

С.К. Игнатов

**КВАНТОВОХИМИЧЕСКОЕ
МОДЕЛИРОВАНИЕ
МОЛЕКУЛЯРНЫХ
СИСТЕМ**

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского»

С.К. Игнатов

КВАНТОВОХИМИЧЕСКОЕ МОДЕЛИРОВАНИЕ МОЛЕКУЛЯРНЫХ СИСТЕМ

Учебное пособие

Рекомендовано методической комиссией химического факультета
для студентов ННГУ, обучающихся по направлению подготовки
04.04.01 «Химия»

Нижний Новгород
2024

УДК 544.18
ББК 24.511.2
И-26

Рецензенты: д.х.н., профессор РАН **С.Ю. Кетков**
д.х.н., профессор **Е.В. Сулейманов**

Игнатов С.К.

И-26 **КВАНТОВОХИМИЧЕСКОЕ МОДЕЛИРОВАНИЕ
МОЛЕКУЛЯРНЫХ СИСТЕМ:** Учебное пособие / С.К. Игнатов. –
Нижний Новгород: Нижегородский госуниверситет им. Н.И.
Лобачевского, 2024. – 106 с.

ISBN 978-5-91326-938-6

В учебном пособии изложены основы описания молекулярных систем, рассматривающиеся в курсе «Молекулярное моделирование». Пособие охватывает материал по темам «Современные методы квантовой химии», «Квантовохимическое описание физико-химических свойств», «Визуализация молекулярных структур».

Для студентов химического факультета ННГУ, обучающихся по направлению подготовки 04.04.01 «Химия» и изучающих курсы «Квантовая химия», «Молекулярное моделирование».

Ответственный за выпуск:
председатель методической комиссии химического факультета ННГУ
д.х.н., профессор **О.В. Нипрук**

ISBN 978-5-91326938-6

УДК 544.18
ББК 24.511.2

© Игнатов С.К., 2024
© Нижегородский государственный
университет им. Н.И. Лобачевского, 2024

Оглавление

	Стр.
Введение	4
Глава 1. Квантовохимическое описание физико-химических свойств и реакционной способности.....	5
1.1. Основные приближения квантовой химии	5
1.2. Современные квантовохимические методы	10
1.3. Выбор базисных функций для квантовохимического расчета ...	16
1.4. Программы для квантовохимического расчета	23
Глава 2. Способы описания структуры химической системы.....	26
2.1. Декартовы координаты	26
2.2. Внутренние координаты	27
2.3. Координаты Z-матрицы	29
2.4. Дробные координаты	33
2.5. Форматы файлов	34
2.6. Молекулярные визуализаторы и молекулярные конструкторы...	39
Глава 3. Способы квантовохимического моделирования для изучения различных физико-химических свойств	41
3.1. Типы компьютерного эксперимента в квантовой химии	41
3.2. Оптимизация молекулярной геометрии	41
3.3. Расчет молекулярных свойств при фиксированной геометрии....	47
3.4. Расчет колебательных частот и термодинамических параметров	53
3.5. Поиск переходных состояний и оценка констант скорости реакций	64
3.6. Учет влияния среды на физико-химические свойства и реакционную способность	73
3.7. Визуализация механизмов реакций, изученных квантовохимическими методами	75
3.8. Изучение реакций методами молекулярной динамики с квантовохимическими потенциалами	82
Список литературы	103
Предметный указатель	105
Приложение	106

Введение

Молекулярное моделирование, которое часто называется атомистическим моделированием – теоретическое направление в химии, описывающее химические свойства вещества и химические превращения на основе моделей, в которых эти свойства проявляются как следствие взаимодействия и движения частиц субатомного, атомного и молекулярного масштаба. Этими частицами могут быть электроны, ядра, атомы, группы атомов или молекулы.

Молекулярное моделирование отличается от квантовой механики, на которой оно частично базируется, во-первых, тем, что описывает, главным образом, химические феномены; во-вторых, тем, что использует не только квантовые, но и классические уравнения. В отличие от других методов моделирования веществ и материалов (метод конечных элементов, термодинамическое моделирование и т.д.), базовыми взаимодействующими структурными элементами при молекулярном моделировании являются частицы не крупнее отдельных молекул.

Кроме того, важным разделом молекулярного моделирования является визуализация химических систем, основанная на применении современных компьютерных алгоритмов и программ для изображения молекул, молекулярных поверхностей и структуры материалов.

В настоящее время молекулярное моделирование базируется на нескольких теоретических методах и подходах, позволяющих предсказывать свойства системы в зависимости от взаимодействий и движений составляющих ее частиц:

квантовая химия – описание взаимодействий атомов и молекул, а также химических превращений методами квантовой механики;

молекулярная механика – описание взаимодействий атомов и молекул на основе заданных (эмпирических) классических потенциалов;

молекулярная динамика – описание свойств вещества и химических превращений путем расчета, отслеживания и усреднения траекторий движения большого ансамбля молекул или атомов. Траектории строятся на основе классических уравнений движения (законов Ньютона) молекул;

моделирование методом Монте-Карло – описание свойств вещества на основе генерации и усреднения большого количества случайно выбранных конфигураций молекул жидкости, раствора, твердого тела.

В настоящем пособии рассматриваются первая группа методов – методы, основанные на квантовохимическом описании атомно-молекулярных процессов.

Пособие предназначено для студентов химического факультета ННГУ, обучающихся по направлению подготовки 04.04.01 «Химия» и изучающих курс «Молекулярное моделирование». Пособие может быть также полезно аспирантам и научным сотрудникам, использующим методы молекулярного моделирования в своей научной работе.

Глава 1

Квантовохимическое описание физико-химических свойств и реакционной способности

1.1. Основные приближения квантовой химии

Квантовая химия описывает свойства атомов или молекул на основе квантовой механики. В основе такого описания лежит описание движения электронов и ядер молекулы стационарным уравнением Шредингера:

$$\hat{H}\Phi = E\Phi, \quad (1)$$

где \hat{H} – гамильтониан атома или молекулы, учитывающий кинетическую энергию движения всех электронов и ядер, потенциальную энергию притяжения электронов к ядрам, потенциальную энергию межэлектронного и межъядерного отталкивания; Φ – многоэлектронная волновая функция, зависящая от координат и спинов всех n электронов, а также координат всех N ядер системы.

$$\Phi = \Phi(q_1, q_2, \dots, q_n; \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N),$$

где $q_i = \{\mathbf{r}_i, \sigma_i\}$ – совокупность пространственных и спиновых координат электрона i ; \mathbf{R}_j – координаты ядра j . Одним из базовых приближений при решении молекулярного уравнения Шредингера является приближение Борна-Оппенгеймера, в соответствии с которым ядра атомов считаются неподвижными, а при изменении их положения электроны мгновенно «подстраиваются» под их новые координаты. Это позволяет исключить из волновой функции движение ядер и представить ее в виде явной зависимости только от координат электронов:

$$\Psi = \Psi(q_1, q_2, \dots, q_n). \quad (2)$$

Такое представление является приближенным, и это надо учитывать при молекулярном моделировании: не для всех химических процессов приближение Борна-Оппенгеймера является хорошим приближением. Существует ряд быстрых реакций (реакции высоковозбужденных молекул, перестройки высокоспиновых состояний молекулярных комплексов, некоторые фотохимические реакции), в которых применение стандартных подходов квантовой химии не позволит получить правильные решения.

Тем не менее, для подавляющего большинства термически активируемых химических реакций погрешности приближения Борна-Оппенгеймера значительно меньше точности экспериментальных измерений в химии.

Еще одним ограничением применения уравнения (1) является его нерелятивистский характер. В силу этого, в системах, где электроны движутся со скоростями, сравнимыми со скоростью света, это уравнение будет давать значительные погрешности. Такими системами являются тяжелые атомы и

молекулы, образуемые ими, поскольку внутренние электроны таких атомов вблизи ядра движутся с релятивистскими скоростями. Погрешность нерелятивистского уравнения Шредингера, таким образом, заметно увеличивается при описании атомов выше 4-5 периодов и весьма заметна при описании соединений лантаноидов, актиноидов и любых трансурановых элементов. В квантовой химии, однако, разработаны и совершенствуются специальные методы для учета релятивистских эффектов (так называемые релятивистские псевдопотенциалы), обсуждение которых выходит за рамки данной работы.

Подходом фундаментальной важности для решения уравнения (1) с волновой функцией (2) является *метод Хартри-Фока*, в котором спиновые переменные (спины α или β) отделяются от пространственных переменных и учитываются за счет представления волновой функции в виде определителя (детерминанта Слейтера):

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(1)\alpha(1) & \varphi_2(1)\alpha(1) & \dots & \varphi_m(1)\alpha(1) & \varphi_{m+1}(1)\beta(1) & \dots & \varphi_n(1)\beta(1) \\ \varphi_1(2)\alpha(2) & \varphi_2(2)\alpha(2) & \dots & \varphi_m(2)\alpha(2) & \varphi_{m+1}(2)\beta(2) & \dots & \varphi_n(2)\beta(2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \varphi_1(n)\alpha(n) & \varphi_2(n)\beta(n) & \dots & \varphi_m(n)\alpha(n) & \varphi_{m+1}(n)\beta(n) & \dots & \varphi_n(n)\beta(n) \end{vmatrix}. \quad (3)$$

Здесь числа в скобках обозначают совокупность пространственных или спиновых переменных электронов 1, 2, ... n. Перестановка любых строк или столбцов определителя соответствует перестановке координат двух электронов и автоматически меняет знак волновой функции. Такое представление удовлетворяет фундаментальным требованиям квантовой механики о перестановочной антисимметрии волновой функции фермионов и, одновременно, сводит сложную многоэлектронную функцию к произведению одноэлектронных орбиталей φ_i . Вариация таких функций с требованием добиться минимума энергии системы E сводит многоэлектронное уравнение Шредингера к системе одноэлектронных уравнений Хартри-Фока:

$$\hat{F}_i \varphi_i = \varepsilon_i \varphi_i, \quad i = 1, 2, \dots, n,$$

где \hat{F}_i – оператор Фока, отличающийся от точного электронного гамильтониана тем, что межэлектронное взаимодействие заменено взаимодействием каждого электрона с усредненным полем всех остальных электронов; ε_i – энергия электрона на орбитали i .

Существенное упрощение решения уравнений Хартри-Фока возникает при использовании приближения линейной комбинации молекулярных орбиталей (МО ЛКАО), согласно которому, одноэлектронные орбитали раскладываются по конечному числу одноэлектронных базисных функций χ_i , centered на ядрах атомов:

$$\varphi_i = \sum_{j=1}^m c_{ij} \chi_j, \quad i=1,2,\dots,n, \quad m \geq n. \quad (4)$$

Это приближение сводит уравнения Хартри-Фока, являющиеся дифференциальными уравнениями второго порядка, к алгебраическому уравнению Хартри-Фока-Рутана (или просто уравнению Рутана):

$$\mathbf{F}\mathbf{c}_i = \varepsilon_i \mathbf{S}\mathbf{c}_i. \quad (5)$$

Здесь каждый вектор \mathbf{c}_i – это набор коэффициентов c_{ij} , $j=1,2,\dots,n$ в разложении (4). С математической точки зрения уравнение (5) является так называемой обобщенной задачей на собственные значения матриц \mathbf{F} и \mathbf{S} размерности $n \times n$, которые называются матрицей Фока (фокианом) и матрицей интегралов перекрывания. Элементы этих матриц даются выражениями:

$$F_{ij} = H_{ij} + \sum_{k=1}^m \sum_{l=1}^m P_{kl} \left((ij | kl) - \frac{1}{2} (il | jk) \right) \quad (6)$$

$$H_{ij} = \int_V \chi_i(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 - \sum_{i=1}^{N_A} \frac{kZ_A e^2}{|\mathbf{r} - \mathbf{R}_A|} \right) \chi_j(\mathbf{r}) d\mathbf{r}$$

$$(ij | kl) = \int_V \int_V \chi_i(1) \chi_j(1) \frac{ke^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_k(2) \chi_l(2) d\mathbf{r}_1 d\mathbf{r}_2$$

$$P_{ij} = \sum_{s=1}^n n_{occ}(i) c_{si} c_{sj}$$

$$S_{ij} = \int_V \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) d\mathbf{r}$$

В этих выражениях $n_{occ}(i)$ – число электронов на i -той орбитали (заселенность), интегрирование проводится по всему пространству. Расселение электронов по орбиталям производится от орбиталей низшей энергии к орбиталям более высоких энергий так, чтобы полностью заселить все низшие по энергии состояния (так называемый *принцип Aufbau*).

Результатом решения уравнения (5) являются энергии орбиталей ε_i и коэффициенты разложения орбиталей по базисным функциям \mathbf{c}_i , то есть выражение одноэлектронных волновых функций в виде линейных комбинаций выбранных базисных функций. Базисные функции при этом не обязательно должны иметь строгий физический смысл или являться решениями какой-либо точной физической задачи. Достаточно, чтобы они были удобными для

вычислений интегралов H_{ij} и $(ij|kl)$, а также обеспечивали правильное разложение (4) при минимальном числе слагаемых.

В формулы входит величина P_{ij} , называемая матрицей плотности и зависящая от коэффициентов разложения c_i , которые являются искомыми результатами при решении уравнений Рутана. В силу этого, решение этих уравнений возможно только в ходе итерационной процедуры, называемой процедурой самосогласования или ССП-процедурой, то есть процедурой достижения самосогласованного поля. В ходе этой процедуры сначала задаются начальные (затравочные) величины c_i , в качестве которых часто используются результаты более простого квантовохимического метода, например расширенного метода Хюккеля. С этими коэффициентами формируется матрица плотности и фокиан. После решения системы (5) получают обновленные значения c_i , которые используются для формирования обновленной матрицы плотности и система решается повторно. Процедура самосогласования повторяется, пока энергии орбиталей и элементы матрицы плотности не перестанут существенно изменяться. В связи с важностью самосогласования, метод Хартри-Фока часто называют методом МО ССП (метод молекулярных орбиталей самосогласованного поля).

Сходимость процесса самосогласования является одним из ключевых факторов успешного выполнения квантовохимического расчета. При быстром достижении самосогласования решение уравнений Рутана требует всего нескольких итераций. В случае плохой сходимости решение может потребовать сотен и тысяч итераций, что увеличивает время расчета в сотни раз, либо же сходимость может быть не достигнута совсем. Для успешного достижения сходимости ССП критически важно правильное задание геометрии молекулы (длины связи и валентные углы не должны иметь нереалистично большие или малые значения), а также правильное задание молекулярного заряда и мультиплетности. Например, расстояния между химически связанными атомами желательно задавать в пределах 0.7–1.5 от суммы ковалентных радиусов связанных атомов. Даже при выполнении этих условий самосогласование может быть медленным или не достигаться. Для его улучшения большинство квантовохимических программ включают в себя так называемые конвергеры – специальные алгоритмы, улучшающие сходимость. В ряде программ конвергеры включаются автоматически, но в большинстве случаев их выбор и использование в расчете ложится на пользователя. Если сходимость ССП не достигнута, результаты расчета нефизичны, и их нельзя использовать в дальнейшем. Поэтому первый вопрос в ходе анализа результатов – убедиться, что в расчете не произошло ошибок при достижении самосогласования. В случае таких ошибок программы обычно выдают сообщения типа «*Convergence criterion not met*» или «*SCF not converged*», и расчет прерывается. В случае возникновения такой ошибки можно предпринять несколько шагов: (1) проверить правильность задания заряда и мультиплетности молекулы, а также команд выбора расчетного метода; (2) проверить и, в случае необходимости, улучшить задание молекулярной геометрии (часто помогает понижение симметрии молекулы и некоторое уменьшение начальных межатомных расстояний); (3) повторить

расчет, используя другой набор конвергеров и/или увеличив число итераций ССП. Например, в программе Gaussian мощный конвергер ХЭС включается командой SCF=(хс,MaxCycle=300).

После того, как самосогласование достигнуто, результатом решения уравнений Рутана являются энергии одноэлектронных орбиталей и коэффициенты их разложения по базисным функциям. Поскольку коэффициенты разложения являются, по сути, представлением волновой функции системы, любые наблюдаемые свойства изучаемой системы могут быть вычислены на основе этих величин.

Представление волновой функции многоэлектронной системы в виде детерминанта Слейтера (3) является одной из ключевых идей большинства методов современной квантовой химии. В этой связи важно помнить, что построение такого детерминанта может проводиться двумя различными способами, называемыми способами спаривания в волновой функции. Во-первых, можно считать, что каждая пространственная орбиталь φ_i описывает движение единственного электрона со спином α или β . Однако расчеты показывают, что в большинстве систем с замкнутыми электронными оболочками (то есть с четным числом электронов и нулевым полным спином, см. рис. 1а) пары электронов α или β имеют очень близкие пространственные распределения. Поэтому в системах с замкнутыми электронными оболочками можно с хорошей степенью точности считать, что каждая орбиталь φ_i описывает движение пары электронов с противоположными спинами. Тогда размерность детерминанта Слейтера и число орбиталей уменьшаются в два раза. Способ спаривания, когда каждому электрону соответствует своя пространственная орбиталь, называют неограниченным методом Хартри-Фока, НХФ (*unrestricted Hartree-Fock method, UHF*). Способ спаривания, когда каждая орбиталь описывает пару электронов с противоположным спином –ограниченным методом Хартри-Фока, ОХФ (*restricted Hartree-Fock method, RHF*). Применение модели RHF значительно упрощает и ускоряет расчет и его использование предпочтительно там, где он может быть применен. Такими системами является большинство молекул с нулевым полным спином, например, большинство органических соединений, вода, CO, CO₂, CH₄, SO₂, N₂, F₂, Cl₂, HCl и т.д. Однако многие молекулы имеют незамкнутую оболочку (см. рис. 1б). Такими системами являются радикалы, бирадикалы, многие возбужденные состояния молекул, высокоспиновые состояния молекулярных комплексов переходных металлов. В частности, основным (т.е. наиболее энергетически выгодным) состоянием молекулярного кислорода O₂ является бирадикальное триплетное состояние с двумя неспаренными электронами с параллельными спинами.

Указание способа спаривания является одним из ключевых параметров квантовохимического расчета. Он должен быть осознанно выбран пользователем на основе анализа электронного состояния системы, то есть анализа ее мультиплетности, заряда, числа электронов и возможности вырождения верхних занятых орбиталей. Неправильный выбор способа спаривания может привести к существенной погрешности или даже качественно неправильным результатам.

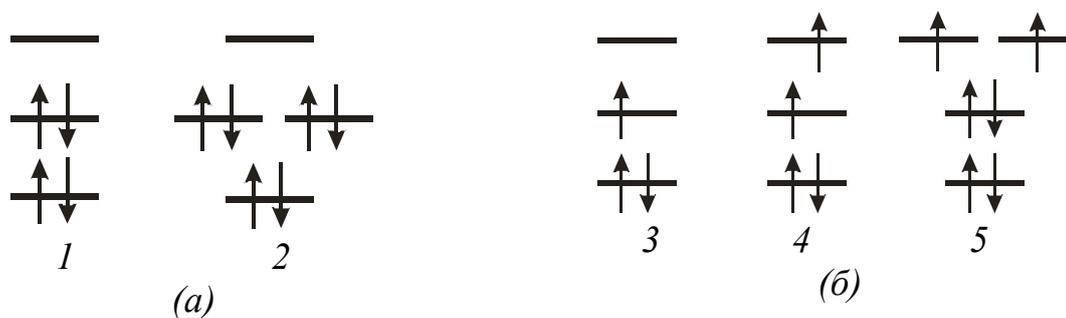


Рис. 1. Примеры молекул с замкнутыми (а) и незамкнутыми (б) электронными оболочками: 1 – молекула с нулевым полным спином и невырожденными орбиталями; 2 – молекула с вырожденными, но полностью заполненными орбиталями; 3 – радикал; 4 – возбужденное триплетное состояние молекулы; 5 – триплетный бирадикал, с вырожденной верхней занятой орбиталью

Следует отметить, что точность метода Хартри-Фока в виде уравнения (5) заметно хуже погрешностей современных экспериментальных измерений и уже не удовлетворяет запросам современной химии. Поэтому для практического применения в молекулярном моделировании этот метод почти никогда не используют в непосредственном виде. Вместо этого используются усовершенствованные теории, в которых метод Хартри-Фока является начальным приближением либо идейной базой для получения более точных решений.

Специально отметим, что некоторые квантовохимические теории, в частности, теория функционала плотности формально базируются на ином описании многоэлектронных систем. В этом описании волновая функция системы может не фигурировать вообще, вместо нее система описывается функцией электронной плотности. Таким образом, исходное уравнение (1) формально заменяется другими уравнениями, решение которых отличается от схемы (5)–(6). Однако фундаментом такой теории все равно остается описание системы методами квантовой механики.

1.2. Современные квантовохимические методы

Современные методы квантовой химии образуют три большие группы: неэмпирические методы, также называемые методами *ab initio*; методы теории функционала плотности (DFT); и полуэмпирические методы. Иногда DFT также относят к неэмпирическим методам, хотя строго говоря, эта теория включает ряд приближений, выбираемых эмпирически. Кроме того, иногда квантовохимические методы подразделяют на методы волновой функции (неэмпирические и полуэмпирические) и методы функционала плотности (DFT).

Неэмпирические методы (методы *ab initio*, от лат. «с начала») основаны на решении исходного уравнения Шредингера путем последовательного применения упрощающих приближений, строгость которых контролируется, а расчетные формулы выводятся дедуктивным путем. Таким путем получают строгие расчетные выражения, для которых можно добиться очень высокого

уровня точности, хотя вычислительные затраты для решения этих строгих уравнений также остаются высокими.

К неэмпирическим методам относятся метод Хартри-Фока (HF), методы теории возмущений Мёллера-Плессета (MP2, MP4), методы связанных кластеров (CC, их варианты CCSD, CCSD(T)), метод бракнеровских пар (BD), методы конфигурационного взаимодействия (CI, CIS, CISD, MCSCF, CASSCF).

Все эти методы, кроме методов конфигурационного взаимодействия, используют исходную волновую функцию в виде одного слейтеровского детерминанта. Этот детерминант может быть построен на основе способа спаривания RHF или UHF. Таким образом, методы теории возмущений, связанных кластеров и бракнеровских пар в случае открытых систем с открытыми оболочками могут функционировать аналогично UHF в вариантах UMP2, UMP4, UCCSD, UBD и т.д.

Метод Хартри-Фока (*HF в вариантах RHF и UHF*) кратко описан выше. Хотя его точность не удовлетворяет современным требованиям, он является основой других неэмпирических методов, поскольку позволяет сформировать волновую функцию, являющуюся начальным приближением для них.

Теория возмущений Мёллера-Плессета второго порядка MP2 значительно повышает точность расчета по сравнению с методом HF. Однако вычислительные затраты этого метода также заметно выше. В целом это метод средней точности со средней производительностью. Точность расчета структурных и энергетических параметров обычно достаточна для решения большинства задач органической, неорганической и металлоорганической химии. Однако поскольку почти такой же уровень точности сегодня обеспечивается методом DFT при заметно более низких затратах, использование MP2 ограничивается расчетом слабосвязанных комплексов, в которых DFT не работает, а также малых молекул, поскольку MP2 дает чуть более высокую точность, чем DFT. Следует отметить, что при расчете колебательных частот MP2 заметно завышает эти величины и для этих целей использование DFT часто остается предпочтительным.

Теория возмущений Мёллера-Плессета четвертого порядка MP4 значительно более затратный метод, чем MP2. Его не имеет смысла использовать для расчета молекул среднего и большого размера. Иногда он используется для оценки свойств ван-дер-ваальсовых комплексов, однако здесь он конкурирует с гораздо более точными методами CCSD и CCSD(T).

Методы связанных кластеров CCSD и, в первую очередь, CCSD(T) на сегодняшний день являются «золотым стандартом» точности в квантовой химии и используются для прецизионных оценок энергии и структурных параметров небольших молекул. Точность в случае CCSD(T) часто попадает в диапазон погрешностей экспериментальных методов. Вычислительные затраты этих методов очень велики и не позволяют проводить рутинные расчеты систем, включающих более чем 10 атомов. Эти методы в полной мере могут быть использованы для изучения слабосвязанных комплексов, в том числе ван-дер-ваальсовых молекул. Они также входят в состав так называемых композитных

экстраполяционных схем для высокоточного определения реакционных энтальпий типа G3, G4, CBS.

Метод бракнеровских пар VD по своим идеям и параметрам очень близок методам связанных кластеров, обладая сравнимо высокой точностью и вычислительными затратами. Для рутинных квантовохимических расчетов они применяются редко, однако используются в экстраполяционных схемах типа W1BD, которые дают рекордно высокую точность определения энтальпий небольших молекул.

Методы конфигурационного взаимодействия – это неэмпирические методы, в которых волновая функция представляется в виде не одного, а комбинации из нескольких или многих слейтеровских детерминантов. Это позволяет в отличие от всех остальных неэмпирических методов описывать системы и состояния с существенной нестационарной корреляцией – сложные системы с открытыми оболочками, например, синглетные бирадикалы и ионные пары; возбужденные и многоэлектронно-возбужденные состояния; комплексы с переносом заряда и процессы переноса электрона; высокоспиновые комплексы переходных металлов. Наиболее простые варианты этих методов CIS и CISD незначительно повышают вычислительные затраты. Однако их недостатком является невысокая точность из-за невозможности описания динамической электронной корреляции. Они применяются для описания электронных спектров молекул среднего размера (10-30 атомов), однако при отсутствии нестационарной корреляции времязависимая теория возмущений (TD-DFT, см. ниже) дает значительно более высокую точность при сравнимых затратах. Методы типа MCSCF и CAS SCF (последний является более удобным, устойчивым и экономичным вариантом MCSCF) обеспечивают более правильное и точное описание большинства систем со сложным распределением электронов при открытых оболочках. Однако вычислительные затраты очень быстро растут с размером системы и одновременно быстро растут запросы на используемое дисковое пространство и оперативную память. Эти методы следует применять только в тех случаях, когда использование однодетерминантных методов и DFT не позволяет даже качественно воспроизвести структуру волновой функции. Вариантом подхода CAS являются многоконфигурационные методы теории возмущений и связанных кластеров (CAS PT2, CAS MP2, CAS CCSD). Хотя эти методы в принципе позволяют одновременно описывать динамическую и нестационарную корреляцию, их применение до сих пор ограничено из-за сложности и вычислительных затрат, а метод CAS CCSD до сих пор находится в стадии экспериментальных разработок.

В основе *теории функционала плотности* лежит отказ от использования волновой функции для описания системы и использование для этой цели функции электронной плотности. В результате основным уравнением становится не уравнение Шредингера, а его аналог – уравнения Кона-Шэма или другие выражения. Выигрышем при этом является высокая вычислительная эффективность при незначительном ухудшении точности. Недостатком DFT является то, что при таком подходе остаются неизвестными точные и строгие формулы, описывающие ряд компонент энергии, что приводит к определенному

произволу при выборе расчетных формул. Такие расчетные формулы в DFT принято называть функционалами DFT. В настоящее время предложено около 130 различных функционалов, каждый из которых обладает определенными достоинствами и недостатками и часто ориентирован на различные системы. Различаются также точность и вычислительная эффективность каждого функционала, причем точность часто зависит от типа рассматриваемой химической системы.

В зависимости от того, какие физические эффекты описывают функционалы (хартри-фоковский обмен или корреляцию электронов), различают обменные или корреляционные функционалы. Большинство современных вариантов DFT включают формулы, описывающие комбинацию обменных и корреляционных эффектов.

Другое важное деление функционалов основано на приближении, используемом для описания электронной энергии. Выделяют следующие группы функционалов:

LDA – функционалы, в которых электронная энергия зависит только от функции электронной плотности (*приближение локальной плотности*, LDA);

GGA – включают зависимость энергии от электронной плотности и ее градиента (*обобщенное градиентное приближение*, GGA). Поскольку GGA – приближение более высокого уровня, чем LDA, большинство современных функционалов DFT используют именно его;

Гибридные – являются комбинацией функционалов LDA или GGA с вкладами в энергию, выражающимися через волновую функцию (например, выражения для обменной энергии метода Хартри-Фока), взятых с определенными весами, причем веса часто подбираются по экспериментальным данным.

Мета-гибридные – гибридные функционалы, в которых функционалом являются не только выражения потенциальной энергии взаимодействия внутри электронной плотности, но и выражения для кинетической энергии электронов.

Часто используются следующие функционалы:

BLYP – очень простой и «быстрый» GGA-функционал, хорошо описывает органические и легкие неорганические молекулы, однако точность уступает многим другим, например, B3LYP и PBE;

B3LYP – очень распространенный гибридный функционал общего применения, дает результаты средней точности на широком классе соединений с очень однородным распределением погрешностей для разных классов соединений. Рекомендуется как метод выбора для молекулярных систем, свойства которых еще не изучены, так как результаты будут достоверными даже для систем, в которых некоторые другие функционалы приведут к значительным погрешностям. Хорошо воспроизводит также свойства водородно-связанных систем и координационных соединений.

PW91, *BPW91*, *B3PW91* – функционалы, рекомендуемые для расчетов неорганических и металлических кристаллов и кластеров.

PBE – более современный, чем B3LYP, GGA-функционал широкого применения (существует его гибридный вариант *PBE0*). Часто рассматривается

как более точная и современная альтернатива B3LYP. Применим как молекулярным, так и кристаллическим системам, органическим, неорганическим и металлическим. Точность на многих классах соединений несколько лучше, чем у B3LYP, однако вычислительные затраты выше. Кроме того, сходимость процедуры SSP также часто хуже, чем у B3LYP.

M05, *M06* и их разновидности – функционалы, ориентированные на водородно-связанные системы. Точность на таких классах соединений значительно выше, чем у B3LYP или PBE. Однако очень часто возникают сложности со сходимостью процедуры SSP, даже в случае относительно простых систем типа кластеров воды. Более современный вариант этих функционалов *M08* часто рассматривается как один из лучших функционалов для соединений элементов главных групп.

TPSS, *SCAN* – современные мета-гибридные функционалы, часто рассматриваемые как образец «строгости» DFT, с минимальным использованием эмпирических поправок.

В последнее время метод DFT часто дополняется полуэмпирической схемой учета дисперсионных взаимодействий, оценка которых в DFT представляет определенную проблему (дисперсионные взаимодействия являются чисто корреляционными эффектами, для которых в DFT нет точной формулы). Обычно дисперсионно-корректированные функционалы обозначаются прибавлением к символу функционала аббревиатуры типа $-D_n$ (n – число, указывающее «поколение» используемого приближения), например, *PBE0-D4*, *B3LYP-D3BJ*. Некоторые функционалы изначально разработаны с учетом дисперсионных поправок, например часто используемый для расчетов возбужденных состояний и энергий электронного возбуждения функционал ω B97XD (символ D означает дисперсионные поправки). Отметим, что дисперсионные поправки, хотя оказываются полезными при расчетах органических реакций, не всегда точно воспроизводят тонкие дисперсионные взаимодействия в ван-дер-ваальсовых комплексах и не всегда улучшают расчет.

В целом, область применения DFT очень широка – органические, неорганические, металлоорганические, координационные и водородно-связанные соединения; молекулярные, кристаллические аморфные, полимерные. Перед неэмпирическими методами преимуществом является значительно более высокая производительность, что дает возможность рассматривать системы значительно большего размера (в 10 и более раз) при вполне сопоставимом уровне точности воспроизведения структурных, термодинамических, электронных и спектральных параметров. В то же время огромным преимуществом DFT перед полуэмпирическими методами является отсутствие необходимости калибровать атомные параметры для данного типа атомов или данного класса систем, а также устойчивость результатов при переходе к неизвестным классам соединений. В последнем случае полуэмпирические методы часто дают нефизические результаты из-за того, что эти соединения не учитывались при калибровке полуэмпирических параметров.

Системами, которые не рекомендуется рассматривать с помощью DFT, являются, во-первых, слабосвязанные и ван-дер-ваальсовы атомные и

молекулярные комплексы, поскольку дисперсионные взаимодействия, лежащие в их основе, не воспроизводятся теорией функционала плотности и могут быть описаны только на основе неэмпирических приближений высокого уровня (методы MP4, CCSD(T), BD). Во-вторых, системы с существенно многодетерминантным характером волновой функции (синглетные бирадикалы, ионные пары) также не могут быть корректно описаны DFT, поскольку в основе уравнений Кона-Шэма лежит однодетерминантное представление волновой функции. В этом случае следует использовать многоконфигурационные неэмпирические методы.

