Расчеты и поиск минимумов энергии межмолекулярных взаимодействий методами ММ проводили для всех попарных сочетаний оснований ДНК и их метилпроизводных (1-метилпиримидинов и 9-метилпуринов), используя в качестве начальных самые разные их взаимные положения. Эти положения выбирали как случайным образом, так и близкими к предполагаемым минимумам энергии, в том числе к найденным нами или другими авторами ранее. При расчетах КМ-методами начальные взаимные положения и геометрия молекул выбирались соответствующими (или близкими) полученным при оптимизации ММ-методами. Мы будем приводить и обсуждать количественные данные, относящиеся к взаимодействию метилпроизводных оснований, молекулярная структура и обозначения атомов которых приведены на рис. 1. Эти молекулы ближе соответствуют основаниям полинуклеотидной цепочки, чем пурины и пиримидины без метильной группы, результаты для которых мы будем упоминать только в некоторых случаях. Мы для краткости используем

однобуквенные обозначения оснований и их метилпроизводных. Еще до начала систематических расчетов было ясно, что для многих сочетаний оснований общее количество минимумов, получаемых каждым из методов, довольно велико. Построение ММ- и КМ-методами поверхностей потенциальной энергии в шестимерном пространстве взаимных положений для каждой из пар оснований и последующий анализ этих поверхностей представляет собой слишком трудоемкую задачу, большую часть результатов которой было бы практически невозможно сравнить с экспериментальными данными или использовать для их интерпретации. Мы рассмотрим здесь наиболее глубокие минимумы, отвечающие трем типам взаимного положения молекул. Первый тип соответствует образованию двух (а для пары G:C и трех) Н-связей N–H...O и/или N–H...N. Обе молекулы во многих, но не во всех таких минимумах расположены почти в одной плоскости. В положениях второго типа плоскости колец двух оснований расположены параллельно (или, при расчетах некоторыми из методов, почти параллельно, угол между плоскостями может превышать  $10^\circ$ ), так называемые стекинг-положения. Положения третьего типа соответствуют почти перпендикулярному (Т-образному) положению молекул и образованию одной Н-связи. Для некоторых сочетаний оснований при расчетах КМ-методами угол между плоскостями колец может отклоняться от  $90^\circ$  и составлять не более  $70^\circ$ . Для взаимных положений молекул, отвечающих первому типу, мы можем быть уверены, что найдены все возможные минимумы, хотя и не все эти минимумы воспроизводятся каждым из использованных нами ММ- и КМ-методов. Для положений второго и особенно третьего типа мы уверены только в том, что найдены наиболее глубокие для этого сочетания оснований минимумы, но не можем исключить существования не найденных нами минимумов, отделенных от полученных невысокими барьерами. Примеры взаимного положения оснований в минимумах энергии взаимодействия, соответствующих этим трем типам, приведены на рис. 2.

Для каждого из найденных минимумов энергии комплексов, отвечающих разным типам взаимного положения оснований, мы приводим значения энергии взаимодействия как разности энергии комплекса и суммы энергий двух свободных оснований в их минимумах энергии (эту энергию часто называют энергией связывания, *binding energy*). Такое «термодинамическое» определение величины энергии межмолекулярных взаимодействий отличается от другого, часто используемого в КМ-исследованиях,

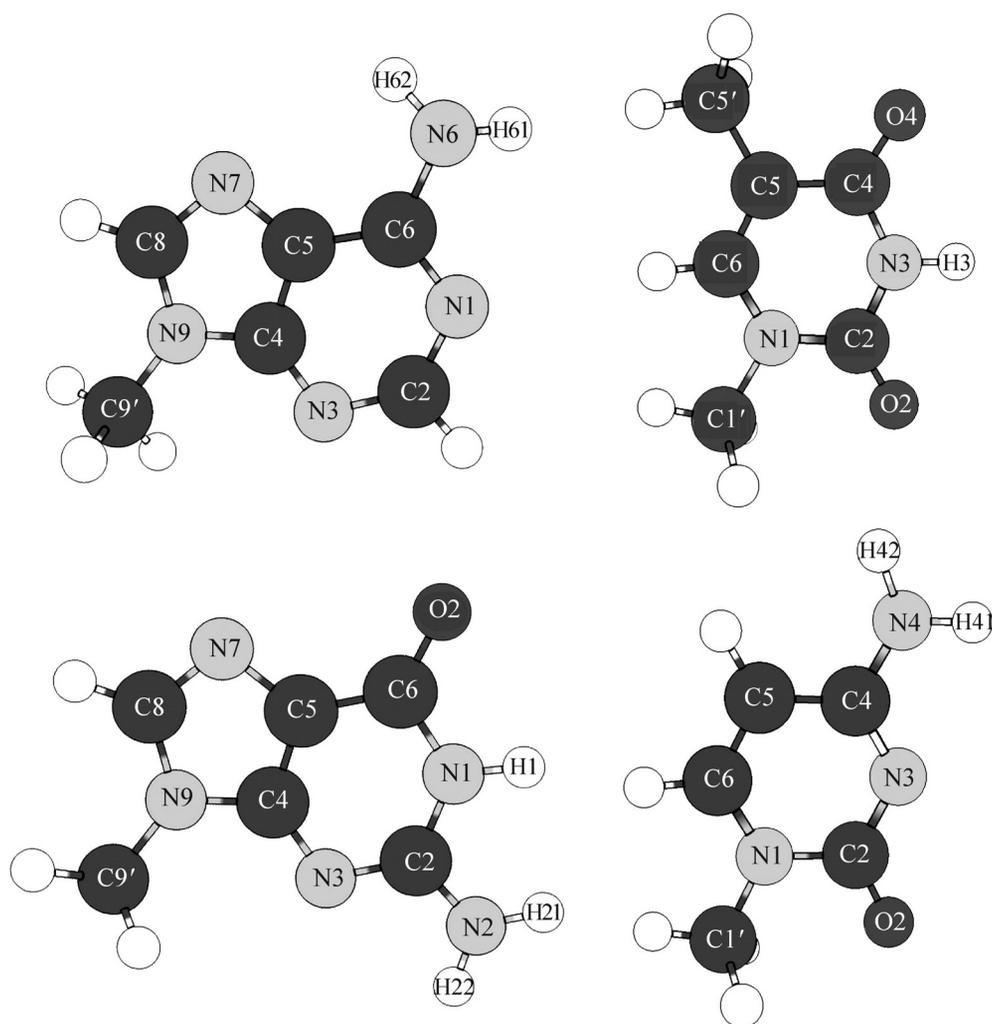


Рис. 1. Молекулярная структура и обозначения атомов обсуждаемых метилпроизводных оснований.

