

УДИВИТЕЛЬНАЯ МАКРОМОЛЕКУЛА ДНК: КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ЕЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ И МНОГООБРАЗИЕ УОТСОН–КРИКОВСКИХ КОНФОРМАЦИЙ ДУПЛЕКСА

© 2023 г. В. Полтев*,#, В. Домингез*, А. Руиз*, А. Дерябина*, Э. Гонзалез*

*Автономный Университет Пуэблы, Пуэбла, 72570, Мексика

#E-mail: poltev@fcfm.buap.mx

Поступила в редакцию 06.12.2022 г.

После доработки 06.12.2022 г.

Принята к публикации 21.12.2022 г.

Рассматривается развитие представлений о пространственной структуре ДНК и механизмах ее формирования, начиная с открытия двойной спирали до наших дней. Прослеживаются пути и методы моделирования пространственной структуры на разных этапах исследования главной молекулы жизни. На основании результатов расчетов внутримолекулярных и межмолекулярных взаимодействий субъединиц макромолекулы обсуждается удивительная целесообразность ее молекулярной структуры, ее приспособленность к выполнению важнейших биологических функций. Приводятся новые данные о существенном вкладе химически однородного конформационно лабильного сахара-фосфатного остова в формирование пространственной структуры ДНК, зависящей от последовательности оснований. Многообразие конформационных возможностей ДНК проявляется в образовании как дуплексов (а также триплексов и квадруплексов) с разной геометрией пар оснований, так и дуплексов с уотсон-криковскими парами нуклеозидов, содержащих локальные конформации, соответствующие разным областям торсионных углов сахара-фосфатного остова (разным конформационным классам). Эти классы, на основании наших расчетов, могут быть разделены на две группы. К первой группе относятся те локальные конформации, в которых торсионные углы близки к одному из минимумов энергии его изолированного минимального повторяющегося фрагмента, а ко второй — конформации, в которых один или несколько из этих углов отличаются от ближайшего минимума более, чем на 30° . Закономерности формирования локальной пространственной структуры этих двух групп существенно отличаются.

Ключевые слова: ДНК, пространственная структура, многообразие конформаций, сахара-фосфатный остов, компьютерное моделирование, молекулярная механика, теория функционала плотности, неэмпирические расчеты.

DOI: 10.31857/S0006302923020102, EDN: CAVHTJ

НЕМНОГО ИСТОРИИ

Открытие в 1953 году Дж. Уотсоном и Ф. Криком пространственной структуры ДНК [1] дало, хотя пока только в принципе, ответ на вопрос о том, как, с точки зрения физики, могут воспроизводиться и развиваться живые системы. Из модели, содержащей две комплементарные полинуклеотидные цепочки, с азотистыми основаниями 4 типов и повторяющимися фрагментами химически однородного сахара-фосфатного остова (СФО), следуют, в общих чертах, механизмы хра-

нения, воспроизведения, использования и эволюционных изменений генетического материала. С тех пор изучению этого биополимера посвящены многие тысячи исследований, выполненных с применением широкого спектра экспериментальных и теоретических методов. К настоящему времени (по состоянию на 9 ноября 2022 г.) в базе данных нуклеиновых кислот NDB [2] собраны координаты атомов более 12432 структур фрагментов нуклеиновых кислот и их комплексов с другими молекулами, построенных по экспериментальным данным и с привлечением расчетно-теоретических методов. Этот постоянно пополняющийся материал позволяет делать все более интересные новые обобщения о Главной Молекуле Жизни и механизмах ее биологического

Сокращения: СФО — сахара-фосфатный остов, ММ — молекулярная механика, дДМФ — дезоксидинуклеозид монофосфат, кдДМФ — комплекс комплементарных дДМФ, УКД — уотсон-криковские дуплексы, КК — конформационный класс дДМФ.

функционирования. Более 60 лет назад, будучи студентом кафедры биофизики физического факультета МГУ, по предложению и под руководством Льва Александровича Блюменфельда первый автор этой статьи начал заниматься исследованием ДНК и ее компонентов [3], что и продолжает делать все последующие годы своей научной деятельности.

Эта статья состоит из двух частей, в первой из которых кратко описывается развитие представлений о ДНК, ее пространственной структуре, конформационной подвижности и вкладе субъединиц в свойства этой удивительной макромолекулы, необходимые для выполнения ею своих функций. На примере некоторых прошлых исследований прослеживается развитие методов молекулярного моделирования и его вклад в формирование современных представлений о ДНК как о конформационно лабильном и структурно многообразном биополимере. Во второй части статьи приводятся некоторые новые результаты исследований, выполненных авторами данной публикации в последние месяцы, как иллюстрация продолжающегося по мере накопления экспериментальных данных и расширения возможностей применения расчетно-теоретических методов уточнения наших представлений о «неисчерпаемой» макромолекуле ДНК.

Первые результаты молекулярного моделирования. Само открытие двойной спирали ДНК в 1953 г. (как и открытие двумя годами ранее структур регулярных α - и β -участков белка) стало возможным благодаря молекулярному моделированию, что в то время означало «ручное» построение модели молекулы из проволочных, картонных и деревянных фрагментов, обозначающих атомы, химические связи и целые субъединицы макромолекулы. Успех в построении модели ДНК объясняется не только «мощной интуицией» авторов этого открытия, но и комплексным использованием при построении модели результатов экспериментальных исследований физическими (рентгеноструктурный анализ волокон ДНК и ее субъединиц), химическими (молекулярная структура мономеров, таутомерные формы оснований и водородные связи между ними) и биологическими (правила Чаргаффа) методами. Такое «ручное» моделирование на новом уровне, с использованием профессионально разработанных и промышленно изготовленных пластиковых моделей атомов и химических связей, потом еще не одно десятилетие (до того, как персональные компьютеры стали общедоступными) широко использовалось для построения и визуализации фрагментов нуклеиновых кислот и их комплексов с другими молекулами.

До конца 70-х годов прошлого века модели пространственной структуры макромолекулы

ДНК основывались на результатах рентгеноструктурных исследований волокон синтетических полинуклеотидов с повторяющимися моно-, ди- и тринуклеотидными последовательностями [4]. По таким данным были построены более детальные, чем по первым, «до-утсон-криковским» данным, модели регулярных дуплексов. Их пространственные структуры оказались зависящими от повторяющейся нуклеотидной последовательности, от содержания в образце воды и противоионов, но для заключения о возможности непрерывного изменения конформационных параметров в зависимости от окружения и об изменении локальной пространственной структуры вдоль цепи прямых структурных данных было недостаточно; этот вопрос и не обсуждался. Начиная с 60-х годов прошлого века необходимость молекулярного моделирования с использованием расчетно-теоретических методов становилась все более очевидной. При этом уже в первые десятилетия после открытия Уотсона и Крика было ясно, что это моделирование должно основываться на приближенных классических (или полуклассических) представлениях. Достаточно строгие квантово-механические методы, по необходимости, приближенные в применении к любым многоатомным системам, еще долго не могли быть успешно применимы к количественным оценкам энергии и конформационных характеристик даже простых фрагментов ДНК. Тем временем метод молекулярной механики (ММ), хоть и под другими названиями, уже успешно применялся к исследованиям органических молекул и их кристаллов. Описание принципов метода ММ, его истории и современного (на тот период) состояния приведено в главе, написанной первым автором этой статьи [5]. Здесь мы только упомянем основные черты этого метода, важные в связи с его ролью в изучении пространственной структуры ДНК от первых приближенных оценок до настоящего времени, когда применение квантово-механических методов стало не только необходимым, но и реально возможным, хотя пока для изучения только сравнительно простых фрагментов ДНК. Метод ММ можно рассматривать как обобщение и количественное представление, с помощью математических выражений, схематического изображения молекул в классической химии. Набор таких выражений вместе с численными коэффициентами для расчета энергии молекулы или надмолекулярного комплекса называют силовым полем. ММ — это метод классической физики, метод «механистического» (то есть с помощью ньютоновской механики) описания квантовых систем. Важной чертой метода ММ является предположение об аддитивности энергий взаимодействия отдельных атомов и взаимодействий разной физической природы. Хотя каждому типу взаимодействий (например, кулоновскому, дис-