Несмотря на эти ограничения, DFT вполне применима и часто используется для расчетов возбужденных состояний и процессов электронного переноса. В основе этого обычно лежит подход времязависимой теории функционала плотности (*time-dependent DFT, TD-DFT*). Точность TD-DFT в этих случаях сопоставима, а во многих случаях выше, чем точность неэмпирических теорий CIS, CISD, и даже CAS SCF (при небольшом активном пространстве). Одновременно, производительность метода во много раз выше, что часто делает эту теорию методом выбора для оценки электронных спектров средних и больших систем. Для исследования процессов электронного возбуждения часто используются специальные модификации функционалов, т.н. *CAM-corrected (Coulomb-attenuated methods)* или *LR-corrected (Long-range corrected functionals)*, в которых формулы обменных или корреляционных вкладов модифицированы для улучшения описания электронных переходов. Примерами таких функционалов являются *CAM-B3LYP* или упоминавшийся выше *ω B97XD*. Эти функционалы, хотя и не гарантируют абсолютного превосходства над другими, часто дают значительно лучшее согласие с экспериментом при расчетах энергий электронного возбуждения, электронных спектров и возбужденных состояний молекул.

Полуэмпирические методы квантовой химии основаны на описании молекулярных систем на основе волновой функции и уравнения Шредингера. В отличие от неэмпирических методов, решение сложного многоэлектронного уравнения Шредингера достигается за счет того, что часть величин в нем заменяется на числовые параметры, калибруемые по экспериментальным данным. Процедура калибровки начинается с выбора реперного набора соединений с известными экспериментальными данными, реперными величинами (энтальпии образования, структурные параметры, дипольные моменты, потенциалы ионизации, колебательные частоты и т.д.). Затем производится расчет реперных соединений полуэмпирическим методом и калибруемые параметры варьируются таким образом, чтобы добиться наилучшего согласия результатов расчета с реперными величинами. В результате получается расчетная схема, производительность которой иногда в десятки и сотни раз выше, чем в случае методов *ab initio* или DFT, а точность расчета в классе соединений, близких к реперным, часто сравнима или даже лучше, чем в случае методов *ab initio* или DFT. Однако недостатком такого подхода является частое отсутствие необходимых полуэмпирических параметров для некоторых атомов. Кроме того, все полуэмпирические схемы страдают

общим недостатком – при переходе к неизвестным классам систем, которые не входили в реперный набор, полуэмпирический метод может приводить к значительно более неточным или даже качественно неправильным результатам. Вследствие этого, их применение сегодня ограничено либо очень большими системами, где никакие другие методы не применимы, либо использованием в качестве начальных приближений для более строгих методов.

Наиболее известными полуэмпирическими методами, используемыми сегодня, являются:

метод PM3 и его разновидности PM6, PM7. Один из наиболее последовательных методов, тщательно калиброванный по широкому спектру соединений, включая органические и неорганические молекулы атомов главных подгрупп и водородно-связанные системы. Методы PM6 и PM7 дополнительно включают параметризацию для соединений переходных металлов. Хорошо передает структуру, термодинамику, дипольные моменты, потенциалы ионизации, колебательные частоты. Для простых органических соединений по точности приближается к DFT, превосходя DFT по производительности в десятки раз. Средняя погрешность определения энтальпий образования порядка 5 ккал/моль.

Метод AM1 – метод, хорошо откалиброванный для описания органических и, прежде всего, водородно-связанных систем, в том числе биоорганических. Хорошо передает структуру, термодинамику, дипольные моменты, потенциалы ионизации, колебательные частоты этих соединений. Не включает параметры большинства переходных элементов. Производительность и характеристики точности примерно такие же, как в случае PM3.

Метод ZINDO и его разновидность ZINDO/S – полуэмпирический метод для оценки структуры и электронных спектров органических и неорганических соединений. Включает параметризацию многих элементов главных групп и некоторых переходных элементов. Производительность несколько выше, чем у PM3 и AM1, но точность оценки энергии и структуры заметно хуже. Его преимущество перед PM3 и AM1 в основном проявляется при оценке электронных, магнитно-резонансных параметров и спектров электронного возбуждения.

На рис. 2 представлена сравнительная характеристика различных методов квантовохимического расчета в координатах точность – производительность.



Рис. 2. Сравнительная характеристика производительности и точности квантовохимических методов

Контрольные вопросы

1. Какие методы подойдут лучше, если требуется рассчитать структуру молекулы SO_2 и сравнить их с высокоточными данными микроволновой спектроскопии?
2. Какие методы лучше использовать для описания структуры и свойств алкильных эфиров бензойной кислоты?
3. Какие методы можно использовать для изучения структуры олигомера целлюлозы, включающего 50 мономерных звеньев?
4. Какие методы подойдут для изучения свойств возбужденных состояний молекулы NO_2 ?

1.3. Выбор базисных функций для квантовохимического расчета

Решение уравнений Хартри-Фока-Рутана, а также аналогичных уравнений Кона-Шэма в DFT подразумевает задание базисных функций в разложении (4) для представления волновой функции молекулы или ее электронной плотности. Выбор базисных функций (базисного набора или базиса) – весьма ответственный этап при квантовохимическом исследовании. От правильного выбора зависит точность решения, время вычислений и даже возможность добиться самосогласования в процедуре SSP.

В настоящее время различают три основных типа базисных функций: слейтеровские, гауссовы и базисы плоских волн.

Слейтеровские базисные функции – функции, в которых радиальная часть функции является экспонентой или линейной комбинацией экспонент от радиальной переменной $r = \sqrt{x^2 + y^2 + z^2}$, например:

$$\chi(x, y, z) = \sum_{k=1}^K d_k x^l y^m z^n \exp(-\alpha_k r) \quad \text{или} \quad \chi(r, \theta, \varphi) = \left(\sum_{k=1}^K p_k(r) \exp(-\alpha_k r) \right) Y_{lm}(\theta, \varphi),$$

где d_k, α_k – постоянные коэффициенты; l, m, n – целые числа; $p_k(r)$ – полиномы от r . Слейтеровские функции требуют значительных затрат для вычисления атомных и молекулярных интегралов. В настоящее время такие базисные наборы применяются почти исключительно в полуэмпирических методах, где число интегралов невелико, причем в разложении используется только одна экспонента ($K = 1$). Другое применение этих функций – специализированные квантовохимические расчеты межмолекулярных взаимодействий и высоковозбужденных состояний атомов (ридберговских атомов), поскольку «хвосты» экспоненциальных функций простираются на значительно большие расстояния вокруг ядер.

Гауссовы базисные функции – функции, в которых радиальная часть функции является гауссовой функцией от расстояния, а угловая часть описывается либо сферической гармоникой (чистые сферические гауссовы функции χ_S), либо полиномом от декартовых координат (декартовы гауссовы функции χ_D), обеспечивающим угловую зависимость:

$$\chi_S(r, \theta, \varphi) = \left(\sum_{p=1}^P c_p \sum_{k=1}^K d_k p_k(r) \exp(-\alpha_k r^2) \right) Y_{lm}(\theta, \varphi)$$

$$\chi_D(x, y, z) = \sum_{p=1}^P c_p \sum_{k=1}^K d_k x^l y^m z^n \exp(-\alpha_k r^2)$$

В этих формулах число отдельных слагаемых (примитивных гауссовых функций $g_{lmn} = x^l y^m z^n \exp(-\alpha_k r^2)$) обычно значительно больше, чем в случае слейтеровского базиса и может достигать нескольких десятков для каждого атома. Однако вычисление молекулярных интегралов с такими функциями значительно проще, чем в случае слейтеровских функций, и может производиться путем очень быстрых рекуррентных процедур. Поэтому в целом гауссовы базисы приводят к значительному выигрышу в производительности, и сегодня большинство квантовохимических расчетов производятся именно с помощью таких базисов. Число K в этих формулах называется показателем контракции базиса, а число P – экспоненциальностью базиса. При этом коэффициенты c_p варьируются в ходе решения уравнений Рутана, а коэффициенты d_k остаются фиксированными. Такое представление позволяет добиться большой гибкости базиса при одновременном сокращении вычислительных затрат. Например, для базисной функции, описывающей p -

орбиталь атома С, можно заранее подобрать коэффициенты d по атомным параметрам, а в ходе процедуры ССП варьировать только коэффициенты c (по сути, эти коэффициенты являются коэффициентами c_{ij} в разложении (4)). Таким образом, например, p -АО будет «разбита» на две части (внутреннюю и внешнюю), варьировать независимо, что повышает гибкость полной волновой функции. Часто такие базисы называют *валентно-расщепленными базисами* или *двух- трех-, и т.д. экспоненциальными базисы*. Например трехэкспоненциальный базис соответствует $P = 3$, то есть каждая валентная оболочка (s -, p - и d -функции) состоит из трех радиальных частей, варьировать независимо.

Во многих случаях гауссовы базисные функции описывают только валентные и внутренние оболочки атомов, незаполненные оболочки не включаются в рассмотрение. Такой базисный набор называется валентным базисом. Точность расчета с таким базисом обычно невысока. Чтобы повысить эту точность, атомный базис дополняется функциями, называемыми поляризационными (или функциями высоких угловых моментов). Они являются обычными гауссовыми функциями, у которых угловые части соответствуют более высоким угловым моментам, чем это требуется для описания данной орбитали атома. Например, к s -АО атома водорода добавляются поляризационные функции p - или d -типа, а к p -АО тяжелых атомов (всех атомов, кроме водорода) добавляются функции d -, f -, g - или более высоких типов. Это позволяет сделать молекулярную орбиталь существенно более гибкой, позволяя ей поляризоваться в направлениях, запрещенных симметрией орбитали в атоме. В обозначениях базисов поляризационные функции обычно обозначаются символом типа функции в скобках после названия базиса, причем первая или единственная группа символов относится к тяжелым атомам (т.е. ко всем атомам, кроме атома водорода), вторая группа – к атомам H. Например, обозначения (d) означают наличие d -поляризационных функций у тяжелых атомов; (d,p) – d -функций у тяжелых атомов и p -функций у атомов водорода; (2d,2p) – по два набора d -функций у тяжелых атомов и p -функций у атомов H; (3df,3pd) – три набора d -функций + набор f -функций у тяжелых атомов и три набора p -функций + один набор d -функций у атомов H. Добавление поляризационных функций обеспечивает значительное улучшение точности, однако заметно увеличивает размер базиса и время расчета, поэтому для достижения оптимального баланса следует внимательно относиться к выбору базиса. В настоящее время применение поляризационных функций считается необходимым почти всегда. Их отсутствие часто приводят к большим погрешностям или даже качественно неправильным результатам. Например, молекула аммиака является плоской в валентном базисе 6-31G, в то время как базис с поляризационными d -функциями азота 6-31G(d) правильно воспроизводит пирамидальную структуру этой молекулы.

Еще одним типом дополнительных базисных функций являются так называемые диффузные базисные функции. Диффузные функции — это обычные гауссовы функции s - или p -типа с очень малой величиной α . В силу этого эти функции очень сильно делокализованы в пространстве, что позволяет

описать взаимодействия между удаленными молекулами. Применение таких функций является обязательным при расчетах межмолекулярных комплексов, водородно-связанных систем, адсорбции, высоковозбужденных состояний и ридберговских атомов, а также анионов. При изучении межмолекулярных взаимодействий диффузные функции существенно сокращают базисную суперпозиционную погрешность, что критически важно при оценке энергии координации слабосвязанных комплексов и энергии адсорбции. При изучении компактных валентно-связанных систем диффузные функции обычно не приводят к улучшению результатов, их использование только увеличивает время расчета. Диффузные функции в обозначении базиса часто обозначаются символами + (диффузные функции p -типа на тяжелых атомах), ++ (диффузные функции sp -типа на тяжелых атомах и s -типа на атомах H), либо префиксом *aug-* (от англ. *augmented*), например: 6-31+G(d), 6-31++G(2d,2p), aug-ccpVTZ.

Наиболее распространенные типы базисных наборов:

STO-3G – валентный базисный набор, в котором каждая орбиталь описывается тремя неконтрактированными ($P = 1$) гауссовыми функциями. Это очень примитивный базис, который не обеспечивает требуемой сегодня точности и применяется только для предварительных расчетов.

6-31G – двухэкспоненциальный базисный набор начального уровня, который обеспечивает среднюю точность вычислений. Для достижения хороших результатов он, как правило, должен дополняться поляризационными функциями: 6-31G(d), 6-31G(d,p). Также возможно его использование с диффузными функциями, например: 6-31+G(d), 6-31++G(d,p).

6-311G – трехэкспоненциальный базисный набор среднего уровня, обеспечивающий несколько более высокую точность вычислений, чем 6-31G. Он также должен дополняться поляризационными функциями: 6-311G(d,p), 6-311G(2d,2p), 6-311G(3df,3pd). Возможно его использование с диффузными функциями, например: 6-311++G(d,p), 6-311++G(2d,2p), 6-311++G(3df,3pd).

Базисы типа *cc-pVnZ* ($n = D, T, Q, 5, 6, 7$ и т.д.) – современные базисные наборы для высокоточных расчетов. Число или буква n – *cardinal number*, то есть экспоненциальность. Символ p в названии означает, что поляризационные функции уже включены в базисный набор. Таким образом, cc-pVDZ примерно соответствует базису 6-31G(d,p), cc-pVTZ – базису 6-311G(2d,2p), остальные являются значительно более широкими. Важным отличием базисов данного типа является то, что они калиброваны по результатам высокоточных расчетов с учетом энергии корреляции, что отражено в их названии (*cc* – *correlation consistent*). В связи с этим такие базисы используются для высокоточных расчетов методами CCSD, CCSD(T), BD. Использование в высокоточных методах малых и средних базисов не рекомендуется, поскольку точность метода будет нивелироваться погрешностями базиса. Диффузные функции могут быть включены в эти базисы путем добавления префикса *aug-*: aug-cc-pVTZ, aug-cc-pVQZ.

Базисы плоских волн – это базисные наборы, в которых базисные функции являются комбинациями гармонических функций. Эти базисы используются в расчетах систем с трансляционной симметрией в специальных программах, ориентированных на расчеты кристаллов, поверхностей, полимеров (программы VASP, Quantum Espresso, Wien2k). Однако в рамках таких программ эти базисы также могут быть применены и для расчетов систем без трансляционной симметрии (например, молекул или кластеров, помещенных в центр достаточно большой элементарной ячейки). Для молекул такое применение не оправдано, поскольку базис плоских волн хорошо описывает длинно-периодические особенности волновых функций, но требует очень большого количества коротковолновых базисных функций для описания локальных особенностей внутри молекул. Поскольку все базисные функции в таком базисе однотипны, задание базиса обычно сводится к указанию так называемой энергетической отсечки, то есть максимальной величины кинетической энергии свободно движущегося электрона (волны де Бройля), которая обратно пропорциональна квадрату длины волны. Чем более детально мы хотим описать нашу систему, тем больше коротковолновых компонент должно быть представлено в базисе, и тем большая величина отсечки должна быть задана. Величина отсечки, обеспечивающая расчеты хорошего уровня составляет 300-400 эВ и выше.

Отдельным вопросом при выборе базиса является описание электронов внутренних оболочек атомов (*остовных электронов*). Хотя эти электроны не определяют валентные свойства атома, количество остовных электронов очень велико, особенно для атомов высших периодов периодической системы. Кроме того, у атомов с высоким атомным номером электроны низших оболочек движутся со скоростями, которые уже нельзя считать нерелятивистскими. При описании таких электронов нерелятивистским уравнением Шредингера возникают значительные погрешности. Таким образом, хотя внутренние электроны не участвуют в образовании химических связей, их учет при квантовохимическом расчете создает значительные трудности. Одним из способов решения этой проблемы является использование так называемых *остовных псевдопотенциалов*, то есть особых функций, которые бы создавали вокруг ядра потенциал, аналогичный потенциалу, создаваемому остовными электронами. В этом случае при решении уравнения Шредингера достаточно учитывать только валентные электроны, движущиеся в остовном псевдопотенциале. Проблема выбора псевдопотенциала состоит в том, что взаимодействие между остовными и внешними электронами не сводится к чисто кулоновским взаимодействиям, а включает квантовые эффекты, поскольку внешние и внутренние электроны неразличимы и между ними имеет место квантовый обмен. Кроме того, усложняет задачу и релятивистский характер движения внутренних электронов. Тем не менее, в современной квантовой химии существует большое число вариантов решения этой задачи и разработано большое число псевдопотенциальных функций практически для всех тяжелых элементов таблицы Менделеева. Современные псевдопотенциалы различаются по сложности и в системах различного типа обеспечивают различный уровень точности. Псевдопотенциалы не имеет смысла применять для легких атомов (для

элементов 1, 2, и возможно, 3 периодов). Для элементов 3-4 периодов их применение оправдано, если требуется сократить время расчетов в больших системах. Однако надо помнить, что полноэлектронный расчет (без псевдопотенциала) для таких атомов почти всегда более точный, чем псевдопотенциальный. Для элементов пятого и более высоких периодов псевдопотенциальные расчеты обычно являются единственно возможными, и для многих тяжелых атомов часто вообще не существует полноэлектронных базисных наборов.

Наиболее известные и часто используемые псевдопотенциалы для молекулярных расчетов неорганических и металлоорганических соединений:

LANL2DZ – псевдопотенциал и соответствующий двухэкспоненциальный базис для большинства элементов 1-5 периодов + U, Np, Pu. По уровню точности сопоставим с базисом 6-31G(d,p).

LANL08 – псевдопотенциал и соответствующий двухэкспоненциальный базис для элементов 3-6 периодов с повышенной по сравнению с LANL2DZ точностью и, как заявлено авторами, лучшими свойствами сходимости. Однако число неконтрактированных базисных функций и, следовательно, скорость расчета заметно ниже.

CRENBL/CRENBS – псевдопотенциалы и соответствующей базисы для всех элементов периодической системы. Скорость расчета и сходимости очень высокие, однако уровень точности ниже, чем у LANL2DZ.

В большинстве квантовохимических программ базисные наборы и псевдопотенциалы заложены в специальные библиотеки внутри программ и могут быть использованы путем указания определенных ключевых слов либо могут быть заданы вручную в виде таблицы числовых значений. Если в программе отсутствует необходимый базис или псевдопотенциал, его можно получить из базы данных *Basis Set Exchange* [Basis Set Exchange: A Community Database for Computational Sciences Schuchardt, K.L., Didier, B.T., Elsethagen, T., Sun, L., Gurumoorthi, V., Chase, J., Li, J., and Windus, T.L. J. Chem. Inf. Model., 47(3), 1045-1052, 2007, doi:10.1021/ci600510j.], которая поддерживается лабораторией *Environmental Molecular Sciences Laboratory* исследовательского центра *Pacific Northwest National Laboratory*, США. Сайт базы данных <https://bse.pnl.gov/bse/portal>. В этой базе данных достаточно указать химический элемент или набор элементов, для которых нужны базис или псевдопотенциал. После этого возникает список доступных базисных наборов и псевдопотенциалов, которые можно скачать в нужном формате, указав программу, используемую для расчета.

Комбинация квантовохимического метода и базисного набора, включая при необходимости и псевдопотенциал, обычно называется уровнем квантовохимической теории. В англоязычной литературе иногда используется выражение *model chemistry* (модель химии). Выбор уровня теории является ключевым фактором правильности квантовохимического исследования.

Для правильного выбора уровня теории принимают во внимание следующее.

1. Какие свойства требуется описывать и с какой целью. Для высокоточной оценки энергии реакций и эталонных расчетов структуры требуют использования методов CCSD, CCSD(T), BD. Расчеты структуры, термодинамических свойств, колебательных спектров и энергетике реакций небольших и средних органических, неорганических и металлоорганических молекул возможны с помощью MP2 или DFT. Расчеты возбужденных состояний могут быть выполнены с помощью TD-DFT, CISD, CASSCF. Расчеты ионных пар, синглетных бирадикалов, высокоспиновых комплексов требуют применения многоконфигурационных методов типа MC SCF или CAS SCF, CASPT2, CAS MP2. Высокоточная оценка энтальпий и энергий обычно производится с помощью экстраполяционных методов G3, G4, CBS, W1, W1BD.

2. Размер описываемой химической системы (количество атомов). Расчеты систем до 5-10 атомов возможны с использованием любого современного метода и можно выбирать максимально высокий. Однако расчеты систем из 10-20 атомов как правило сужают круг неэмпирических методов до MP2. Системы 10-100 атомов доступны для расчета методом DFT. Разупорядоченные системы из более 100 атомов, как правило, удается рассчитывать только полуэмпирическими методами.

3. Наличие тяжелых атомов и атомов, для которых следует применять псевдопотенциалы. Расчеты атомов 5-6 периодов почти всегда требуют использования псевдопотенциалов. Псевдопотенциалы могут быть использованы и для сокращения времени расчета больших систем с участием элементов 3-4 периодов, даже если для таких элементов есть полноэлектронные базисные наборы.

4. Оценка энергии координации, энергии адсорбции, а также расчеты анионов значительно улучшаются при дополнении базиса диффузными функциями. Следует избегать базисов без поляризационных функций. Для высокоточных методов рекомендуется использование корреляционно-согласованных базисных наборов.

Контрольные вопросы

1. Требуется определить наиболее выгодную конформацию молекулы $C_6H_5C(O)OCH_2CH_3$ в вакууме. Какой уровень теории можно использовать?

2. Для интерпретации данных микроволновой спектроскопии необходимо установить длину связи в молекулах BeH и O_3 . Какой уровень теории следует использовать? Какой способ спаривания?

3. Необходимо установить энергию образования молекулярного комплекса воды с SiF_4 . Какой уровень теории может быть использован? Какие требования к базису?

4. Необходимо рассчитать энергию образования кислородной вакансии в кристалле ZrO_2 . Какой тип базиса наиболее часто используется в таких случаях?

5. Необходимо оценить энергию адсорбции молекул глиоксаля (НОС-СОН) на ледяной микрочастице в атмосфере. Микрочастица моделируется

кластером из 1000 молекул воды со структурой кристалла льда. Какой метод может подойти для расчета?

1.4. Программы для квантовохимического расчета

В настоящее время известно несколько десятков бесплатных и коммерческих квантовохимических программ для расчетов молекул и периодических систем, см., например, https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software. Среди них выделяются несколько наиболее широко используемых и надежно протестированных программ:

Gaussian – одна из наиболее известных универсальных программ квантовой химии. Реализует большинство известных современных квантовохимических методов расчета молекул, включая HF, MP2, MP4, CCSD/CCSD(T), BD, G2-G4, CBS, W1, CIS/CISD, CASSCF, CASMP2, полуэмпирические методы PM3, PM6, AM1, большое количество функционалов DFT, в том числе в варианте TDDFT, и другие методы. Позволяет рассчитывать большое количество молекулярных свойств, включая структуру, энергию, термодинамические характеристики, колебательные частоты, энергии электронных возбуждений, свойства молекул в возбужденных состояниях, влияние среды, исследовать ППЭ химических реакций и другие свойства. Имеются также возможности проведения молекулярно-динамических расчетов и расчетов периодических систем, хотя эти возможности в данной программе имеют ограничения по сравнению со специализированными программами. Для программы существует собственный визуализатор/конструктор GaussView, позволяющий визуально редактировать молекулярные структуры и запускать программу в диалоговом режиме. Gaussian – коммерческая (платная) программа, существуют версии для Windows, Linux, MacOS и суперкомпьютерных кластеров. Сайт <http://gaussian.com>

GAMESS – бесплатная универсальная квантовохимическая программа для молекулярных расчетов. Набор методов несколько уже, чем в случае Gaussian, однако имеется возможность проведения расчетов в рамках теорий HF, ROHF, GVB, MP2-MP4, CC/EOMCC, CASSCF, CIS/CISD, PM3, AM1. Существует три основных разновидности программы, поддерживаемых различными группами: GAMESS US, GAMESS UK, и Firefly (PC GAMESS), которые несколько различаются по своим возможностям. Версия Firefly разработана в МГУ (автор – А. Грановский). Версия GAMESS UK не является бесплатной для пользователей за пределами Великобритании. Возможности для расчета молекулярных свойств также несколько уже, чем в случае Gaussian, однако они включают геометрическую оптимизацию, расчет колебательных частот, расчет возбужденных состояний. Программа имеет ряд возможностей для расчетов комбинированными методами квантовой и молекулярной механики (QM/MM), учета влияния среды (метод эффективных фрагментных потенциалов) и интерфейсы сопряжения к другим молекулярно-динамическими и молекулярно-механическими программами. Особенностью программы Firefly является очень эффективная реализация методов MP2/MP4, которая позволяет проводить профессиональные расчеты этими методами на персональных компьютерах под управлением Windows. Кроме того, в программе Firefly реализован очень

эффективный метод ХМСQDPT2, представляющий собой версию CAS SCF с учетом электронной корреляции на уровне MP2. Метод позволяет гибко и устойчиво моделировать сложные возбужденные системы с открытыми оболочками, обеспечивая высокую точность и скорость расчета. Программа имеет версии для Windows, Linux и суперкомпьютерных систем. Сайты программы:

<https://www.msg.chem.iastate.edu/games/> (GAMESS UK),

<http://www.cfs.dl.ac.uk/index.shtml> (GAMESS UK),

<http://classic.chem.msu.su/gran/games/index.html> (Firefly).

NWchem – бесплатная универсальная квантовохимическая программа, реализующая большое количество основных квантовохимических методов для расчетов молекул. Список реализованных методов квантовой химии включает HF, MP2-MP4, CCSD/CCSD(T), DFT/TDDFT, CI, CASSCF. Программа позволяет оптимизировать структуру молекул, рассчитывать колебательный спектр и термодинамические характеристики, энергии и структуры возбужденных состояний. Программа также позволяет выполнять молекулярно-динамические расчеты с классическими потенциалами и расчеты периодических неметаллических систем в базисе плоских волн. Сайт программы <http://www.nwchem-sw.org>

HyperChem – программа для квантовохимических и молекулярно-динамических расчетов молекул, ориентированная на начинающих пользователей и обучения квантовой химии. Одна из немногих программ, имеющих собственный интегрированный оконный интерфейс и ориентированная на работу в экранном режиме. В этом режиме она позволяет нарисовать молекулу на экране, выбрать в меню уровень теории и тип расчета, выполнить расчет и проанализировать результаты. Расчеты могут проводиться методами молекулярной механики, полуэмпирическими методами, методами DFT, HF и MP2. Тип расчета включает расчеты при фиксированной геометрии, оптимизацию геометрии, расчет колебательных частот, поиск переходных состояний. При анализе можно строить различные молекулярные поверхности, включая орбитали, электронную и спиновую плотности, электростатический потенциал. После расчета частот можно визуализировать ИК спектр и анимировать моды нормальных колебаний. Программа работает под управлением ОС Windows, LINUX, MacOS. Хотя программа платная, доступна бесплатная пробная версия. Сайт программы: <http://www.hyper.com>.

Для проведения расчетов периодических систем (кристаллы, поверхности, полимеры), включая металлы, рекомендуется использовать специализированные программы. Среди таких программ выделяются:

VASP – коммерческая программа для проведения профессиональных расчетов периодических систем. Реализует большинство современных алгоритмов для проведения периодических расчетов, включая эффективные алгоритмы ускорения сходимости и параллелизации. Позволяет проводить расчеты в рамках теорий HF, DFT (LDA, GGA, гибридный функционалов и meta-GGA) с использованием нормосохраняющих и ультрамягких псевдопотенциалов в базисе проектированных дополненных плоских волн (PAW). Позволяет

расчеты большого количества свойств, включая статические и динамические диэлектрические свойства, пьезоэлектрические свойства, частотно-зависимые оптические свойства и энергии электронного возбуждения, фононные спектры, константы упругости, магнитные свойства и спин-орбитальные поправки. Изучение ППЭ периодических систем возможно на основе оптимизации геометрии, поиска переходных состояний и борн-оппенгеймеровской молекулярной динамики. Имеет версии для LINUX и суперкомпьютерных кластеров. Сайт программы: <https://www.vasp.at>.

QuantumEspresso – бесплатная программа для периодических расчетов DFT в базисе плоских волн. Позволяет проведение расчетов кристаллов, поверхностей и полимеров в основном состоянии, включая оптимизацию геометрии, исследование ППЭ и борн-оппенгеймеровскую динамику на основе DFT потенциала, квантово-классическую динамику Кара-Парринелло, поиск переходных структур методом скорректированной эластичной нити (NEB), расчеты фононных частот, диэлектрических свойств, ИК и КР активностей, расчеты спектральных свойств на основе TDDFT, а также расчеты параметров электронного переноса. Расчет электронной структуры включает различные варианты DFT (LDA, GGA, meta-GGA) с использованием нормосохраняющих и ультрамягких псевдопотенциалов в базисе проектированных дополненных плоских волн (PAW) с возможностью дополнительных поправок для оценки дисперсионных взаимодействий (поправки Гримме) и эффектов коррелированных электронов (поправки DFT+U). Сайт программы: <https://www.quantum-espresso.org>.

CRYSTAL – коммерческая программа для расчетов периодических систем, однако использующая не базис плоских волн, а обычный базис атомных орбиталей (*приближение КО ЛКАО*, кристаллическая орбиталь – линейная комбинация атомных орбиталей). Такой подход очень нетипичен для современных программ и имеет свои достоинства и недостатки. Достоинством является лучшее описание дефектов и других локальных свойств кристаллической ячейки. Недостаток состоит в том, что расчеты свойств кристалла как целого требуют несколько больших затрат, часто хуже сходятся и требуют специальных псевдопотенциалов, не совместимых с другими программами. Программа позволяет проводить расчеты в рамках методов HF и DFT с несколькими функционалами, включая оптимизацию геометрии, расчеты зонной структуры, статических и динамических (частотно-зависимых) диэлектрических и оптических свойств, расчеты фононного спектра. Сайт программы: <http://www.crystal.unito.it>.

Контрольные вопросы

1. Требуется определить наиболее выгодную конформацию молекулы $C_6H_5C(O)OCH_2CH_3$ в вакууме. Какие программы вы порекомендуете?
2. Необходимо рассчитать энергию образования кислородной вакансии в кристалле ZrO_2 . Какие программы лучше использовать?

3. Необходимо оценить энергию адсорбции молекул глиоксаля (НОС-СОН) на ледяной микрочастице в атмосфере. Микрочастица моделируется кластером из 1000 молекул воды со структурой кристалла льда. Какое программное обеспечение может подойти для этой цели?

Глава 2

Способы описания структуры химической системы

Молекулярное (атомистическое) моделирование базируется на описании свойств химической системы в зависимости от координат составляющих ее частиц. Таким образом, описание системы начинается с задания координат ядер образующих ее атомов, групп атомов или молекул. В зависимости от задачи различают два основных типа систем:

- системы, не обладающие периодичностью в пространстве: атом, молекула, комплекс¹ (кластер) двух или нескольких атомов или молекул;
- системы, которые могут быть описаны набором периодически повторяющихся образов в пространстве: кристалл, аморфное тело, раствор, жидкость, газ, поверхность между двумя фазами.

Для описания атомно-молекулярных систем существует несколько способов задания координат, каждый из которых имеет свои преимущества и недостатки: (1) декартовы координаты; (2) внутренние координаты; (3) координаты Z -матрицы; (4) дробные координаты.

2.1. Декартовы координаты

Декартовы координаты – это пространственные координаты в выбранной пользователем декартовой системе координат, дополненные информацией о типе атома (типе химического элемента). Таким образом, в молекулярном моделировании декартовы координаты обычно записывается строкой из четверки чисел или символов « $N X Y Z$ », где N – химический символ атома или его атомный номер, а X , Y , Z – собственно декартовы координаты атома. Разделителями между числами или символами могут быть пробелы, запятые, или другие знаки в зависимости от используемой программы. Такой способ записи будем называть *$NXYZ$ -координатами* (или *$NXYZ$ -форматом* декартовых координат).

Часто в зависимости от используемой программы или метода расчета, $NXYZ$ -координаты записываются с некоторыми дополнительными данными, которые не имеют отношения к пространственному положению атома. Ими могут быть имя атома, выбранное для данной системы, его заряд и т.д. В этом случае запись для каждого атома записывается в виде строк « $N D X Y Z$ » или « $D N X Y Z$ », где D – дополнительная характеристика атома. Такие форматы будем называть $NDXYZ$ - или $DNXYZ$ -форматами декартовых координат.

В большинстве современных программ молекулярного моделирования порядок следования атомов при описании декартовых координат не имеет значения. Однако он может стать важным, если на определенные атомы имеются ссылки при дальнейшем описании молекулы. Поэтому часто при задании типа

¹ Комплексом называют молекулу, образованную из двух или более фрагментов (молекул или атомов), связанных невалентными межмолекулярными силами или водородными связями.