например в работах [2,3], соответствующего разности энергии комплекса и суммы энергий его компонентов в конфигурации, соответствующей минимуму энергии комплекса (в программе GAUSSIAN [13] такую энергию называют энергией комплексообразования, complexation energy). Разница этих двух энергий зависит от того, насколько геометрия (а, значит и энергия) отдельных молекул изменяется при образовании комплекса и может, по нашим оценкам, превышать для некоторых минимумов 1,5 ккал/моль.

Мы не обсуждаем здесь целый ряд других, менее глубоких минимумов энергии, например, минимумы, отвечающие положению оснований в одной (или почти одной) плоскости с образованием только одной Н-связи N–H...O или N–H...N и еще одной слабой Н-связи C–H...O или C–H... N (ранее [14] мы рассматривали некоторые из них), минимумы с одной слабой Н-связью C–H...O или C–H... N, а также многие

минимумы, которые обнаруживаются только одним или двумя из использованных методов.

**Метод молекулярной механики.** Этот полупырический метод классической физики сейчас редко используется для расчетов столь простых систем, как рассматриваемые в этой работе. В последние годы с его помощью обычно рассматриваются фрагменты биополимеров и их комплексы, содержащие тысячи атомов. Обзор метода, его история, ограничения и перспективы были опубликованы в 2012 и 2017 гг. одним из авторов этой работы в виде глав в справочниках по вычислительной химии [15,16]. Здесь мы возвращаемся к применению этого метода для расчетов простых систем, чтобы обсудить возможности и ограничения разных силовых полей при расчетах взаимодействий между основаниями ДНК и сравнить результаты таких расчетов с результатами, полученными разными КМ-методами. Мы используем здесь

только «аддитивные» силовые поля, не учитывающие поляризационные (индукционные) взаимодействия. Общая формула энергии системы при этом (формула 1) содержит слагаемые, описывающие попарные взаимодействия атомов, не связанных химическими связями между собой или с третьим атомом ( $\Sigma E_{nb}$ ), члены, отвечающие за изменения всех длин связей ( $E_b$ ), валентных и торсионных углов ( $E_a$  и  $E_t$  соответственно):

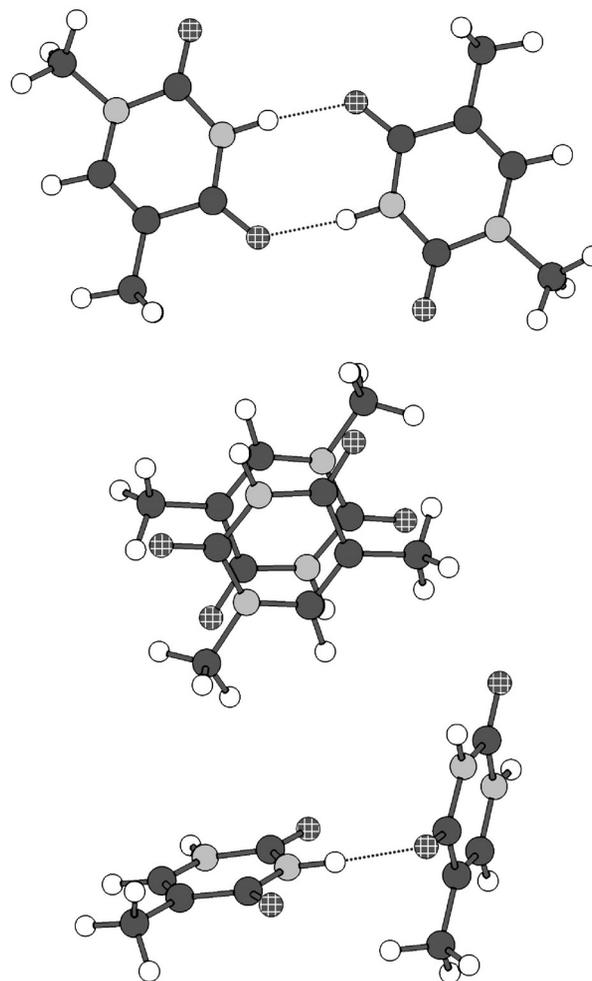
$$E = \Sigma E_{nb} + \Sigma E_b + \Sigma E_a + \Sigma E_t. \quad (1)$$

Каждое слагаемое невалентных атом-атомных взаимодействий содержит кулоновский член (взаимодействие эффективных зарядов на атомах  $q_i$  и  $q_j$ ) и ван-дер-ваальсовы члены притяжения и отталкивания в виде 1-6-12 потенциала (формула 2):

$$E_{ij}(r_{ij}) = Kq_iq_j/r_{ij} - A_{ij}/r_{ij}^6 + B_{ij}/r_{ij}^{12}. \quad (2)$$

В этой формуле  $r_{ij}$  – расстояние между атомами  $i$  и  $j$ ,  $A_{ij}$  и  $B_{ij}$  – полуэмпирические параметры. Разные силовые поля отличаются друг от друга небольшими изменениями в формулах для расчета слагаемых, иногда введением дополнительных слагаемых и, что наиболее существенно, полуэмпирическими параметрами в выражениях для расчета каждого из слагаемых. Эти различия параметров (а также величин отдельных слагаемых энергии) могут быть довольно большими (иногда в несколько раз), но суммарные энергии для разных систем параметров обычно отличаются не столь сильно, но часто довольно существенно для обсуждения свойств системы.

Мы проводили расчеты с использованием наиболее часто применяемых для расчетов биополимеров силовых полей AMBER [11] и CHARMM [12], а также нашего силового поля, специально подобранного для расчетов оснований нуклеиновых кислот и родственных соединений, последняя модификация которого была опубликована в 2002 г. [14]. Использование этого силового поля предполагает жесткую структуру молекулы (кроме безбарьерного вращения метильной группы вокруг C–C- или C–N-связей), т.е. в формуле (1) учитываются только слагаемые  $E_{nb}$ . Обоснованием этому является тот факт, что геометрия оснований в разных кристаллах мономеров и изменения геометрии оснований при всех биологически существенных взаимодействиях незначительны. Отличительной чертой этого силового поля (мы будем обозначать его в дальнейшем MM1, а поля AMBER и CHARMM – MM2 и MM3 соответственно) является потенциал 1-10-12 (формула 3) для расчетов взаимодействий атомов водо-



**Рис. 2.** Три типа взаимных положений двух молекул оснований в локальных минимумах энергии взаимодействия на примере двух молекул 1-метилтимина. Атомы углерода обозначены черным цветом, водорода – белым, азота – серым, кислорода – клетчатými кружками; пунктиром показаны водородные связи. Сверху – пара с двумя Н-связями, в центре – стекнинг-положение, справа – перпендикулярное положение.