персионному или изменению длин связей и валентных углов) соответствует свой член суммарной энергии, физический смысл имеет только общая энергия, которая может быть сопоставлена с экспериментальными данными, а не отдельные слагаемые, которые могут существенно отличаться в разных силовых полях. Этот метод, внесший существенный вклад в первые десятилетия изучения ДНК, до сих пор остается единственно возможным для моделирования крупных фрагментов биополимеров, а также для предварительных оценок конформационных состояний и конформационных возможностей их фрагментов перед расчетами более строгими, но все равно приближенными, и более затратными квантово-механическими методами.

Первые оценки энергии взаимодействий в ДНК.

Несмотря на появившийся сразу же после открытия двойной спирали интерес физиков, физико-химиков и математиков к деталям строения ДНК, к изучению взаимного влияния ее субъединиц и взаимодействий разной физической природы, первая публикация, содержащая попытку оценить количественно, хоть и очень приближенно, вклад некоторых взаимодействий в формирование и стабильность двойной спирали, появилась только в конце первого десятилетия после открытия Уотсона и Крика. В 1962 г. Де Во и Тиноко [6] опубликовали первые количественные оценки энергии взаимодействия между основаниями в модели ДНК, используя приближение молекулярных точечных диполей, то есть расчет взаимодействий точечных диполей, помещенных в центры оснований, расположенных как в модели В-формы двойной спирали. Суммарная энергия взаимодействия между двумя основаниями в паре или в стопке предполагалась равной сумме взаимодействий постоянных диполей (кулоновские электростатические взаимодействия), постоянного диполя одной молекулы с индуцированным диполем другой (индукционные взаимодействия) и между двумя индуцированными диполями (лондоновские дисперсионные взаимодействия). Такое приближение не позволило получить даже качественно правильные результаты (например, энергия взаимодействия в комплементарной паре А:Т получалась положительной), но сама проблема количественной оценки изменения энергии при образовании или при изменении конформации ДНК была сформулирована и привлекла внимание целого ряда исследователей (включая первого автора этой статьи). Вскоре в работе Д. Бредли с соавторами [7] было показано, что такое «дипольное» приближение непригодно для оценки взаимодействия оснований, независимо от локализации диполей и замены точечных диполей «реальными», то есть имеющими размеры. Авторы работы [7] предложили «монопольное приближение», они вычисляли энергию взаимо-

действия молекул, предположив, что главный вклад в энергию взаимодействия оснований вносят взаимодействия постоянных зарядов, помещенных на каждом из атомов одной молекулы с зарядами на атомах другой молекулы, а сами заряды при этом вычислялись с помощью полупирических методов квантовой химии. Это было первым указанием на необходимость расчета атом-атомных взаимодействий в теоретических исследованиях пространственной структуры ДНК. В работе Нэша и Бредли [8] это приближение было использовано для поиска минимумов энергии взаимодействия между двумя основаниями, находящимися в одной плоскости. При этом учитывались и ван-дер-ваальсовы взаимодействия в приближении модели «твердых тел», то есть атомы не могли располагаться ближе друг к другу, чем сумма их ван-дер-ваальсовых радиусов. Полученные в работе [8] минимумы энергии соответствуют взаимным положениям оснований с водородными связями, которые практически совпадают с полученными позже самыми разными расчетными методами (включая наиболее строгие из возможных квантово-механических методов и последние версии силовых полей ММ). В следующих работах, которые можно рассматривать как продолжение работ Бредли с соавторами [7, 8], для расчетов взаимодействия между основаниями ДНК были предложены модели, представляющие каждую молекулу в виде совокупности точечных зарядов, диполей и квадрупольей, расположенных не только на атомах, но и на связях [9], которые рассчитывались методами квантовой механики, но представляли собой хоть и усложненное, но все равно классическое представление квантовой системы. Результаты таких расчетов (см., например, работы [9, 10]) уже тогда не имели преимуществ по сравнению с результатами «классической» молекулярной механики, а теперь могут представлять интерес только для истории изучения внутримолекулярных взаимодействий в ДНК.

Применение метода молекулярной механики на первых этапах изучения пространственной структуры ДНК. Широкое использование и выработка параметров метода ММ для применения к расчетам нуклеиновых кислот начались через несколько лет после того, как этот метод уже успешно применялся к исследованию пространственной структуры и взаимодействий белковых фрагментов. Эту ситуацию можно объяснить «специальной ролью» ДНК как носителя генетической информации, для которого можно было бы ожидать «специфических» сил между основаниями, которые являются сопряженными молекулами с делокализованной π -электронной системой. В то время бытовало мнение, что электронный обмен между основаниями в комплементарных парах и обменные взаимодействия в стопках параллель-

ных оснований вносят решающий вклад в биологическое функционирование ДНК. Потребовалось несколько лет, прежде чем А. Пуллман и Б. Пуллман, пионеры применения квантово-механического подхода к биологическим молекулам, написали: «Возможные вклады резонансной энергии стабилизации посредством делокализации электронов через водородные связи для горизонтальных взаимодействий и перекрывание их электронных облаков или перенос заряда для вертикальных взаимодействий, оказывается, много меньше по порядку величины, чем стабилизация, обусловленная силами Ван-дер-Ваальса—Лондона» [11]. Однако интересно отметить, что несмотря на «отставание» в применении ММ к нуклеиновым кислотам по сравнению с его применением к белкам, некоторые, теперь кажущиеся очевидными, нововведения в ММ для исследования биополимеров были сначала предложены для расчетов нуклеиновых кислот и только затем использованы в расчетах белков. Например, это уже упоминавшийся выше учет взаимодействия между эффективными зарядами на атомах и разные параметры для разных электронных конфигураций одного и того же типа атомов.