атома N символ атома комбинируют с порядковым номером элемента: H1, O5, C34, Ti27 и т.д.

Единицы измерения координат X, Y, Z могут быть различны в зависимости от выбранной программы, но обычно это ангстремы (\AA), атомные единицы длины (боры) (a.u., Bohr), нанометры (нм). Соотношения между этими единицами приведены в Приложении.

Примеры

1. Координаты молекулы воды в $NXYZ$ -формате, где N – символ химического элемента и порядковый номер атома в молекуле:

H1	0.000000	0.756686	-0.475344
O2	0.000000	0.000000	0.118836
H3	0.000000	-0.756686	-0.475344

2. Координаты молекулы воды в $DNXYZ$ -формате, где D – обозначение атома в данной задаче, а N – атомный номер элемента:

Oху	8.0	0.000000	0.000000	0.118836
Hyd	1.0	0.000000	0.756686	-0.475344
Hyd	1.0	0.000000	-0.756686	-0.475344

3. Координаты молекулы воды в $NDXYZ$ -формате, где N – символ химического элемента, а D – заряд атома в молекуле:

O	-2.50	0.000000	0.000000	0.118836
H	1.25	0.000000	0.756686	-0.475344
H	1.25	0.000000	-0.756686	-0.475344

2.2. Внутренние координаты

Внутренними координатами называют любой способ описания молекулярной систем, кроме декартовой системы, если он в любой момент эволюции системы позволяет взаимно однозначно выразить декартовы координаты всех атомов. Наиболее часто в качестве внутренних координат выбирают межатомные расстояния r , углы между тройками атомов α (обычно называемые валентными углами, хотя эти углы могут быть заданы и между атомами, не связанными химической связью). Кроме того, это могут быть *диэдральные (торсионные)* углы θ между двумя плоскостями ABC и BCD, образованными четверками атомов A, B, C, D, или неправильные торсионные углы, которые часто используются в молекулярной динамике.

Выражения внутренних координат через декартовы координаты $\mathbf{r}_s = (x_s, y_s, z_s)$, $s = i, j, k, l$ атомов i, j, k, l , даются формулами:

$$r = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

$$\cos \alpha = \frac{(\mathbf{r}_i - \mathbf{r}_j) \cdot (\mathbf{r}_k - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j| |\mathbf{r}_k - \mathbf{r}_j|}$$

$$\cos \theta = \frac{\mathbf{n}_{ijk} \cdot \mathbf{n}_{jkl}}{|\mathbf{n}_{ijk}| |\mathbf{n}_{jkl}|}. \quad (7)$$

В последней формуле \mathbf{n}_{ijk} , \mathbf{n}_{jkl} – нормали к плоскостям, образованным атомами i, j, k и j, k, l . Эти нормали могут быть выражены через векторные произведения соответствующих разностей радиус-векторов атомов:

$$\mathbf{n}_{ijk} = [\mathbf{r}_i - \mathbf{r}_j, \mathbf{r}_k - \mathbf{r}_j], \quad \mathbf{n}_{jkl} = [\mathbf{r}_j - \mathbf{r}_k, \mathbf{r}_l - \mathbf{r}_k].$$

Интервалы изменения межатомных расстояний и валентных углов: $0 \leq r < \infty$, $0 \leq \alpha \leq \pi$. В случае торсионных углов, вычисляемых по формуле (7), однако, интервал значений $0 \leq \theta \leq \pi$. Это означает, что вычисляется минимальный угол между плоскостями ijk и jkl , что создает неоднозначность при выборе координат. Поэтому в большинстве программ используется другой метод определения двугранных углов, при котором двугранный угол изменяется в пределах $0 \leq \theta \leq 2\pi$:

$$\mathbf{a} = \mathbf{r}_i - \mathbf{r}_j, \quad \mathbf{b} = \mathbf{r}_j - \mathbf{r}_k, \quad \mathbf{c} = \mathbf{r}_l - \mathbf{r}_k$$

$$\mathbf{x} = [\mathbf{a}, \mathbf{b}], \quad \mathbf{y} = [\mathbf{c}, \mathbf{b}], \quad \mathbf{z} = [\mathbf{b}, \mathbf{x}] \quad (8)$$

$$\cos \theta = \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| |\mathbf{y}|}, \quad \sin \theta = \frac{(\mathbf{z}, \mathbf{y})}{|\mathbf{z}| |\mathbf{y}|}$$

Сопоставление знаков $\cos \theta$ и $\sin \theta$ в формуле (8) позволяет однозначно определить угол θ в пределах $0 \leq \theta \leq 2\pi$ или $-\pi \leq \theta \leq \pi$. Для целей определения молекулярных координат следует использовать именно этот метод. Отметим, что в разных программах часто по-разному определяется направление изменения θ : в некоторых программах положительное направление соответствует вращению атома i вокруг вектора $j \rightarrow k$ до совпадения с плоскостью jkl по левому винту (математическое определение угла, соответствует формулам (8)), в других, в том числе в известной программе Gaussian – по правому винту (в этом случае в формулах (8) вектор \mathbf{b} определяется с обратным знаком).

Обратное преобразование из внутренних к декартовым координатам представляет собой более сложную процедуру и обычно проводится последовательно, от первого к последующим атомам. При этом первые три атома должны быть определенным образом привязаны к выбранной декартовой системе координат. Наиболее часто для такой привязки используются следующие дополнительные предположения, которые в дальнейшем мы будем называть стандартной привязкой:

– первый по порядку атом системы располагается в центре декартовой системы координат (в точке O): $\mathbf{r}_1 = (0, 0, 0)$;

- второй атом системы располагается на оси OX: $\mathbf{r}_2 = (r_{21}, 0, 0)$;
- третий атом располагается в плоскости OXY $\mathbf{r}_3 = (x_3, y_3, 0)$.

Конкретные значения координат x_3, y_3 зависят от атомов, с которыми атом 3 образует связь и валентный угол.

При описании структуры на основе декартовых координат используется $3N$ координат (N – число атомов). При использовании внутренних координат и способа привязки, описанного выше, для описания структуры нелинейной молекулы требуется $3N-6$ координат, так как координаты $x_1, y_1, z_1, y_2, z_2, z_3$ заданы равными нулю.

2.3. Координаты Z-матрицы

Особенно часто используется выбор внутренних координат, записываемый в виде так называемой *Z-матрицы*. Z-матрица – это описание атомных координат в виде n записей (для каждого атома i в данной n -атомной системе):

$$N_i \ r_{ia} \ \alpha_{iab} \ \theta_{iabc} \ N_a \ N_b \ N_c,$$

где N_i – тип атома i (в виде атомного номера, имени химического элемента или в виде какой-либо комбинации, аналогично *NXYZ*-координатам, описанным выше); $r_{ia}, \alpha_{iab}, \theta_{iabc}$ – межатомное расстояние, валентный и торсионный углы, образуемые данным атомом с атомами a, b, c называемыми опорными атомами. Опорные атомы указываются их именами или порядковыми номерами в данной молекулярной системе N_a, N_b, N_c , причем в момент описания атома i , координаты этих атомов должны быть уже описаны, то есть номер атома i в системе должен быть больше, чем номера этих атомов: $i > N_a, N_b, N_c$.

При использовании Z-матрицы обычно используется стандартная привязка к декартовым координатам. Таким образом, для первых трех атомов достаточно описать только часть координат:

- для первого атома: N_1
- для второго атома: $N_2 \ r_{21}$
- для третьего атома: $N_3 \ r_{3a} \ \alpha_{3ab} \ N_a \ N_b$

Порядок, в котором указываются координаты $r_{ia}, \alpha_{iab}, \theta_{iabc}$ и опорные атомы N_a, N_b, N_c может быть различен в различных программах. Наиболее часто используются два формата: Z-матрица в формате программы Gaussian (*GZ*-матрица) и Z-матрица в формате программы MORAC:

В первом случае (*GZ*-формат) N_a, N_b, N_c указываются в виде *имен атомов* перед соответствующим значением координаты:

$$N_i \ N_a \ r_{ia} \ N_b \ \alpha_{iab} \ N_c \ \theta_{iabc}$$

Например, для описания молекулы H₂O₂ (группа симметрии C₂), в которой длины связи ОН равны 0.95Å, длина связи О–О 1.21 Å, валентные углы НОО = 105°, а диэдральный угол НООН = 120°, Z-матрица может иметь вид:

```

H1
O2  H1  0.95
O3  O2  1.21  H1  105.
H4  O3  0.95  O2  105.  H1  120.

```

В случае формата МОРАС, эта же Z-матрица записывается в виде:

$$N_i \ r_{ia} \ I_1 \ \alpha_{iab} \ I_2 \ \theta_{iabc} \ I_3 \ N_a \ N_b \ N_c,$$

где дополнительные параметры I_1 , I_2 , I_3 представляют собой числа 1 и 0, указывающие на необходимость вариации данного геометрического параметра ($I=1$) или фиксации его значения ($I=0$) при расчетах данной молекулы.

Например, для описания аналогичной молекулы H₂O, что и в предыдущем примере, при оптимизации всех структурных параметров Z-матрица в формате МОРАС имеет вид:

```

1
8  0.95  1
8  1.21  1  105.  1
1  0.95  1  105.  1  3  2  1

```

В настоящее время GZ-формат является наиболее часто применяемым и позволяет очень гибко задавать структуру молекул, используя, в том числе, информацию о симметрии. Поэтому в последующем изложении будет рассматриваться именно этот формат. Следует отметить, что в различных программах часто используются варианты этих форматов, которые зависят от конкретной программной реализации. Тем не менее, общая идея, объединяющая все варианты использования Z-матрицы состоит в использовании внутренних координат с указанием соответствующих опорных атомов.

При использовании формата Z-матрицы следует помнить о важном правиле: валентный угол не должен принимать значения 0 и 180° (несмотря на то, что с точки зрения внутренних координат эти значения допустимы). Это обусловлено тем, что задание торсионных параметров последующих атомов зависит от нормалей к плоскостям опорных атомов. В случае, когда валентный угол между атомами ijk принимает значения 0 или 180°, использование плоскости ijk для определения торсионного угла становится невозможным. Возникает вопрос, как описывать линейные молекулы, где валентные углы равны 180°. Для этой цели, в большинстве программ, использующих формат Z-матрицы, предусмотрено использование псевдоатомов (*dummy atoms*). *Псевдоатомы* – это точки, которые не участвуют в квантовохимическом или другом расчете физических величин, но позволяют описать вспомогательные положения в пространстве, чтобы избежать запрещенных значений параметров, или делают описание геометрии более удобным. В формате GZ-матрицы,

псевдоатомы обозначаются символами X или XX. Например, чтобы описать линейную молекулу CO_2 с длинами связи C–O 1.2Å, можно использовать Z-матрицу:

```

O1
C2  O1  1.20
XX3 C2  1.0   O1  90.
O4  C2  1.20  XX3 90.   O1  180.

```

В этом примере вместо описания второго атома O через валентный угол $\text{OCO} = 180^\circ$, используется псевдоатом XX3, образующий угол 90° с атомами O1 и O4. Таким образом избегается запрещенное значение валентного угла. Длина связи между XX3 и C2 не важна и может быть принята за 1 Å.

Еще одной особенностью GZ-формата является возможность использования переменных и констант, описывающих значения геометрических параметров при оптимизации. Например, описание молекулы CO_2 и H_2O_2 из предыдущих примеров может быть выполнено следующим образом:

Молекула CO_2

```

O1
C2  O1  rCO
XX3 C2  1.0   O1  90.
O4  C2  rCO  XX3 90.   O1  180.
Variables:
rCO 1.20

```

Молекула H_2O_2

```

H1
O2  H1  rOH1
O3  O2  rOO   H1  aНОН
H4  O3  rOH4   O2  aНОН   H1  120.
Variables:
rOH1 0.95
rOH4 0.95
rOO 1.20
Constants:
aНОН 105.

```

В случае CO_2 задана одна переменная rCO, которая будет варьироваться в ходе расчета – длина связи C–O, начальное значение которой 1.20 Å. Значения остальных геометрических параметров в этом случае (длины связи C2-XX3, валентных и двугранных углов считаются фиксированными (*константами*) и в ходе расчета не варьируются. Обратите внимание, что при вариации длин связей, они будут всегда принимать одинаковые значения, так как обе связи описаны с помощью одной переменной. Таким образом, использование переменных позволяет описать симметрию молекулы.

В случае молекулы H_2O_2 заданы две переменные rOH1 и rOH4, что означает, что изменение длин этих связей будет происходить независимо друг от друга. Углы НОН описаны как константы aНОН = 105 и не варьируются в ходе

расчета. Кроме того, все величины, заданные числовыми параметрами также считаются фиксированными параметрами.

Особенностью программы Gaussian, которая изначально использует такой тип Z-матрицы, является требование, чтобы все действительные значения параметров (длины связи и углы) были записаны с десятичной точкой: например, угол не 120, а 120. (с точкой). Кроме того, в этой программе признаком окончания Z-матрицы является пустая строка, которая обязательно должна присутствовать после описания молекулы. При нарушении этих требований Gaussian выдает сообщение об ошибке. Эти требования не обязательно распространяются на другие программы.

Еще несколько примеров задания Z-матриц:

1. Z-матрица молекулы CH₄ учитывающая симметрию T_d:

```

H1
C2  H1  rCH
H3  C2  rCH  H1  109.47
H4  C2  rCH  H1  109.47  H3  120.
H5  C2  rCH  H1  109.47  H3  -120.
Variables:
rCH 1.09

```

2. Z-матрица молекулы NH₃, учитывающая группу симметрии молекулы C_{3v}:

```

N1
XX2  N1  1.0
H3  N1  rNH  XX2  85.
H4  N1  rNH  XX2  85.  H3  120.
H5  N1  rNH  XX2  85.  H3  240.
Variables:
rNH 1.01

```

3. Z-матрица молекулы циклобутана C₄H₈, учитывающая группу симметрии молекулы D_{2d}:

```

XX1
XX2  XX1  rxx
C3  XX1  r1  XX2  90.
C4  XX1  r1  XX2  90.  C3  180.
C5  XX2  r1  XX1  90.  C3  90.
C6  XX2  r1  XX1  90.  C3  270.
H7  C3  rCH  xx2  aHCC  c5  90.
H8  C3  rCH  xx2  aHCC1  c5  -90.
H9  C4  rCH  xx2  aHCC2  c5  90.
H10 C4  rCH  xx2  aHCC1  c5  -90.
H11 C5  rCH  xx1  aHCC  c3  90.
H12 C5  rCH  xx1  aHCC1  c3  -90.
H13 C6  rCH  xx1  aHCC2  c4  90.
H14 C6  rCH  xx1  aHCC1  c4  -90.
Variables:
rxx 0.12
r1 1.11
rCH 1.1
aHCC 128.

```

Обратите внимание, что двугранные углы могут задаваться в виде положительных или отрицательных значений: двугранные углы $\theta = -120.$ и $\theta = 240.$ эквивалентны. Величина 109.47 или 109.4712 – это тетраэдрический угол в градусах, который часто записывается как 109°28', однако при задании молекулярных координат требуется использование градусов и долей градуса.

Контрольные задания

1. Запишите Z-матрицы для молекул: CO, H₂O, C₂H₄, CH₃OH.
2. Используя псевдоатомы, запишите Z-матрицы для молекул C₂H₂, H₂C=C=C.
3. Запишите Z-матрицы, учитывающие симметрию молекул SiF₄ (*T_d*), SF₆ (*O_h*), C₆H₆ (*D_{6h}*), Fe(C₅H₅)₂ (*D_{5d}*).

2.4. Дробные координаты

Дробные координаты используются для описания координат системы с трансляционной симметрией в одном или нескольких направлениях, которая характеризуется элементарной ячейкой размерами *a*, *b*, *c*. В этом случае дробные координаты каждого атома *i* представляются в виде отношения декартовых координат к размерам ячейки (*x_i/a*, *y_i/b*, *z_i/c*). Таким образом, кроме задания собственно положений атома, должны быть заданы характеристики ячейки. Эта информация может быть в виде группы симметрии системы (например, пространственной группы кристалла) или в виде координат направляющих векторов **a**, **b**, **c**, образующих ячейку. В самом простом случае тетрагональной симметрии, когда вектора **a**, **b**, **c** взаимно ортогональны, достаточно задания параметров ячейки *a*, *b*, *c*, то есть длин векторов **a**, **b**, **c**.

Часто при описании кристаллической системы используется дополнительное понятие *суперячейки*, которая представляет собой вырезку из кристалла размером (*n_a*, *n_b*, *n_c*) элементарных ячеек. В этом случае величины (*n_a*, *n_b*, *n_c*) также должны быть сообщены программе. Такой способ значительно упрощает расчет больших кристаллических или аморфных систем.

При использовании дробных координат система не обязательно должна соответствовать определенной идеальной симметрии кристалла. Часто на систему накладываются периодические условия, позволяющие учесть влияние окружения, однако система внутри элементарной ячейки остается неупорядоченной. Примером такого подхода является задание начальных структур разупорядоченных систем при молекулярно-динамических расчетах в так называемом приближении минимального отображения.

2.5. Форматы файлов

Описанные выше форматы записи координат используются в различных квантовохимических программах для описания геометрии в файлах входных данных или файлах результатов расчета. Одним из наиболее простых форматов, который понимается большим количеством программ, является файл типа *.xyz. В этом файле первые две строки (иногда от 1 до 3) отводятся для служебной информации о молекуле. В некоторых программах требуется, чтобы на первой строке указывалось число атомов, вторая строка отводится под произвольный комментарий. После этого следует *n* строк, описывающих декартовы координаты *n* атомов молекулы. Однако в большинстве современных программ не требуется указывать число атомов, программа определяет это число автоматически. Важно только, чтобы по окончании описания геометрии в файле присутствовала пустая строка, что является признаком завершения описания структуры. В этом случае первые две строки в файле могут быть заняты произвольной информацией. Некоторым программам, например, программе Moltran можно сообщить, сколько строк в начале файла пропустить перед началом описания геометрии, для этого используется ключ в командной строке /x<число пропускаемых строк>, например:

```
C:>moltran file.xyz /x5
```

Пример файла *.xyz показан на рис. 3.

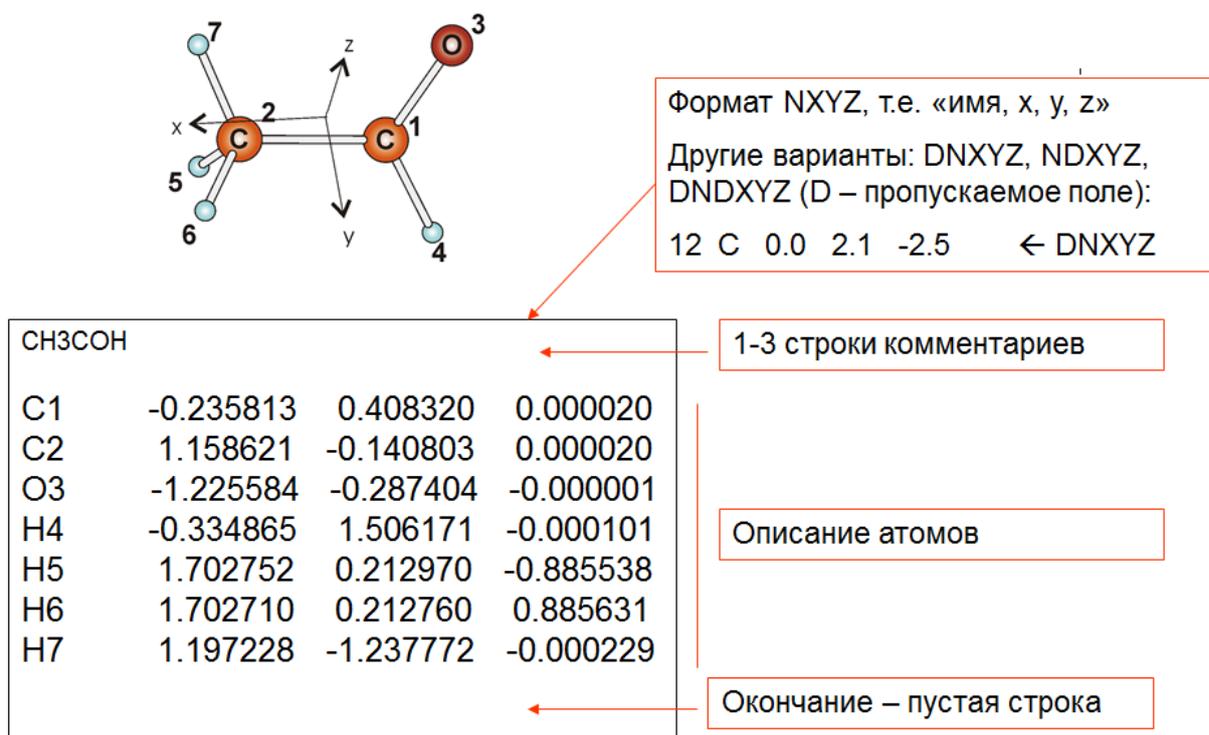


Рис. 3. Пример файла *.xyz, в котором описывается молекула CH₃COH с в формате NXYZ

В отличие от файлов типа *.xyz входные файлы квантовохимических программ обычно предусматривают специальное расположение в файле

геометрии и ключевых слов для описания типа и параметров квантовохимического расчета.

Входные файлы программы Gaussian обычно имеют расширение *.gjf, их структура показана на рис. 4. В начале файла располагаются команды, начинающиеся с символа %, информирующие программу о способе расчета, например:

%Nproc=2	Использовать для расчета не более 2 процессоров
%Mem=6000MB	Использовать для расчета не более 6000 МБ оперативной памяти
%rwf=file.rwf	использовать файл file.rwf для записи информации, необходимой для рестарта
%NoSave	После успешного завершения расчета удалить все временные файлы, описанные выше этой команды
%chk=file.chk	Использовать для расчета файл со служебными данными под названием file.chk. Он содержит информацию, которая не выведена в файл и может использоваться для передачи этой информации программе при последующих запусках. Кроме того, он в ряде случаев позволяет провести рестарт, то есть продолжить оборвавшийся расчет.

Команды, начинающиеся с %, необязательно указывать для расчета, однако рекомендуется использовать команду %chk, поскольку chk-файл, имея относительно небольшой объем, часто позволяет выполнить рестарт и содержит ценную дополнительную информацию после успешного расчета.

После команд % следует основная командная строка (или несколько строк), начинающаяся с символа #. На этой строке или строках указывается метод расчета и тип компьютерного эксперимента. Конкретные ключевые слова для описания этих данных будут рассмотрены в следующих разделах.

Командные строки, начатые символом #, должны обязательно заканчиваться пустой строкой. За ней следуют одна или несколько строк комментариев, признаком окончания которых также является пустая строка.

После этого в файле *.gjf располагается описание молекулярной системы. Оно начинается со строки, содержащей заряд молекулы и спиновую мультиплетность системы, после которой следует описание молекулярной геометрии в формате GZ-матрицы или NXYZ. Завершение описания геометрии – пустая строка.

На рис. 4 и 5 приведены примеры файлов *.gjf, подготовленных для расчета радикала $\text{CH}_3\text{-C=O}$ (заряд 0, мультиплетность 2, используется Z-матрица) и молекулы цис-глиоксаля $(\text{COH})_2$ (заряд 0, мультиплетность 1, используются декартовых координат).

```

%chk=MeCO.chk
# B3LYP/6-311++G(2d,2p) FOpt Freq

MeCO

O 2
O1
C2 O1 1.20
C3 C2 1.54 O1 120.
H4 C3 1.09 C2 109.5 O1 0.
H5 C3 1.09 C2 109.5 O1 120.
H6 C3 1.09 C2 109.5 O1 -120.

```

Тип и параметры расчета
 Комментарии
 Заряд и мультиплетность молекулы
 Z-матрица в формате Gaussian
 Признак окончания – пустая строка

Рис. 4. Файл *.gjf для для оптимизации геометрии и расчета частот радикала $\text{CH}_3\text{-C=O}$ на уровне B3LYP/6-311++G(2d,2p) в программе Gaussian

```

%chk=(COH)2.chk
# b3lyp/6-311++G(2d,2p) FOpt Freq

cis-OHC-CHO -- H2O

O 1
C1 0.220108 0.078700 0.006679
C2 1.737374 0.217855 -0.005520
O3 2.266991 1.296561 -0.023236
H4 2.294585 -0.735741 0.002464
O5 -0.490083 1.054375 0.002581
H6 -0.172834 -0.950511 0.019692
O7 -2.898248 -0.562652 -0.013651
H8 -3.820275 -0.308372 0.067079
H9 -2.393365 0.260892 -0.001538

```

Тип и параметры расчета
 Комментарии
 Заряд и мультиплетность молекулы
 NXYZ-матрица
 Признак окончания – пустая строка

Рис. 5. Файл *.gjf для оптимизации геометрии и расчета частот молекулы $(\text{COH})_2$ на уровне B3LYP/6-311++G(2d,2p) программе Gaussian

Программа GAMESS требует для своей работы файлы, которые обычно имеют расширение *.inp. Структура входного файла, подготовленного для расчета молекулы воды, приведена на рис. 6. Файл состоит из нескольких секций, каждая из которых начинается с выражения \$<имя секции> и заканчивается выражением \$END. Каждая секция включает одну или несколько команд, описывающих тип волновой функции и уровень теории MP_n (SCFTYP и MPLEVL), способ оптимизации геометрии (RUNTYP), лимиты на время расчета, количество используемой памяти и шагов оптимизации (TIMLIM, MEMORY, NSTEP), а также используемый базис (секция \$BASIS). Отдельная секция

\$DATA описывает симметрию и геометрию молекул. Если группа симметрии выше C_1 , после нее требуется дополнительная строка, которая может быть оставлена пустой. Если группа симметрии выше C_1 , указываются координаты только уникальных атомов, то есть тех, которые при преобразованиях симметрии не переходят друг в друга.

Программа NWchem использует входной файл *.nw, формат которого показан на рис. 7. Геометрия и различные типы команд указываются в секциях, начинающихся с названия секции и заканчивающегося словом end (требуется только если секция занимает несколько строк). Геометрия может задаваться в формате NXYZ или в формате Z-матрицы.

Программа HureChem использует экранный интерфейс для задания координат молекулы и описания типа расчета. Поэтому ее файл входных данных *.hin используется только для хранения служебной информации и выглядит значительно менее информативно, см. рис.8.

```

$CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE MPLEVEL=2
! EXETYP=CHECK
$END
$$SYSTEM TIMLIM=3000 MEMORY=15000000 $END
$$SCF DIRSCF=.TRUE. $END
$$STATPT NSTEP=200 HSEND=.TRUE. $END
$BASIS GBASIS=N311 NGAUSS=6 NDFUNC=2 NPFUNC=2
  DIFFSP=.TRUE. DIFFS=.TRUE. $END
$DATA
H2O
CNV 2

O      8.0  0.0000000000  0.0000000000  0.0361667310
H      1.0 -0.7583272833  0.0000000000 -0.5498469164
$END

```

Тип и параметры расчета, включая заряд и мультиплетность

Комментарии

Группа симметрии

DNXYZ-матрица с учетом симметрии

Признак окончания

Рис. 6. Формат входного файла программы GAMESS, подготовленный для оптимизации геометрии молекулы воды с симметрией C_{2v} на уровне MP2/6-311++G(2d,2p)

```

start h2o
memory 16000MB

title "H2O pbe0/cc-pVTZ opt+freqs"
print low

geometry
H 0.000000 0.756686 -0.475344
H 0.000000 0.000000 0.118836
O 0.000000 -0.756686 -0.475344
end

basis
H library cc-pVTZ
O library cc-pVTZ
end

dft
xc pbe0
end

task dft optimize
task dft freq

```

Задание имени временного файла и выделение памяти

Комментарий и объем печати

Геометрия в формате NXYZ или Z-матрицы

Выбор базиса для всех атомов

Выбор функционала DFT

Тип расчета: оптимизация геометрии + расчет частот

Рис. 7. Формат входного файла программы NWchem, подготовленный для оптимизации геометрии и расчета частот молекулы воды на уровне PBE0/cc-pVTZ

```

;Ethanol, or ethyl alcohol
forcefield mm+
sys 0
view 40 0.381 55 15 0.09885125 -0.8386148 -0.5356806 0.9827889 -0.002147554 0.1847199
mol 1
atom 1 - C C4 - 0 0.6910678 -0.8604058 -4.251571e-007 4 2 s 4 s 5 s 6 s
atom 2 - C C4 - 0 0.6910678 0.6795942 -4.251571e-007 4 1 s 3 s 7 s 8 s
atom 3 - O O2 - 0 -0.6571509 1.156256 -4.251571e-007 2 2 s 9 s
atom 4 - H H - 0 1.718731 -1.223735 -4.251571e-007 1 1 s
atom 5 - H H - 0 0.1772445 -1.223735 0.8899869 1 1 s
atom 6 - H H - 0 0.1772333 -1.223743 -0.8899781 1 1 s
atom 7 - H H - 0 1.204891 1.042923 0.889987 1 2 s
atom 8 - H H - 0 1.204902 1.042932 -0.8899781 1 2 s
atom 9 - H HO - 0 -0.6571509 2.116256 -4.251571e-007 1 3 s
endmol 1

```

Заголовок: название, комментарий, метод расчета, заряд

Параметры изображения на экране

Геометрия + атомные заряды + параметры расчета

Рис. 8. Файл *.hin программы NureChem, содержащий служебную информацию о структуре рассчитываемой молекулы и способе расчета

2.6. Молекулярные визуализаторы и молекулярные конструкторы

Использование Z-матрицы для задания молекулярных координат делает возможным описание весьма сложных структур без каких-либо дополнительных расчетов. Обычно достаточно лишь нарисовать набросок структуры молекулы, пронумеровать на чертеже атомы и определить необходимые внутренние координаты. Однако существуют программы, которые отображают структуру молекулы на экране и позволяют построение или редактирование такой структуры в визуально-диалоговом режиме. Программы, которые отображают на экране структуру, описанную в какой-либо системе координат, называются визуализаторами молекулярной структуры (*molecular viewers*). Программы, которые позволяют создать или изменить структуру в диалоговом режиме называются молекулярными конструкторами (*molecular builders*). Большинство программ объединяют в себе средства визуализации и редактирования структуры, но некоторые имеют большую ориентацию на визуализацию, чем на редактирование. Несколько известных визуализаторов и конструкторов:

VMD – профессиональный свободно распространяемый визуализатор, ориентированный на очень большие структуры, в т.ч. белки и полимеры, результаты профессионального молекулярно-динамического моделирования с сотнями тысяч атомов и тысячами наборов координат. Прежде всего ориентирован на ОС LINUX, но есть версия для Windows. В первую очередь это визуализатор, хотя есть некоторые возможности редактирования структуры. Сайт программы: <https://www.ks.uiuc.edu/Research/vmd/>

Avogadro – бесплатный конструктор-визуализатор. Позволяет создание и редактирование на экране небольших и средних молекул, в том числе с генерацией Z-матриц и входных файлов для программ Gaussian, Gamess, NWchem и др. Позволяет работать с дробными координатами. Имеет довольно большую библиотеку структур различных химических соединений и фрагментов. Есть версии для Windows, Linux, MacOS. Сайт: <http://avogadro.cc/>

ChemCraft – коммерческий визуализатор, позволяющий редактировать молекулярные структуры и генерировать XYZ-координаты и Z-матрицы, а также генерировать входные файлы программ GAMESS и MOLPRO. Имеет большое количество полезных средств визуализации не только структур, но и молекулярных поверхностей, орбиталей, уровней энергии, анимации молекулярных колебаний, средства создания высококачественных иллюстраций в формате JPEG. Имеется версия для Windows и Linux, а также пробная бесплатная версия и бесплатная версия Chemcraft-Light. Платная версия значительно дешевле аналогов. Сайт <http://www.chemcraftprog.com/>

GaussView – коммерческий конструктор и визуализатор, ориентированный на работу с программой Gaussian. Позволяет редактировать структуры в экранном режиме, генерировать координаты и входные файлы Gaussian в различных форматах, а также сразу же запускать квантовохимические расчеты для построенных структур, вызывая программу Gaussian, и анализировать результаты расчета, читая выходные файлы. Имеет большое число средств

анализа квантовохимических результатов, включая анимацию, изображение орбиталей и других молекулярных поверхностей. Сайт <https://gaussian.com/>

Moltran – бесплатный визуализатор результатов квантовохимических расчетов для Windows с возможностью создания и редактирования структур, разработанный на химическом факультете ННГУ. Позволяет читать и визуализировать выходные файлы различных квантовохимических программ, в т.ч. Gaussian, GAMESS, NWChem, Priroda, CFOUR, DFTB, MOPAC. Кроме визуализации и редактирования позволяет проводить термодинамические расчеты при различных условиях на основе данных квантовохимического расчета с учетом вкладов внутренних вращений, визуализировать молекулярные колебания, генерировать Z-матрицы и координаты, скорректированные на величины колебательных векторов. Сайт: <http://www.qchem.unn.ru/moltran/>

Глава 3

Способы квантовохимического моделирования для изучения различных физико-химических свойств

3.1. Типы компьютерного эксперимента в квантовой химии

Под типом компьютерного эксперимента понимается режим работы программы квантовохимического расчета для получения той или иной информации в зависимости от цели квантовохимического исследования. Различают несколько основных типов компьютерного эксперимента:

- оптимизация молекулярной геометрии;
- расчет молекулярных свойств при фиксированной молекулярной геометрии;
- расчет колебательных частот и термодинамических параметров;
- поиск переходных состояний реакций и констант скорости;
- молекулярно-динамическая эволюция системы.