рода, способных образовывать водородные связи, с потенциальными атомами-акцепторами водорода:

$$E_{ij}(r_{ij}) = Kq_iq_j/r_{ij} - A_{ij}/r_{ij}^{10} + B_{ij}/r_{ij}^{12}. \quad (3)$$

Этот тип потенциала использовался также в ранних версиях силового поля AMBER и, на наш взгляд, обеспечивает лучшее, чем потенциал 1-6-12, воспроизведение зависимости энергии Н-связей от расстояния. Для расчетов с использованием силового поля AMBER применяли разные версии программы AMBER (например AMBER 2016 [17]), а с использованием силового поля CHARMM – программу NAMD [18].

**Квантовомеханические расчеты методом *ab initio*.** При выборе уровня теории для расчетов *ab initio* мы исходили из того, чтоб расчеты были достаточно строгими и, с другой стороны, не требовали слишком больших компьютерных ресурсов для проведения многих сотен оптимизаций из разных начальных положений. Этому критерию удовлетворяет стандартный уровень теории MP2/6-31G(d, p), учитывающий корреляцию электронов и достаточно широкий набор базисных функций. Мы будем обозначать этот метод как MP2. Выборочные пробные расчеты с более расширенными наборами базисных функций (до aug-cc-pvtz) не улучшали соответствие характеристик минимумов энергии экспериментальным данным. С другой стороны, учет поправок на суперпозиционную ошибку базисного набора по Бойс и Бернарди [19] (учет basis set superposition error, BSSE) существенно улучшал согласие с экспериментальными данными рассчитанных значений энергии в минимумах и их геометрии. Этот метод расчета в дальнейшем будем обозначать как MP2/BSSE. Все обсуждаемые нами характеристики относятся к истинным минимумам энергии, что для КМ-методов каждый раз подтверждалось расчетом матриц вторых производных энергий.

**Квантовомеханические расчеты методом теории функционала плотности.** Этот неэмпирический квантово-механический метод [4] можно рассматривать как разумный компромисс между точностью получаемых результатов и затратами на их проведение. Вследствие замены многоэлектронной волновой функции  $3N$  переменных ( $N$  – число электронов) функцией электронной плотности (три пространственных переменных) метод ТФП требует существенно меньших компьютерных ресурсов, поэтому в последние десятилетия он стал наиболее популярным методом расчета электронной структуры и свойств биоорганических молекул. Мы не проводили систематических расчетов энергии взаимодействия для всех попарных сочетаний оснований методом ТФП. Разные минимумы энергии для избранных попарных сочетаний оснований рассчитывались для сравнения с результатами расчетов методами ММ и *ab initio*, а также для сравнения с результатами расчетов дДМФ и кдДМФ разными методами. В наших расчетах использован ряд наиболее популярных функционалов этого метода (PW91, PBE, B3LYP).

В связи с трудностями описания дисперсионных взаимодействий методом ТФП широкое распространение получил метод ТФП-Д, который предполагает включение в энергию системы дополнительного члена, обратно пропорцио-

нального шестой степени расстояния [20]. Мы рассматриваем этот метод и его недавнюю модификацию, называемую обычно ДФТ-Д3, когда для параметризации дисперсионного члена используются характеристики атомов, рассчитанные методом *ab initio* [21], и обозначаем его как «гибридный квантово-классический». Для метода ТФП-Д3 предложен целый ряд функционалов, мы рассчитали ряд минимумов энергии взаимодействия оснований (а также конформации дДМФ и кдДМФ, которые здесь не описываются), используя те из них, что наиболее часто применяются для расчетов биомолекулярных структур. Расчеты методом ТФП проведены по программе GAUSSIAN 16 [13].

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

**Минимумы энергии взаимодействия оснований, соответствующие образованию между ними двух или трех водородных связей.** Образование двух (для пары G:C) и трех Н-связей между основаниями определяет специфичность взаимодействий между нуклеотидами при биосинтезе нуклеиновых кислот, при формировании пространственной структуры ДНК и РНК, при взаимодействиях кодон-антикодон. Ранее [14] нами методом ММ1 были найдены 28 таких минимумов энергии межмолекулярных взаимодействий 9-метилпуринов и 1-метилпиримидинов. Положения, близкие ко многим из этих минимумов, обнаружены экспериментально в кристаллах компонентов или фрагментов ДНК и РНК. В этой работе мы провели оптимизацию геометрии каждого из этих минимумов с использованием разных вычислительных методов. Взаимные положения молекул в каждом из 28 минимумов, найденных с помощью силовых полей ММ1 и ММ2, оказались близки, а 26 из этих положений близки и к минимумам, найденным методом ММ3 и КМ-методами *ab initio* и ТФП.

Для неметилированных оснований все эти минимумы тоже существуют и близки (по энергии и по взаимному положению молекул) к соответствующим минимумам, полученным тем же методом для 9-метилпуринов и 1-метилпиримидинов, но для них существуют также минимумы, в которых образуются Н-связи с участием групп N1–Н пиримидинов и/или N9–Н пуринов. Положения, близкие к таким минимумам, имеются в кристаллах некоторых оснований, но в нуклеиновых кислотах они реализоваться не могут, и мы их здесь не рассматриваем.

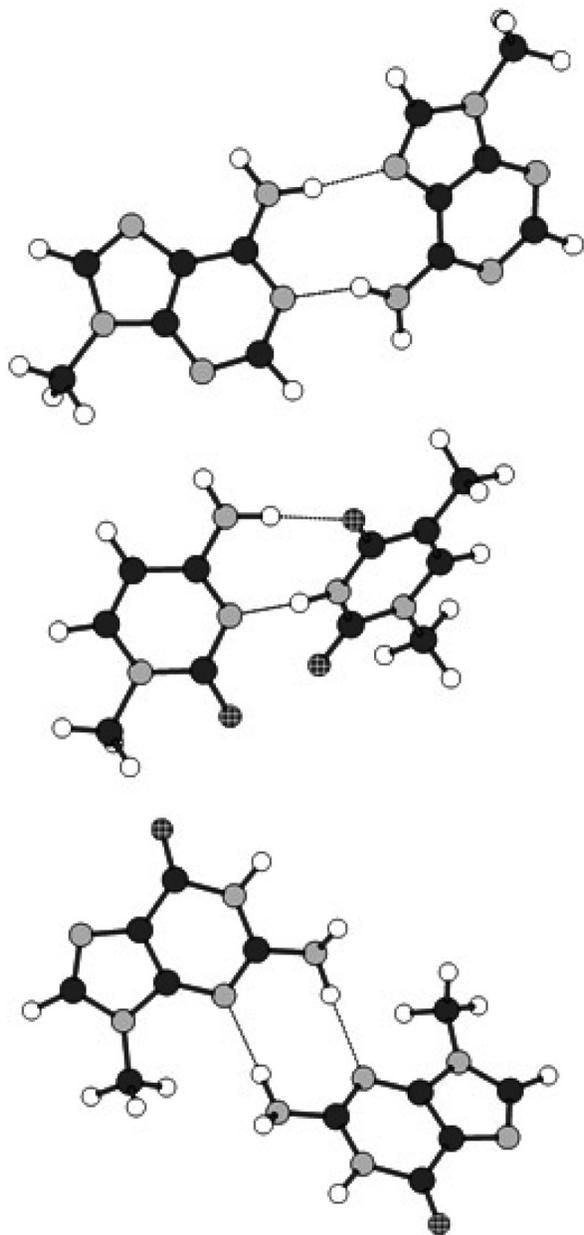
Для разных сочетаний метилпроизводных оснований количество минимумов, соответствующих образованию Н-связей, меняется от од-