Первый автор данной статьи оказался первым исследователем, начавшим в 1967 г. применять молекулярную механику (этот термин, как и термин «силовое поле», не применялся тогда к подобным расчетам) к теоретическому изучению взаимодействий между субъединицами (сначала только между азотистыми основаниями) ДНК [12] и предложившим для этого самосогласованный набор потенциальных функций. При этом предполагалась усредненная по экспериментальным данным геометрия каждого из оснований. Для сравнения, в предложенных несколько позднее параметрах Шераги для расчетов белков [13] предполагалась фиксированная геометрия пептидной группы и некоторых аминокислот. Энергия каждого из атом-атомных взаимодействий содержала электростатический член, индукционную составляющую (вскоре исключенную по причине малого вклада для наших систем и для избегания учета неаддитивности атом-атомных и фрагмент-фрагментных взаимодействий) и члены, соответствующие лондоновской дисперсионной энергии и близкодействующему отталкиванию. При этом заряды на атомах рассчитывались простейшими полуэмпирическими методами квантовой химии (параметризованными так, чтобы воспроизводить дипольные моменты рассматриваемых и подобных им молекул). Коэффициенты члена дисперсионной энергии первоначально оценивались по приближенной формуле Лондона (и поэтому изначально зависели от электронной конфигурации атома), а для оценки близкодействующего отталкивания использовались несколько увеличенные (в отличие от мно-

гих существовавших тогда силовых полей) ван-дер-ваальсовы радиусы атомов. На следующем этапе эти коэффициенты были уточнены по результатам расчетов молекулярных кристаллов родственных соединений. Предложенный набор коэффициентов был затем дополнен нами путем включения параметров для учета взаимодействий сахаро-фосфатного остова [14] и взаимодействий с молекулами воды [15]. Эта простая атом-атомная схема, коэффициенты которой уточнялись (как и для всех силовых полей ММ) по мере появления новых экспериментальных данных, широко использовалась в 70-х и 80-х годах прошлого века не только нами и нашими коллегами, но и другими исследователями. С помощью этой схемы удалось объяснить ряд экспериментальных данных, а также предсказать некоторые свойства ДНК до появления данных эксперимента. К таким предсказаниям можно отнести предсказание возможности значительных изменений конформационных параметров двойной спирали без существенного изменения энергии [16], путей возникновения всех транзиций и трансверсий при биосинтезе ДНК [17] и структуры фрагментов ДНК с «неправильными» парами оснований [18]. Первые экспериментальные данные о структуре фрагментов дуплекса с «неправильными» парами появились вскоре после наших публикаций на эту тему. До этого основным путем включения в ДНК «неправильных» оснований при биосинтезе считался переход одного из оснований образующейся пары в редкую таутомерную форму с сохранением геометрии СФО каждой цепочки и взаимного положения двух цепочек [19], как было предложено еще Уотсоном и Криком [1].

Теперь это силовое поле, не учитывающее некоторые слагаемые суммарной энергии (что было вполне оправданно для наших систем, предполагающих фиксированными длины химических связей и многие валентные углы), только иногда используется нами и нашими коллегами. Это связано не с тем, что разработаны более адекватные методы, а с переходом практически всех исследователей, применяющих методы ММ (а ранее и квантовой механики), к использованию стандартных компьютерных программ со встроенными силовыми полями, позволяющих проводить расчеты широкого спектра молекулярных комплексов и не предполагающих возможность изменений или дополнений силовых полей каждым пользователем. В те годы прошлого века, когда предлагались и широко использовались наши параметры, компьютерные программы, как правило, составлялись и модифицировались самими исследователями и их коллегами применительно к конкретной задаче. Предложенные нами более полувека назад параметры метода ММ конечно же требуют (как и параметры других силовых полей) дополнительного уточнения с использова-

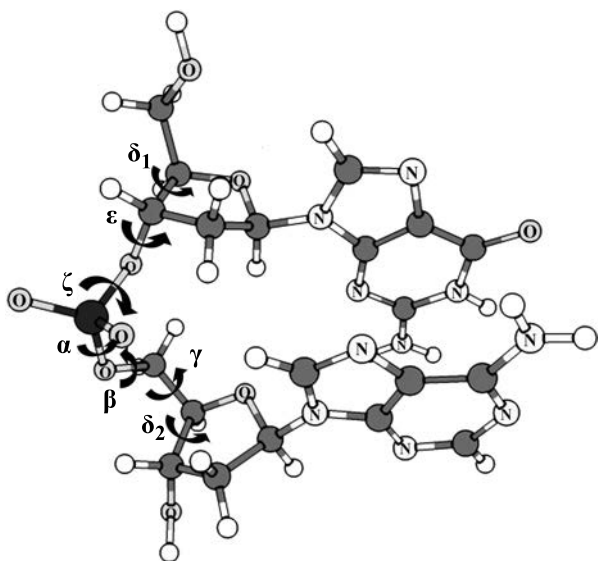


Рис. 1. Молекулярная структура и обозначения торсионных углов сахаро-фосфатного остова дДМФ (dGpдA).

нием более широкого набора экспериментальных данных и наиболее строгих результатов квантово-механических расчетов, но и в прежнем виде они до сих пор дают в некоторых случаях более адекватные результаты расчетов простых фрагментов ДНК, чем силовые поля, разработанные и уточненные в этом веке [20]. Но наше силовое поле не может быть использовано в новых версиях современных, разработанных коллективами программистов-профессионалов, наиболее распространенных программ, таких как AMBER [21] или CHARMM [22].

Экспериментальное изучение пространственной структуры фрагментов ДНК на атомарном уровне и ее конформационная подвижность. Новый этап исследований пространственной структуры нуклеиновых кислот начался в 1980-х годах прошлого века с появлением возможности получения и рентгенографического исследования монокристаллов фрагментов двойной спирали (впоследствии и других конформаций полинуклеотидов). Результаты для первого такого фрагмента, самокомплементарного додекануклеотида dCdGdC-dGdAdAdTdCdGdCdG, часто называемого додекамером Дрю–Дикерсона [23], сначала показались неожиданными и вызывающими сомнения. Оказалось, что конформационные параметры (как торсионные углы СФО и упаковка колец дезоксирибозы, так и взаимные положения оснований в соседних парах) изменяются в широких пределах и зависят от последовательности оснований. В последующие годы усилиями исследователей разных стран были кристаллизованы и расшифрованы десятки полинуклеотидных фраг-

ментов, анализ которых демонстрировал широкий диапазон изменения локальных конформационных параметров двойной спирали в пределах «канонических» В- и А-семейств конформаций. Возможность таких согласованных изменений конформационных параметров (в частности, изменений торсионных углов в пределах 30°) и зависимость локальной конформации двойной спирали от последовательности оснований, ранее предсказанные расчетами методом ММ (например, [16]) и казавшиеся очевидными из общих теоретических соображений, становятся общепринятыми. С тех пор количество описанных на атомарном уровне структур фрагментов нуклеиновых кислот и их комплексов увеличивается с каждым годом (и месяцем), также расшифровываются все более сложные структуры. Анализ этих данных (собранных в базе данных NDB [2]) позволяет не только делать новые обобщения о пространственной структуре ДНК в связи с ее биологической функцией, но и ставит перед исследователями новые вопросы. В следующих разделах мы обсуждаем часть этих вопросов, основываясь на результатах наших расчетов элементарных фрагментов ДНК, выполненных в последние пятнадцать лет, и приведем некоторые наши новые данные о вкладе отдельных субъединиц ДНК в ее биологическое функционирование.

МЕТОДИКА РАСЧЕТОВ

Рассматриваемые системы. Рассматриваются элементарные фрагменты одиночной цепи ДНК (нейтрализованные ионом Na^+ дезоксидинуклеозид-монофосфаты (дДМФ) и двойной спирали (комплементарные комплексы дДМФ – кдДМФ), а также молекулы, соответствующие элементарным фрагментам сахаро-фосфатного остова, то есть дДМФ, в которых оба основания заменены на атомы водорода. Молекулярная структура одного из дДМФ и обозначения его торсионных углов представлены на рис. 1. Геометрические характеристики дДМФ соответствуют таковым для динуклеотидных фрагментов, рассматриваемых Шнейдером и др. [24] в их анализе и классификации конформаций элементарных фрагментов цепи ДНК.