Как видно из приведенного списка, эти типы компьютерного эксперимента различаются способами контролируемого изменения молекулярной структуры в ходе расчета. Ниже рассмотрены особенности этих способов компьютерного моделирования.

3.2. Оптимизация молекулярной геометрии

Оптимизация молекулярной геометрии является основным и наиболее часто используемым типом компьютерного эксперимента в квантовой химии. Она заключается в минимизации полной энергии молекулы E_{tot} при вариации координат атомов. Поскольку зависимость E_{tot} от координат ядер является поверхностью потенциальной энергии (ППЭ), оптимизация геометрии является поиском точек локальных минимумов на ППЭ. Вариация координат может проводиться в разных системах координат (декартовых, координат Z-матрицы или различных типов внутренних координат). Большинство программ позволяют выбрать тип координат для оптимизации, и от этого выбора сильно зависит скорость сходимости оптимизации геометрии. Во многих профессиональных квантовохимических программах оптимизация по умолчанию проводится в *избыточных внутренних координатах (ИВК, redundant internal coordinates)*. Это внутренние координаты молекулы (межатомные расстояния, валентные и диэдральные углы), среди которых по специальному алгоритму отобраны те, вариация которых приводит к наибольшему изменению энергии. Вариация этих координат дает наиболее быстрое приближение к минимуму энергии. Хотя выбор такой системы различен для разных молекул и определение ИВК является сложным алгоритмом, временные затраты на выбор такой системы при

квантовохимическом расчете многократно окупаются за счет сокращения числа итераций.

Минимизация энергии производится по специальным численным алгоритмам безусловной оптимизации функций. Большинство программ имеют несколько вариантов выбора такого алгоритма. В настоящее время считается, что одним из наиболее быстрых математических алгоритмов безусловной оптимизации является метод BFGS (*алгоритм Бroyдена-Флетчера-Голдфарба-Шанно*) или его разновидности. Другой часто используемый алгоритм это более простой метод сопряженных градиентов. Оба алгоритма являются методами градиентной оптимизации, то есть требуют расчета в каждой точке ППЭ расчета энергии и ее градиента, который выполняется квантовохимическим методом. После того как расчет произведен, алгоритм оптимизации вычисляет направление движения по ППЭ (вектор изменения координат) и делает шаг или несколько шагов в этом направлении, контролируя понижение энергии. В новой точке процесс повторяется. Оптимизация останавливается при одновременном выполнении всех или нескольких *критериев останова*. Этими критериями обычно являются следующие условия: (1) величина максимальной компоненты вектора градиента энергии (то есть вектора сил на атомах) меньше пороговой величины g_{max} ; (2) среднеквадратичная величина компонент градиента энергии (сил) меньше пороговой g_{rms} ; (3) максимальная компонента вектора изменения координат относительно предыдущей точки ППЭ (то есть вектора шага Δx) меньше порогового значения x_{max} ; (4) среднеквадратичная величина изменения координат относительно предыдущей точки ППЭ меньше порога x_{rms} . Кроме того, оптимизация останавливается при превышении заданного максимального числа шагов или превышении допустимого времени расчета. Пороговые значения критериев останова заложены в программе, но могут быть изменены пользователем. Алгоритм оптимизации геометрии показан на рис. 9.

Ход оптимизации, а также причина останова обязательно должна контролироваться пользователем. Если останов произошел не по причине выполнения заданных критериев, а например, по причине превышения заданного числа шагов, оптимизация должна быть продолжена.

Оптимизация геометрии может быть *полной или частичной*. В первом случае варьируются все молекулярные координаты, во втором – некоторые из них считаются фиксированными (*замороженными*). Такая «заморозка» координат может быть выполнена с помощью Z-матрицы путем объявления некоторых элементов матрицы константами, или в случае ИВК – путем специальных команд, объявляющих постоянными некоторые из внутренних координат. При полной оптимизации геометрии N -атомной нелинейной молекулы число оптимизируемых параметров равно $N_{\text{опт}} = 3N - 6$ ($N_{\text{опт}} = 3N - 5$ в случае линейной молекулы), при частичной оптимизации $N_{\text{опт}} < 3N - 6$ ($N_{\text{опт}} < 3N - 5$ в случае линейной молекулы).

Оптимизация молекулярной геометрии



Рис. 9. Блок-схема алгоритма оптимизации молекулярной геометрии

Следует помнить, что задание Z -матрицы позволяет контролировать симметрию молекулы в ходе оптимизации. Если некоторые внутренние координаты молекулы описаны через одну и ту же переменную, они будут оставаться равными в ходе всей оптимизации, даже если это невыгодно с точки зрения энергии. Таким образом, задание симметричной Z -матрицы может сократить число шагов за счет сокращения числа варьируемых параметров в симметричной молекуле, но одновременно может привести к ошибке, если асимметричная структура молекулы энергетически более выгодна.

В программе Gaussian оптимизация геометрии задается ключевым словом FOpt (полная оптимизация) или POpt (частичная оптимизация) в строке команд.

Примеры

1. Полная оптимизация молекулы воды, заданной декартовыми координатами, Оптимизация по умолчанию будет проходить в ИВК, декартовы координаты используются только для задания начальной геометрии.

```
# B3LYP/6-31G(d,p) FOpt

H2O

0 1
H1  0.000000    0.756686   -0.475344
O2  0.000000    0.000000    0.118836
H3  0.000000   -0.756686   -0.475344
```

2. Частичная оптимизация молекулы воды на уровне CCSD(T)/cc-pVTZ. Для уменьшения времени расчета учтена симметрия молекулы - длины связей О-Н заданы эквивалентными. Варьироваться будут координаты Z -матрицы (POpt=(Z-matrix)), угол НОН не будет варьировать, так как он задан в виде константы.

```
# CCSD(T)/cc-pVTZ POpt=(Z-matrix)

H2O

0 1
H1
O2 H1 rOH
H3 O2 rOH H3 105.
Variables:
rOH 0.95
```

3. Частичная оптимизация молекулы воды в избыточных внутренних координатах (ИВК). Угол НОН будет оставаться фиксированным (105°) вследствие команды заморозки ИВК F (freeze): F <atom1> [<atom2>[<atom3>[<atom4> [<значение>]]]]. Если значение не указано, будет использована величина из заданных координат.

```
# MP2/6-31G(d,p) POpt

H2O

0 1
H1 0.000000 0.756686 -0.475344
O2 0.000000 0.000000 0.118836
H3 0.000000 -0.756686 -0.475344

F 1 2 3
```

Для выполнения квантовохимического расчета с этими файлами с помощью программы Gaussian03 достаточно выполнить команду (предполагается, что имя файла с входными данными h2o.gjf): g03 h2o.gjf.

После запуска программа создаст выходной файл (файл с результатами расчета) h2o.out (ОС Windows) или h2o.log (ОС LINUX), который можно просматривать во время расчета (но нельзя редактировать). Просмотр и анализ выходной информации возможен также во время работы программы некоторыми визуализаторами, например, с помощью программы Moltran (ОС Windows).

Оптимизация геометрии включает в себя итерации, или циклы, обозначенные на рис. 9 пунктирной рамкой. Каждый цикл в программе Gaussian начинается таблицей текущих координат и заканчивается сообщением:

Item	Value	Threshold	Converged?
Maximum Force	0.000228	0.000450	YES
RMS Force	0.000167	0.000300	YES
Maximum Displacement	0.017077	0.001800	NO
RMS Displacement	0.007140	0.001200	NO
Predicted change in Energy=-2.231553D-06			
GradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGr			

Эта таблица показывает текущее состояние критериев сходимости. В данном случае видно, что два из них выполнено, так как значения максимального и среднеквадратичного градиента ниже пороговых значений. Однако два других показателя (максимальное и среднеквадратичное смещение атомов на данном шаге оптимизации) превышают пороговые. Поэтому оптимизация будет

продолжена. Оптимизация заканчивается, как правило, когда все четыре критерия получают статус YES.

В случае успешного завершения работы программы в выходном файле должно появиться сообщение об этом, например:

```
Normal termination of Gaussian 03 at Tue Jan 10 19:56:09 2017.
```

Хотя работа программы завершилась успешно, это не означает, что цель расчета достигнута. Признаком успешного завершения оптимизации является сообщение:

```
Optimization completed.  
-- Stationary point found.
```

После этого сообщения программа печатает оптимизированные координаты во внутренней координатной системе (ИВК), в координатной системе, соответствующей координатам во входном файле (Input orientation) а также в координатной системе, в которой программа производила расчет (Standard orientation). Если при оптимизации требовалось проводить оптимизацию в виде Z-матрицы, то будут выданы оптимизированные координаты в этом формате.

Оптимизированную структуру удобнее анализировать, используя программы-визуализаторы. Визуализация в программах ChemCraft и GaussView выполняется в диалоговом режиме путем открытия необходимого выходного файла с помощью меню. Для визуализации оптимизированной структуры и структур всех промежуточных шагов оптимизации из выходных файлов Gaussian, Gamess, Nwchem с помощью программы Moltran в ОС Windows достаточно выполнить команду: C:>moltran h2o.out.

На рис. 10 показаны результаты визуализации молекулы воды и германийорганического соединения в программах ChemCraft и Moltran.

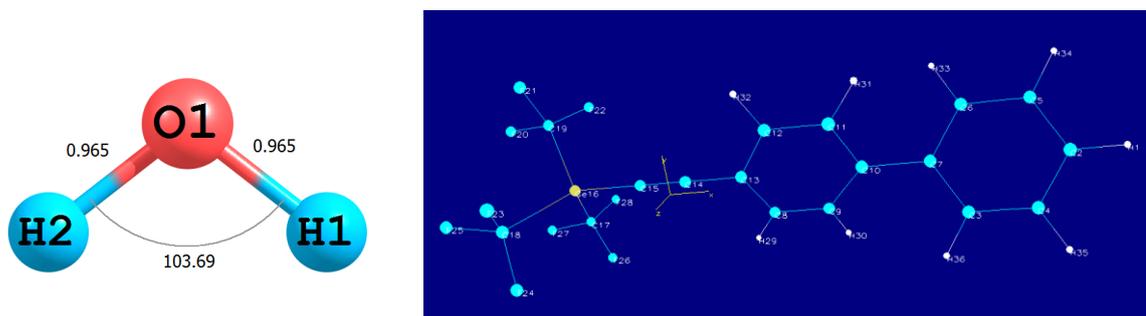


Рис. 10. Визуализация молекулярной структуры после оптимизации геометрии с помощью программы ChemCraft (слева) и программы Moltran (справа)

В конце выходного файла печатается краткая сводка результатов, предназначенная для архивирования:

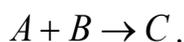
```

1|1|UNPC-Q5|Freq|RB3LYP|6-31G(d,p)|H2O|USER|31-Mar-2019|0||#N Geom=A1
lCheck Guess=TCheck SCRF=Check GenChk RB3LYP/6-31G(d,p) Freq||H2O||0,1
|H,0.0014810062,0.,-0.0079501897|O,-0.0095436124,0.,0.9573230713|H,0.9
256921412,0.,1.1965052112||Version=IA32W-G09RevA.02|State=1-A1|HF=-76.
4197366|RMSD=2.716e-011|RMSF=4.921e-005|ZeroPoint=0.0213648|Thermal=0.
0241998|Dipole=0.637882,0.,-0.4894641|DipoleDeriv=0.2776884,0.,0.06343
57,0.,0.3567232,0.,0.0307497,0.,0.0875804,-0.3544573,0.,-0.0403492,0.,
-0.7134464,0.,-0.0403492,0.,-0.3760803,0.0767688,0.,-0.0230865,0.,0.35
67232,0.,0.0095995,0.,0.2885|Polar=6.283576,0.,2.9936344,0.9402582,0.,
6.7874589|PG=C02V [C2(O1),SGV(H2)]|NImag=0||0.04995153,0.,-0.00008206,
0.01229217,0.,0.53470177,-0.04500822,0.,-0.02371771,0.53614849,0.,0.00
002646,0.,0.,-0.00005291,0.04068384,0.,-0.51862505,0.10257459,0.,0.591
11805,-0.00494330,0.,0.01142554,-0.49114027,0.,-0.14325844,0.49608357,
0.,0.00005561,0.,0.,0.00002646,0.,0.,-0.00008206,-0.05297601,0.,-0.016
07672,-0.07885688,0.,-0.07249300,0.13183289,0.,0.08856972||0.00007652,
0.,0.00002087,-0.00007616,0.,0.00005844,-0.00000036,0.,-0.00007931|||@

```

Среди этой информации выражение HF=-76.4197366 означает полную энергию молекулы, полученную в результате оптимизации (соответствующую оптимизированной геометрии). Это выражение может несколько варьироваться в зависимости от метода расчета. При расчетах на уровнях HF или DFT энергия для систем с замкнутой оболочкой будет выдана как HF=, для систем с незамкнутой оболочкой – UHF=. При расчетах на уровнях MP2 или UMP2: MP2= или UMP2= и т.д.

Эта энергия выдается в атомных единицах (Хартри). Перевод в другие единицы измерения можно выполнить с помощью Приложения. Эту энергию можно использовать для расчета энергии реакций. Например, если рассматривается реакция присоединения



расчет энергии всех участников позволяет рассчитать энергию реакции по закону Гесса как разность полных энергий продукта и суммы энергий реагентов. Важным условием правильности такого расчета является то, что расчет для всех участников реакции должен проводиться строго на одном и том же уровне теории. Даже небольшое отклонение от этого правила (например, использование разных базисов или разных функционалов DFT для различных участников реакции) приводит к ошибке в десятки и сотни ккал/моль.

Контрольные вопросы

1. Создайте файл для оптимизации геометрии молекулы SiF₄ на уровне B3LYP/6-31G(d,p) так, чтобы во время расчета поддерживалась тетраэдрическая структура молекулы.

2. Создайте файл для оптимизации геометрии молекулы молекулярного комплекса SiF₄·H₂O на уровне B3LYP/6-31G(d,p) так, чтобы во время расчета структура молекулы SiF₄ могла отклоняться от тетраэдрической. Создайте несколько возможных конформаций этого комплекса, различающихся способами координации атомов.

3. Оптимизируйте геометрию молекул H_2O , SiF_4 и $\text{SiF}_4 \cdot \text{H}_2\text{O}$ описанных в примерах выше и заданиях 1-2. Какая конформация комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$ самая выгодная?

4. Чему равна энергия образования комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$ в единицах Хартри, эВ, ккал/моль, кДж/моль?

3.3. Расчет молекулярных свойств при фиксированной геометрии

Расчет при фиксированной геометрии (*single-point calculation, SP*) используется только в том случае, когда есть уверенность, что структура молекулярной системы физически обоснована, например, найдена в ходе предварительной оптимизации геометрии или выбрана на основе экспериментальных данных. В последнем случае, однако, следует помнить, что экспериментальная структура почти никогда не является минимумом ППЭ, соответствующей данному методу расчета.

Целью расчета при фиксированной геометрии является оценка *статических молекулярных свойств*, которые не требуют вариации геометрии. Такими свойствами являются, например: коэффициенты разложения молекулярных орбиталей по базисным функциям; орбитальные энергии; потенциалы ионизации; атомные заряды; дипольные и высшие мультипольные моменты молекулы; поверхности электростатического потенциала молекулы, электронной плотности, молекулярных орбиталей; вертикальные энергии электронно-возбужденных состояний и силы осцилляторов электронных возбуждений, позволяющие воспроизвести УФ/оптический электронный спектр; поправки к энергии молекулы, обусловленные влиянием среды; поправки к энергии молекулы, обусловленные влиянием базисной суперпозиционной погрешности (BSSE).

Расчет при фиксированной геометрии в программе Gaussian выполняется при задании в командной строке ключевого слова SP. Этот режим является режимом работы по умолчанию, поэтому полное отсутствие команды оптимизации или расчета частот также интерпретируется как команда SP. Если задана только команда SP, выполняется только расчет энергии молекулы и волновой функции. Для расчета и выдачи в файл других молекулярных свойств требуется указание дополнительных ключевых слов. Ниже приводятся типичные примеры таких команд для расчета различных молекулярных свойств.

Примеры расчета различных свойств при фиксированной геометрии

1. Распечатка коэффициентов разложения МО по базисным функциям.

```
# MP2/6-311++G(2d,2p) SP Pop=Full
```

Команда Pop=Full означает проведение полного анализа заселенности МО и выводу в файл всех коэффициентов МО. Она приводит к выдаче очень большого объема данных, сократить этот объем можно используя команду

Pop=Normal, однако в этом случае будут выведены коэффициенты не для всех МО. Отметим, что при расчете методами теории возмущений (MP2), связанных кластеров (CCSD), бракнеровских пар (BD) выдается волновая функция, соответствующая невозмущенной системе, то есть она не полностью соответствует той энергии, которая получается при расчете данным методом.

Расчет при фиксированной геометрии в программе Gaussian полезно проводить с повышенной точностью ССП, т.к. по умолчанию программа понижает эту точность по сравнению с режимом оптимизации геометрии. Иногда это приводит к незначительным, но заметным погрешностям, например нефизической асимметрии коэффициентов МО. Для увеличения точности ССП используется команда SCF=Tight, возможно, с одновременным увеличением ССП-итераций. Например:

```
# MP2/6-311++G(2d,2p) SP Pop=Full SCF=(Tight,MaxCycle=300)
```

Расчет коэффициентов орбиталей очень часто необходим для визуализации молекулярных поверхностей – форм МО, электронной плотности или электростатического потенциала. Такая визуализация возможна с помощью различных программ: ChemCraft, GaussView, MolDen, HyperChem. Одной из наиболее удобных программ является программа ChemCraft, которая позволяет визуализировать поверхности, читая выходной файл программы Gaussian, если в нем присутствуют коэффициенты МО. На рис. 11 представлен пример визуализации фронтальных орбиталей германийорганического соединения $\text{PhMe}_2\text{Ge}-\text{C}\equiv\text{C}-\text{C}_6\text{H}_4-\text{C}_6\text{H}_4-\text{C}\equiv\text{C}-\text{GeMe}_2\text{Ph}$ с помощью программы Chemcraft после расчета на уровне B3LYP/6-31G(d,p). Визуализация молекулярных поверхностей (формы орбиталей, электронной плотности, электростатического потенциала) с помощью программы HyperChem выполняется в диалоговом режиме сразу же после проведения расчета (рис. 12).

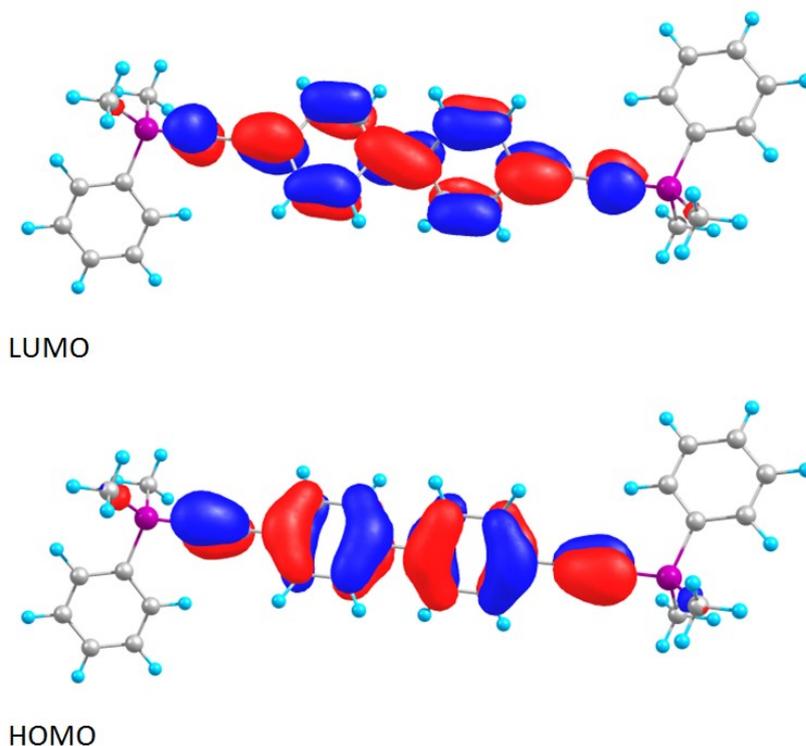


Рис. 11. Фронтальные орбитали молекулы $\text{PhMe}_2\text{Ge-C}\equiv\text{C-C}_6\text{H}_4\text{-C}_6\text{H}_4\text{-C}\equiv\text{C-GeMe}_2\text{Ph}$, визуализированные программой Chemcraft после расчета на уровне B3LYP/6-31G(d, p)

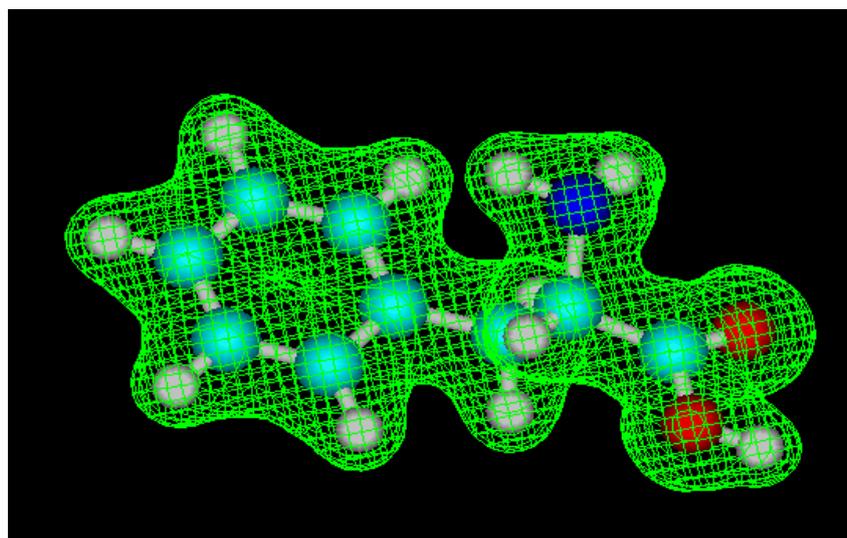


Рис. 12. Электронная плотность молекулы фенилаланина, рассчитанная на уровне B3LYP/6-31G(d,p) и визуализированная с помощью программы HyperChem

Очень часто практической задачей является визуализация одноэлектронных уровней энергии. Такая визуализация также может быть выполнена с помощью программ ChemCraft, GaussView, HyperChem. Однако иногда удобнее получить рисунок в формате векторной графики, для того, чтобы его можно было масштабировать, добавлять дополнительную информацию или внедрять в другие рисунки. Такую возможность предоставляет специальный скрипт PltDraw11 для графического редактора, выполняющий построение

рисунка электронных уровней по данным, подготовленным вспомогательной программой Levels (скрипт и программа размещены на сайте <http://www.qchem.unn.ru>). Программа Levels читает выходные файлы Gaussian, GAMESS, NWchem, в которых находятся коэффициенты МО и, визуализируя эти уровни на экране, позволяет выбрать необходимый диапазон электронных уровней и масштаб энергий. При завершении работы, она создает файл *.plt, который читается скриптом Pltdraw11, выполняемым из графического редактора CorelDraw. Скрипт создает в графическом редакторе изображение уровней одноэлектронных энергий в векторном формате, которое можно легко редактировать и комбинировать с другими изображениями, создавая весьма сложные рисунки. Результаты такой визуализации представлены на рис. 13, где рисунок одноэлектронных уровней энергии совмещен с визуализированными структурами композитных кластеров (использована программа ChemCraft) с добавлением необходимой поясняющей информации.

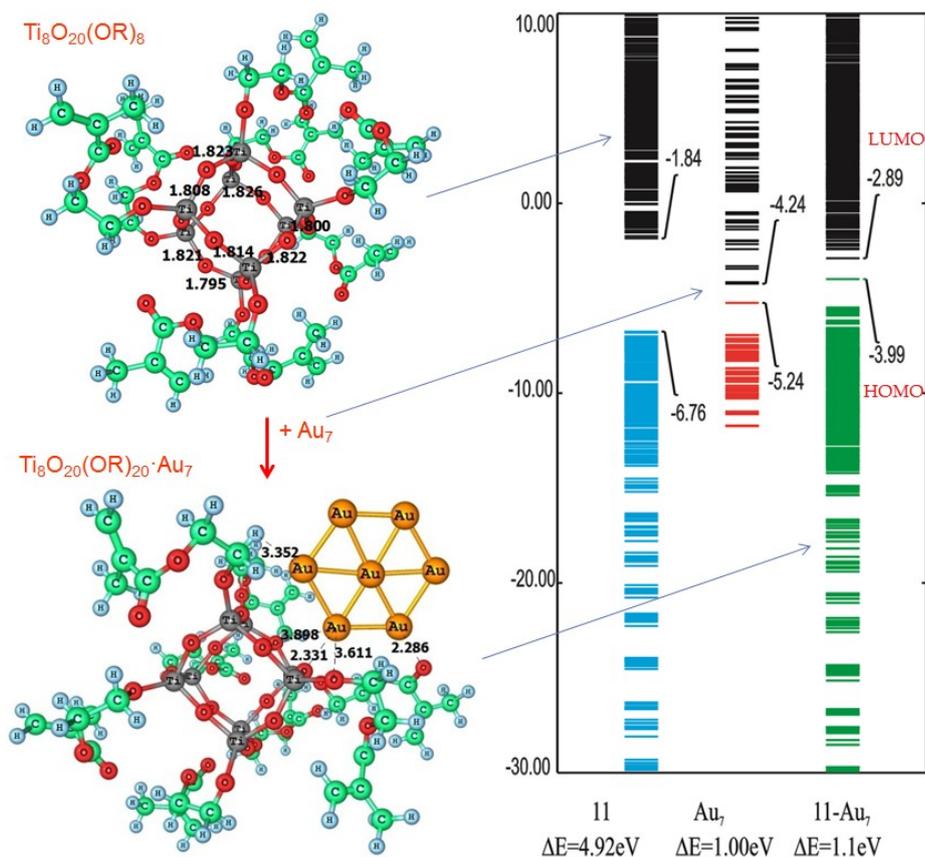


Рис. 13. Визуализация одноэлектронных уровней энергии кластеров $\text{TiO}_2/\text{HEMA}/\text{Au}_7$, моделирующих структуру композитного фотокаталитического наноматериала. Расчет энергий МО проведен в программе Gaussian на уровне UB3LYP/6-31G(d,p). Расчет демонстрирует влияние золотых наночастиц на ширину запрещенной зоны наноматериала

2. Расчет индексов реакционной способности и других статических свойств молекул

Важное направление использования коэффициентов МО, рассчитанных при фиксированной геометрии – оценка различных *индексов реакционной способности*, т.е. молекулярных характеристик, направленных на предсказание активности молекулы в различных химических процессах без непосредственного моделирования этих процессов. Некоторые из таких характеристик приведены в табл. 1.

Таблица 1

Электронные индексы реакционной способности, рассчитываемые на основе орбитальных энергий

Характеристика	Название	Физический смысл
ε_{HOMO}	Энергия ВЗМО	Энергия высшей занятой орбитали (ВЗМО, <i>HOMO</i>)
ε_{LUMO}	Энергия НСМО	Энергия низшей свободной орбитали (НСМО, <i>LUMO</i>)
ε_{SOMO}	Энергия ОЗМО	Энергия однократно занятой орбитали (ВЗМО радикала, <i>SOMO</i>)
$\varepsilon_{gap} = \varepsilon_{LUMO} - \varepsilon_{HOMO}$	Энергетическая щель, запрещенная зона	Разница энергий между ВЗМО и НСМО, дает приближенную величину энергии минимального электронного возбуждения
$I = -\varepsilon_{HOMO}$	Потенциал ионизации по Купмансу	Энергия, необходимая для удаления электрона из химической системы
$A = -\varepsilon_{LUMO}$	Сродство к электрону по Купмансу	Энергия, выделяющаяся при акцептировании электрона химической системой
$\mu = -(I + A) / 2$	Электронный химический потенциал	Понижение энергии системы при акцептировании бесконечно малого отрицательного заряда, $\mu \equiv -\delta E / \delta N$
$\chi = (I + A) / 2$	Электроотрицательность по Малликену	Склонность атома или молекулы принимать электроны, $\chi = -\mu$
$\eta = (I - A) / 2$	Химическая жесткость	Способность противостоять деформации электронной оболочки под влиянием окружающих реагентов, $\eta \equiv \delta \mu / \delta N = \delta^2 E / \delta N^2$
$S = 1 / (2\eta)$	Химическая мягкость	Поляризуемость электронной оболочки под влиянием окружающих реагентов
$\omega = \mu^2 / (2\eta)$	Электрофильность	Характеризует реакционную способность в реакциях электрофильного замещения
$\varepsilon = 1 / \omega$	Нуклеофильность	Характеризует реакционную способность в реакциях нуклеофильного замещения
$\omega^- = \frac{I^2}{2(I - A)}$	Электродонирующая сила	Способность выступать донором электронов
$\omega^+ = \frac{A^2}{2(I - A)}$	Электроакцепторная сила	Способность выступать акцептором электронов

3. Расчет энергий электронных возбуждений (электронного спектра) молекулы:

```
# TD-DFT (Nstates=30, Singlets) CAM-B3LYP/6-31G(d,p) SP  
# TD-DFT (Nstates=10, Triplets) CAM-B3LYP/6-31G(d,p) SP  
# TD-DFT (Nstates=20, 50-50) PBE1PBE/6-311++G(2d,2p) SP
```

Команда TD-DFT означает расчет электронного спектра методом TD-DFT. В первом случае будут рассчитаны энергии 30 низших синглетных электронно-возбужденных состояний, во-втором – 10 низших триплетных состояний, в третьем будут рассчитаны 20 синглетных и 20 триплетных состояний. Отметим, что расчет электронного спектра значительно улучшается при использовании функционала типа CAM-B3LYP (CAM – *Coulomb attenuated method*), в котором используется коррекция зависимости дальнедействующей части кулоновского потенциала, значительно улучшающая результаты расчета энергий возбуждения. В качестве примера на рис. 14. приведено сравнение рассчитанных и экспериментальных УФ-спектров двух германийорганических молекул, рассчитанных методом TD-DFT с четырьмя различными функционалами.