**Таблица 1.** Значения энергии взаимодействия (ккал/мол) в локальных минимумах, соответствующих образованию двух или трех Н-связей, для некоторых попарных сочетаний 9-метилпуринов и 1-метилпиримидинов

Основания	№	Н-связи	ММ1	ММ2	ММ3	МР2	МР2/BSSE
m <sup>9</sup> A : m <sup>1</sup> T	1	N6-H62...O4, N7...H3-N3	-12,6	-13,3	-13,8	-18,9	-13,4
	2	N7...H3-N3, N6-H62...O2	-12,5	-13,5	-13,6	-18,7	-13,2
	3	N6-H61...O4, N1...H3-N3	-11,9	-11,9	-12,0	-18,0	-12,6
	4	N1...H3-N3, N6-H61...O2	-11,4	-12,0	-12,0	-17,6	-12,4
m <sup>9</sup> A : m <sup>9</sup> A	5	N6-H61...N1, N1...H61-N6	-11,3	-9,1	-9,3	-15,4	-11,2
	6	N7...H61-N6, N6-H62...N1	-11,3	-10,1	-11,7	-14,8	-10,7
	7	N7...H62-N6, N6-H62...N7	-11,2	-10,9	-12,9	-14,2	-10,0
m <sup>9</sup> G : m <sup>1</sup> C	8	O6...H41-N4, N1-H...N3, N2-H21...O2	-25,1	-24,6	-24,7	-30,4	-23,8
	9	N1-H1...O2, N2-H21...N3	-17,7	-16,4	-	-	-
	10	N2-H22...N3, N3...H41-N4	-16,8	-13,5	-16,6	-17,4	-12,8
m <sup>9</sup> G : m <sup>9</sup> G	11	O6...H1-N1, N1-H1...O6	-20,5	-24,0	-22,1	-28,5	-21,9
	12	N7...H1-N1, O6...H21-N2	-18,6	-15,4	-17,8	-22,4	-17,0
	13	N2-H22...N3, N3...H22-N2	-14,0	-10,4	-13,2	-13,7	-9,2
m <sup>1</sup> C : m <sup>1</sup> C	14	N3...H4-N41, N4-H41...N3	-17,1	-15,7	-18,0	-22,2	-17,4
m <sup>1</sup> C : m <sup>1</sup> T	15	N3...H3-N3, N4-H41...O4	-12,2	-11,6	-10,4	-16,4	-11,3
	16	N3...H3-N3, N4-H41...O2	-11,3	-11,6	-10,9	-15,6	-10,8

ного (для пары m<sup>1</sup>C : m<sup>1</sup>C) до четырех (для пар m<sup>9</sup>A : m<sup>1</sup>T и m<sup>9</sup>A : m<sup>9</sup>G). Примеры значений энергии в минимумах, полученных методами ММ и *ab initio*, приведены в табл. 1. Положения, отвечающие образованию Н-связей N1-H1...O2 и N2-H21...N3 в паре m<sup>9</sup>G : m<sup>1</sup>C (№ 9 в табл. 1), а также Н-связей N2-H21...N7 и N1-H1...O6 в паре m<sup>9</sup>G : m<sup>9</sup>G соответствуют минимумам энергии при расчетах методами ММ1 и ММ2, но при расчетах методами *ab initio* и с помощью силового поля ММ3 минимумами не являются. Большинство пар оснований в минимумах с двумя Н-связями – практически плоские при расчетах любым из используемых здесь методов. При расчетах КМ-методами некоторых из этих пар (в том числе уотсон-криковских пар m<sup>9</sup>A : m<sup>1</sup>T и m<sup>9</sup>G : m<sup>1</sup>C) минимумы энергии соответствуют небольшому углу между плоскостями колец оснований. Например, этот угол в минимуме для уотсон-криковской пары m<sup>9</sup>G : m<sup>1</sup>C, рассчитанном методом МР2/BSSE, составляет 10,6°. Небольшой (5,2°) угол в этой паре получается и в расчетах методом ММ1, но не другими ММ-методами. Этот сам по себе интересный результат, видимо, вряд ли существенен для структуры ДНК, поскольку взаимодействия с другими субъединицами нуклеиновых кислот часто вызывают большие изменения углов между основаниями в Н-связанных парах при небольшом изменении энергии взаимодействия в парах.

Вместе с тем для ряда пар имеются минимумы, соответствующие значительному углу между плоскостями колец оснований, который, при расчетах некоторыми методами, может превышать 45°. На рис. 3 представлены примеры таких пар. Непланарность некоторых из этих пар является следствием слишком коротких межатомных контактов при положении двух молекул в одной плоскости, например, между атомами O2 в парах m<sup>1</sup>T : m<sup>1</sup>C (рис. 3), когда угол между плоскостями оснований составляет при расчетах ММ методами от 42° до 45°, а при расчетах методами МР2 и МР2/BSSE равен 30° и 34° соответственно. Другие минимумы, в которых основания не копланарны, не могут быть объяснены просто стерическим отталкиванием атомов, а являются следствием взаимодействия зарядовых распределений молекул. Значения угла между плоскостями молекул зависят от метода расчета, изменяясь, например, для пары m<sup>9</sup>A : m<sup>9</sup>A от 31° (метод МР2/BSSE) до 48° (метод ММ3). В минимумах для пар с Н-связями, найденных КМ-методами, аминокислотные группы оснований обычно сохраняют несколько неплоскую конфигурацию, характерную для отдельных оснований в минимумах энергии, найденных корреляционными КМ-методами. ММ-расчеты для всех сочетаний оснований, как и можно было предположить, приводят к минимумам, соответствующим плоским аминокислотным группам.



**Рис. 3.** Примеры взаимных положений молекул в минимумах энергии, отвечающих образованию двух Н-связей и расположению колец оснований в разных плоскостях. Проекция на плоскость кольца правого основания. Обозначения те же, что на рис. 2. Сверху – пара m9A : m9A (№ 6 в табл. 1), в центре – пара m1C : m1T (№ 15 в табл. 1), внизу – пара m9G : m9G (№ 13 в табл. 1).