В таком фрагменте 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^+). Элементарные фрагменты двойной спирали, кдДМФ, включают также взаимодействия между субъединицами комплементарных цепей, а элементарные фрагменты СФО исключают влияние оснований

и позволяют оценить вклад монотонного остова в формирование локальных конформаций полинуклеотида. При подготовке структур к расчетам начальную конфигурацию каждого из фрагментов строили по координатам атомов соответствующих фрагментов экспериментально определенных структур комплементарных дуплексов, приведенных в банке данных нуклеиновых кислот. При этом атомы водорода и ионы натрия помещались в их наиболее вероятные положения.

Методы расчетов. Описание и обоснование методов расчета энергии систем и используемых при этом программ приводится в каждой из наших предыдущих публикаций. Здесь мы только упомянем методы и программы, с помощью которых были получены неопубликованные ранее результаты. Для расчетов энергии и поиска ее локальных минимумов методом ММ был использован программный комплекс AMBER [21] с силовым полем BSC1 [25]. Полученные при этом результаты обозначены в таблицах как ММ. Расчеты квантово-механическими методами проводили с использованием программного комплекса Gaussian09 [26]. Результаты, полученные с их применением обозначены как РВЕ (метод теории функционала плотности с функционалом РВЕР-ВЕ и базисным набором волновых функций 6-31G*), М05-2Х (метод теории функционала плотности с функционалом М05-2Х, учитывающим дисперсионные взаимодействия, с тем же базисным набором) и МР2 (квантово-механический *ab initio* метод с использованием базисного набора 6-31G**). Все минимумы энергии, найденные квантово-механическими методами, были подтверждены расчетами частот колебаний.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Удивительная целесообразность молекулярной структуры ДНК и многообразие ее конформаций. Эволюционная приспособленность структуры ДНК к выполнению ее генетических функций была отмечена сразу же после открытия Уотсоном и Криком двойной спирали. Сочетание в макромолекуле практически жестких и плоских субъединиц (двух пуриновых (pur) и двух пиримидиновых (pyr) оснований, образующих комплементарные пары) с химически однородным и конформационно лабильным сахаро-фосфатным остовом позволяет как хранить информацию (образовывать достаточно стабильную двойную спираль), так и считывать и копировать ее (при расплетании спирали). В первые же годы после построения модели двойной спирали была отмечена важность различия в стабильности двух пар оснований (Ade:Thy с двумя водородными связями и Gua:Cyt с тремя водородными связями), а также отличий в форме и размерах между двумя типами оснований (между пуринами и пирими-

динами) в уотсон-криковских парах. Уотсон-криковские дуплексы (УКД) имеют ряд интересных конформационных особенностей, некоторые из которых мы проанализируем здесь с привлечением результатов наших расчетов [27–34], выполненных в течение последних 15 лет. Эти данные иллюстрируют вклад пространственной структуры УКД в функционирование генетического материала и вклад молекулярной структуры субъединиц полинуклеотидной цепи в формирование ее уникальной пространственной структуры. Мы будем обозначать как «УКД» любые антипараллельные полинуклеотидные дуплексы, содержащие комплементарные уотсон-криковские пары оснований и оба нуклеозида в анти-ориентации основания относительно сахара. Это определение отличается от принятого по умолчанию со времен открытия двойной спирали представления о УКД как о канонической А- или В-конформации (или, потом, о соответствующих семействах конформаций) тем, что включает возможность существования в УКД разных областей торсионных углов СФО.

Уже первые наши расчеты [27], проведенные для нескольких дДМФ с использованием в качестве начальной стандартной В-конформации, позволили выявить ряд интересных особенностей формирования пространственной структуры ДНК. В следующих работах эти результаты были распространены на дДМФ всех 16 возможных последовательностей нуклеозидов в VI- и AI-конформациях, а затем и в AII- и VII-конформациях (в последнем случае это оказалось преждевременным; отсутствие экспериментальных данных для многих последовательностей не позволило тогда обнаружить, что для VII-конформаций эти закономерности более сложны [34]).

Результаты расчетов показали, что для дДМФ с любой нуклеотидной последовательностью есть локальные минимумы энергии с конформационными параметрами, характерными для AI- и VI-семейств конформаций, то есть существование этих конформаций предопределено молекулярной структурой элементарных фрагментов одичной цепи. Этот вывод объясняет эволюционный выбор полинуклеотидной химической структуры как вещества наследственности. В этих же работах было показано, что в минимумах энергии дДМФ с последовательностями оснований «pur–pur» или «pur–pyr» наблюдается существенное перекрытие колец соседних по цепи оснований, а для дДМФ с последовательностями «pyr–pur» и «pyr–pyr» перекрытие незначительно или совсем отсутствует. Эта закономерность, обнаруженная Р. Дикерсоном в 1998 г. [35] в тогда еще не столь многочисленных построенных по экспериментальным данным структурах и наблюдаемая для «канонических» VI- и AI-конформаций УКД (а также, как мы обнаружили

впоследствии, и некоторых других конформаций), определяется взаимодействиями в элементарном фрагменте одной цепи дуплекса, дДМФ. Анализ результатов расчетов показал, что направленность сахаро-фосфатного остова и предпочтительные области его торсионных углов вместе с различиями геометрии пуринов от пиримидинов определяют зависимость пространственной структуры этих конформаций УКД от нуклеотидной последовательности. Расчеты элементарного фрагмента СФО показали, что торсионные углы экспериментально наблюдаемых VI- и AI-конформаций УКД близки (в пределах 30°) к локальным минимумам СФО. Мы подробнее обсудим конформационные возможности СФО ниже, используя наши новые данные. Здесь же укажем только на тот вывод, что химически однородный СФО вносит существенный вклад не только в формирование различных конформаций УКД, но и в зависимость пространственной структуры УКД от последовательности оснований. Другие полинуклеотидные дуплексы, не-УКД, такие как левоспиральные Z-формы, параллельные и хугстиновские дуплексы или дДМФ с не-уотсон-криковскими парами, не обеспечивают эту столь важную для биологической роли ДНК зависимость [32].

Уже в 80–90-х годах прошлого века рассмотрение конформационных характеристик фрагментов ДНК, депонированных в базе данных NDB, показало, что конформационное разнообразие минимальных фрагментов одиночной цепи ДНК не ограничивается только A-, B- и Z-семействами (последнее семейство не относится к УКД, поскольку имеет син-ориентацию сахара относительно основания, то есть другую область гликозидного торсионного угла χ пуриновых нуклеозидов, поэтому здесь мы такие конформации не рассматриваем). B-семейство было подразделено на два – VI («классическую» B-форму) и VII, отличающиеся между собой областями двух торсионных углов ζ и ϵ (рис. 1). Если для семейства VI эти углы находятся в областях гош[–] и транс соответственно, то для семейства VII эти области изменяются на транс (ζ) и гош[–] (ϵ). После обнаружения конформаций VII были выявлены фрагменты, соответствующие дДМФ, с локальной конформацией AII (отличающейся от AI областями углов α и γ), а также целый ряд других локальных конформаций, в том числе такие, которые содержат один нуклеозид, имеющий гликозидный торсионный угол χ и упаковку дезоксирибозы, характерную для формы B (C2'-эндо или близкие), а другой нуклеозид, имеющий эти характеристики, типичные для формы A (упаковка сахара C3'-эндо). Б. Шнейдер с соавторами (см., например, работы [24, 36, 37]) уже более десяти лет проводят анализ и классификацию конфор-