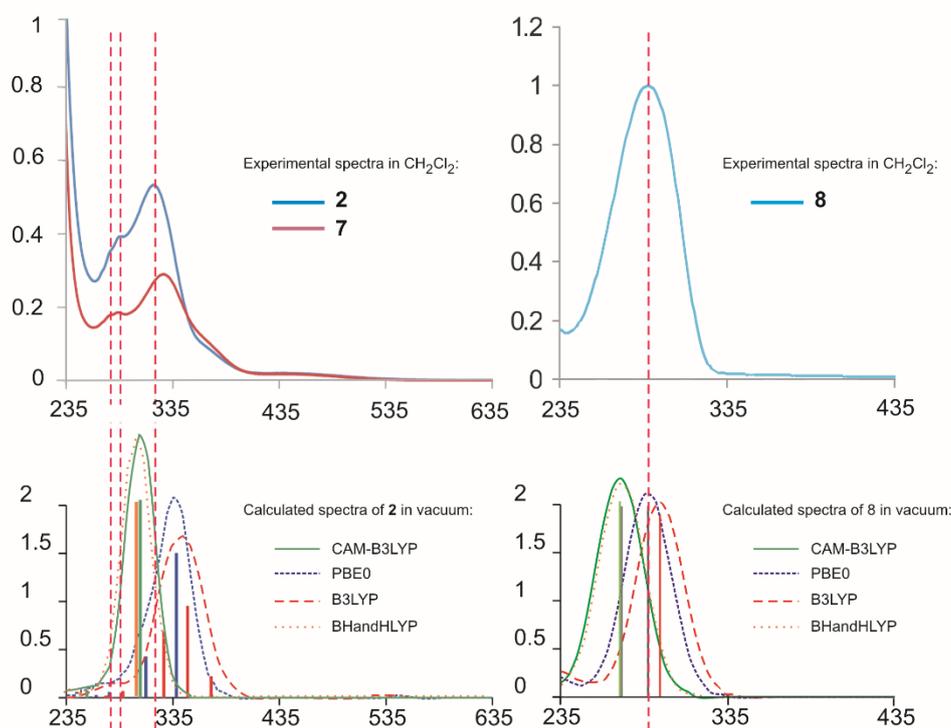


Рис. 14. Электронные спектры трех германийорганических соединений в вакууме (обозначены **2**, **7**, **8**), рассчитанные методом DFT с использованием различных функционалов в сравнении с экспериментальным спектром в CH_2Cl_2 . Горизонтальная ось – длина волны в нм. Вертикальная ось – поглощение, условные единицы. Рисунок демонстрирует типичный разброс значений предсказываемых полос поглощения в УФ-спектре при использовании различных уровней теории и согласие теоретических спектров с экспериментальными данными

3. Расчет базисной суперпозиционной погрешности (BSSE) в предварительно оптимизированной молекуле димера воды:

```
# mp2/6-311++G(2d,2p) SP SCF=Tight Counterpoise=2

H2O dimer

0 1
O      0.000324    1.523438    0.000000    1
H     -0.076645    0.561690    0.000000    1
H     -0.903959    1.837060    0.000000    1
O      0.000324   -1.393689    0.000000    2
H      0.487708   -1.718371    0.759516    2
H      0.487708   -1.718371   -0.759516    2
```

Базисная суперпозиционная погрешность – погрешность, возникающая при расчете энергии координации слабых комплексов. Она возникает из-за того, что при расчете мономеров по отдельности, а затем димера базисный набор формально становится в два раза больше. Так как энергия сильно зависит от базиса, энергия димера кроме физических причин (энергии координации) понижается за счет сложения базисов мономеров. Это понижение – нефизическая величина, и ее следует вычитать из рассчитанной энергии координации. Одним из методов оценки BSSE (предпочтительным) является расширение базиса до такой степени, что он перестает влиять на результаты. Однако этот метод сложно осуществить на практике и часто применяется упрощенный метод, называемый *методом противовеса* или *методом Бойса*. В рамках этого метода программой производятся несколько дополнительных расчетов комплекса и мономеров без оптимизации геометрии. В частности, рассчитываются энергия мономера А без мономера В (E_A); энергия мономера В без мономера А (E_B); энергия мономера А с дополнительными базисными функциями мономера В, центрированными в точках ядер мономера В (но без самих ядер В) (E_{A+}); аналогичная энергия мономера В E_{B+} . Оценка BSSE в этом случае дается формулой:

$$\Delta_{BSSE} = (E_{A+} - E_A) + (E_{B+} - E_B).$$

Все четыре расчета программой Gaussian производятся автоматически, при условии, что координаты атомов соответствуют оптимизированной структуре комплекса. Программе необходимо указать, какие атомы образуют тот или иной мономер. В приведенном примере команда Counterpoise=2 сообщает программе, что надо произвести расчет BSSE, и комплекс состоит из двух мономеров. Отнесение атомов комплекса к одному из двух мономеров указывается на основе дополнительных индексов справа от координат каждого атома (числа 1 или 2 для двух мономеров в комплексе). Дополнительная команда SCF=Tight означает, что при выполнении процедуры ССП следует добиться повышенной точности. Необходимость этого возникает из-за того, что по умолчанию при расчете SP

точность самосогласования ниже, чем при оптимизации геометрии. Чтобы обеспечить точность расчета энергии равной точности при оптимизации, следует указывать команду SCF=Tight.

В случае удачного завершения расчета в конце выходного файла появится сообщение о величине BSSE, например:

```
Counterpoise: corrected energy =      -152.599679704621
Counterpoise: BSSE energy =          0.001455426198
```

Первое значение представляет собой полную энергию молекулы, скорректированную на величину BSSE, второе – саму величину BSSE в атомных единицах.

Контрольные вопросы

1. Рассчитайте величину BSSE методом противовеса для комплекса SiF₄·H₂O оптимизированного ранее на уровне B3LYP/6-31G(d,p).

2. Оптимизируйте ту же молекулу на уровне B3LYP/6-31++G(d,p) и сравните величину BSSE с базисом 6-31G(d,p). Как влияют диффузные функции на величину BSSE?

3.4. Расчет колебательных частот и термодинамических параметров

Расчет колебательных частот производится при задании команды Freq в командной строке. Команда Freq может быть задана как единственный режим работы или может быть скомбинирована с командой FOpt. В последнем случае комбинация FOpt Freq означает, что вначале программа выполнит оптимизацию геометрии, после которой будет произведен расчет частот для оптимизированной структуры. Важно помнить, что вычисление колебательных частот, как правило, не имеет смысла для неоптимизированной геометрии, поскольку частоты сильно зависят от отклонения от точки минимума. Исключением является расчет частот перед началом поиска переходного состояния (см. ниже).

При выполнении этого типа расчета программа вначале вычисляет матрицу вторых производных энергии по координатам атомов, называемую также гессианом (*матрицей Гессе*) или матрицей силовых констант. Этот процесс является весьма времязатратной процедурой. Он выполняется по внутреннему алгоритму программы, который может быть аналитическим или численным. В случае аналитического расчета (он выполняется по умолчанию, если это позволяет квантовохимический метод) силовые константы вычисляются без явной вариации координат атомов по формулам, заданным в программе. Это наиболее быстрый способ расчета, который, однако, возможен не для всех квантовохимических методов. Если это невозможно (как, например, в случае расчета CCSD(T)), выполняется *численный расчет* силовых констант

(его можно всегда реализовать командой `Freq=Numeric`). При численном расчете каждая декартова координата всех атомов по очереди изменяется на величину $+\Delta x$ и $-\Delta x$, и в полученных точках вычисляются энергия и ее градиенты. Комбинация этих энергий дает вторую производную по данной координате. Такой способ расчета требует намного больше времени и к нему следует прибегать только при невозможности аналитического расчета.

После вычисления матрицы силовых констант, эта матрица диагонализуется и собственные значения, деленные на массы атомов, дают квадраты частот колебаний молекулы. Колебательные частоты, кроме собственного значения, являются важным критерием устойчивости молекулы и характеристикой точки ППЭ. Если все колебательные частоты оптимизированной структуры действительны (в программе действительные частоты соответствуют положительным значениям частоты колебания), структура молекулы соответствует локальному минимуму ППЭ. Если одна (и только одна) частота мнимая (в программе мнимые частоты показываются как отрицательные значения частоты), структура соответствует седловой точке ППЭ, то есть переходному состоянию некоторой реакции. Если мнимыми (отрицательными) являются несколько частот, это говорит о том, что структура не оптимизирована, или оптимизация не дошла до конца.

Возможен также случай, когда одна или несколько частот близка к нулю. При расчете в декартовых координатах любая N -атомная нелинейная молекула обладает 6 модами (движениями) нулевой частоты (линейная молекула – пятью такими модами). Эти моды называются тривиальными или нулевыми. Три из них соответствуют поступательному движению центра масс молекулы, три оставшиеся (две в случае линейной молекулы) – вращению молекулы как целого. Оставшиеся $3N-6$ ($3N-5$ в случае линейной молекулы) мод соответствуют собственно колебательному движению. В программе Gaussian тривиальные моды автоматически отделяются и не анализируются. Среди оставшихся мод, однако, могут проявиться движения с низкой частотой ($<100 \text{ см}^{-1}$). Такие моды называются высокоамплитудными колебаниями, они представляют физически наблюдаемые движения. Могут также появиться положительные или отрицательные частоты очень близкие к нулю ($|\nu| < 10 \text{ см}^{-1}$). Такие движения, скорее всего, являются артефактами расчета. Они могут свидетельствовать о том, что оптимизация не дошла до конца, либо о том, что структура нестабильна. Хотя единого алгоритма действий при появлении таких частот нет, факты их появления следует отслеживать, они могут свидетельствовать, что оптимизация геометрии могла привести к неправильной структуре.

Появление мнимых частот после оптимизации молекулы однозначно свидетельствует, что локальный минимум не достигнут, и найденная структура нестабильна. В этом случае стандартным действием является продолжение оптимизации структуры, скорректированной на величину колебательного вектора. Для этого к координатам оптимизированной структуры \mathbf{r} следует прибавить и вычесть колебательный вектор $\Delta \mathbf{r}$ колебания мнимой частоты,

умноженный на некоторую величину (амплитуду смещения s), которая обычно выбирается в пределах 0.05–0.3. Полученная пара структур ($\mathbf{r} + s\Delta\mathbf{r}$ и $\mathbf{r} - s\Delta\mathbf{r}$) подвергается дальнейшей оптимизации. В большинстве случаев правильный выбор амплитуды позволяет уйти от точки с мнимыми частотами и после дополнительной оптимизации найти структуру истинного минимума. Во многих квантовохимических программах (например, в программе NWchem) коррекция на величину колебательного вектора производится автоматически после обнаружения мнимых частот и пользователю достаточно скопировать сгенерированные координаты для продолжения оптимизации. Однако в программах Gaussian и GAMESS такого режима работы нет, и генерация координат выполняется пользователем. Это можно сделать вручную, например, в программе MS Excel. Более удобно использовать программу Moltran, которая генерирует исправленные координаты после указания номера моды и выбранной величины амплитуды.

Другим способом борьбы с мнимыми частотами после оптимизации геометрии является применение процедуры построения внутренней координаты реакции (*IRC, intrinsic reaction coordinate*). При этом типе расчета программа читает информацию о гессiane, определяет направления вектора с мнимой частотой и движется в обе стороны от начальной точки вдоль этого вектора, пока не достигнет минимумов, в которых все частоты положительны. Выполнение процедуры IRC является обязательным при поиске переходных состояний (см. ниже). В случае простой оптимизации геометрии, хотя такой способ устранения мнимых частот является более последовательным и строгим, он значительно более затратный и сложный, чем оптимизация структур, скорректированных на колебательный вектор.

Следует помнить, что рассчитываемые квантовохимическим методом частоты колебаний на любом уровне теории в принципе отличаются от экспериментально регистрируемых. Причина состоит в том, что рассчитываемые частоты рассчитываются в гармоническом приближении, то есть соответствуют чисто квадратичному потенциалу $kx^2/2$. Частоты реальной молекулы всегда являются ангармоническими, поскольку межатомный потенциал отличается от квадратичного. Гармонические частоты всегда выше ангармонических, вследствие чего частоты, рассчитываемые квантовохимическим методом обычно на 5-15% завышены по сравнению с экспериментальными и улучшение уровня теории не приведет к улучшению согласия между ними. Чтобы учесть этот факт, используются несколько приемов. Во-первых, для каждого уровня теории можно выработать масштабирующий коэффициент, умножение на который гармонических частот значительно улучшает согласие рассчитанных и экспериментальных частот. Для уровня HF/6-31G(d) такой коэффициент равен 0.89, для B3LYP/6-31G(d,p) 0.95. Значения масштабирующих коэффициентов могут быть найдены в специализированных обзорах или выработаны для каждой системы самостоятельно на основе расчетов молекул с известными колебательными частотами. Гармонические частоты, умноженные на соответствующий коэффициент, называются масштабированными частотами. Второй метод заключается в том, что при изучении ИК-спектров участников

какой-либо реакции анализируются не абсолютные величины частот, а их изменение в ходе реакции. Например, при изучении комплексообразования двух соединений методом матричной изоляции, в низкотемпературной матрице твердого инертного газа замораживаются молекулы короткоживущего комплекса, ИК-спектр которого можно изучать и сравнивать с ИК спектром отдельных реагентов. Координация двух молекул приводит к смещению полос поглощения в ИК-спектре на величины от единиц до нескольких десятков см^{-1} . Величина и знак смещения частоты – очень характерная величина и она слабо зависит от способа расчета. Поэтому сравнение экспериментальных и рассчитанных сдвигов полос поглощения – значительно более надежный метод анализа спектров, чем анализ абсолютных частот. Пример такого сравнения приведен в табл. 1. Левые столбцы таблицы показывают экспериментальные частоты, которые сильно отличаются от абсолютных значений рассчитанных частот (правые столбцы). Однако наблюдаемые сдвиги полос при координации молекул SiF_4 и H_2O очень хорошо согласуются между собой. В частности, согласие знаков и характерной величины сдвигов -44 , 17 и -12 см^{-1} с экспериментальными значениями -39 , 5 и -8 см^{-1} позволяет однозначно утверждать, что изолированная в эксперименте молекула имеет ту же структуру, что и оптимизированная квантовохимическим методом. Отсутствие в экспериментальном спектре ряда колебательных частот, которые есть в теоретическом ИК-спектре, объясняется тем, что часть колебаний имеют низкую ИК-интенсивность, а другая часть в экспериментальном спектре маскируется сильными полосами остаточных концентраций реагентов.

Одновременно с частотами и модами нормальных колебаний, квантовохимическая программа выдает значения ИК-интенсивности I_i (в физике называемой силой линии) каждого колебания, позволяющие построить колебательный спектр (ИК-спектр) молекулы. Эти величины в большинстве программ выдаются в единицах км/моль , которые соответствуют интегральному представлению интенсивности полосы поглощения в виде:

$$I_i = \int_{\bar{\nu}_i - \Delta\bar{\nu}}^{\bar{\nu}_i + \Delta\bar{\nu}} A(\bar{\nu}) d\bar{\nu}.$$

Здесь $A(\bar{\nu})$ – абсолютный коэффициент экстинкции из закона Бугера-Ламберта-Бэра на дине волны, соответствующей волновому числу $\bar{\nu} = 1/\lambda$:

$$i = i_0 \exp(-ACl).$$

Здесь i_0, i – интенсивность излучения, падающего на образец и прошедшего через образец, соответственно. C – концентрация поглощающего вещества в образце, моль/м^3 ; l – длина пути света в образце, м .

Вычисление ИК-интенсивностей при квантовохимическом расчете (табл. 2) обычно производится в двойном гармоническом приближении, в котором интенсивность ИК-полосы пропорциональна квадрату производной

дипольного момента μ по смещению атомов в направлении вектора колебания (масс-взвешенному вектору нормального колебания Q_i):

$$I_i = \frac{N_A}{12c^2 \varepsilon_0} \left(\frac{\partial \mu}{\partial Q_i} \right)^2.$$

Здесь N_A , c и ε_0 – число Авогадро, скорость света и диэлектрическая проницаемость вакуума.

Таблица 2

Сравнение экспериментальных сдвигов ИК-полос поглощения комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$, изолированного в матрице твердого аргона, со сдвигами полос поглощения этого комплекса, рассчитанными методом MP2/6-311++G(d,p)

Эксперимент $\nu_{\text{exp.}} \text{ см}^{-1}$		Наблюдаемый сдвиг, см^{-1}	Расчет $\nu_{\text{exp.}} \text{ см}^{-1}$		Предсказываемый сдвиг, см^{-1}
$\text{SiF}_4 + \text{H}_2\text{O}$	$\text{SiF}_4 \cdot \text{H}_2\text{O}$		$\text{SiF}_4 + \text{H}_2\text{O}$	$\text{SiF}_4 \cdot \text{H}_2\text{O}$	
				97	
				110	
				112	
				160	
(264)			256	228	-29
(264)			256	240	-17
				266	
				325	
(388)			373	369	-3
(388)			373	376	3
(388)			373	386	14
(801)			767	757	-10
1023	984, 991 ^c	-39, -32	999	955	-44
1023			999	998	-1
1023			999	1009	10
1593	1598	5	1630	1647	17
(3657)			3868	3858	-10
3730	3722	-8	3992	3980	-12

На основе рассчитанных колебательных частот и ИК-интенсивностей может быть построен ИК-спектр соединения, который гораздо удобнее сравнивать с экспериментальными данными. Построение спектра, как правило, выполняется отдельной программой. Спектр может быть представлен в виде *бар-графа*, то есть набора линий высотой I_i на абсциссе ν_i или в виде набора непрерывных огибающих $I(\nu)$, которые обычно рассчитываются в виде комбинации гауссовых функций:

$$I(\nu) = \sum_i I_i \exp\left(-\frac{(\nu - \nu_i)^2}{\sigma^2}\right),$$

где σ – полуширина полосы поглощения, величина которого 10–100 см^{-1} .

Получаемые графики позволяют прямое сравнение рассчитанного и экспериментального спектра, как это показано на рис. 15.

Визуализацию спектров в виде, показанном на рис. 15, можно проводить с помощью различных программ-визуализаторов. Одним из простых способов визуализации является использование графического редактора со специальным скриптом *Spectr11*, который размещен на сайте группы квантовой химии ННГУ <http://www.qchem.unn.ru>.

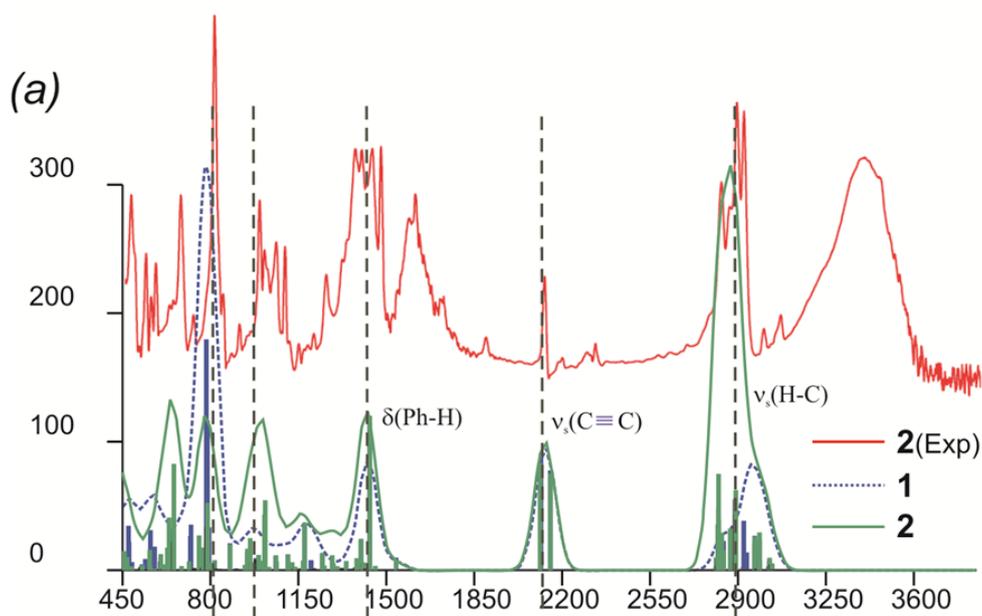


Рис. 15. Сравнение ИК-спектра двух элементоорганических соединений (1 и 2), рассчитанного квантовохимическим методом B3LYP/6-31G(d,p) (огibaющие и отрезки внизу) с экспериментальным спектром (линия вверху). Масштабирующий коэффициент для частот 0.93, полуширина огibaющих $\sigma = 50 \text{ см}^{-1}$

Перед запуском скрипта файл с результатами расчета частот необходимо открыть программой Moltran. После завершения ее работы следует запустить скрипт и указать имя файла *.мои, созданного Moltran. Скрипт нарисует ИК-спектр в соответствии с настройками диапазона, масштабирующего коэффициента, ориентации осей, параметров огibaющей и т.д. Все эти настройки содержатся в самом файле скрипта, снабжены пояснениями и могут быть отредактированы с помощью любого текстового редактора.

Если все колебательные частоты действительны и нет «почти нулевых» частот, которые вызывают подозрения на неправильность оптимизации, можно оценить термодинамические характеристики системы, то есть ее термодинамические функции U , H , S , F , G при заданных температуре и давлении. Квантовохимическая программа вычисляет эти значения по формулам, включающим (а) оптимизированную структуру молекулы и соответствующую этой структуре энергию; (б) мультиплетность электронного состояния M ; (в) колебательные частоты ν_i . Конечные формулы могут несколько отличаться, один из способов расчета дается следующими выражениями:

$$Q_{elec} = M, \quad Q_{tran} = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} \frac{RT}{P},$$

$$Q_{rot} = \frac{\sqrt{\pi}}{\sigma} \left(\frac{8\pi^2 kT}{h^2} \right)^{3/2} \sqrt{I_1 I_2 I_3} \quad (\text{нелинейная молекула}),$$

$$Q_{rot} = \frac{8\pi^2 kT}{\sigma h^2} I \quad (\text{линейная молекула}),$$

$$Q_{vib} = \prod_{i=1}^{N_{vib}} \frac{1}{1 - \exp(-h\nu_i / kT)}, \quad E_{ZPE} = \frac{h}{2} \sum_{i=1}^{N_{vib}} \nu_i$$

$$Q = Q_{elec} Q_{tran} Q_{rot} Q_{vib},$$

$$F = E_{ZPE} - kT \ln Q, \quad S = -\frac{\partial F}{\partial T} = k \ln Q + kT \frac{\partial \ln Q}{\partial T},$$

$$G = F + kT, \quad H = G + TS, \quad U = F + TS$$

В этих формулах Q_{elec} , Q_{tran} , Q_{rot} , Q_{vib} – статистические суммы электронных, поступательных, вращательных и колебательных движений молекулы, а Q – полная молекулярная статсумма; $M = 2S + 1$ – электронная спиновая мультиплетность (S – полный спин молекулы); σ – *вращательное число симметрии*, которое определяется как порядок подгруппы вращений группы симметрии молекулы (для групп C_n , C_{nv} , C_{nh} σ равно n , для групп D_n , D_{nh} , D_{nd} – $2n$, для $D_{\infty h}$, T_d , O_h , I_h – 2, 12, 24, 60, для остальных групп – 1). I_1, I_2, I_3 – главные моменты инерции молекулы, то есть собственные числа матрицы момента инерции. Если молекула линейная, в формулу входит только один главный момент инерции ($I_1 = I_2 \equiv I$, $I_3 = 0$). Величины ν_i – частоты нормальных колебаний, E_{ZPE} – энергия нулевых колебаний.

Значения U , H , G , F , получаемые по этим формулам, представляют собой вклады термодинамических функций индивидуальных соединений в расчете на одну молекулу относительно энергии молекулы E_{tot} в состоянии идеального газа. Величина S по третьему закону термодинамики представляет собой абсолютную энтропию соединения в расчете на 1 молекулу. Молярные величины U , H , G , F , S получаются путем умножения на число Авогадро. Поскольку формулы дают энергетические величины U , H , G , F относительно минимума полной энергии молекулы (E_{tot}), для сравнения термодинамических функций двух молекул к этим величинам необходимо прибавить E_{tot} .

В выходном файле программы Gaussian термодинамические величины выводятся в виде:

```
Zero-point correction= 0.107523 (Hartree/Particle)
Thermal correction to Energy= 0.111166
```

Thermal correction to Enthalpy=	0.112110
Thermal correction to Gibbs Free Energy=	0.081790
Sum of electronic and zero-point Energies=	-157.006337
Sum of electronic and thermal Energies=	-157.002694
Sum of electronic and thermal Enthalpies=	-157.001750
Sum of electronic and thermal Free Energies=	-157.032070

Первая часть представляет собой термодинамические вклады E_{ZPE} , U , H , G (их часто называют *термические поправки*). Вторая часть представляет собой суммы E_{tot} с термическими поправками. В обеих частях таблицы единицы измерения – Хартри в расчете на 1 молекулу (на одну рассчитываемую систему). Для перевода в ккал/моль и кДж/моль надо воспользоваться переводными коэффициентами, приведенными в Приложении.

По умолчанию программа Gaussian производит расчет термодинамических характеристик для температуры $T = 298.15\text{K}$ и $P = 1\text{ атм}$ (101.325 кПа). Обратите внимание, что в настоящее время по рекомендациям IUPAC величина стандартного давления составляет 100 кПа. Величины температуры и давления, а также изотопный состав молекулы, влияющий на частоты и массу молекулы, можно изменить, проведя повторный расчет без расчета силовых констант. Этот способ сокращает время повторного расчета до нескольких секунд при условии, что был сохранен *chk*-файл от расчета колебательных частот. В *gjf*-файле можно задать следующие команды:

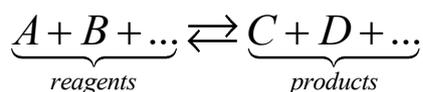
```
%chk=h2o.chk          <-Старый chk-файл с сил.константами
# B3LYP/6-31++G(d,p) Freq=ReadIso <-Команда ReadIsotopes
# Geom=AllCheck Guess=Read <-Эти команды извлекают из chk-файла
                             геометрию и матрицу плотности
300. 2.0              <-Новые T(K) и P(атм)
8                     <-Изотопные массы всех атомов в том же
2                     порядке, что и в сиходной геометрии
2                     молекулы (в данном случае для D2O)
```

Другим способом провести расчет термодинамики при заданных T и P является применение программы MOLTRAN. Выполнение команды

```
moltran h2o.out /tt300 /pp2.0 /0
```

позволит извлечь из файла *h2o.out* результаты расчета частот, выполненных при каких-либо термодинамических условиях и пересчитать термодинамические функции при других температуре и давлении, заданных ключами */tt*<температура, K> и */pp*<давление, атм>. Ключ */0* означает отказ от диалоговой работы, то есть программа выполнит расчет термодинамических функций и завершит работу. Новые термодинамические функции будут выведены в созданный программой файл *h2o.mou*.

Расчет термодинамических величин позволяет вычисление термодинамических функций химических реакций и оценку константы равновесия. Предположим, что в идеальном газе протекает химическая реакция между несколькими веществами,



Термодинамические параметры реакции и стандартная константа равновесия при данной температуре находятся по формулам:

$$\Delta_r E = \sum_i^{\text{products}} E_{\text{tot},i} - \sum_j^{\text{reagents}} E_{\text{tot},j}$$

$$\Delta_r H(T) = \sum_i^{\text{products}} H_i(T) - \sum_j^{\text{reagents}} H_j(T)$$

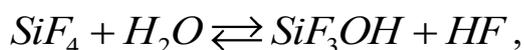
$$\Delta_r G(T) = \sum_i^{\text{products}} G_i(T) - \sum_j^{\text{reagents}} G_j(T)$$

$$K^0(T) = e^{-\Delta_r G(T)/RT}$$

Здесь H_i и G_i подразумевают суммы полной энергии молекулы i и термической поправки (например, $H + E_{\text{tot}}$).

Пример

Рассчитать методом DFT термодинамические функции химической реакции



протекающей в газовой фазе при стандартных условиях, и оценить константу равновесия этой реакции.

Решение

Выберем уровень теории B3LYP/6-31G(d,p) (B3LYP – надежный функционал среднего уровня точности, подходящий для большинства задач, базис 6-31G(d,p) – минимальный базис, с которого можно начинать исследование и затем улучшать его, наличие поляризационных функций обязательно, диффузные функции не требуются).

Проведем расчет всех участников реакции (SiF_4 , H_2O , SiF_3OH , H_2O), используя строку команд

```
# B3LYP/6-31G(d,p) FOpt Freq
```

Из файлов *.out извлекаем сумму полной энергии и термических поправок, например для воды:

```
Sum of electronic and zero-point Energies= -76.398372
Sum of electronic and thermal Energies= -76.395537
Sum of electronic and thermal Enthalpies= -76.394593
Sum of electronic and thermal Free Energies= -76.416029
```

Результаты заносим в таблицу Excel. В таблицу заносим также имена файлов, откуда взяты данные, число вращательной симметрии и число мнимых частот (0, если их нет). Эти данные нужны, чтобы понять, как проводился расчет. Очень часто ошибки в расчете возникают, когда используются результаты неполной оптимизации (с мнимыми частотами) или структуры симметричной молекулы рассчитаны как асимметричной. В данном случае молекулы SiF₄ и H₂O симметричные, они должны иметь числа вращательной симметрии 12 (группа T_d) и 2 (группа C_{2v}), соответственно. Если в ходе расчета эти молекулы считались асимметричными, числа симметрии будут равны 1, и это может привести к довольно заметной погрешности. Получаем таблицу:

3								
4	B3LYP/6-31G(d,p)	SiF4	H2O	=	SiF3OH	HF		
5	Nneg		0		0		0	0
6	Sigma		12		2		1	1
7	Etot		-689.1085625		-76.41973658		-665.0809173	-100.4274616
8	Etot+ZPE		-689.0958637		-76.39837123		-665.0573985	-100.4181548
9	Etot+U		-689.0909072		-76.39553616		-665.0516231	-100.4157942
10	Etot+H		-689.089963		-76.39459195		-665.0506789	-100.41485
11	Etot+G		-689.1222175		-76.41602892		-665.0870492	-100.4345681
12	File	sif4.out	h2o.out		sif3oh.out		hf.out	
13								

Используя средства Excel, по вышеприведенным формулам вычисляем термодинамические функции реакции в а.е., переводим их в ккал/моль и кДж/моль. Вычисляем значение стандартной константы K^0 :

	B3LYP/6-31G(d,p)	SiF4	H2O	=	SiF3OH	HF	dE, a.u.	kcal/mol	kJ/mol	K0	
5	Nneg		0		0		0				
6	Sigma		12		2		1				
7	Etot		-689.1085625		-76.41973658		-665.0809173	-100.4274616	0.01992	12.50	52.3
8	Etot+ZPE		-689.0958637		-76.39837123		-665.0573985	-100.4181548	0.01868	11.72	49.0
9	Etot+U		-689.0909072		-76.39553616		-665.0516231	-100.4157942	0.01903	11.94	49.9
10	Etot+H		-689.089963		-76.39459195		-665.0506789	-100.41485	0.01903	11.94	49.9
11	Etot+G		-689.1222175		-76.41602892		-665.0870492	-100.4345681	0.01663	10.43	43.7
12	File	sif4.out	h2o.out		sif3oh.out	hf.out					

Термодинамические параметры реакции, рассчитанные методом DFT в небольшом базисе, обычно значительно отличаются от экспериментальных результатов. Для вычисления высокоточных термодинамических характеристик следует использовать специальные методы, которые обычно называют *высокоточными экстраполяционными методами определения энтальпий*. Таких методов несколько, наиболее известные и современные – G3, G4, CBS, W1BD. Все эти методы работают по принципу экстраполяции: для данной молекулы проводится расчет среднего уровня точности в среднем базисе и в большом базисе. Разница в энергиях между ними ΔE_1 позволяет оценить вклад расширения базиса. Затем в базисе среднего уровня проводится расчет на среднем и на высоком уровне теории (например, MP2 и CCSD(T)). Разница в

энергиях между ними ΔE_2 позволяет оценить вклад повышения уровня теории. Тогда величина $\Delta E_1 + \Delta E_2$ позволит получить экстраполированное значение энергии, как если бы проводился расчет в большом базисе очень точным методом, что на практике невозможно. Реальные схемы экстраполяции несколько сложнее, но принцип работы соответствует описанному. Подобные схемы позволяют получить очень хорошие оценки энергетических характеристик, а специально выработанные масштабирующие факторы позволяют достаточно точно оценить вклады в энтальпии колебательных частот, поскольку основной вклад в энтальпии дают высокочастотные колебания. К сожалению, точность расчета энтропии таким методом улучшить практически невозможно, поскольку в энтропию основной вклад вносят низкочастотные колебания, и в частности, высокоамплитудные движения, которые невозможно правильно учесть в рамках гармонического приближения. Тем не менее, для большинства достаточно жестких молекул точность расчета энтальпий и других термодинамических параметров такими методами сравнима с точностью определения параметров экспериментальными методами. Точность экстраполяционной схемы W1BD, как заявляют ее авторы, составляет порядка 0.5 ккал/моль, что попадает в интервал экспериментальной погрешности при измерении энтальпий, статистически оцененной для большого количества химических соединений. Для выполнения расчета этими методами достаточно указать в строке команд единственную команду, например, G3 (G4 и W1BD реализованы в программе Gaussian начиная с версии 2009 года Gaussian09):

```
# G3
```

После успешного завершения расчета результаты появляются в конце выходного файла в виде:

```
Temperature=          298.150000 Pressure=          1.000000
E(ZPE)=              0.020511 E(Thermal)=          0.023347
E(QCISD(T))=        -76.207892 E(Empiric)=         -0.025544
DE(Plus)=           -0.012978 DE(2DF)=            -0.074489
E(Delta-G3)=        -0.081653 E(G3-Empiric)=         -0.025544
G3(0 K)=            -76.382044 G3 Energy=          -76.379209
G3 Enthalpy=        -76.378265 G3 Free Energy=       -76.399641
```

Здесь G3(0 K) это $E_{tot} + ZPE$, G3 Energy — $E_{tot} + U$, G3 Enthalpy — $E_{tot} + H$, G3 Free Energy $E_{tot} + G$. Собственно E_{tot} , соответствующая данному методу, может быть найдена как разность G3(0 K) и величиной E(ZPE).