Для большинства как плоских, так и неплюских пар с Н-связями взаимные положения молекул, полученные разными методами расчета, довольно близки и соответствуют почти линейным Н-связям (настолько линейным, насколько взаимное положение атомов в молекулах позволяет образование двух Н-связей). Наиболее существенные отличия для каждого из

минимумов относятся к длинам Н-связей, которые могут варьировать в пределах  $\pm 0,15$  Å. Приведем только один пример изменения длин Н-связей при использовании разных методов расчета для минимума, отвечающего образованию пары m<sup>1</sup>T : m<sup>1</sup>T с двумя почти линейными Н-связями N3–H3...O4. Это расстояние при расчетах методами ММ составляет 1,91 Å (потенциалы ММ1), 1,85 Å (ММ2) и 1,84 Å (ММ3). Расчеты методом МР2 дают растяжение 1,84 Å, а МР2/BSSE – 1,92 Å. Этот минимум, как и многие другие, воспроизводится и в расчетах методом ТФП как в классическом варианте с использованием разных функционалов, так и ТФП с дополнительным членом, учитывающим дисперсионные взаимодействия. Например, значения расстояния Н3...О4 (1,79 Å – с функционалом PBE, 1,84 Å – с функционалом B3LYP, 1,81 Å – B3LYP-D3) близки к найденным методами ММ и *ab initio*. Подобные небольшие, но систематические, различия наблюдаются и при сравнении результатов, полученных разными методами расчета, для других минимумов с двумя Н-связями. Таким образом, все использованные в нашей работе вычислительные методы воспроизводят положения оснований в кристаллах, соответствующие образованию между ними Н-связей, но абсолютные величины энергий в каждом из минимумов и относительные энергии в разных минимумах зависят от метода расчета.

Немногочисленные количественные экспериментальные данные, с которыми можно сравнивать рассчитанные величины энергии взаимодействия, – это результаты масс-спектрометрических исследований ассоциации метилпроизводных оснований в вакууме [22–24]. Оцененные по этим данным величины энтальпии ассоциации можно непосредственно сравнивать со значениями энергии в наиболее глубоких минимумах, отвечающих образованию Н-связей между основаниями. Стекинг-ассоциаты, энтальпия образования которых существенно меньше, чем для пар с Н-связями, экспериментально наблюдались в вакууме только для оснований, метилированных по всем протон-донорным группам, т.е. когда пары с Н-связями невозможны (например, 1,4,4-триметилцитозин или 6,6,9-триметиладенин) [23]. К сожалению, экспериментальные данные об энтальпиях ассоциации имеются только для шести сочетаний метил-производных, точность оценок при этом не превышает 1,5–2,0 ккал/моль, а отнесение экспериментальной величины к определенной конфигурации пары возможно только с привлечением теоретических рассуждений.

В табл. 2 приведены значения энергии взаимодействия в наиболее глубоких минимумах

**Таблица 2.** Сравнение наиболее глубоких минимумов энергии Н-связанных пар, рассчитанных разными методами, с экспериментальными значениями энтальпии ассоциации (ккал/моль)

Метод	Пары					
	m <sup>1</sup> T : m <sup>1</sup> T	m <sup>9</sup> A : m <sup>1</sup> T	M <sup>9</sup> A : m <sup>1</sup> C	M <sup>9</sup> G : m <sup>1</sup> T	m <sup>1</sup> C : m <sup>1</sup> C	M <sup>9</sup> G : m <sup>1</sup> C
ΔН	-9,5	-13,0	-13,5	-14,5	-17,0	-21,0
ММ1	-10,1	-12,6	<u>-13,5</u>	<u>-13,4</u>	-17,1	-25,1
ММ2	-11,7* <sup>+</sup>	<u>-13,5</u> <sup>+</sup>	<u>-12,4</u> * <sup>+</sup>	-15,3	-15,7	-24,6
ММ3	-10,2*	<u>-13,8</u> *	<u>-13,1</u> * <sup>+</sup>	-13,4 <sup>+</sup>	-18,0*	-24,7
МР2	-15,4*	<u>-18,9</u>	<u>-18,0</u>	-19,7	-22,2	-30,4
МР2/BSSE	-10,4	-13,0	-13,4	-14,1	-17,2	-23,6

Примечание. Курсивом выделены значения энергии, превышающие по абсолютной величине экспериментальные оценки более чем на 2 ккал/моль. Подчеркнуты величины энергии для пар, относительная стабильность которых не соответствует экспериментальным данным. Отмечены значения, не соответствующие глобальному минимуму энергии для данного сочетания оснований, полученному этим методом (\*), и значения, соответствующие другим парам с Н-связями, чем для ассоциатов, рассчитанных остальными методами (<sup>+</sup>).

для этих шести сочетаний, найденные методами ММ и *ab initio*, и экспериментально оцененные энтальпии ассоциации, которые не зависят от температуры во всем интервале измерений. Только две группы рассчитанных значений (столбец m<sup>1</sup>G : m<sup>1</sup>C и строка МР2) существенно отличаются от экспериментальных данных. В обоих этих случаях результаты расчетов дают значения энергии, заметно превышающие экспериментальные оценки. Для уотсон-криковской G:C-пары эти расхождения могут быть объяснены экспериментальными погрешностями, связанными с сильными взаимодействиями гуанина и его метил-производных в твердой фазе. Мы обсуждали эти трудности ранее [14]. Для результатов расчетов методом МР2 эти расхождения могут быть объяснены известным фактом переоценки энергии взаимодействия молекул этим методом для самых разных наборов базисных функций. Почти все остальные рассчитанные значения энергии взаимодействия в парах с Н-связями отличаются от экспериментальных энтальпий ассоциации менее чем на 2 ккал/моль, т.е. остаются в пределах экспериментальной погрешности. Но это не может рассматриваться как свидетельство хорошей воспроизводимости экспериментальных данных. Для многих сочетаний оснований глобальный минимум энергии, найденный методами ММ2 и ММ3, соответствует одному из стекинг-минимумов с энергией, иногда существенно превышающей энтальпию ассоциации. Кроме того, эти методы не воспроизводят экспериментально определенную относительную стабильность пар m<sup>1</sup>A : m<sup>1</sup>C и m<sup>1</sup>A : m<sup>1</sup>T. Метод ММ1 не воспроизводит относительную стабильность пар m<sup>1</sup>A : m<sup>1</sup>C и m<sup>1</sup>G : m<sup>1</sup>T, но в этом случае различие энергий двух пар составляет всего 0,1 ккал/моль. Для многих пар относительные

энергии разных конфигураций одной и той же пары, рассчитанные методами ММ2 и ММ3, отличаются от рассчитанных другими методами. Наименьшие отклонения от экспериментальных оценок энтальпии ассоциации были получены при расчетах методом МР2/BSSE, который воспроизводит также относительные стабильности для всех шести пар. Значения энергии в минимумах и относительная стабильность разных конфигураций пар, полученные методом ММ1, довольно близки к таковым для метода МР2/BSSE.