маций элементарных фрагментов полинуклеотидной цепочки. В последних публикациях они обозначают классы этих фрагментов как NtC, динуклеотидный конформационный класс, который соответствует рассматриваемым нами конформациям дДМФ. Мы будем далее обозначать эти конформационные классы как КК и анализировать, основываясь на результатах расчетов, отличие геометрических и энергетических характеристик экспериментально обнаруженных и оптимизированных конформаций элементарных фрагментов, принадлежащих разным классам. Оказалось, что многие из экспериментально наблюдаемых КК имеют области торсионных углов СФО, совместимые с УКД. В основном эти КК мы рассматриваем в наших последних работах. Не останавливаясь на деталях отнесения конформаций дДМФ к тому или иному классу, отметим только, что VI-конформация относится к классу VB00, VII – к классу VB07, AI – к классу AA00. Встречается также целый ряд других КК, обозначаемых как BB (оба нуклеозида в B-подобной конформации), например, BB04 (промежуточный между VB00 и VB07) или BB12 (где торсионные углы α , β и γ находятся в областях, отличных от VB00 и VB07) и классы с обозначениями AB (первый нуклеозид в форме A, второй – B) и BA (первый нуклеозид имеет конформацию B, второй – A). Различия конформаций дДМФ, принадлежащих к трем КК (VB00, VB07 и BB12) и торсионные углы, вызывающие это отличие, показаны на рис. 2. На рис. 3 различия структуры СФО этих КК продемонстрированы их наложением, так, чтобы гликозидные связи в 5'-нуклеозиде всех трех структур совпадали.

Важно отметить, что, если на первых этапах исследования пространственной структуры ДНК рассматривались различия характеристик фрагментов двойной спирали, состоящих из нескольких нуклеотидных пар (включая число пар на виток, положение оснований одной пары относительно другой и относительно оси спирали, взаимное положение оснований в паре), то экспериментально обнаруженное многообразие КК относится к элементарным фрагментам одной полинуклеотидной цепи. Многие структуры с фрагментами УКД, координаты которых имеются в базе данных NDB, содержат элементарные фрагменты полинуклеотидных цепей, относящиеся к нескольким КК. Это ставит задачу исследования причин такого многообразия и вклада взаимодействий субъединиц цепи в образование разных КК, а также задачу выяснения роли этого многообразия и отдельных КК в биологическом функционировании ДНК. Некоторые последние наши результаты на пути решения этих задач обсуждаются в следующем разделе.

Роль сахаро-фосфатного остова в формировании и конформационных особенностях разных

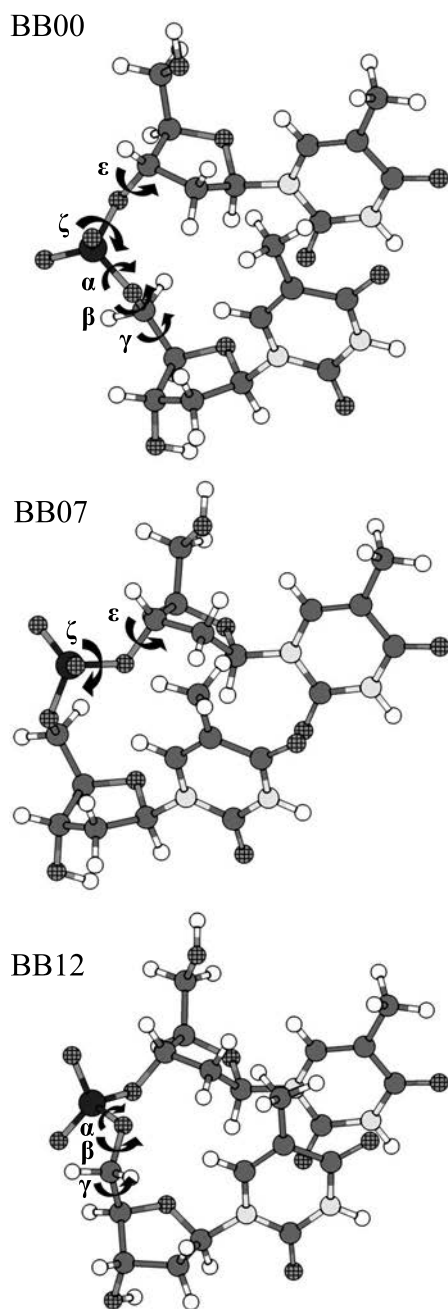


Рис. 2. Конформации дДМФ (dTpdT), принадлежащие трем КК. Стрелками обозначены торсионные углы, области значений которых различаются в разных КК.

классов уотсон-криковских дуплексов. Само существование разных конформационных классов, то есть разных наборов торсионных углов СФО дуплексов, соответствующих В- или А- структурам ДНК, наводит на мысль о важной роли СФО. Мы сделали вывод об этой роли в первой же работе, посвященной квантово-механическому исследованию минимальных фрагментов полинуклеотидной цепи [27], до того, как многообразие КК

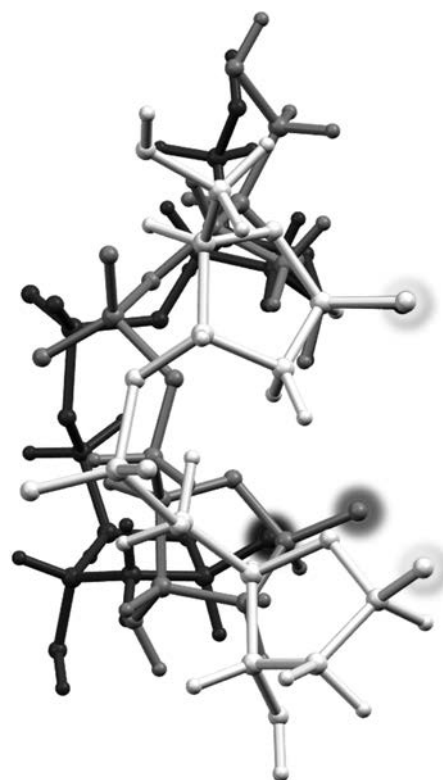


Рис. 3. Конформации СФО трех КК УКД (ВВ00 — черный цвет, ВВ07 — светлый, ВВ12 — серый), наложенные так, чтобы совпадали положения атомов С1 и направления гликозидных связей в полинуклеотидной цепи С5' нуклеозида. Сферами обозначены атомы водорода, на которые заменены основания цепи при построении СФО по координатам соответствующего дДМФ.

было систематизировано и классифицировано в работах Б. Шнейдера с соавторами (см., например, работу [24]). В последующих работах [30, 31] мы показали, что направленность сахаро-фосфатного остова и предпочтительные области его торсионных углов вместе с различиями геометрии пуринов от пиримидинов определяют зависимость пространственной структуры ВІ- и АІ-конформаций УКД от нуклеотидной последовательности. Многообразие экспериментально наблюдаемых КК УКД ставит вопрос о механизмах их формирования, о роли взаимодействий субъединиц ДНК в том, что в УКД реализуются разные наборы торсионных углов, совместимые с такой структурой. Можно было бы предположить, что это многообразие связано с внешними факторами (например, с взаимодействиями с другими молекулами в кристалле), но анализ конформационных предпочтений разных нуклеотидных последовательностей указывает на то, что это не может быть не только единственной, но и главной причиной существования разных КК. Опре-

деляющей является роль молекулярной структуры самой ДНК.