При проведении любых термодинамических вычислений следует помнить, что, во-первых, все расчеты относятся к идеально-газовому состоянию вещества. Протекание реакции в растворе или конденсированном состоянии может приводить к существенному расхождению между результатами такого расчета и экспериментальных величин. Для учета влияния среды и конденсированного состояния следует применять специальные расчетные схемы, например метод РСМ (см. ниже).

Во-вторых, проведением такого рода вычислений, как правило, невозможно получить стандартные энтальпии образования из простых веществ,

которые обычно указываются в справочниках. Причина этого состоит в том, что многие простые вещества, включая углерод и металлы, в стандартном состоянии представляют собой кристаллы. Расчет соединений в кристаллическом состоянии с термодинамической точностью в настоящее время невозможен. Если все-таки стоит задача оценки функций образования из простых веществ, следует провести расчет этих функций по закону Гесса из вспомогательных газообразных соединений. Для этого выбирается несколько соединений, которые в стандартном состоянии газообразны и имеют хорошо определенные табличные значения функций образования из простых веществ. Вычисление функций образования из этих молекул с использованием закона Гесса позволяет вывести значения функций образования исследуемого соединения из простых веществ. При подборе вспомогательных реакций для уменьшения погрешности рекомендуется выбирать *изодесмические реакции*, то есть реакции, в которых не изменяется количество химических связей (то есть любые типы химических связей представлены в одинаковом количестве в реагентах и продуктах).

Контрольные задания

1. Рассчитайте колебательные частоты молекул SiF_4 и H_2O и комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$ на уровне B3LYP/6-31G(d,p). Чему равны ИК сдвиги наиболее интенсивных колебаний? Можно ли зарегистрировать образование этого комплекса (при условии достаточной концентрации), если разрешающая способность спектрометра 1 см^{-1} .

2. Визуализируйте рассчитанный ИК-спектр комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$.

3. Рассчитайте термодинамические параметры $\Delta H, \Delta G$ образования этого комплекса с учетом BSSE. Чему равна константа равновесия этого комплекса при стандартных условиях в газовой фазе?

4. Рассчитайте термодинамические параметры $\Delta H, \Delta G$ образования этого комплекса с учетом BSSE в газовой фазе при следующих условиях (а) $T=100\text{K}$, $P=10 \text{ мм рт.ст.}$; (б) $T=800\text{K}$, $P=2 \text{ атм.}$

3.5. Поиск переходных состояний и оценка констант скорости реакций

Переходное состояние или седловая точка – это точка ППЭ, в которой градиент энергии равен нулю (так называемая *стационарная точка ППЭ*), а среди колебательных частот *одна и только одна является мнимой*, в то время как остальные – действительные. Седловая точка не является точкой минимума или максимума энергии. Поэтому поиск переходных состояний заключается в минимизации нормы градиента энергии с последующим контролем единственности мнимой частоты. Поиск таких точек может осуществляться с помощью нескольких алгоритмов, наиболее простой и часто используемый из которых – метод следования по собственному вектору (*eigenvector following*). В этом методе в начальной точке оптимизации вычисляются колебательные частоты и в случае единственности отрицательной частоты начинается движение

в направлении повышения энергии вдоль собственного вектора гессиана, соответствующего мнимой частоте. Движение происходит шагами варьируемой длины, причем на каждом шаге гессиан корректируется по специальным формулам или пересчитывается заново, что позволяет скорректировать направление. После достаточного количества шагов такая процедура приводит к седловой точке. К сожалению, правильная сходимость в таком процессе не гарантируется и в ходе движения мнимая частота может исчезать, или появляться две или несколько мнимых частот, и процедура будет сходиться к неправильным конечным точкам. Кроме того, сам выбор начальной точки, в которой есть правильная мнимая мода, далеко не прост и требует определенного навыка и интуиции. То же относится и к другим алгоритмам поиска переходных состояний. В связи с этим, поиск переходных состояний до настоящего времени остается весьма сложной задачей, часто требующей большого количества попыток при выборе начального приближения, химической интуиции и долгого времени для достижения сходимости. Блок-схема алгоритма поиска переходных состояний показана на рис. 16.



Рис. 16. Блок-схема алгоритма поиска переходных состояний

Хотя в программе Gaussian поиск переходного состояния можно выполнить в ходе одного запуска, рекомендуется разбить эту задачу на три этапа.

На первом этапе нужно выбрать начальную геометрию молекулы близкой к структуре переходного состояния. Для этого лучше использовать Z-матрицу или воспользоваться одной из программ-визуализаторов и изменить связи молекулы (молекул), чтобы разрываемые связи были заметно удлинены, а

межатомные расстояния, соответствующие образуемым в ходе реакции связям, приближались к удлиненным значениям этих связей. Степень удлинения связей заранее предсказать сложно, но часто это удлинение попадает в интервал 50-70% от нормальной длины связи.

Протекание бимолекулярной реакции предполагает сближение двух молекул перед образованием переходного состояния. В этом случае удобно начать поиск переходного состояния с оптимизации *предреакционного комплекса*. Поскольку молекулы испытывают взаимное притяжение за счет межмолекулярных сил, обычно оптимизации пары таких молекул приводит к более стабильной структуре комплекса. Начальную структуру для оптимизации его структуры желательно выбирать такой, чтобы она обеспечивала легкое протекание изучаемой реакции (без необходимости выполнения сложной переориентации мономеров в ходе дальнейшей реакции). Когда такой комплекс оптимизирован, его можно рассматривать как единую молекулу, в которой происходит мономолекулярная реакция через искомое переходное состояние. Именно в этой структуре необходимо провести искажение структуры, которая бы напоминала предполагаемую структуру переходного состояния.

Для подготовленной искаженной структуры необходимо выполнить расчет частот без оптимизации геометрии, например:

```
# b3lyp/6-31g(d,p) Freq=NoRaman
```

Ключевое слово NoRaman позволяет ускорить расчет, поскольку отключает вычисление активностей спектра комбинационного рассеяния. Как только расчет завершен, необходимо проанализировать полученные частоты. В идеале должна присутствовать одна мнимая частота и, что очень важно, форма колебания этой частоты должна соответствовать направлению изучаемой реакции. Например, если изучается реакция переноса атома водорода $O-H...O \rightarrow O...H-O$, колебание с мнимой частотой должно включать движение атома H от одного атома кислорода к другому и обратно с заметной амплитудой. Если мнимых частот нет, или форма колебания совершенно не соответствует изменению структуры в ходе реакции, структура должна быть скорректирована и расчет частот повторен заново. Если имеются две и более мнимых частот, следует проанализировать формы их колебаний. Допустимо, если низшая частота имеет колебание, соответствующее требуемому направлению реакции, и при этом по своей величине заметно отличается от других мнимых частот. В противном случае следует скорректировать структуру и повторить расчет колебательных частот. Chk-файл после расчета частот должен быть сохранен.

Если начальная структура выбрана правильно, можно перейти ко второму этапу и запустить оптимизацию переходного состояния командой

```
# b3lyp/6-31g(d,p) FOpt=(TS,RCFC,NoEigenTest,MaxCycle=100)
```

Ключевое слово TS означает оптимизацию переходного состояния методом следования по собственному вектору, RCFC означает, что силовые константы для расчета направления движения берутся из предыдущего расчета

(из сохраненного на первом этапе chk-файла, имя которого должно быть указано в команде %chk=). Ключевое слово NoEigenTest означает, что оптимизация не должна останавливаться при возникновении второй мнимой собственной частоты, MaxCycle задает максимальное число циклов оптимизации). Если процесс оптимизации завершился успешно (выполнены все или большинство критериев сходимости), можно перейти к третьему этапу. Если нет (например, превышено число циклов), оптимизацию следует продолжить. Иногда перед началом или продолжением оптимизации может возникнуть необходимость заново рассчитать гессиан, например, если chk-файл утерян или поврежден. Это можно сделать, указав вместо RCFC ключевое слово CalcFC. Иногда в особо сложных случаях полезно рассчитывать гессиан на каждом шаге, для этого следует указать ключ CalcAll. Chk-файл после завершения второго этапа также полезно сохранить.

Если второй этап завершился успешно, на третьем этапе необходимо провести расчет частот в найденной точке. Для этого полезно использовать сохраненный chk-файл второго этапа и использовать gjf-файл в виде:

```
%chk=ts.chk  
# b3lyp/6-31g(d,p) Freq=NoRaman Geom=AllCheck Guess=Read
```

(конец файла)

Команды Geom=AllCheck и Guess=Read означают, что оптимизированная геометрия и самосогласованная матрица плотности будут прочитаны из chk-файла, что экономит время на достижение ССП. Структуру молекулы и другие данные после команды Geom=AllCheck включать в файл не требуется. NoRaman, как и на первом этапе, экономит время на расчет частот. После завершения расчета следует тщательно проанализировать полученные частоты. Во-первых, структура должна иметь только одну мнимую частоту. Во-вторых, ее величина должна заметно отличаться от нуля, иначе нет гарантии, что это действительно колебание, а не артефакт расчета. В-третьих, форма колебания должна соответствовать изменению структуры в ходе реакции.

Даже после выполнения всех этих условий, часто остаются сомнения, что найденная структура является тем самым переходным состоянием, которое связывает продукты и реагенты изучаемой реакции. Кроме того, реакция может оказаться многостадийной. В этом случае, переходное состояние будет не единственным на пути реакции. Поэтому после каждого успешного поиска переходного состояния, как правило, следует выполнять построение пути реакции.

Установление пути реакции, протекающей через данное переходное состояние, выполняется с помощью процедуры построения внутренней координаты реакции (*IRC, intrinsic reaction coordinate*). IRC-процедура начинается в точке переходного состояния и двигается в направлении понижения энергии в двух направлениях от седловины – в прямом и обратном направлениях реакции до достижения точек минимума, то есть продуктов и реагентов данной реакции. Она выполняется командой

b31yp/6-31g(d,p) IRC(RCFC) Guess=Read Geom=AllCheck

Движение по координате реакции довольно долгий и сложный процесс, однако он очень важен для понимания механизма реакций. Так, если найденные конечные локальные минимумы не являются реагентами или продуктами изучаемой реакции, это возможно означает, что на пути реакции имеются интермедиаты, а сама реакция многостадийная. В этом случае следует продолжить изучение возможных стадий, включив в рассмотрение найденные интермедиаты и переходное состояние. Кроме того, может оказаться, что найдено переходное состояние не изучаемой реакции, а какого-то параллельного процесса. Все эти возможности следует внимательно проанализировать. Делать выводы о механизме реакции следует только после того, как от реагентов до продуктов процедурой IRC проведены непрерывные пути реакции, на котором каждый локальный минимум соединяется с другим через переходное состояние.

В случае успешного построения полного пути реакции и переходных состояний, можно оценить скорость лимитирующей стадии и других стадий реакции. Расчет констант скорости наиболее просто выполнить, используя стандартную теорию переходного состояния (*conventional transition state theory*, *CTST*). Для оценки константы скорости прямой реакции этим методом достаточно знания термодинамических функций реагентов и переходного состояния реакции. Поскольку для переходного состояния расчет частот уже выполнен, достаточно оптимизировать геометрию и рассчитать частоты реагентов (в случае успешного построения IRC геометрия реагентов также уже известна).

Для оценки константы скорости надо заполнить таблицу, аналогичную табл. 1, в которой вместо продуктов реакции (или вместе с ними) будет фигурировать переходное состояние (образец такой таблицы приведен в примере ниже). В этой таблице наиболее важной является рассчитанная величина ΔG^\ddagger , которая позволяет оценить вероятность реакции по формуле:

$$k(T) = \frac{kT}{h} \frac{Q^\ddagger}{Q_A Q_B} \exp\left(-\frac{\Delta H^\ddagger(0)}{RT}\right) = \frac{kT}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right). \quad (9)$$

В этой формуле Q – статсуммы переходного состояния и реагентов А и В (для бимолекулярной реакции), а величина $\Delta H^\ddagger(0) = \Delta(E_{tot} + ZPE)^\ddagger$ – энтальпия переходного состояния относительно реагентов при абсолютном нуле (энергия активации, рассчитанная между нулевыми колебательными уровнями реагентов и переходного состояния). Величина ΔG^\ddagger – изменение функции Гиббса при образовании переходного состояния из реагентов при данной температуре и давлении. Оба варианта формулы эквивалентны, но выражение с ΔG^\ddagger использовать удобнее. Такая формула дает вероятность протекания реакции в смеси с единичными вероятностями встречи реагентов в газовой фазе. Для сопоставления этой величины с измеряемыми значениями констант скорости, следует учесть в этой формуле стандартную концентрацию:

$$k(T) = \frac{kT}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) \left(\frac{P^0}{RT}\right)^{\Delta\nu}. \quad (10)$$

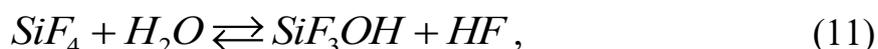
В этой формуле P^0 – это стандартное давление (100 кПа); $\Delta\nu$ – изменение числа молекул при образовании переходного состояния, для бимолекулярной реакции оно равно -1 . Следует помнить, что хотя формула (10) формально подходит для реакций различной молекулярности, стандартная теория переходного состояния изначально разработана как теория бимолекулярных реакций. В случае мономолекулярных реакций механизм активации может быть другим и зависеть, например, от концентрации молекул инертного буферного газа, активирующего молекулу реагента. Поэтому применять вышеприведенные формулы к реакциям иной молекулярности следует с осторожностью. Кроме того, при протекании реакции в конденсированном состоянии эти формулы также будут давать большую погрешность. В случае газозависимых реакций эти формулы обычно позволяют предсказать скорость реакции с точностью 1-2 порядков.

Следует также помнить, что многие реакции диссоциации протекают без образования седловой точки на ППЭ. Профиль ППЭ, соответствующий увеличению длины диссоциирующей связи, представляет собой монотонно возрастающую функцию с выходом на плато в области продуктов диссоциации. Иногда такие реакции называют реакциями с разрыхленным переходным состоянием (*loose TS*). К таким реакциям применять стандартную теорию переходного состояния нельзя, а поиск седловой точки ППЭ описанным выше методом не имеет смысла. В качестве первого приближения можно считать, что энергия активации такой реакции равна энергии диссоциации.

Для более точного описания диссоциативных мономолекулярных реакций, многостадийных процессов с очень низкими энергиями активации, а также тонких зависимостей скоростей химических реакций от концентраций реагентов и буферного газа используются более сложные варианты теории переходного состояния, например, *вариационная теория переходного состояния (variational transition state theory, VTST)* и ее варианты. Расчеты в рамках VTST значительно более сложные, включают большое число входных параметров, которые требуют тщательной настройки и калибровки по экспериментальным данным. Такие расчеты могут быть реализованы, например, с помощью программ *PolyRate* (<https://comp.chem.umn.edu/polyrate/>) или *MultiWell* (<http://clasp-research.engin.umich.edu/multiwell/>).

Пример

Найти переходное состояние бимолекулярной химической реакции



протекающей в газовой фазе при стандартных условиях, и оценить константу скорости этой реакции, используя стандартную теорию переходного состояния.

Решение

Для получения начальной структуры переходного состояния оптимизируем вначале предреакционный комплекс между реагентами – $\text{SiF}_4 \cdot \text{H}_2\text{O}$. Выше мы уже нашли структуру этого комплекса, воспользуемся этим результатом. В оптимизированной ранее структуре предреакционного комплекса с помощью любой программы-визуализатора (ChemCraft, Avogadro, GaussView, MolTran) искажим связи так, как показано на рис. 17.

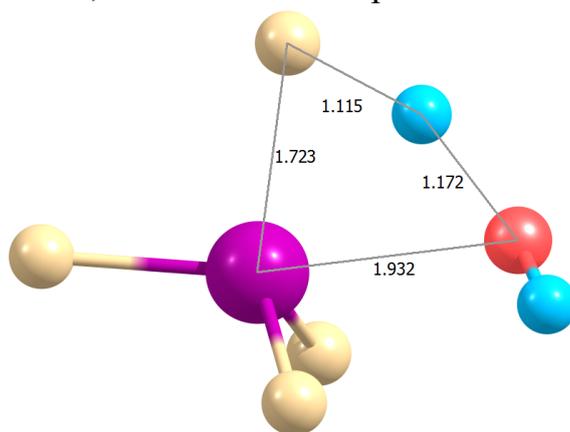


Рис. 17. Искаженная структура предреакционного комплекса $\text{SiF}_4 \cdot \text{H}_2\text{O}$ для поиска переходного состояния реакции (11)

Для полученной начальной структуры проводим расчет колебательных частот, используя команды

```
%chk=sif4h2o-tsl.chk  
# b3lyp/6-31g(d,p) Freq=NoRaman
```

После завершения расчета в файле находим, что структура имеет две мнимые частоты: $381i \text{ см}^{-1}$ и $202i \text{ см}^{-1}$ (записанные в файле как -318 и -202). Анализ формы нормальных колебаний, соответствующих этим частотам показывает, что первая имеет форму, соответствующую движениям атомов в реакции (11), то есть одновременное сближение атомов Si и O, удаление F и Si и перенос H от O к F. В то же время второе мнимое колебание не соответствует изучаемой реакции: оно представляет собой крутильное колебание группы OH вокруг направления Si...O, то есть не приводит к образованию или разрыву связи. Поскольку колебание с низшей частотой имеет правильную форму и почти в два раза отличается по частоте от второго колебания, есть надежда, что движение по этому колебательному вектору приведет к переходному состоянию. Если этого не произойдет, следует вернуться к выбору начальной структуры и, изменяя геометрию, добиться устранения второй мнимой частоты.

Полезно после каждого этапа поиска переходного состояния делать копии chk-файла. Иногда случайные неправильные действия пользователя могут

привести к запуску расчета с неправильными командами и chk-файл будет испорчен. Чтобы не повторять все расчеты заново, полезно перед началом расчета сохранять копию chk-файла предыдущего этапа.

Используя chk-файл `sif4h2o-ts1.chk`, запускаем оптимизацию переходного состояния с той же самой структурой, для которой рассчитывались частоты, используя строку команд:

```
%chk=sif4h2o-ts1.chk  
# b3lyp/6-31G(d,p) FOpt=(TS,NoEigenTest,RCFC,MaxCycle=300)
```

Оптимизация завершается за 25 циклов:

	Item	Value	Threshold	Converged?
	Maximum Force	0.000049	0.000450	YES
	RMS Force	0.000024	0.000300	YES
	Maximum Displacement	0.000384	0.001800	YES
	RMS Displacement	0.000161	0.001200	YES
	Predicted change in Energy=-4.425608D-08			
	Optimization completed.			
	-- Stationary point found.			

Визуализируя оптимизированную структуру, обнаруживаем, что она заметно отличается от начальной (рис. 18).

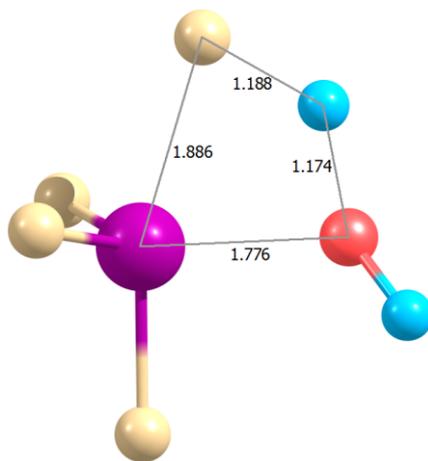


Рис. 18. Оптимизированная структура переходного состояния реакции (11)

Используя тот же chk-файл `sif4h2o-ts1.chk`, запускаем расчет частот в найденной структуре:

```
%chk=sif4h2o-ts2.chk  
# b3lyp/6-31G(d,p) Freq Geom=AllCheck Guess=Read
```

Анализируя выходной файл, обнаруживаем единственную мнимую частоту $811i \text{ см}^{-1}$. Это говорит о том, что найденная структура является переходным состоянием. Анимация формы колебания с мнимой частотой показывает, что оно имеет форму, соответствующую движениям атомов в реакции (11), то есть одновременное сближение атомов Si и O, удаление F и Si и

перенос H от O к F. Тем не менее, полезно убедиться, что механизм реакции не является многостадийным, и найдено именно то переходное состояние, которое соединяет на ППЭ реагенты и продукты реакции (11). Для этого выполним построение пути реакции с помощью процедуры IRC.

Используя тот же chk-файл `sif4h2o-ts1.chk` (необходимо сделать его копию), запускаем процедуру IRC в разных направлениях реакции по отдельности. В одном файле запускаем IRC в направлении прямой реакции (ключевое слово `Forward`), в другом файле – в направлении обратной (ключевое слово `Reverse`):

```
# b3lyp/6-31G(d,p) IRC(RCFC,Forward,MaxPoints=100)
# b3lyp/6-31G(d,p) IRC(RCFC,Reverse,MaxPoints=100)
```

Команда `MaxPoints=100` означает, что программа должна сделать не более 100 шагов в направлении минимума. Если минимум будет достигнут раньше, программа завершит работу.

По окончании расчета программа в выходной файл будут выданы координаты точек пути реакции и соответствующие энергии. Структуры конечных точек показаны на рис. 19. Видно, что конечное состояние напоминает комплекс продуктов, объединяемых межмолекулярными силами (*постреакционный комплекс*).

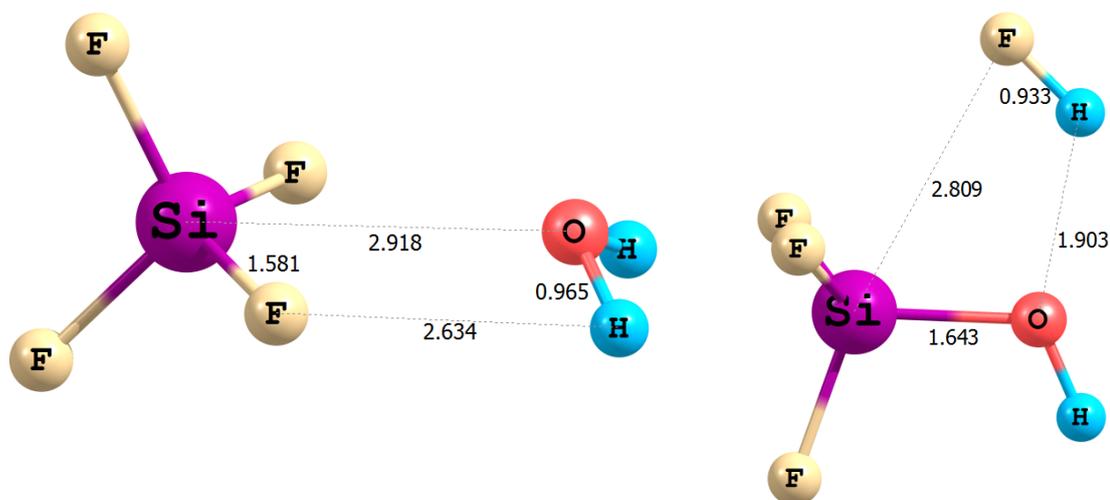


Рис. 19. Конечные структуры процедуры IRC для реакции (11): слева – предреакционный комплекс, справа – постреакционный комплекс

Энергию вдоль пути реакций удобно изобразить на графике после обработки файлов IRC программой *IRCtracer*. Эта программа выбирает из файлов энергии и структуры, а также рассчитывает значения выбранных пользователем структурных параметров (длины связи, валентные и диэдральные углы), что позволяет построить трехмерные графики пути реакции (рис. 20).

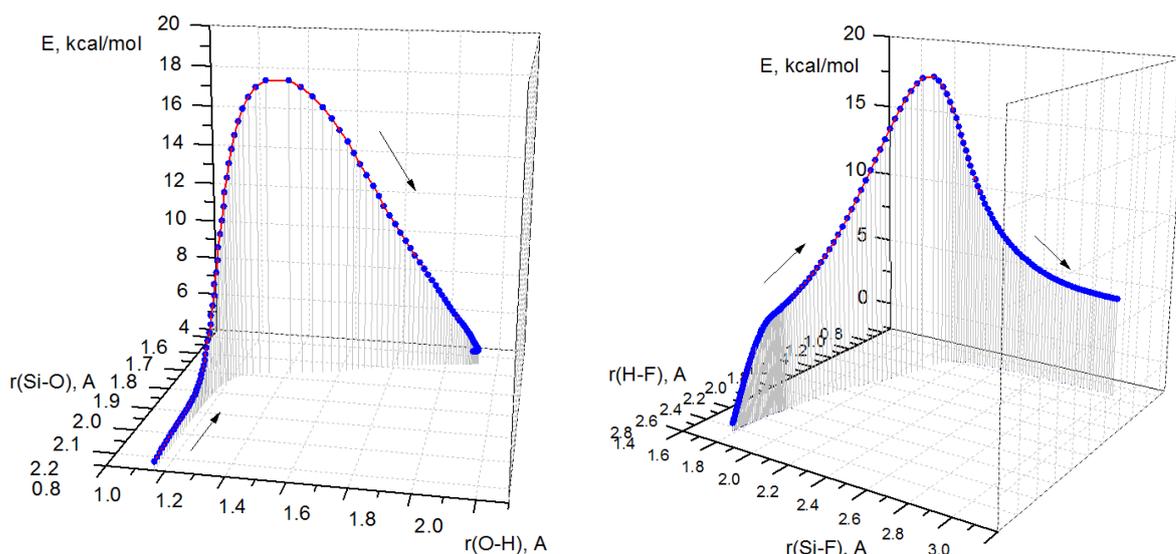


Рис. 20. Энергетические профили пути реакции (11) в зависимости от изменения различных длин связей. По вертикальной оси отложена энергия относительно изолированных реагентов. Стрелки показывают направление реакции

Анализируя путь реакции и конечные структуры процедуры IRC, мы делаем вывод, что реакция одностадийная бимолекулярная и не имеет на своем пути каких-либо интермедиатов, кроме слабосвязанных предреакционного и постреакционного комплексов. Это означает, что найденное переходное состояние позволяет оценить константу скорости реакции. Для этой оценки из файла с расчетом частот оптимизированного переходного состояния выбираем его термодинамические функции аналогично тому, как мы это делали при расчете термодинамики реакции. Результаты заносим в табл. 3. В отдельной колонке проводим расчет константы скорости по формуле (9).

Таблица 3

Расчет константы скорости реакции

B3LYP/6-31G(d,p)	SiF4	H2O	= TS	Ea, a.u.	kcal/mol	kJ/mol
Nneg	0	0	1			
Sigma	12	2	1			
Etot	-689.1085625	-76.41973658	-765.499149	0.02915	18.29	76.5
Etot+ZPE	-689.0958637	-76.39837123	-765.464633	0.029601	18.57	77.7
Etot+U	-689.0909072	-76.39553616	-765.458282	0.028161	17.67	73.9
Etot+H	-689.089963	-76.39459195	-765.457338	0.027217	17.08	71.5
Etot+G	-689.1222175	-76.41602892	-765.495138	0.043108	27.05	113.2
File	<i>sif4.out</i>	<i>h2o.out</i>	<i>sif4h2o-ts3.out</i>			
				k(T)		
				k(298) по формуле (9)	9.2E-08	
				k(298) по формуле (10)	3.7E-12	cm ³ /molec/s

3.6. Учет влияния среды на физико-химические свойства и реакционную способность

До сих пор неявно предполагалось, что рассчитываемые молекулы находятся в вакууме или образуют идеальный газ. Однако в большинстве случаев на практике оцениваются свойства веществ в растворах, в конденсированном состоянии или в состоянии реального газа. Поэтому при квантовохимическом моделировании часто необходимо учитывать влияние конденсированной фазы на физико-химические свойства молекул-участников, термодинамические характеристики реакций или их кинетические константы.

Наиболее простым и часто применяемым методом такого учета является метод самосогласованного реакционного поля (SCRF), который включает несколько моделей влияния среды. Наиболее простая из них – модель Онзагера. В этой модели влияние среды (растворителя или конденсированной фазы) сводится к электростатическому взаимодействию молекулы, помещенной в воображаемую сферическую полость внутри конденсированной фазы с поляризованной средой на границах полости. На этих границах под влиянием атомных зарядов молекулы возникают связанные заряды, которые создают электрическое поле внутри полости. Это поле создает дополнительный потенциал, который учитывается при квантовохимическом расчете, и приводит к перераспределению атомных зарядов в молекуле. Это, в свою очередь, изменяет поляризацию среды вокруг полости и такие итерации повторяются до самосогласования. Параметрами метода являются размер полости и диэлектрическая проницаемость среды. Зависимость от размера полости создает определенный произвол и делает модель Онзагера неудобной для расчетов. Гораздо более точной и строгой является модель поляризуемого континуума (*polarizable continuum model, PCM*), а также ее разновидности IPCM и SCIPCM. В модели PCM полость выбирается путем описывания вокруг атомов сферических полостей с радиусами равными ван-дер-ваальсовым радиусам атомов. Получаемая полость сложного размера значительно правильнее описывает контакты окружающей среды и молекулы. В модели IPCM полость выбирается на основе изоповерхности постоянной плотности вокруг молекулы, а в модели SCIPCM ее конфигурация находится итерационным методом и варьируется одновременно с поляризацией среды. Кроме того, разработан также специализированный метод для учета влияния электролитов CPCM. Несмотря на большую гибкость, модели IPCM и SCIPCM не дают однозначного преимущества перед PCM, и последний является сегодня одним из наиболее часто используемых вариантов SCRF для неэлектролитов. Для его использования в программе Gaussian достаточно задать команду SCRF и указать модель и растворитель, в котором находится молекула растворенного вещества, например:

```
# b3lyp/6-31G(d,p) FOpt Freq SCRF=(PCM,Solvent=Benzene)
```

Вместо РСМ могут быть указаны ІРСМ, SCІРСМ, СРСМ или Dipole (для выбора модели Онзагера). Программа хранит информацию о большом количестве растворителей. Неизвестный программе растворитель также можно описать дополнительно с помощью специальных ключевых слов.

Влияние среды на структуру молекулы и ее полную энергию обычно не очень велико. Однако учет влияния среды на энергии и функции Гиббса химических реакций в растворе может быть весьма заметно, особенно если энергия реакции невысока, а продукты и реагенты сильно различаются своим дипольным моментом. Еще более заметно влияние растворителя на возбужденные состояния. В этом случае полосы поглощения и флуоресценции в электронном спектре могут сдвигаться на десятки нанометров.

3.7. Визуализация механизмов реакций, изученных квантовохимическими методами

Механизмы многих реакций представляют собой сложные многостадийные и разветвленные процессы. Квантовохимическое исследование таких реакций позволяет найти все интермедиаты, переходные состояния и соединяющие их реакционные пути на ППЭ. Кроме того, квантовохимический расчет дает энергии, другие термодинамические параметры этих точек ППЭ, высоты активационных барьеров и соотношение энергий элементарных реакций. В результате у исследователя появляется большой массив данных о структуре ППЭ и протекающих элементарных реакциях, который часто очень трудно осмыслить. Для этого очень полезно визуализировать полученные фрагменты ППЭ и представлять реакцию в виде графических схем. На рис. 21 показана схема ППЭ процесса миграции атома водорода на поверхности кластера платины Pt_{24} , который играет роль модели реальной наночастицы платины, участвующей во многих каталитических реакциях. Этот процесс начинается с адсорбции молекулярного водорода на поверхности кластера. Молекулярный водород диссоциирует на атомы, и каждый атом мигрирует по поверхности наночастицы, попадая в локальные минимумы и преодолевая активационные барьеры между этими минимумами. Поскольку число локальных минимумов разного типа на поверхности достаточно велико, возникает большое количество возможных путей такого процесса, которые трудно изобразить в виде обычных химических уравнений, но которые легко отразить в виде схемы ППЭ.

Схема на рис. 21 построена в графическом редакторе с помощью программ `edx2edg` и скрипта `EdiagPlt`. Первая программа использует полные энергии локальных минимумов и переходных состояний, рассчитанные квантовохимической программой, и записывает их в специальном формате, который читается скриптом `EdiagPlt`, выполняющим построение изображения. Полученное изображение может быть отредактировано, дополнено или скомбинировано с другими изображениями.