Таким образом, если положения оснований в минимумах с Н-связями, рассчитанные разными методами, неплохо воспроизводят экспериментальные данные о структурах с Н-связями, то рассчитанные энергетические характеристики этих минимумов сильно зависят от метода расчета и в ряде случаев противоречат экспериментальным данным. Приведенные в этом разделе результаты расчетов уже позволяют сделать предварительные заключения о способности разных методов количественно описывать взаимодействия между основаниями, но мы обсудим возможности, недостатки и ограничения различных методов после рассмотрения других типов взаимных положений молекул в минимумах энергии взаимодействия оснований.

**Минимумы энергии взаимодействия оснований, соответствующие почти параллельному расположению их плоскостей.** Стекинг-взаимодействия между расположенными друг над другом парами оснований вносят существенный и зависящий от нуклеотидной последовательности вклад в стабильность двойной спирали ДНК. Расчеты с использованием разных методов показали, что для каждого сочетания оснований имеется несколько таких минимумов энергии,

соответствующих двум типам ориентации плоскостей их колец. Первый тип ориентации соответствует сдвигу одной из молекул в направлении, перпендикулярном плоскости ее кольца, и сдвигам в этой плоскости, т.е. молекулы обращены разными сторонами друг к другу (мы будем называть такие положения «стекинг 1»). Другой тип ориентации соответствует повороту одной из молекул на  $180^\circ$  вокруг оси, находящейся в ее плоскости, и сдвигу в перпендикулярном плоскости направлении, т.е. обе молекулы обращены друг к другу той же стороной («стекинг 2»). Стекинг-минимумы, рассчитанные любым из методов, существенно более пологие, чем минимумы с Н-связями между основаниями, т.е. возможны заметные согласованные изменения шести переменных, определяющих взаимное положение молекул (до нескольких  $\text{Å}$  и нескольких десятков градусов), при изменении энергии в пределах 0,5 ккал/моль. Полученные в результате расчетов положения минимумов энергии взаимодействия оснований не могут быть количественно сравнены с экспериментальными данными. В полинуклеотидных структурах и в кристаллах их компонентов или фрагментов положение молекул существенно сдвинуто относительно минимума энергии взаимодействия двух оснований под влиянием других взаимодействий. Однако сравнение относительных величин энергий стекинг-минимумов и минимумов с Н-связями, а также взаимных положений двух молекул в стекинг-минимумах и в экспериментальных структурах существенно для оценки достоверности разных методов расчета.

Все стекинг-минимумы характеризуются почти параллельным положением колец оснований и частичным их перекрыванием. Степень перекрывания различна для разных минимумов, а расстояние между почти параллельными плоскостями зависит от метода расчета. Различны (иногда в два и более раз) и величины энергии взаимодействия в минимумах, похожих по взаимному положению оснований, но полученных разными методами. Наибольшие по абсолютной величине значения энергии получаются при расчетах методом MP2, что соответствует уже упомянутому факту переоценки энергии взаимодействий этим методом. Эта переоценка энергии, характерная для всех типов взаимного расположения молекул, в случае стекинг-минимумов сопровождается сокращенным расстоянием между плоскостями молекул и, как следствие, нереалистично короткими межатомными расстояниями, иногда существенно меньшими сумми ван-дер-ваальсовых радиусов (до 2,9  $\text{Å}$  для С...N- и С...С-расстояний). Поэтому применение этого метода для количественных оценок взаи-

модействий между основаниями и конформационных расчетов фрагментов ДНК нельзя считать уместным. Ранее [7,8] мы уже приходили к такому выводу, проведя пробные расчеты нескольких дДМФ и кдДМФ. Расчеты методом MP2/BSSE в значительной мере лишены этого недостатка, энергия взаимодействия уменьшается почти в два раза, и только некоторые минимумы сохраняют умеренно сокращенные межатомные расстояния (до 3,2  $\text{Å}$  для атомов С и N). Параметризация современных ММ-методов исключает возникновение сокращенных межатомных контактов при расчетах стекинг-взаимодействий. Все минимумы, рассчитанные методами ММ1, ММ2 и ММ3 для каждого из сочетаний оснований, соответствуют расстоянию между практически плоскими основаниями около 3,4  $\text{Å}$ . Во многих случаях применение методов ММ1 и ММ2 (а часто и ММ3) приводит к минимумам с довольно близким взаимным положением оснований. Из-за отсутствия экспериментальных данных по этому критерию невозможно отдать предпочтение одному из ММ-методов. Однако абсолютные значения энергии в стекинг-минимумах, большие чем значения энергии в минимумах с Н-связями (некоторые из них упомянуты в предыдущем разделе), наводят на мысль о переоценке стекинг-взаимодействий при расчетах методами ММ2 и ММ3.

Расчеты методом ТФП с наиболее популярными функционалами (например, PW91, PBE, B3LYP), не учитывающими в полной мере дисперсионные взаимодействия, приводят, как правило, к нереально большому расстоянию между плоскостями оснований. В некоторых случаях при таких расчетах вообще не удается найти минимумы энергии для положений, близких к стекинг-ассоциатам, рассчитанным методами ММ и *ab initio*. Поэтому эти методы не могут быть использованы для количественной оценки энергии взаимодействия оснований в стекинг-положениях, но оказались полезны для анализа конформационных особенностей разных нуклеотидных последовательностей дуплексов [5–9], когда вклад таких взаимодействий не может изменить результатов, predeterminedных другими взаимодействиями в полинуклеотидных цепочках. Существующие функционалы полупэмпирического метода ТФП-Д, использующие дополнительное «классическое» слагаемое энергии для учета дисперсионных взаимодействий, также не могут правильно воспроизвести расстояние между плоскостями оснований, приводя, в противоположность «классическому» методу ТФП, к слишком коротким межатомным расстояниям. Мы обнаружили это несоответствие при расчетах (проведенных по настояниям

рецензентов некоторых из наших работ [6–9]) методом ТФП-Д ряда дДМФ и кдДМФ, используя в числе прочих и функционалы, предлагаемые для расчетов таких взаимодействий молекул, когда дисперсионный вклад существенен.