В течение последних нескольких лет мы провели расчеты и оптимизацию широкого набора дДМФ, кдДМФ и СФО, соответствующих 16 часто встречающимся КК, в которых представлены различные наборы торсионных углов СФО, совместимые с УКД. Предварительные результаты анализа этих расчетов были нами недавно опубликованы [34], подробное описание значительной части этой работы приведено в диссертации В. Домингез и будет опубликовано позднее, после более углубленного анализа и дополнительных расчетов. Рассматриваемые КК включают как наиболее изученные ранее конформации, такие как ВВ00 (канонические структуры семейства ВI), ВВ07 (ВII), АА00 (АI) и АА01 (АII), так и ряд других, достаточно часто встречающихся в банке данных. Среди них — семь КК с обоими сахарами, характерными для В-конформаций (ВВ02, ВВ04, ВВ07, ВВ12, ВВ13, ВВ10 и ВВ15), а также пять КК с одним из сахаров, характерным для В-конформаций, а другим — характерным для А-конформаций (ВА01, ВА05, ВА08, ВА09 и ВА16). Количество дДМФ, относящихся к рассмотренным классам, превышает 70% от их общего числа в банке данных, то есть рассмотренная выборка КК может считаться вполне представительной, учитывая, что часть дДМФ в структурах фрагментов ДНК, представленных в банке данных, не соответствует УКД (например, Z-конформации, квадруплексы, одноцепочечные участки), а еще 18.9% дДМФ не отнесены ни к одному из КК [24].

Здесь мы сосредоточимся на результатах, относящихся к сахаро-фосфатному остову и приведем дополнительные соображения о важном вкладе СФО в удивительную целесообразность молекулярной структуры ДНК, в ее приспособленность к выполнению важнейших биологических функций. Кроме важности этого вклада в структуру и функционирование ДНК, рассмотрение СФО позволяет на время обойти некоторые трудности оценки взаимодействий во фрагментах ДНК, рассмотренные нами недавно [20], в частности, трудности, связанные с учетом стекин-г-взаимодействий оснований и относительных вкладов стекинга и Н-связей. В результате оптимизации геометрии фрагментов СФО нескольких дДМФ для каждого из рассматриваемых КК, используя разные расчетные методы, было найдено по несколько локальных минимумов энергии, отличающихся между собой значениями торсионных углов (иногда значительно, а иногда в пределах всего нескольких градусов) и энергии (от нескольких десятых до нескольких ккал/мол). Было выявлено, что минимумы энергии СФО для одного КК могут соответствовать торсионным углам другого КК, а иногда торсионным углам, не соответствующим ни одному из классифициро-

ванных в работе [24] КК. По этим результатам каждый из рассмотренных (и, можно предположить, всех) КК, допускающих образование УКД, может быть отнесен к одной из двух групп — КК, имеющие средние экспериментальные значения торсионных углов, отличающиеся от одного из локальных минимумов энергии отдельного фрагмента СФО этого класса не более чем на 30°, и КК, для которых хотя один (а большей частью два или три) из средних торсионных углов экспериментальных структур отличается от его значения в любом локальном минимуме энергии отдельного СФО этого КК больше, чем на 30° [34]. Хотя выбор точной величины отличий (30°) в некоторой степени субъективен, его обоснованием можно считать то, что соответствующие торсионные углы экспериментальных структур дДМФ, принадлежащих каждому из рассмотренных КК, могут отличаться друг от друга до 60°, а от средней величины этого угла — до 30°. Поэтому конформации первой группы КК будем называть соответствующими, а второй группы не соответствующими минимумам энергии СФО. Результаты анализа расчетов дДМФ и кдДМФ указывают на то, что в подавляющем большинстве этих фрагментов, относящихся к первой группе, соблюдаются закономерности формирования пространственной структуры, выявленные нами ранее [27, 31] для «канонических» ВI- и АI-конформаций и упомянутые в предыдущем разделе. Рассмотрение большего числа экспериментальных структур и оптимизация геометрии СФО некоторых из них выявили исключения из этих правил; для некоторых КК эти исключения составляют до 10% проанализированных структур. Для понимания причин существования этих исключений предстоит как более детальный анализ экспериментальных структур, так и оптимизация разными методами большего числа дДМФ, кдДМФ и СФО. Для КК, относящихся ко второй группе, наблюдается более сложная картина — хотя для ряда дДМФ соблюдаются такие же закономерности, как для КК первой группы, для других КК второй группы дДМФ и кдДМФ не могут быть подходящими моделями для предсказания закономерностей формирования их пространственной структуры, которая зависит и от взаимодействий с соседними фрагментами и/или с другими молекулами. Составление процента дДМФ каждого из КК (по данным [24] от общего количества дДМФ во фрагментах ДНК, имеющих в базе данных NDB, с их принадлежностью к одной из двух групп позволяет заключить, что более половины дДМФ всех рассмотренных нами КК относятся к первой группе. Это более чем в три раза превышает процент дДМФ в КК, относящихся ко второй группе. По нашим уточненным данным, конформационные семейства ВI, АI и АII относятся к первой группе. Большинство ВII (ВВ07) конфор-

маций дДМФ, как показало наше рассмотрение, относятся ко второй группе КК.

К первой группе относятся девять рассмотренных КК (AA00, AA01, BB00, BB01, BB10, BB12, BA01, BA05, BA09), ко второй — семь (BB02, BB04, BB07, BB13, BB15, BA08 и BA16). Торсионные углы СФО, отвечающие одному или нескольким наиболее глубоким минимумам энергии, рассчитанным разными методами, приведены в нашей недавней работе [34]. Более глубокий анализ этих и полученных нами в самое последнее время результатов расчетов СФО позволяет сделать дополнительные выводы, которые мы иллюстрируем здесь на примере нескольких выбранных КК. Среди этих классов — рассмотренные ранее BB00, BB07, AA00 и классы, отнесенные к первой группе, — BB12, отличающийся от класса BB00 торсионными углами α , β и γ , а также BA05 и BA01, отличающиеся от BB00 более, чем на 30° , только углом δ_2 , связанным с упаковкой кольца дезоксирибозы второго нуклеотида.

В табл. 1 приведены средние значения торсионных углов СФО для дДМФ этих КК (по данным работы [24]) и отличия от них рассчитанных нами разными методами углов отдельных молекул СФО в наиболее глубоком минимуме, для класса BB07 — в двух наиболее глубоких минимумах энергии. Первый из этих минимумов, полученный оптимизацией СФО одного из дДМФ, отвечающего КК BB07, соответствует наименьшему из всех минимумов отклонению от средних экспериментальных значений торсионных углов (хотя и большему, чем 30° для одного из них), а второй — наименьшим относительным значениям энергии. Анализ расчетных данных показал, что иногда два различных КК могут считаться одним с точки зрения энергии соответствующих СФО, дДМФ, кДМФ. Например, к каноническому семейству конформаций VI следует отнести два КК — BB00 и BB01, торсионные углы СФО которых отличаются в среднем всего на 7° , а конформационные параметры в минимумах энергии СФО, дДМФ и кДМФ для одинаковых нуклеотидных последовательностей совпадают. Одним классом с точки зрения внутримолекулярных взаимодействий можно также считать BA01 и BA05, имеющие такое же среднее отличие значений соответствующих углов СФО одного КК от другого, близкие значения гликозидных торсионных углов и общие минимумы энергии СФО. Из анализа результатов этой таблицы можно заключить, что, в подавляющем большинстве случаев, отличия средних экспериментальных значений торсионных углов от соответствующих значений в минимумах энергии СФО, полученных разными методами, достаточно малы и часто имеют одинаковое направление (в сторону больших или меньших значений по сравнению со средним экс-