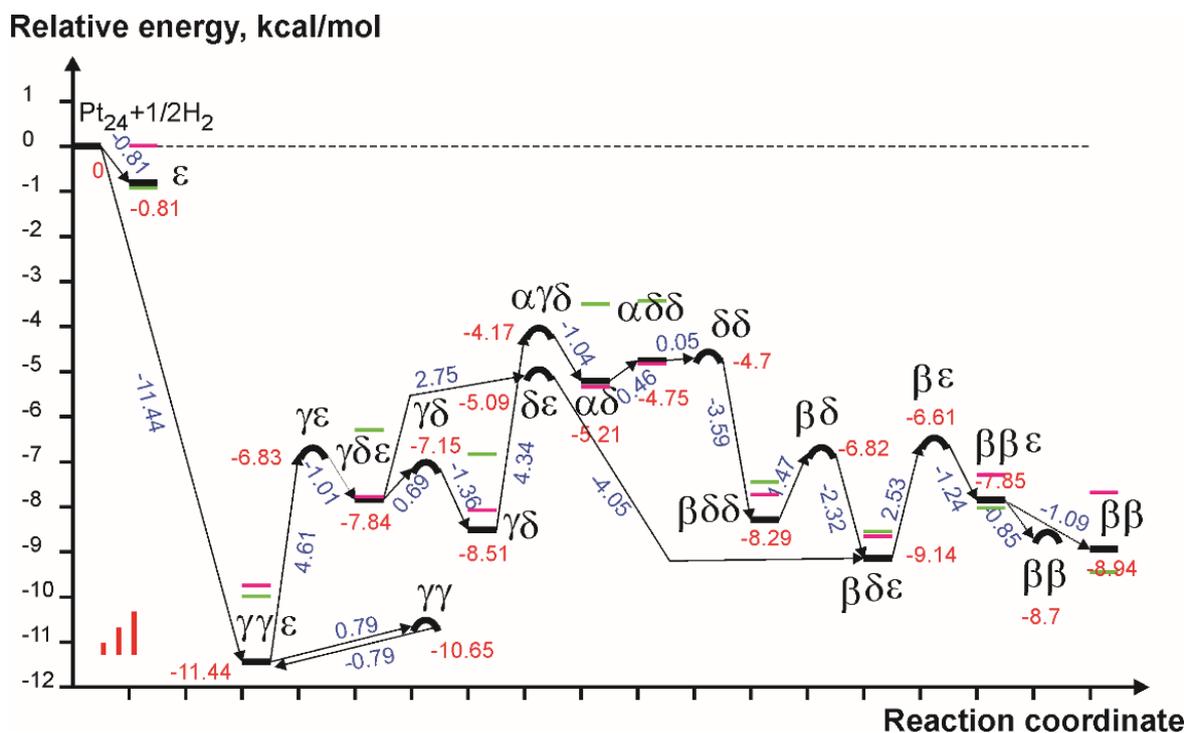


Рис. 21. Схема ППЭ поверхностной миграции атомов Н на поверхности кластера Pt_{24} . Расчет DFT с функционалом BLYP с базисом 6-31G(p) для атомов Н и с псевдопотенциальным базисом CRENBS для атомов Pt. Горизонтальные площадки – локальные минимумы, дуги – переходные состояния. Значения у площадок (красные) – энергии стационарных точек ППЭ относительно исходной системы $Pt_{24}+1/2H_2$ (в ккал/моль), значения у стрелок (синие) – разность энергий между стационарными точками, то есть энергии реакций или энергии активации. Греческими буквами обозначены стационарные точки различной структуры. Столбики в левом нижнем углу – величины тепловой энергии при 100, 300, 500 К. Видно, что этой энергии не хватает для преодоления большинства активационных барьеров, и движение атомов водорода по поверхности нельзя считать свободным – они испытывают редкие скачки между отдельными устойчивыми состояниями.

3.8. Изучение реакций методами молекулярной динамики с квантовохимическими потенциалами

Изучение механизмов многостадийных разветвленных реакций с невысокими энергетическими барьерами часто затрудняется тем, что на их ППЭ сложно найти многочисленные переходные состояния, а стандартная теория переходного состояния (CTST), описанная в разделе 3.5, дает только грубые оценки констант скорости. Уравнения CTST основаны на представлении реакции в виде бимолекулярного столкновения с четко выраженным энергетическим барьером, и используют ряд других предположений о распределении энергии по степеням свободы. Эти предположения далеко не всегда соответствуют реальному поведению молекул в ходе реакции на ППЭ с неглубокими минимумами, когда быстрое изменение геометрии молекул приводит к иным энергетическим распределениям. Кроме того, в диссоциативных и ассоциативных реакциях седловая точка ППЭ может вообще отсутствовать. Как было указано выше, в этом случае могут быть применены более сложные варианты теории переходного состояния, например VTST. Однако эти теории обычно требуют задания большого числа дополнительных параметров, которые часто неизвестны исследователю, требуют тщательной калибровки, а при неточном задании не гарантируют правильности решения. Одним из способов прямого описания таких процессов, который позволяет избежать описанных трудностей, является метод *молекулярной динамики*.

Молекулярная динамика (МД) – это подход к описанию атомно-молекулярных систем и процессов, в котором временная эволюция системы изучается путем отслеживания классических траекторий движения взаимодействующих между собой частиц (атомов или молекул). Подразумевается, что в любой момент времени нам известны механические силы взаимодействия между частицами системы. Если потенциальная энергия системы в момент времени t равна $U(t)$, то сила, действующая в этот момент на частицу i с координатами (x_i, y_i, z_i) :

$$\mathbf{F}_i(t) = -\nabla U(t) = \left(-\frac{\partial U(t)}{\partial x_i}, -\frac{\partial U(t)}{\partial y_i}, -\frac{\partial U(t)}{\partial z_i} \right). \quad (12)$$

Зная силы, действующие на частицу, а также первоначальные скорости и положения частиц, можно определить их скорости и положения в момент времени $t+\Delta t$, например, по формулам ньютоновской механики:

$$\begin{aligned} \mathbf{V}_i(t + \Delta t) &= \mathbf{V}_i(t) + (\mathbf{F}_i(t) / m) \Delta t \\ \mathbf{r}_i(t + \Delta t) &= \mathbf{r}_i(t) + \mathbf{V}_i(t) \Delta t + \frac{\mathbf{F}_i(t)(\Delta t)^2}{2m_i}. \end{aligned}$$

Хотя на практике используются несколько видоизмененные выражения, лучше подходящие для определения термодинамических характеристик, общая стратегия остается той же – исходя из начальных положений частиц системы, надо получить классическую *траекторию движения* всех частиц, т.е. последовательность координат и скоростей $\{\mathbf{r}_i, \mathbf{V}_i\}$ в моменты времени $i\Delta t$,

$i=0,1,2,\dots,n$. Величина Δt называется *шагом интегрирования (шагом по времени)*, чем она меньше, тем точнее траектория описывает реальное движение частиц. Знание траектории позволяет, в принципе, определить любое термодинамическое свойство A системы, если оно может быть представлено в виде зависимости от координат и скорости частиц. Это определение основано на усреднении исследуемого свойства по точкам траектории (т.н. *усреднение по времени*):

$$\langle A \rangle_t = \frac{1}{n} \sum_i A(\mathbf{r}_i, \mathbf{V}_i)$$

С другой стороны, можно интересоваться не средним по времени, а вероятностью какого-либо события, происходящего в момент времени t_i (например, произошла ли реакция или нет за время T). Тогда следует рассчитать m различных траекторий (*ансамбль траекторий*) и усреднить характеристики событий в точках t_i разных представителей ансамбля a :

$$\langle A \rangle_a = \frac{1}{m} \sum_a A_a(t_i)$$

Такое усреднение называется *усреднением по ансамблю*. В статистической физике доказывается, что при бесконечной длине траектории и бесконечном количестве копий системы в равновесном термодинамическом ансамбле средние по времени и средние по ансамблю равны. К сожалению, на практике такое равенство достигается только при длительном моделировании, причем для расчета термодинамических характеристик следует прилагать специальные усилия, добываясь термодинамической равновесности системы.

Результаты, получаемые при усреднении точек траектории, сильно зависят от условий обмена энергией и веществом между системой и окружающей средой. Если система изолирована от внешней среды, т.е. не обменивается с ней энергией и веществом, то в ней сохраняется число частиц N , объем V и энергия E . Ансамбль, построенный из бесконечного числа копий таких систем, называется *NVE-ансамблем*. Если система обменивается тепловой энергией с окружающей средой (телом большой теплоемкости с температурой T), образующийся ансамбль называется *NVT-ансамблем* (сохраняется число частиц N , объем V и температура T). Если система обменивается с окружающей средой и тепловой и механической энергией (работой расширения $P\Delta V$), ансамбль называется *NPT-ансамблем* (сохраняется число частиц N , давление P и температура T). Эти три типа систем наиболее часто рассматриваются при моделировании методом МД. Следует отметить, что большим преимуществом метода МД является то, что можно моделировать не только равновесные системы, описываемые термодинамическими характеристиками (например, энергию Гиббса растворения или ассоциации молекул при постоянных T , V или P), но и быстрые неравновесные процессы (например, возникновение и распространение детонационных волн в веществе при быстрой реакции распада).

Потенциальная энергия системы U , необходимая для расчета сил, часто представляется в виде суммы парных потенциалов U_{ij} , описывающих

взаимодействия между парами частиц, или более сложных функций, учитывающих взаимодействия нескольких атомов. В простых случаях потенциалы, описывающие взаимодействия частицы i с другими частицами системы, могут быть заданы в виде замкнутых формул (например, закона Кулона, эмпирического потенциала Леннард-Джонса или потенциала гармонического осциллятора), а атомы или молекулы рассматриваются как классические материальные точки с соответствующими массами. Такой подход называется *классической молекулярной динамикой с эмпирическими потенциалами*, он исключительно широко применяется в настоящее время для описания супрамолекулярных и биологических систем. Его ограничением, однако, является то, что, во-первых, требуется правильное описание потенциала взаимодействия между атомами, связанными в различные химические группировки, как внутри молекулы, так и между молекулами различных типов. Задание таких потенциалов весьма нетривиальная задача, которая требует тщательной калибровки большого числа потенциальных параметров, объединяемых в т.н. *силовые поля* – наборы потенциальных параметров для определенной группы молекул. Во-вторых, такой подход в большинстве случаев исключает изучение химических реакций между частицами, поскольку в ходе реакции потенциальные функции сами должны изменяться (например, из-за того, что атомы изменяют свои заряды в ходе реакции и потенциалы их электростатических взаимодействий изменяются сложным образом). Хотя в настоящее время достигнуты определенные успехи в разработке т.н. *реакционных силовых полей* (например, силовое поле ReaxFF, потенциалы AIREBO, COMB), их применение до сих пор не позволяет полностью решить задачу описания химических реакций на основе классической МД.

Одним из способов описания сложных химических процессов является метод *неэмпирической молекулярной динамики (ab initio molecular dynamics, AIMD)*, или, точнее, *классической молекулярной динамики с квантовохимическими потенциалами*. Этот подход основан на том, что вместо задания межатомных потенциалов в виде известных функций, их расчет производится квантовохимическим методом в каждой новой точке МД траектории, «на лету». При этом уровень квантовохимической теории может варьироваться – от полуэмпирических методов для расчета систем из десятков и сотен атомов до высокоточных теорий типа CCSD для прецизионного расчета динамики малых молекул. В большинстве случаев различные варианты DFT в базисах среднего размера являются предпочтительным выбором, обеспечивая хорошую точность расчета при возможности моделирования систем из десятков атомов.

Если потенциалом в формуле (12) выступает полная (в квантовохимическом смысле) энергия молекулы $E_{tot}(\mathbf{r}_i)$, то такой вариант AIMD обычно называется *борн-оппенгеймеровской молекулярной динамикой БОМД (Born-Oppenheimer molecular dynamics, BOMD)*. Выбор названия обусловлен тем, что траекторией движения системы является кривая на поверхности потенциальной энергии (ППЭ). Возможность описания системы с помощью ППЭ основано на приближении Борна-Оппенгеймера: каждой точке траектории

соответствует решение электронного уравнения Шредингера, описывающего движение электронов в поле ядер с фиксированными координатами. Большим преимуществом такого подхода является то, что не требуется какая-либо сложная калибровка потенциальных параметров, достаточно указать только квантовохимический метод, который будет использован для расчета энергии. При этом возможно описание систем, в которых проходят химические реакции, в том числе сложные разветвленные и многостадийные процессы с неглубокими минимумами, нахождение переходных состояний в которых затруднительно. Недостатком метода BOMD, как и AIMD в целом, является значительно более высокие вычислительные затраты по сравнению с классической МД, поскольку теперь расчет сил производится не по простым формулам типа закона Кулона, а требует полноценного квантовохимического расчета в каждой точке.

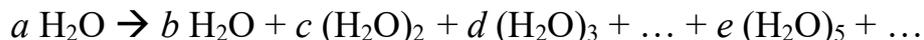
Метод AIMD может быть значительно ускорен, если вместо BOMD применять более изощренные подходы для квантовохимической оценки потенциальной функции. Одним из них является метод *распространения атом-центрированной матрицы плотности* (*atom-centered density matrix propagation, ADMP*). Этот метод является развитием идеи, впервые реализованной в т.н. *методе Карра-Парринелло* (*Carr-Parrinello molecular dynamics, CPMD*). В этом методе для устранения необходимости самосогласования молекулярной волновой функции в каждой точке траектории и ускорения расчета рассматривается динамика не только ядерных координат, но и самих одноэлектронных волновых функций. Получающиеся уравнения динамики включают не только координаты и скорости частиц, но и уравнения, описывающие изменение во времени коэффициентов волновой функции. Хотя теоретическая строгость такого подхода не до конца ясна, долгий опыт применения подобных методов показывает, что во многих случаях они дают весьма разумные с химической точки зрения и достаточно точные результаты. Например, они позволяют описывать перераспределение электронной плотности при протекании реакций с разрывом и образованием химических связей, причем результаты хорошо согласуются с экспериментальными данными.

Методы BOMD и ADMP обычно не требуют задания каких-либо сложных физических параметров. Для начала расчета достаточно указать только уровень квантовохимической теории (метод и базис), шаг по времени, длину траектории (количество точек на ней), начальные координаты атомов и начальную энергию системы. В то же время расчет ADMP значительно быстрее BOMD, при хорошем согласии результатов их решений. Следует четко понимать, что под согласием результатов понимается сходство величин, усредненных по траекториям или ансамблям, хотя сами траектории могут значительно различаться между методами BOMD и ADMP или даже при различных запусках в рамках одного и того же метода.

Методы AIMD (в вариантах BOMD) реализованы в программах Gaussian, NWchem, VASP, QuantumEspresso. Метод CPMD реализован в программах CPMD, CP2K. Метод ADMP реализован в программах Gaussian09, Gaussian16.

Рассмотрим расчет методом ADMP в программе Gaussian09 динамики молекулярных кластеров воды, в частности, пентамера (H₂O)₅. Кластеры воды

представляют собой молекулярные комплексы из n молекул воды в газовой фазе, которые образуются молекулами H_2O за счет водородных связей между отдельными мономерами:



Коэффициенты a , b , c , d , e в этом уравнении, определяющие равновесные концентрации кластеров, зависят от температуры и парциального давления воды в системе. При высоких температурах и невысоких парциальных давлениях воды молекулы кластеров неустойчивы, время их жизни невелико, и они присутствуют в равновесной смеси в очень небольших концентрациях. Однако по мере снижения температуры и приближении системы к состоянию конденсации, доля кластеров $(\text{H}_2\text{O})_n$ с большим n увеличивается. В конечном итоге большие кластеры становятся ядрами нуклеации (зародышами каплеобразования), притягивая к себе все больше молекул воды, после чего конденсация становится необратимой. Отдельные кластеры $(\text{H}_2\text{O})_n$, в которых молекулы мономеров связаны водородными связями, являются нежесткими молекулами, их геометрия постоянно флюктуирует. ППЭ такой системы является исключительно сложной, содержит большое количество неглубоких минимумов и невысоких переходных состояний. Исчерпывающее изучение изомеризации кластеров методом поиска переходных состояний практически невозможно. В то же время изучение процессов образования и временной эволюции кластеров важно для атмосферной химии, поскольку присутствие кластеров сильно влияет на поглощение и баланс энергии в атмосфере Земли (в т.ч. оказывая влияния на парниковый эффект и изменение климата). Кроме того, кластеры являются своего рода «газофазными катализаторами», обеспечивая высокую скорость протекания реакций гидролиза и гидратации различных атмосферных примесей, которые очень медленно реагируют с мономерами воды.

Для генерации молекулярно-динамической траектории пентамера воды методом ADMP в программе Gaussian можно использовать файл `w5-admp.gjf`, содержащий следующую информацию:

```
%NProc=16
%Mem=24000MB
%rwf=w5-admp.rwf
%chk=w5-admp.chk
# x3lyp/6-31+G(d,p) ADMP=(MaxPoints=100000,StepSize=2000,NKE=19935)
```

Water pentamer - optimized geo

```
0 1
O1      2.340725   -0.964001   -1.068642
O2     -2.340709    0.967245   -1.065719
O3      2.492047    1.020858    0.905580
O4     -2.492062   -1.023603    0.902471
O5      0.000012   -0.000160    0.116275
H6      2.767059   -1.818344   -0.974087
H7      2.743838   -0.375476   -0.406585
```

H8	-2.767050	1.821301	-0.968631
H9	-2.743846	0.376745	-0.405440
H10	1.530434	0.873536	0.841910
H11	2.709245	1.006436	1.839967
H12	-1.530448	-0.876100	0.839324
H13	-2.709313	-1.012211	1.836887
H14	0.601255	-0.504051	-0.462306
H15	-0.601271	0.505449	-0.460764

Ключевые слова `%Nproc` и `%Mem` определяют количество процессорных ядер и памяти, выделяемых для расчета, их значения не должны превышать возможностей используемого компьютера. Команды `%rwf` и `%chk` определяют возможность рестарта в случае аварийного прерывания расчета, их задание важно, поскольку расчет длительный. Командная строка описывает квантовохимический метод расчета энергии. В данном случае используется метод DFT с функционалом X3LYP, который хорошо себя зарекомендовал при расчетах водородно-связанных комплексов и кластеров воды. При выборе базисного набора следует учесть, что водородные связи – это межмолекулярные взаимодействия, и для правильного их описания следует обязательно использовать базис с диффузными функциями (+ или ++). Также очень полезно использовать наборы поляризованных функций на всех атомах (d,p) или (2d,2p). При наличии вычислительных возможностей в качестве варианта выбора можно было бы порекомендовать базис 6-311++G(2d,2p), обеспечивающий хорошую точность при описании энергии взаимной координации молекул воды. Однако следует помнить, что МД траектория будет включать тысячи и десятки тысяч точек, в каждой из которых проводится квантовохимический расчет. Поэтому даже небольшое расширение базиса будет приводить к резкому увеличению времени расчета. Кроме того, расширение базиса может приводить к численной неустойчивости решений и обрыву расчета. В связи с этим в нашем примере мы используем «минимально приемлемый» для такого рода расчетов базис 6-31+G(d,p). Даже в этом случае размер выходного файла `w5-admp.log` составляет около 1 Гб, а время расчета около двух суток.

Команда `ADMP=(MaxPoints=100000,StepSize=2000,NKE=19935)` запрашивает расчет ADMP-траектории длиной (`MaxPoints`) 100000 точек с шагом по времени 2000 единиц (`StepSize`). Каждая единица времени в программе Gaussian равна 10^{-4} фемтосекунды, т.е. запрашиваемый шаг равен $0.2 \text{ фс} = 0.2 \cdot 10^{-15} \text{ с}$. Это достаточно малый шаг, и в последующем можно попробовать его увеличить с целью сокращения длительности расчета. При заданном шаге длительность траектории составляет 100000 шагов $\times 0.2 \text{ фс} = 20 \text{ пс}$. Это небольшой интервал времени по сравнению с временами, которые сейчас доступны для изучения методами МД с эмпирическими потенциалами - в настоящее время длительность траектории при изучении белковых молекул обычно составляет сотни наносекунд, а в рекордных случаях до миллисекунды. Однако в методе ADMP расчет каждой точки траектории производится, по сути, методом DFT без какой-либо эмпирической калибровки, что делает невозможным расчеты таких длительных траекторий. Если увеличить шаг

интегрирования в два раза, то длительность траектории увеличится до 40 пс. С другой стороны, увеличив шаг, можно сохранить общую длительность (20 пс), но уменьшить число точек траектории и сократить время расчета. Однако увеличение шага интегрирования следует производить очень осторожно. Во-первых, это сразу же сказывается на точности траектории. Например, ряд особенностей ППЭ (неглубокие минимумы или седловины) система может «проскакать», не замечая их, что не может не сказаться на результатах усреднения. Во-вторых, очень часто даже небольшое увеличение шага приводит к проблемам сходимости при расчете волновой функции, и расчет может обрываться без каких-либо внешних причин. Наконец, может возникнуть ситуация, когда большой шаг приводит к тому, что атом наталкивается на другой атом, что вызывает огромный скачок энергии и также приводит к обрыву расчета.

При выборе шага интегрирования и длительности траектории полезно оценить характерные периоды молекулярных колебаний в исследуемой системе. Колебания жестких связей тяжелых атомов в органических молекулах обычно характеризуются волновыми числами порядка 500-2000 см⁻¹, что соответствует периодам колебаний 15-60 фс. Колебания связей С-Н и О-Н порядка 2900-3700 см⁻¹ соответствует периоду около 10 фс. Колебания водородных и межмолекулярных связей значительно медленнее и могут составлять 100-500 фс и более. Очевидно, что при изучении процессов между несколькими молекулами траектории должны охватывать периоды как минимум 10-100 межмолекулярных колебаний. Чтобы убедиться в выполнении этого условия, достаточно провести оптимизацию геометрии исследуемой системы в какой-то характерной конформации и рассчитать колебательные частоты, как описано в разделе 3.4. Низшие действительные частоты ν (частоты «мягких» колебаний), взятые из выходного файла, будут определять характерный период «мягкого» колебания по формуле:

$$\tau[\text{фс}] = \frac{33356}{\nu[\text{см}^{-1}]}$$

Соответственно, длительность траектории, которую необходимо сгенерировать, должна составлять не менее 10τ . Шаг интегрирования траектории, очевидно, должен быть как минимум в 10 раз меньше периода колебаний максимальной частоты. Таким образом, для системы с О-Н связями, он, как правило, не должен превышать 1 фс.

Ключевое слово `NKE=<целое_число>` в командной строке задает величину начальной кинетической энергии ядер атомов системы в единицах микроХартри ($1 \text{ мкХ} = 10^{-6} \text{ а.е.}$, см. «Приложение» для перевода в другие единицы измерения). Выбор этой величины определяется требуемой температурой моделируемой системы T , в расчете, что каждый атом имеет три поступательных степени свободы. На каждую из них приходится энергия теплового движения $kT/2$ (k – постоянная Больцмана), причем центр масс системы фиксирован, что уменьшает число степеней свободы на 3. Таким образом, величина NKE (в микроХартри)

для системы из N атомов при температуре T (выраженной в К) задается выражением

$$NKE[\mu H] = 3(N-1)\frac{kT}{2} = 4.74665(N-1)T[K]. \quad (13)$$

При моделировании NVE -ансамбля (что подразумевается по умолчанию) в системе сохраняется полная механическая энергия, т.е. сумма кинетической и потенциальной энергии, в то время как сами эти компоненты взаимно переходят друг в друга. Следует учитывать, что в полную механическую энергию дает вклад начальная потенциальная энергия рассматриваемой молекулы. Если структура молекулы изначально соответствует минимуму энергии, ее начальная потенциальная энергия равна нулю. Однако, если какие-то связи молекулы растянуты, начальная потенциальная энергия может быть весьма велика, и эта энергия будет давать вклад в полную механическую энергию. Этот факт часто используют для моделирования динамики колебательно-возбужденных молекул, например молекул, одно или несколько колебаний которых возбуждены с помощью короткого лазерного импульса ИК диапазона. В этом случае в качестве начальной структуры используется молекула, координаты атомов которой являются суммой координат минимума ППЭ и вектора нормального колебания, возбужденного лазером. Амплитуда этого вектора определяет энергию, переданную молекуле лазерным импульсом.

Следует также учесть, что если начальная система соответствовала минимуму энергии (оптимизированная геометрия), то при движении атомов их средние положения будут отличаться от равновесных и часть начальной кинетической энергии NKE перейдет в потенциальную энергии растянутых связей. Вследствие этого средняя энергия тепловых колебаний (и, следовательно, средняя температура системы) будет примерно в два раза ниже, чем следует из формулы (13).

После командной строки и обычной строки комментария в файле данных приводится описание начальной геометрии системы, которая может быть задана любым способом, описанным в разделе 2, в данном случае используется $NXYZ$ матрица декартовых координат. В описываемом примере эти координаты были получены оптимизацией геометрии одного из изомеров $(H_2O)_5$ на том же уровне теории ($X3LYP/6-31+G(d)$), структура этого изомера показана на рис. 22. Можно использовать и другие способы описания начальной структуры, однако рекомендуется избегать слишком грубого задания геометрии с сильно удлиненными или сильно укороченными валентными связями. Это может привести к очень большой начальной энергии системы, что приведет к большим скоростям атомов, неустойчивости траектории и аварийному останову. Если в файле не указаны другие параметры, начальные скорости атомов генерируются программой автоматически, исходя из случайного гауссова распределения проекций скоростей ядер при температуре, соответствующей заданной величине NKE . Начальные скорости атомов можно также задать в явном виде, используя в команде `ADMP` ключевые слова `ReadVelocity` или `ReadMWvelocity`. В первом случае пользователем задаются декартовы скорости атомов в единицах Бор/с, во

втором – масс-взвешенные скорости $V_{mi} = m_i^{1/2}V_i$ в единицах аем^{1/2}Бор/с. Для задания скоростей во входном файле после всех координат атомов (15 в данном примере) следует добавить пустую строку и еще 15 строк, каждая из которых содержит три проекции начальной скорости соответствующего атома $V_x V_y V_z$ в научном формате (когда D или E дескрипторы означают степень десяти (например, 3.25d-6).

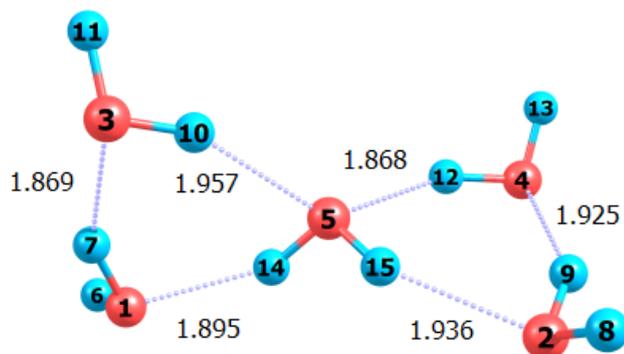


Рис. 22. Начальная геометрия кластера $(H_2O)_5$, использованная для ADMP расчета

Расчет запускается обычной командой для запуска Gaussian, например,
`g16 w5-admp.gjf ,`

после чего ход вычислений можно отслеживать в файле `w5-admp.log`, используя стандартные средства просмотра (но не редактирования!) программ `Far` (Windows) или `mc` (Linux). В среде Linux можно также использовать команду

```
tail -f w5-admp.log ,
```

которая выводит на экран конечную часть просматриваемого файла, постоянно обновляя ее. Поскольку расчет занимает значительное время, в среде Linux рекомендуется запускать расчет в фоновом режиме командой

```
g16 w5-admp.gjf &
```

или перевести запущенный расчет в фоновый режим обычными средствами Linux (`Ctrl-Z`, `bg`, `disown`).

В случае обрыва расчета по внешним причинам (отключение питания, перезагрузка компьютера), расчет может быть продолжен. Однако обычные средства рестарта программы Gaussian (командная строка `# restart`) в этом случае не помогут. Для продолжения прерванного расчета необходимо включить в команду ADMP ключевое слово `restart`, не изменяя другие опции и команды:

```
# x3lyp/6-31+G(d,p)
# ADMP=(Restart,MaxPoints=100000,StepSize=2000,NKE=19935)
```

(a) $t = 736$ фс

(б) $t = 914$ фс

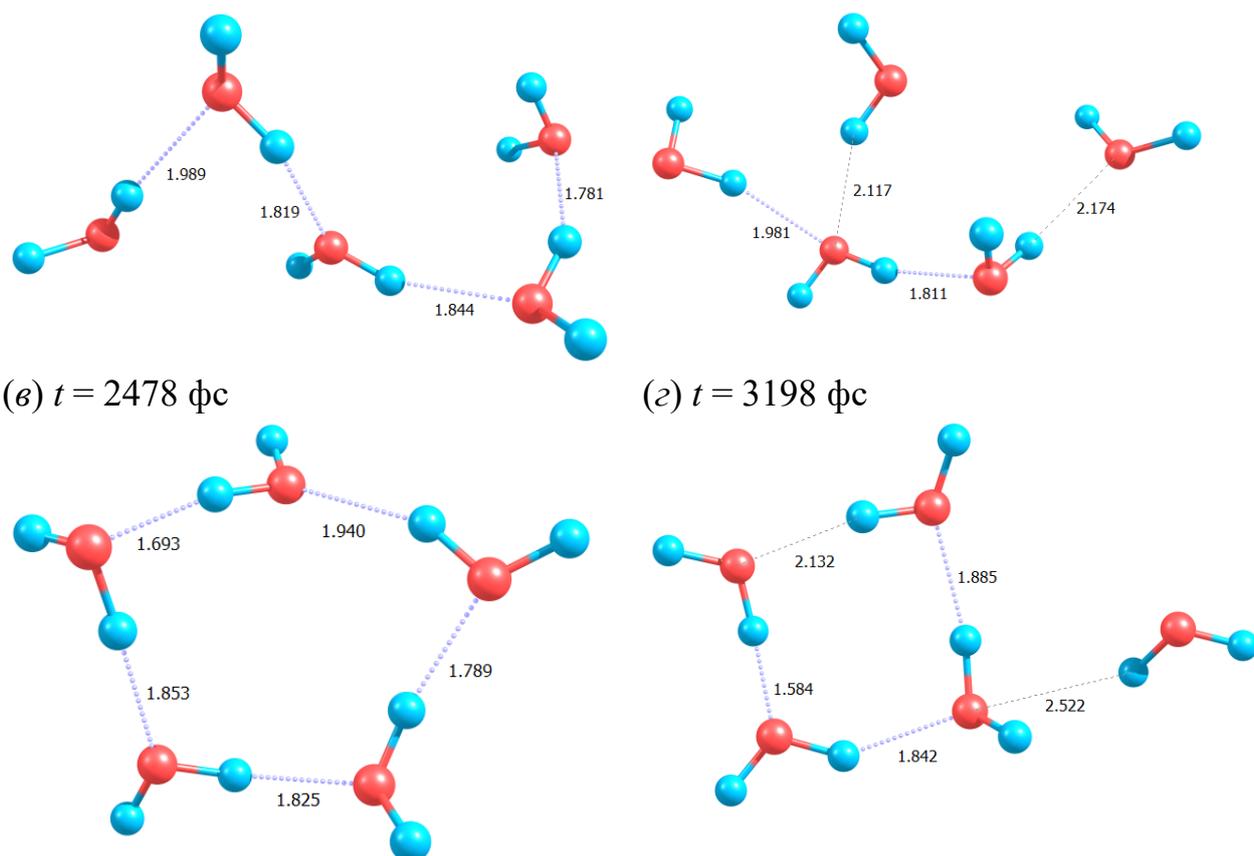


Рис. 23. Структуры кластера $(\text{H}_2\text{O})_5$ в различные моменты МД траектории

После завершения расчета (когда сгенерировано запрошенное число точек траектории) удлинить траекторию таким способом нельзя. Требуется начать новый расчет, являющийся продолжением предыдущего. В качестве начальных координат атомов надо указать последнюю рассчитанную геометрию предыдущего расчета и использовать ключевое слово `ReadMWvelocity` для ввода масс-взвешенных скоростей последней точки предыдущей траектории.

В ходе динамики структура кластера может значительно изменяться. На рис. 23 показаны структуры, которые возникают в различных точках МД траектории. Кроме того, если кинетическая энергия ядер системы высока, тепловая энергия может превысить энергию водородных связей. В этом случае одна или несколько молекул могут отрываться от кластера и удаляться на большое расстояние. Если требуется анализ связанной структуры, необходимо повторить расчет с другой начальной точкой или понизить значение НКЕ. Усредненный по нескольким запускам период жизни в связанном состоянии (в состоянии, когда все молекулы связаны хотя бы одной водородной связью) является оценкой времени жизни кластера. Если тепловая энергия не превышает энергию диссоциации, одна или несколько молекул могут удаляться на значительное расстояние, но затем возвращаются, и связанное состояние восстанавливается.