В этой работе мы провели систематические расчеты для некоторых стекинг-ассоциатов методом ТФП и ТФП-Д и сравнили результаты таких расчетов с данными, получаемыми методами ММ и *ab initio*. Расчеты показали, что использование функционалов, специально разработанных для систем, в которых дисперсионные взаимодействия существенны, во многих случаях приводит к слишком коротким межатомным контактам. В качестве примера отметим результаты для стекинга двух молекул 1-метилтимина. Если для минимума «стекинг 1» применение целого ряда функционалов, специально учитывающих дисперсионные взаимодействия (M05-2X, PWB6K, B3LYP-D3, B97-D3, PBE-D3BJ) не приводит к межатомным контактам, меньшим  $3,2 \text{ \AA}$ , то для минимума «стекинг 2» не удается найти функционал, не приводящий к слишком коротким контактам. Схема взаимного положения молекул в этом минимуме приведена на рис. 4. Например, расстояния между атомами C2 и N3 составляют 2,99, 3,02, 3,04, 3,07 и  $3,06 \text{ \AA}$  соответственно при использовании перечисленного выше ряда функционалов. Мы заметили, что в минимуме «стекинг 1» двух молекул  $m^1T$  плоскости их колец практически параллельны при расчетах с использованием всех перечисленных функционалов (угол менее  $1^\circ$ ), а в минимуме «стекинг 2» отклонение от параллельности достигает  $15^\circ$ . В целом ряде стекинг-минимумов, рассчитанных с разными функционалами, включающими дисперсионные взаимодействия, наблюдаются столь близкие C...C- и C...N-контакты между атомами колец. Анализ взаимных положений молекулярных групп в таких минимумах и сравнение с минимумами, рассчитанными тем же методом, но не содержащими укороченных контактов, позволили нам обнаружить общую черту, характерную для всех таких минимумов. Слишком короткие контакты возникают, если атомная группа одного основания, содержащая акцептор Н-связи (например, C=O), находится над или под группой, содержащей донор Н-связи. Не являются при этом исключением и группы C–H оснований, такие как C8–H пуринов или C6–H пиримидинов, способные образовывать слабые Н-связи C–H...O или C–H...N. При этом наиболее короткий контакт возникает не между атомами, которые могли бы (при другом взаимном положении молекул) участвовать в Н-связи (например, N3–H...O2),

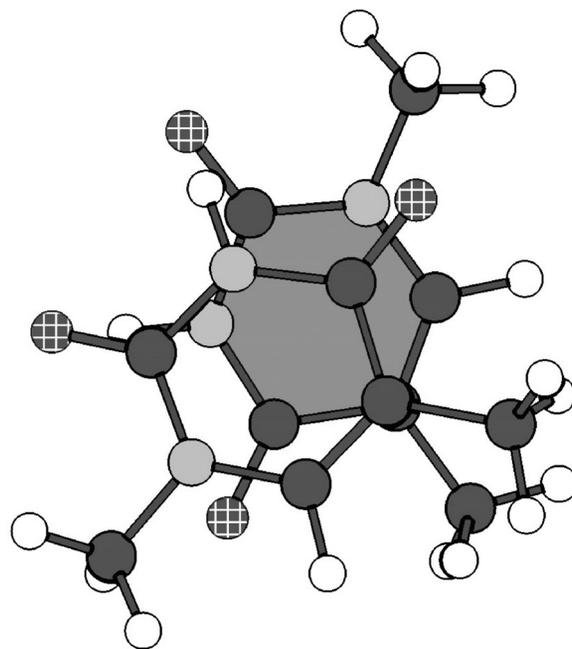


Рис. 4. Схема взаимного положения двух молекул 1-метилтимина для минимума «стекинг 2».

а между соседними к ним атомами кольца. Анализ минимумов для ряда структур дДМФ и кдДМФ, рассчитанных такими методами, демонстрирует ту же закономерность. Умеренно сокращенные контакты в стекинг-минимумах, рассчитанных методом MP2/BSSE, наблюдаются в основном в таких же положениях. Можно предположить, что эти КМ-методы переоценивают тенденцию образования Н-связей между протон-донорными и протон-акцепторными группами оснований на расстояниях, больших, чем характерные для Н-связей. Возможно, это наблюдение отражает и некоторую физическую закономерность. В любом случае мы надеемся, что это наблюдение может быть полезным при выработке метода, позволяющего корректнее воспроизводить невалентные взаимодействия в ДНК и в биомолекулярных системах вообще.

**Минимумы энергии взаимодействия оснований, соответствующие почти перпендикулярному расположению их плоскостей.** Минимумы энергии, соответствующие почти перпендикулярному положению колец оснований с образованием одной Н-связи, были получены еще в 2001 г. [25] и 2002 г. [14] с использованием методов ММ2 и ММ1 соответственно. Такие взаимные положения оснований могут возникать в неспиральных участках нуклеиновых кислот или при взаимодействиях ДНК с ее фрагментами. До недавнего времени можно было бы сомневаться, не являются ли такие минимумы артефактом силовых полей молекулярной механики.

В этой работе мы провели систематические расчеты, направленные на поиск таких минимумов всеми используемыми нами методами, и показали, что эти минимумы могут быть получены как каждым из трех использованных силовых полей, так и КМ-методами *ab initio* и ТФП. Сравнение результатов расчетов для таких минимумов с экспериментальными данными еще более затруднительно, чем для других типов минимумов, но правильная оценка взаимодействий оснований в подобных положениях существенна для моделирования ДНК.

Мы не ставили целью найти все такие минимумы, их количество существенно больше, чем мы исследовали, а для некоторых сочетаний оснований (например, гуанина с цитозином) минимумы с почти перпендикулярным положением оснований, воспроизводимые хотя бы несколькими методами, не были найдены. Поэтому мы обсудим здесь только общие характеристики этих минимумов и их воспроизводимость разными ММ- и КМ-методами. Каждый из этих минимумов характеризуется почти линейной Н-связью между атомом-акцептором водорода одного основания и донорной группой другого основания, например N–H...N. Расстояния между атомами Н-связи лежат примерно в тех же пределах, что и для минимумов с двумя Н-связями, обычно они близки к таковым для соответствующего сочетания оснований, найденным тем же методом. Например, для минимумов  $m^1T-m^1T$  с Н-связью N3–N3...O4, найденным всеми примененными нами методами ММ, *ab initio* и ТФП, расстояние N3...O4 лежит в пределах  $1,9 \pm 0,06$  Å. При этом взаимные ориентации двух оснований, т.е. угол между плоскостями колец и угол поворота вокруг Н-связи, могут существенно изменяться. Обычно положение колец оснований в минимумах, найденных ММ методами и методом ТФП-Д, ближе к перпендикулярному, чем в соответствующих минимумах, найденных методами *ab initio* и ТФП без дисперсионной поправки. Глубина минимумов этого типа обычно (а методами ММ1 и MP2/BSSE – для всех найденных нами минимумов) меньше, чем у минимумов с двумя Н-связями и у наиболее глубоких стекинг-минимумов.