периментальным значением для этого класса). Имеющиеся существенные отличия (например, угла δ_2 класса BB12, полученного методом теории функционала плотности с функционалом M05-2X) связаны с тем, что мы приводим данные для наиболее глубокого минимума, а другой, несколько менее глубокий минимум, полученный этим методом, имеет близкие к полученным другими методами значения углов. Более близкого соответствия значений, полученных разными методами, трудно и ожидать, учитывая различия методов расчета (*ab initio*, теория функционала плотности, ММ). При отнесении КК к одной из двух групп мы старались учесть результаты всех методов, в большинстве случаев это отнесение соответствует результатам, полученным каждым из примененных методов. В некоторых случаях, правда, приходилось игнорировать результат, полученный одним из методов для одной из нескольких рассмотренных конформаций данного класса. Например, оптимизация СФО методом MP2 одного из рассмотренных дДМФ, принадлежащего к BB00 классу, дает структуру, соответствующую минимуму энергии класса BA05, тогда как оптимизация других СФО этим методом и этого СФО другими методами приводит к минимуму, соответствующему КК BB00. Подобная ситуация наблюдалась при оптимизации СФО двух дДМФ, принадлежащих к КК AA00, а оптимизация СФО одного из дДМФ, принадлежащих к классу BA05, приводила к минимуму, соответствующему BB00.

Представляет интерес сравнение энергий молекул, соответствующих минимальным фрагментам СФО разных КК. При расчетах каждым из применяемых нами методов наиболее глубокий из минимумов соответствует VI-семейству конформаций, наиболее часто встречающемуся во фрагментах ДНК и ее комплексов, представленных в банке данных. Это относится к найденным минимумам, как соответствующим определенным КК, так и к минимумам, не соответствующим ни одному из КК; как для рассматриваемых здесь КК, так и других, в том числе не соответствующих УКД (например, Z-конформациям). В табл. 2 приведены значения энергии в наиболее глубоких минимумах для каждого из рассматриваемых здесь КК, рассчитанные разными методами и отнесенные к энергии минимума для класса BB00 (и BB01). Интересной характеристикой набора торсионных углов СФО, соответствующих этому классу, является то, что такие же наборы реализуются в других конформациях дДМФ, например, в хугстиновских и параллельных дуплексах, а также в квадруплексах.

Стоит отметить, что мы не проводили систематический конформационный анализ фрагмента СФО. Эта задача требует слишком много ком-

Таблица 1. Средние значения торсионных углов СФО рассмотренных КК

КК	Метод	δ_1	ϵ	ζ	α	β	γ	δ_2
BV00, BV01	<BV00> [24]	138	183	258	304	180	44	138
	<BV01> [24]	131	181	266	301	176	49	120
	PBE	-1	+24	+18	-9	-11	+5	-2
	M05-2X	+7	+10	+16	-9	-17	+5	+2
	MP2	+5	+14	+18	-11	-16	+4	+2
	MM	+5	+11	+24	-13	-9	+5	0
BA05, BA01	<BA05> [24]	131	184	269	296	169	52	104
	<BA01> [24]	136	189	255	300	161	53	88
	PBE	+6	+24	+14	+5	+4	+4	-11
	M05-2X	+9	+5	+5	-1	-2	+6	-9
	MP2	+9	+24	+17	-8	+4	+4	-4
	MM	+3	+14	+17	-7	+5	0	+8
BV12	<BV12> [24]	140	196	280	257	76	171	140
	PBE	-2	+9	-6	+19	+7	+6	-8
	M05-2X	+1	+7	-2	+14	+3	+7	+25
	MP2	+1	+5	-5	+20	+5	+7	-7
	MM	0	+2	-4	+24	-3	+15	-2
AA00	<AA00> [24]	82	206	288	293	173	55	82
	PBE	+20	-19	-12	0	-5	0	+10
	M05-2X	+20	-24	-11	+1	-3	+2	+19
	MP2	+20	-24	-11	+2	-5	+1	+17
	MM	+16	-21	-4	-3	-1	-2	+29
BV07	<BV07> [24]	144	247	169	297	141	46	141
	PBE	-15	+ <u>31</u>	-12	-5	-16	-3	-3
		-8	- <u>49</u>	0	-6	+ <u>45</u>	+2	-4
	M05-2X	-17	+ <u>31</u>	-19	-6	-25	-3	+3
		-3	- <u>52</u>	-3	-2	+ <u>48</u>	+2	+1
	MP2	-14	+ <u>33</u>	-14	-5	-22	-6	+3
		-5	- <u>52</u>	-3	-2	+ <u>40</u>	+1	0
	MM	-12	+29	-12	-2	- <u>37</u>	0	-3
		-7	- <u>37</u>	-6	-6	- <u>47</u>	-1	-12

Примечание. Приведены значения углов в градусах (по данным работы [24]) и отклонения от этих значений соответствующих углов в наиболее глубоких минимумах энергии, рассчитанных разными методами. В случаях, когда два КК имеют одинаковый минимум энергии СФО, приведены отклонения от средних значений первого из них. Жирным шрифтом выделены средние значения торсионных углов каждого КК, отличающиеся от соответствующих значений для класса BV00. Для КК BV07 приведены отличия для двух минимумов энергии. Подчеркнуты отличия торсионных углов от средних для КК большие, чем 30°.

пьютерного времени и не очень интересна с точки зрения исследования ДНК. Мы оптимизировали структуры СФО, соответствующие выбранному КК УКД. Многие другие минимумы СФО могут не соответствовать не только УКД, но и вообще полинуклеотидной цепочке из-за слишком коротких контактов атомов оснований. Такими могут быть и минимумы, получае-

мые при оптимизации геометрии СФО, соответствующих второй группе КК.

Отсутствие среди полученных нами минимумов структур с более выгодной, чем для VI-конформаций, энергией не исключает их существования, хотя и делает его мало вероятным для КК, соответствующих УКД. Интересно отметить, что порядок увеличения энергии СФО при переходе

Таблица 2. Относительные значения энергии (ккал/мол) в наиболее глубоких минимумах, полученные при оптимизации фрагментов СФО рассматриваемых КК

КК	PBE	M05-2X	MP2	ММ
BB00, BB01	0.0	0.0	0.0	0.0
BA05, BA01	1.51	2.38	2.09	2.75
BB12	2.30	2.72	3.23	4.51
AA00	2.85	3.60	3.71	3.49
BB07	5.14	5.46	5.44	8.91
BB07	2.18	3.05	3.30	8.51

Примечание. Первый минимум для СФО BB07 соответствует наименьшему отклонению торсионных углов от среднего значения для КК, а второй — несколько большему отклонению, но менее положительной относительной энергии.

от одного КК к другому сохраняется при расчетах всеми тремя используемыми нами квантово-механическими методами (а также в отдельных, не-систематических расчетах с функционалом PBE и расширенным набором базисных функций, 6-311G(d)), но не во всех результатах, полученных методом ММ.