**Сравнение характеристик разных минимумов и пути совершенствования методов расчета взаимодействий между основаниями.** Результаты расчетов энергии взаимодействия между основаниями и характеристики локальных минимумов, найденных разными ММ- и КМ-методами, свидетельствуют о трудностях количественного описания невалентных взаимодействий нуклеиновых кислот. Несмотря на большой и все увеличивающийся объем экспериментальной

информации о структуре нуклеиновых кислот, их фрагментов разной степени сложности и отдельных их компонентов, имеется довольно мало данных, допускающих непосредственное количественное сравнение с результатами расчетов. Вместе с тем уже сравнение относительных энергетических и структурных характеристик, следующих из экспериментальных данных, с результатами расчетов разными ММ- и КМ-методами указывает на недостатки всех имеющихся методов. Ни один метод, допускающий расчеты хотя бы минимальных фрагментов двойной спирали ДНК или одной цепочки (кдДМФ или дДМФ) не позволяет количественно воспроизвести все даже относительные экспериментальные данные. Применение все более строгих КМ-методов не позволяет в ближайшей перспективе получить количественные результаты для дДМФ или кдДМФ. Применение теории возмущений на уровне MP2 даже с довольно широким набором базисных функций систематически переоценивает стекинг-взаимодействия и ведет к недопустимо коротким значениям для межатомных контактов. Наиболее строгие корреляционные КМ-методы не могут быть применены даже для поиска минимумов попарных взаимодействий оснований. Использование метода ТФП позволило провести расчеты минимальных фрагментов ДНК и получить интересные закономерности формирования пространственной структуры [5, 9], но наиболее распространенные функционалы переоценивают расстояния между соседними по цепи основаниями как следствие неспособности правильно описывать дисперсионные взаимодействия. Ни один из опробованных нами функционалов «гибридного квантово-классического» метода ДФТ-Д (в том числе специально разработанные для учета дисперсионных взаимодействий) не позволяет правильно, без слишком коротких межатомных контактов, описать не только конформации минимальных фрагментов ДНК, но и многие стекинг-минимумы энергии попарных взаимодействиях оснований. Более совершенная параметризация метода ДФТ-Д может в дальнейшем привести к лучшему согласию с экспериментом, но, на наш взгляд, более прямой путь к адекватному моделированию взаимодействия оснований лежит через подбор параметров (и, может быть, некоторую модификацию формул) силовых полей ММ. Этот путь должен включать не только использование наиболее строгих КМ-расчетов, но и непосредственный подбор коэффициентов по количественным экспериментальным (структурным и энергетическим) данным, в том числе о комплексах и кристаллах компонентов и простых фрагментах нуклеиновых кислот.

## СПИСОК ЛИТЕРАТУРЫ

1. H. DeVoe and I. Tinoco, *J. Mol. Biol.* **4**, 500 (1962).
2. P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006).
3. J. Řezáč and P. Hobza, *Chem. Rev.* **116**, 5038 (2016).
4. В. Кон, *Успехи физ. наук* **172**, 336 (2002).
5. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *J. Biomol. Struct. Dyn.* **24**, 660 (2007).
6. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *Comput. Theor. Chem.* **975**, 69 (2011).
7. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., *Biopolymers* **101**, 640 (2014).
8. В. И. Полтев, В. М. Анисимов, К. Санчес и др. *Биофизика* **61**, 259 (2016).
9. V. Poltev, V. M. Anisimov, V. Dominguez, et al., *J. Mol. Model.* **24**, 46 (2018).
10. H. M. Berman, W. K. Olson, D. L. Beveridge, et al., *Biophys. J.* **63**, 751 (1992).
11. T. E. Cheatham, III and D. A. Case, *Biopolymers* **99**, 969 (2013).
12. K. Vanommeslaeghe and A. D. MacKerell, Jr., *Biochim. Biophys. Acta: General Subjects* **1850**, 861 (2015).
13. M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., *Gaussian 16* (Gaussian Inc., Wallingford CT, 2016).
14. В. И. Полтев, А. С. Дерябина, Э. Гонсалес и Т. И. Грохлина, *Биофизика* **47**, 996 (2002).
15. V. Poltev, In: *Handbook of Computational Chemistry*, Ed. by J. Leszczynski (Springer Science, 2012), pp. 259–291.
16. V. Poltev, In: *Handbook of Computational Chemistry*, Ed. by J. Leszczynski et al. (Springer International Publishing Switzerland, 2017), pp. 21–67.
17. D. A. Case, R. M. Betz, D. S. Cerutti, et al., *AMBER 2016* (University of California, San Francisco, 2016).
18. J. C. Phillips, R. Braun, and W. Wang, *J. Comput. Chem.* **26**, 1781 (2005).
19. S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
20. S. Grimme, *J. Comput. Chem.* **25**, 1463 (2004).
21. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
22. I. K. Yanson, A.B. Teplitsky, and L. F. Sukhodub, *Biopolymers* **18**, 1149 (1979).
23. L. F. Sukhodub, *Chem. Rev.* **87**, 589 (1987).
24. V. I. Poltev, N. V. Shulyupina, V. I. Bruskov, et al., *J. Biomol. Struct. Dyn.* **9**, 101 (1991).
25. M. Kabelac and P. Hobza, *J. Phys. Chem. B* **105**, 5804 (2001).

## Some Problems of Computer Simulation of Nonbonded Interactions in DNA

V. Poltev\*, A. Deriabina\*, V. Dominguez\*, C. Sanchez\*, E. Gonzalez\*, and N.A. Polteva\*\*

\*Autonomous University of Puebla, Puebla, 72570 Mexico

\*\*Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences,  
Institutskaya ul. 3, Pushchino, Moscow Region, 142290 Russia

Some complications that arise in application of various computer simulation methods are analyzed to study regularities of a three-dimensional structure of DNA. In view of the important role of nitrous bases we performed computations of intermolecular interaction energy for ten pairwise combinations of methylated bases (1-methylpyrimidines and 9-methylpurines) in various mutual base positions. Using different molecular mechanics force fields, *ab initio* methods in quantum mechanics and density functional theory, local energy minima that correspond to different mutual molecule positions have been found. The energy minima that correspond to three types of mutual molecule positions, namely, (1) base positions with two N–H...O and/or N–H...N hydrogen bonds; (2) nearly parallel arrangements of base ring planes, i.e. base stacking; and (3) T-shaped (nearly perpendicular) base ring positions and hydrogen bond formation were extensively studied. The majority of these minima were obtained by using the methods of different complexity, but the method of calculation determines which minimum is the most deep (global) for certain base combination, and relative depths of various minima. Analysis of these data, of our previous results on simple DNA fragments, as well as of extensive reference data suggests that till now there is no method suitable for quantitative description of all the experimental data on non-bonded interactions in DNA. The comparison of energy and structure characteristics of complexes calculated by various methods demonstrates their possibilities and shortcomings. A valid description of non-bonded interactions of nucleic acids requires additional investigations of their simple fragments and combined use of various methods.

*Keywords: DNA, computer simulation, molecular mechanics, non-empirical calculations, density functional theory*