Анализ результатов оптимизации СФО позволяет заключить, что многие, но не все дДМФ в составе имеющихся в банке данных двуспиральных участков ДНК имеют торсионные углы СФО близкие к локальным минимумам его отдельного минимального фрагмента. Анализ показывает, что многие часто встречающиеся КК дДМФ имеют конформацию СФО, близкую к наиболее глубоким минимумам. Эти результаты можно рассматривать, как дополнительное подтверждение удивительной целесообразности молекулярной структуры главной биомолекулы. Вместе с тем они требуют дополнительных, более широких и глубоких исследований физической природы и вклада взаимодействий отдельных субъединиц ДНК в образование конформаций УКД, не соответствующих минимумам энергии СФО (отвечающих КК второй группы). Такие исследования нами сейчас проводятся.

БЛАГОДАРНОСТИ

Авторы благодарят Национальную суперкомпьютерную лабораторию Юго-востока Мексики (LNS), члена сети национальных лабораторий Национального совета по науке и технологиям (CONACYT), за представление компьютерных ресурсов и техническое сопровождение и поддержку. Авторы также благодарят В.М. Анисимова (Национальная лаборатория Аргонн, США) за интерес к работе и полезное обсуждение результатов.

ФИНАНСИРОВАНИЕ РАБОТЫ

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

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

СПИСОК ЛИТЕРАТУРЫ

1. J. D. Watson and F. H. C. Crick, *Nature*, **171**, 964 (1953).
2. H. M. Berman, W. K. Olson, D. L. Beveridge, et al., *Biophys. J.*, **63**, 751 (1992).
3. Б. И. Сухоруков, В. И. Полтев и Л. А. Блюменфельд, Докл. АН СССР, **149**, 1380 (1963).
4. A. G. W. Leslie, S. Arnott, R. Chandrasekaran, and R. L. Ratliff, *J. Mol. Biol.*, **143**, 49 (1980).
5. V. Poltev, In *Handbook of Computational Chemistry*, Ed. by J. Leszczynski et al. (Springer Int. Publ., Switzerland, 2017), p. 21.
6. H. DeVoe and I. Tinoco, *J. Mol. Biol.*, **4**, 500 (1962).
7. D. F. Bradley, S. Lifson, and B. Honig, In *Electronic Aspects of Biochemistry*, Ed. by B. Pullman (Acad. Press Inc., New York, 1964), p. 77.
8. H. A. Nash and D. F. Bradley, *J. Chem. Phys.*, **45**, 1380 (1964).
9. J. Langlet, P. Claverie, F. Caron, and J. C. Boeue, *Int. J. Quant. Chem.*, **20**, 299 (1981).
10. N. Gresh, P. Claverie, and A. Pullman, *Int. J. Quant. Chem.*, **29**, 101 (1986).
11. A. Pullman and B. Pullman, *Adv. Quant. Chem.*, **4**, 267 (1968).
12. В. И. Полтев и Б. И. Сухоруков, Биофизика, **12**, 763 (1967).
13. F. A. Momany, R. McGuire, A. Burgess, and H. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).

14. В. Б. Журкин, В. И. Полтев и В. Л. Флорентьев, Молекуляр. биология, **14**, 1116 (1980).
15. V. I. Poltev, T. I. Grokhilina, and G. G. Malenkov, J. Biomol. Struct. Dyn., **2**, 413 (1984).
16. V. E. Khutorsky and V. I. Poltev, Nature, **264**, 483 (1976).
17. В. И. Полтев и В. И. Брусков, Молекуляр. биология, **11**, 661 (1977).
18. V. P. Chuprina and V. I. Poltev, Nucl. Acids Res., **11**, 5205 (1983).
19. M. Topal and J. Fresco, Nature, **263**, 285 (1976).
20. В. Полтев, А. Дерябина, В. Домингес и др. Биофизика, **64**, 243 (2019).
21. D. A. Case, H. M. Aktulga, K. Belfon, et al., *AMBER 2021* (University of California, San Francisco, 2021).
22. K. Vanommeslaeghe and A. D. MacKerell, Jr., Biochim. Biophys. Acta, **1850**, 861 (2015).
23. R. E. Dickerson and H. R. Drew, J. Mol. Biol., **149**, 761 (1981).
24. J. Černý, P. Božíková, J. Svoboda, and B. Schneider, Nucl. Acids Res., **48**, 6367 (2020).
25. I. Ivani, P. D. Dans, A. Noy, et al., Nat. Methods, **13**, 55 (2016).
26. M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., *Gaussian-09 Revision D.01* (Gaussian Inc, Wallingford, CT, USA, 2009).
27. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., J. Biomol. Struct. Dyn., **25**, 563 (2008).
28. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., Int. J. Quant. Chem., **110**, 2548 (2010).
29. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., Comput. Theor. Chem., **975**, 69 (2011).
30. V. I. Poltev, V. M. Anisimov, V. I. Danilov, et al., Biopolymers, **101**, 640 (2014).
31. В. И. Полтев, В. М. Анисимов, К. Санчес и др., Биофизика, **61**, 259 (2016).
32. V. Poltev, V. M. Anisimov, V. Dominguez, et al., J. Mol. Model., **24**, 46 (2018).
33. V. Poltev, V. M. Anisimov, V. Dominguez, et al., In *Proc. of Advances in Quantum Systems in Chemistry, Physics, and Biology*, Ed. by L. Mammì et al. (Springer: Cham, Germany, 2020), p. 233.
34. V. Poltev, V. M. Anisimov, V. Dominguez, et al., Computation, **9**, 98 (2021).
35. 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), p. 17.
36. D. Schneider, P. Božíková, I. Nečasová, et al., Acta Crystallogr. D, **74**, 52 (2018).
37. D. Svozil, J. Kalina, M. Omelka, and B. Schneider, Nucl. Acids Res., **36**, 3690 (2008).

The Marvelous DNA Macromolecule: Computer Modeling of DNA 3D Structure and Diversity of Watson-Crick Duplex Conformations

V. Poltev*, V. Dominguez*, A. Ruiz*, A. Deriabina*, and E. Gonzalez*

**Autonomous University of Puebla, Puebla 72570, Mexico*

The work reviews the development of ideas about the 3D structure of DNA and the mechanisms of its formation, from the discovery of the Double Helix to the present day. It tracks the methods of modeling the 3D structure at different stages of the study of the main molecule of life. The discussion underscores a marvelous expedient in DNA molecular structure and its adaptability to important biological functions based on the results of calculations of the intra- and intermolecular interactions of macromolecule subunits. The work presents new data about substantial contribution of chemically monotonous and conformationally flexible sugar-phosphate backbone to the formation of sequence-dependent 3D structure of DNA. The diversity of the conformational possibilities of DNA is visible both in the formation of duplexes (as well as triplexes and quadruplexes) with different geometric parameters for base pairs and in the formation of duplexes with Watson-Crick nucleoside pairs containing local conformations corresponding to different regions of the torsion angles of the sugar-phosphate backbone (different conformational classes). Based on our calculations, these classes can be divided into two groups. The first group includes local conformations in which the torsion angles are close to one of energy minima of its isolated elemental repeating fragment, and the second group consists of conformations with one or more of these angles deviating from that of the nearest energy minimum by more than 30°. The regularities of the formation of the local 3D structure of these two groups differ significantly.

Keywords: DNA, 3D structure, diversity of conformations, sugar-phosphate backbone, computer modeling, molecular mechanics, density functional theory, nonempirical calculations