Поскольку в ходе расчета генерируется большой объем данных, визуализация и анализ полученной траектории представляет собой весьма сложную задачу. Как правило, требуется использовать дополнительные

программы для получения химической информации на основе полученной траектории. Одна из таких программ – программа GaussView, которая представляет собой стандартное средство Gaussian для анализа результатов расчетов этой программы. Однако программа GaussView является коммерческой и, кроме того, вследствие своего многоцелевого назначения, не всегда подходит для получения именно той информации, которая необходима пользователю. В этом случае рекомендуется использовать собственные программы или скрипты, извлекающие информацию из выходного файла Gaussian. Эти программы могут быть написаны на любом языке программирования (в т.ч. Python, C, Fortran, VisualBasic, Pascal..). Их задача – открыть текстовый файл с результатами (w5-admp.log) и прочитать оттуда координаты атомов в различных точках траектории, энергии в этих точках, и, если требуется, скорости атомов. Траектория в файле представлена в виде последовательности шагов, для каждого из которых выводятся различные виды энергии, угловой момент, координаты атомов и их скорости в масс-взвешенной декартовой системе координат, например:

```

Summary information for step      500
Time (fs)      50.000000
EKinC          =      0.1425727; EKinPA =      0.0002189; EKinPB =      0.0002189
EKin           =      0.1430105; EPot  = -2600.8910909; ETot   = -2600.7480803
ETot-EKinP     = -2600.7485181
Adiabaticity Index =      2.5453D-05
Integrand for the Velocity-Velocity Auto-Correlation Function =
1.380281616032D-01
Angular momentum
  JX = -0.1251811926D-12  JY = -0.8260244321D-12  JZ =  0.1521314658D-11
Jtot =  0.1735622361D-11 H-BAR;  J (Quantum Number) =  0.0000000000D+00
Total energy -2.600748080D+03  Delta-E -1.331728D-04 A.U.
Total angular momentum  1.735622D-12  Delta-A -2.851581D-13 h-bar
Tr(PA)-NA =  0.000D+00
Tr(PA**2-PA) =  2.392D-11
Tr[(PA**2-PA)**2] =  1.163D-13
Cartesian coordinates:
I=   1 X= -1.341103359793D+00 Y= -1.080206809868D+00 Z=  6.675831982333D-01
I=   2 X= -4.953076589801D-02 Y= -3.851011238330D+00 Z=  7.241488396178D+00
I=   3 X= -2.051715427312D+00 Y= -6.220825222130D+00 Z=  2.912535301371D+00
I=   4 X= -2.330568318083D+00 Y= -7.175939275204D+00 Z= -2.679277126512D+00
I=   5 X=  4.491782439671D-01 Y= -4.640810317120D+00 Z= -7.227856141784D+00
I=   6 X=  8.825136816142D-01 Y=  7.405337413132D+00 Z= -2.612002073615D+00
I=   7 X= -2.624488676930D+00 Y=  8.676308387599D+00 Z=  4.337307748936D+00
I=   8 X= -2.700739007151D+00 Y=  4.354638140250D+00 Z= -3.912445698862D-01
I=   9 X= -1.155141160922D+00 Y=  2.135385216197D+00 Z= -4.597357343275D+00
I=  10 X= -7.852184651157D-03 Y=  2.120916991436D+00 Z=  4.706276385507D+00
I=  11 X=  3.739413935417D+00 Y=  2.985401479846D+00 Z= -7.929158398743D-01
I=  12 X=  3.953132341146D+00 Y= -2.802214862425D+00 Z=  2.232803789308D+00
I=  13 X=  3.236900698596D+00 Y= -1.906979903381D+00 Z= -3.797341724586D+00
MW Cartesian velocity:
I=   1 X=  3.488323604203D+13 Y= -3.198065126643D+12 Z=  1.971770516226D+13
I=   2 X=  1.079676854694D+13 Y= -2.769805538653D+13 Z= -1.642265887468D+13
I=   3 X= -1.009032775324D+14 Y= -1.004679679879D+13 Z= -2.037442830426D+13
I=   4 X= -1.240967387789D+14 Y= -6.476130693264D+13 Z=  3.659722330719D+13
I=   5 X=  1.206955913278D+13 Y= -1.492917460376D+14 Z= -2.856880175241D+12
I=   6 X=  8.781118155384D+13 Y=  7.020123009086D+13 Z=  2.564574814266D+13
I=   7 X= -2.890502164826D+14 Y=  1.205374446919D+14 Z=  1.104698130341D+14
I=   8 X= -4.047379246900D+13 Y=  5.136980482919D+13 Z=  9.174789096278D+13

```

```

I= 9 X= 2.964042812610D+13 Y= -2.020987779413D+13 Z= -1.008164714823D+14
I= 10 X= 1.176358519862D+14 Y= 6.555315440907D+13 Z= 3.197345144092D+13
I= 11 X= 1.018427339649D+14 Y= -3.235487526221D+13 Z= -6.115053777900D+13
I= 12 X= 9.070068574476D+13 Y= -5.395280857092D+13 Z= -9.655470109377D+13
I= 13 X= 6.914358016533D+13 Y= 5.385189788844D+13 Z= -1.797615434068D+13

```

Наиболее важными энергетическими характеристиками являются величины E_{Pot} - потенциальная (квантовохимическая) энергия, т.е. энергия, которая получается при квантовохимическом расчете данной конфигурации неподвижных атомов), а также E_{Kin} (кинетическая энергия, включающая сумму кинетической энергии ядер E_{KinC} и электронов E_{KinPA} и E_{KinPB} . Полная энергия системы $E_{Tot}=E_{Kin}+E_{Pot}$ должна сохраняться на протяжении всей траектории, если не используется алгоритм термостатирования (см. ниже). При использовании термостата должна сохраняться E_{kin} . Поскольку эти величины всегда флуктуируют, сохранение энергий подразумевает, что не должно быть заметного дрейфа их средних значений. Величины J_X, J_Y, J_Z, J_{TOT} , показывают величины проекций углового момента и его модуль в единицах приведенной постоянной Планка. Поскольку начальная структура была неподвижной, угловой момент должен сохраняться близким к нулю на протяжении всей траектории. Декартовы координаты атомов приводятся в а.е. и соответствуют координатам в Å, приводимым в таблице Input Orientation. Скорости (MW Cartesian velocity) приводятся в масс-взвешенных координатах $\mathbf{r}_{mi} = m_i^{1/2} \mathbf{r}_i$ и выражены в единицах $amu^{1/2} \cdot Bohr/s$. Для выражения абсолютной скорости атома i в м/с следует использовать формулу

$$V_{i[m/c]} = \frac{0.529177 \cdot 10^{-10}}{\sqrt{m_i[a.e.m.]}} V_{mi}$$

В многоатомной системе кинетическая энергия изолированной системы претерпевает значительные осцилляции с большой амплитудой и однозначно указать температуру системы становится очень трудно. Довольно часто при значительной длине траектории и большом шаге интегрирования средняя кинетическая энергия имеет также склонность к дрейфу, т.е. неконтролируемому медленному изменению за счет накопления малых погрешностей интегрирования на каждом шаге расчета. Чтобы избежать этого, можно использовать специальные алгоритмы, устраняющие дрейф. Эти алгоритмы называются термостатами, один из наиболее простых из них состоит в том, чтобы равномерно масштабировать скорости всех атомов на определенном шаге траектории, добиваясь постоянства полной кинетической энергии. В программе Gaussian вызов алгоритма термостатирования (метод масштабирования скоростей) осуществляется с помощью команды IOp, которую надо указать в командной строке входного файла вместе с командой ADMP, например:

```

# x3lyp/6-31+G(d,p) ADMP=(MaxPoints=100000,StepSize=2000,NKE=19935)
# IOp(1/80=1000000,1/82=300,1/89=5)

```

Команда `IOP` изменяет внутренние параметры Gaussian с номерами 1/80, 1/82, 1/89 (подробнее о внутренних параметрах см. <https://gaussian.com/iop/>). Выражение 1/80=1000000 включает алгоритм термостатирования, выражения 1/82=300 и 1/89=5 означают, что температура системы должна поддерживаться в пределах (300 ± 5) К. С термодинамической точки зрения сохранение постоянной температуры и объема системы соответствует моделированию ансамбля NVT , свойства которого отличаются от свойств ансамбля NVE .

Следует помнить, что использование термостата не всегда оправдано с физической точки зрения. При температуре 298 К и стандартном полном давлении среднее время между столкновениями молекул идеального газа составляет около 1 нс. Таким образом, на протяжении примерно 10^6 фс пентамер находится в свободном состоянии, и его кинетическая и потенциальная энергия свободно флюктуируют, переходя друг в друга. С другой стороны, если применяется алгоритм термостатирования, он корректирует кинетическую энергию системы, как только она выходит за пределы установленного интервала температур, и это происходит с типичным периодом в 1-100 фс. Корректировка температуры соответствует добавлению или убыли дополнительной энергии в/из системы со стороны внешней среды. Столкновения и обмен энергией между системой и окружающей средой с такой частотой могут происходить в жидкости или на поверхности твердого тела (колебания и столкновения молекул жидкости происходят с периодом порядка долей пикосекунды, т.е. 10-100 фс), но эта частота совершенно нехарактерна для газовой фазы. Таким образом, моделирование кластера в газовой фазе правильнее проводить без алгоритма термостатирования. Также следует помнить, что оценка равновесных свойств системы (например, свободной энергии) требует и весьма длительной траектории. Поведение системы на временных интервалах порядка 10-20 пс в газовой фазе может существенно отличаться от равновесного состояния системы.

Для анализа результатов динамики кластера полезно проанализировать его структуру в различных точках траектории. Например, можно проанализировать наличие водородных связей между отдельными атомами кислорода. Простейший способ для этого – рассчитать расстояния между атомами кислорода i и j . Если это расстояние, например, менее 3 Å, можно считать, что между этими атомами есть водородная связь (будем называть это условие – *критерий 1*). Альтернативный способ состоит в том, чтобы находить расстояния H...O между атомами *различных* молекул воды, которые оказываются короче, например, 2.1 Å (далее будем называть это *критерием 2*). Существуют и более строгие способы выявления водородных связей, но в нашем примере мы не будем на них останавливаться. Определив наличие водородных связей между всеми атомами кислорода, мы получаем матрицу (т.н. *матрицу смежности*), в которой элемент a_{ij} равен 1, если между атомами i и j есть водородная связь, и 0 в противном случае. Матрица симметрична, ее диагональные элементы могут принимать произвольные значения, например, 0. Вследствие этого достаточно рассматривать только нижний или только верхний треугольники этой матрицы.

Например, для кластера на рис. 22 матрица смежности и ее верхний наддиагональный треугольник записываются как

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 1 \\ & 0 & 1 & 1 \\ & & 0 & 1 \\ & & & 1 \end{pmatrix}$$

Каждая матрица смежности описывает *молекулярный граф*, который характеризует химическую структуру. Кроме матрицы смежности, граф можно описать *списком ребер*, т.е. списком пар атомов (*вершин*), соединенных связью (*ребром*). Например, для обсуждаемого кластера список ребер: (1,2), (1,5), (2,4), (2,5), (3,5), (4,5). Молекулярный граф характеризуется *связностью*, т.е. количеством фрагментов, которые можно нарисовать, не отрывая карандаша от бумаги. Например, в соответствии с критерием 2, структуры (а) и (в) на рис. 23 односвязны, структура (б) трехсвязна, а структура (г) двусвязна.

Поскольку связанные атомы в молекуле могут иметь разные номера, нельзя установить химическую идентичность структур, сравнивая элементы матрицы смежности напрямую (при различной нумерации атомов строки и столбцы двух одинаковых структур не будут совпадать). Вместо этого следует проверить два молекулярных графа на *изоморфность*, поскольку две молекулы одинаковы, если их молекулярные графы *изоморфны* друг другу. Два графа называются *изоморфными*, если один можно перевести в другой только перемещая на бумаге вершины и растягивая (но не разрывая!) ребра. Такую операцию относительно просто провести графически, но при сравнении матриц смежности или списков ребер, требуются специальные алгоритмы. Подробнее об изоморфности и использовании теории графов для описания молекул и определения химически идентичных структур см. [22, 23]. Хотя установление изоморфизма графов довольно непростая математическая задача, существуют большое число готовых программ и библиотек, которые выполняют такое сравнение и многие другие операции с графами. В частности, такими программами являются библиотека *networkx* языка Python (<https://networkx.org/>), библиотека *igraph* на языках C и Python (<https://igraph.org/>), набор программ *Nauty* для работы с графами на сайте проф. В.МсКау (<https://users.cecs.anu.edu.au/~bdm/>, <https://users.cecs.anu.edu.au/~bdm/nauty/>) и многие другие. Эти программы анализируют матрицы смежности или списки ребер и определяют изоморфизм молекулярных графов. Изоморфные структуры являются идентичными в химическом смысле, т.е. имеют одинаковое расположение связей между атомами. Проанализировав изоморфизм точек МД траектории, можно определить набор уникальных изомеров и каждой точке траектории приписать номер уникального изомера, связность структуры и другие свойства.

Для быстрого предварительного анализа ADMP траектории можно использовать программу `admp2xyz` (<https://github.com/skignatov/admp2xyz>), которая автоматически выбирает из выходного файла нужные результаты и готовит их для визуализации, а также выполняет предварительную обработку этих результатов и расчет ряда полезных свойств, включая связность молекулярного графа. Для обработки результатов, программу следует запустить в директории, где находится файл входных данных `admp2xyz.inp`. Этот текстовый файл может содержать следующую информацию (каждая команда на отдельной строке):

```
admp d:\admp\w5-admp4-pm3.log
out w5-admp4-pm3
iBeg 0
iEnd 1000000
Each 10
Bond 1 8 2.5
```

Команда `admp` является обязательной и указывает имя и расположение входного файла (т.е. `log`-файла с результатами расчета ADMP). Команда `out` задает имя файлов с результатами обработки (при ее отсутствии имя этого файла повторяет имя входного файла). Команды `iBeg` и `iEnd` указывают начальные и конечные точки рассчитанной траектории, которые нужно проанализировать (номера шагов ADMP начинаются с нуля, поэтому `iBeg` равен 0, а `iEnd` может быть заведомо большим числом). По умолчанию `iBeg=0`, `iEnd=1000000`. Команда `Each <n>` указывает, что для обработки необходимо использовать только каждую n -ую (в данном примере – каждую десятую) точку траектории (по умолчанию `Each=1`). Поскольку временной шаг траектории очень мал (0.2 фс), для достижения приемлемой точности вполне достаточно анализировать только каждую десятую или даже двадцатую точку. Необязательная команда `Bond <NA> <NB> <r>` указывает, что молекулярный граф надо строить, учитывая дополнительные связи между атомами с атомными номерами NA и NB (в данном случае – атомами Н и О) при расстоянии между такими атомами r не более 2.5 Å. Это обеспечивает учет водородных связей в дополнение к обычным связям с длинами, рассчитанными на основании сумм стандартных радиусов атомов.

После запуска программы `admp2xyz.exe` в директории запуска появятся файлы `w5-admp4-pm3.xyz`, `w5-admp4-pm3.dat`, `w5-admp4-pm3.vib`. Первый из них содержит координаты атомов в формате NXYZ для выбранных точек траектории. Этот файл позволяет визуализировать траекторию с помощью программы Moltran, используя команду

```
moltran w5-admp4-pm3.xyz /nxyz0@geostep /all /c10000 /nosym /hb2.5
```

Ключ `/c<n>` позволяет просмотр до n точек траектории (в данном случае $n=10000$), этот ключ нужен для просмотра длинных траекторий, т.к. по умолчанию Moltran резервирует массивы под 2000 структур. Ключ `/hb2.5` нужен

для визуализации водородных связей (в данном случае длиной меньше 2.5 Å по критерию 2). После появления на экране молекулярной структуры нажатие клавиш «0», «.» (точка), «N», «B» на любом регистре латинской раскладки позволяет просматривать начальную, конечную, следующую или предыдущую точку траектории.

Файл `w5-admp4-pm3.dat` содержит данные о кинетической, потенциальной и полной энергиях системы в каждой точке траектории, а также связности молекулярного графа. В конце таблицы этих данных выводится информация о средних (Mean), минимальных и максимальных значениях этих энергий, а также их дисперсии (MSD) и стандартном отклонении (SD) для всей рассматриваемой выборки (множестве точек траектории, описываемом параметрами `iBeg`, `iEnd`, `Each`).

Файл `w5-admp4-pm3.vib` содержит данные о колебательном спектре системы, полученном путем Фурье-преобразования автокорреляционной функции кинетической энергии системы (см. ниже).

Рассмотрим в качестве примера две траектории ADMP-динамики (H₂O)₅, различающиеся начальными структурами. Система А соответствовала начальной точке, показанной на рис. 22, начальная структура системы В соответствовала оптимизированной циклической структуре, близкой к рис. 23в. Для ускорения расчета динамика проводилась полуэмпирическим методом PM3 без термостатирования, со средней температурой 300-350К, длина траектории 20 пс с шагом 0.2фс, начальная энергия 55061 мкХ:

```
# pm3 ADMP=(MaxPoints=100000,NKE=55061,StepSize=2000)
```

Начальная кинетическая энергия была выбрана примерно вдвое большей той, чем требуется для создания температуры 400К из-за того, что начальные структуры представляли собой оптимизированные точки локальных минимумов с нулевой потенциальной энергией. Как было описано выше, при тепловых колебаниях гармонического осциллятора половина механической энергии расходуется на поддержание средней потенциальной энергии, равной средней кинетической. Хотя в кластере колебания являются ангармоническими, начальная механическая энергия перераспределится между потенциальной и кинетической энергиями примерно пополам. Таким образом, для невозбужденных начальных структур установившаяся средняя температура будет примерно вдвое меньше, чем ее начальное значение. На соотношение средних потенциальной и кинетической энергии также влияет выбор начальной структуры, потенциальная энергия которой дает вклад в полную механическую энергию.

Полезно построить график зависимости E_{Rot} от времени для всех точек траектории, отмечая цветом индекс связности, в случае пентамера он может быть от 1 (единый кластер) до 5 (пять диссоциировавших молекул воды). Часто число точек траектории очень велико, поэтому достаточно анализировать и откладывать на графике только каждую десятую или даже каждую сотую точки. Рис. 24 представляет энергии (H₂O)₅ в ходе ADMP-динамики систем А и В, цвет

обозначает связность. Эти результаты были получены путем обработки сгенерированной траектории программой admp2xyz.

Видно, что односвязный граф (т.е. единая молекула пентамера) существует только весьма ограниченное время. Уже через 1-3 пс происходит выброс одной из молекул воды, в результате чего кластер долгое время эволюционирует как разделенные мономер и тетрамер. Для системы А довольно быстро происходит элиминирование еще одной молекулы воды и система превращается в тример и два мономера (зеленый цвет). Аналогичный процесс начинает происходить в системе В только примерно через 18 пс. Визуальный анализ структур в файлах *.xyz показывает, что выброс молекул воды происходит с удалением этой молекулы на расстояния примерно 40 Å, что практически означает необратимость этого процесса.

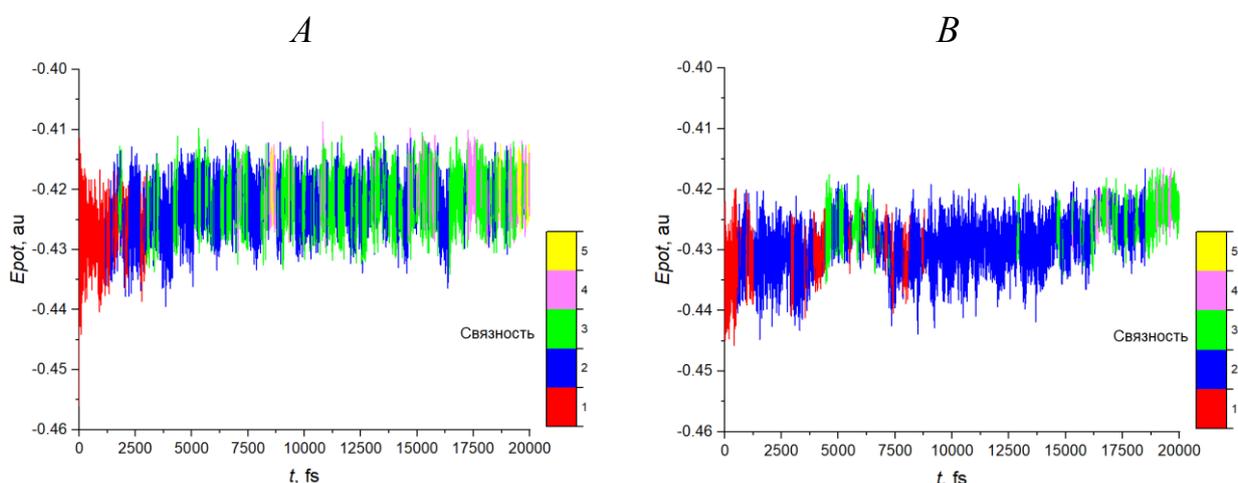


Рис. 24. Потенциальная энергия точек ADMP траектории пентамера воды (системы А и В) и связность кластера в каждой точке траектории.

В соответствии с классической молекулярно-кинетической теорией кинетическая энергия движения N атомов при фиксированном центре их масс однозначно связана с температурой системы по формуле

$$T = \frac{2E_{kin}}{3k(N-1)}$$

Таким образом, зависимость кинетической энергии системы в ходе динамики можно описывать как флуктуации локальной температуры системы. Усреднение ее временной зависимости приводит к среднему по времени значению этой температуры. В выходном файле *.dat программы admp2xyz.exe мгновенные значения температуры приводятся в виде отдельного столбца, в конце файла также приводятся соответствующие среднее и стандартное отклонение. В случае системы А средняя температура за время моделирования составляла 313 К, системы В – 287 К. Мгновенные значения температуры весьма сильно флуктуируют и имеют склонность к заметному дрейфу, который связан как с перераспределением энергии по степеням свободы, так и с химическими

изменениями в системе (изменением числа водородных связей). На рис. 25 показаны изменения мгновенных значений температуры (черные осциллирующие линии). Белая кривая показывает сглаживание коротких флуктуаций и дает т.н. *скользящее среднее* значений локальной температуры. Постепенное понижение температуры означает протекание эндотермического процесса элиминирования молекул из кластера, в результате чего часть первоначальной кинетической энергии была потрачена на разрыв водородных связей.

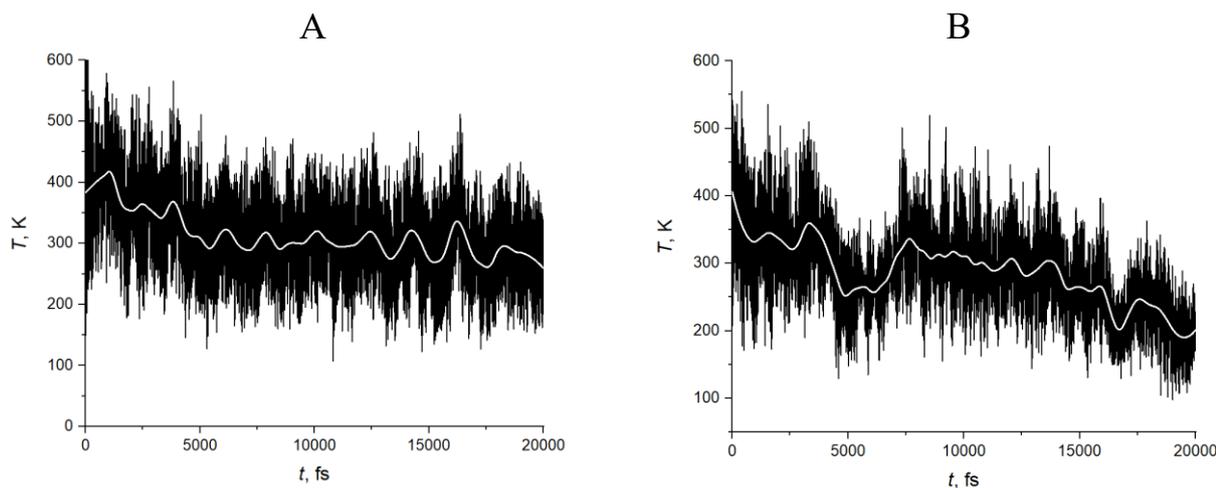


Рис. 25. Мгновенная локальная температура пентамера воды (системы А и В) за время ADMP-динамики.

Движение атомов молекулярной системы в ходе ADMP динамики складывается из медленных высокоамплитудных движений (раскрытие и образование циклов, удаление молекул на большие расстояния, образование новых конформаций) и быстрых колебательных движений валентных связей О-Н, валентных углов Н-О-Н и водородных связей Н...О. Хотя частоты колебаний устойчивых конформаций (минимумов ППЭ) можно провести квантовохимическим методом как описано в разделе 3.4, представляют интерес и частоты колебаний сильно искаженных, возникающих на короткое время структур МД траектории. Анализ МД траектории позволяет разделить и проанализировать эти движения. Для этого используется т.н. фурье-анализ автокорреляционных функций скорости. *Автокорреляционной функцией (АКФ)* временного ряда $X(t_i) = x(t_i) - \langle x \rangle$ ($i = 1, 2, \dots, m$) отклонений какой-либо величины $x(t_i)$ от своего среднего $\langle x \rangle$ называется функция $A(\tau)$ вида

$$A(\tau) = \frac{1}{m} \sum_{i=1}^m X(t_i)X(t_i + \tau)$$

АКФ показывает, насколько во временном ряду коррелируют между собой величины, разделенные временным промежутком τ . Величину τ часто называют *временным лагом*. Поскольку длина МД траектории n является конечной величиной, m не может превосходить ее: $m < n$. Более того, если m будет сравнима с n , точки АКФ с большим τ будут усредняться по меньшему количеству точек траектории, что приведет к потере точности в них. В связи с

этим величину m обычно выбирают значительно меньше n . Например, для нашей выборки с $n = 10000$, можно выбрать $m = 1000$.

Дискретное преобразование Фурье АКФ, вычисленной для скоростей частиц в точках МД траектории, отражает спектр частот колебаний этих частиц. Преобразование Фурье, выполняемое над дискретным набором m величин A_k , $k = 0, 1, \dots, m-1$, определяется формулой:

$$F_j = \sum_{k=0}^{m-1} A_k e^{-\frac{2\pi jk}{m}i}$$

Здесь A_k это значения АКФ $A(\tau)$ в точках $\tau_k = (k+1)\Delta t$, Δt – шаг интегрирования траектории, i обозначает мнимую единицу. Последовательность комплексных величин F_j , $j = 0 \dots m$, называется дискретным фурье-образом АКФ ($FT-ACF$), т.е. набором периодических компонент, колеблющихся с частотами $\omega_j = j/(2m\Delta t)$. Относительную амплитуду соответствующего колебания характеризует модуль фурье-образа $A(\omega_j) = (2/m)[\text{Re}(F_j)^2 + \text{Im}(F_j)^2]^{1/2}$, причем j здесь пробегает только половину своих значений $[1, m/2-1]$.

Таким образом, чтобы разложить траектории атомов системы на колебательные движения, происходящие с различными частотами, необходимо рассчитать АКФ проекций скоростей отдельных частиц и выполнить их фурье-преобразование. Эти операции также выполняет программа `admp2xuz`, за исключением того, что АКФ строится не для отдельных компонент скорости, а для мгновенной кинетической энергии системы (т.е. суммы квадратов компонент скорости). Хотя такая методика не полностью соответствует строгому алгоритму и вносит в расчет в определенные погрешности, в целом ее результаты согласуются с расчетом на основе проекций скорости. На рис. 26 показаны фурье-образы АКФ для систем А и В. Легко видеть, что системы характеризуются несколькими максимумами в области $3700-3900 \text{ см}^{-1}$. Самый высокочастотный среди них соответствует колебаниям свободных ОН связей молекул воды (валентное асимметричное колебание), ниже находится пик симметричных колебаний. Самый нижний максимум – это колебания ОН-групп, связанных водородными связями О-Н...О. В области $1600-2000 \text{ см}^{-1}$ находятся два максимума, соответствующих деформационным колебаниям Н-О-Н связанных и свободных молекул воды. Резкий подъем кривой вблизи нуля анализировать не требуется, он не является отражением каких-либо физических закономерностей. Различие в остроте максимумов для систем А и В, по-видимому, объясняются тем, что в системе А интервал времени, в котором ОН-группы молекул воды проводят в свободном состоянии значительно длиннее, чем в случае системы В. Связанное состояние характеризуется более размытым и менее интенсивным максимумом колебательного спектра.

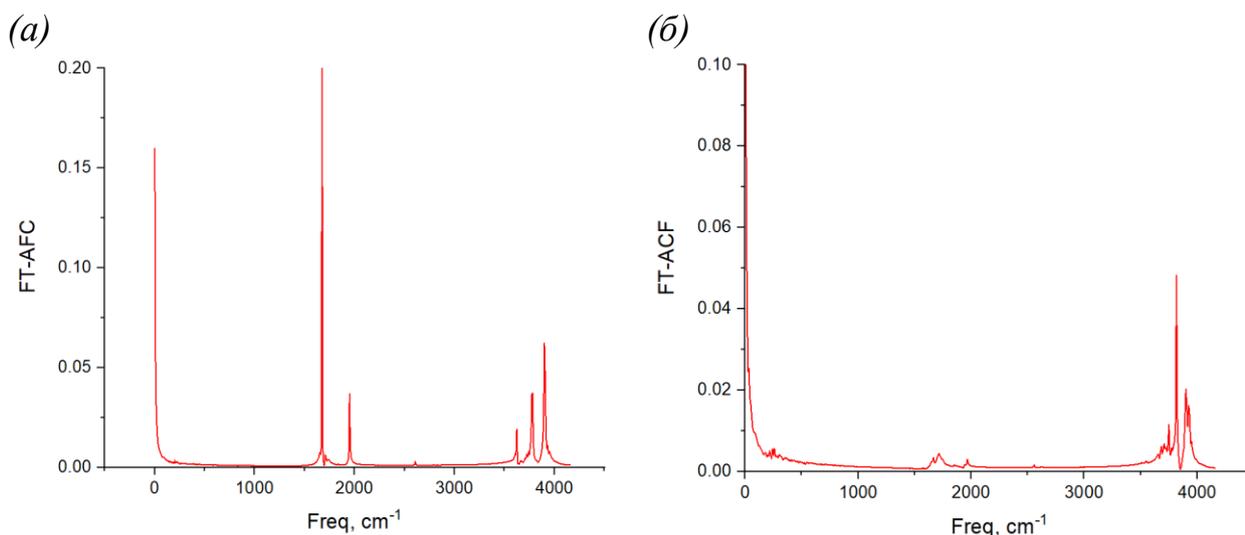


Рис. 26. Колебательные спектры, полученные на основе фурье-преобразования автокорреляционной функции для траекторий А и В, показанных на рис. 24.

Следует учитывать, что полученный фурье-образ АКФ нельзя напрямую сопоставлять с ИК-спектром системы. Дело в том, что интенсивность пиков в ИК-спектре определяется не только амплитудой колебаний атомов, но и изменением дипольного момента молекулы в процессе колебания. Большое изменение дипольного момента приводит к высокой интенсивности полос ИК-спектра, малое – к низкой интенсивности этих полос. Изменение дипольного момента никак не учитывается в FT-ACF и высокие максимумы могут не соответствовать заметным максимумам в наблюдаемом ИК-спектре. Максимумы FT-ACF свидетельствуют, что эти полосы *могут проявиться* либо в ИК, либо в КР спектрах, но не обязательно будут наблюдаться в них. Тем не менее, анализ FT-ACF на рис. 26 показывает, что рассчитанные колебательные спектры довольно хорошо согласуются с ИК спектрами водных комплексов, в которых обычно наблюдаются полосы в области $3700\text{--}3900\text{ см}^{-1}$, широкие полосы в области $3200\text{--}3600\text{ см}^{-1}$, а также полосы в области $1500\text{--}1700\text{ см}^{-1}$ средней интенсивности. Важным преимуществом расчета FT-ACF перед стандартными расчетами гармонических частот колебаний в приближении ЖРГО (см. раздел 3.4), является то, что такой подход учитывает влияние ангармонизма молекулярных колебаний и высокоамплитудные коллективные движения молекулярных фрагментов. Этот факт очень важен при описании слабосвязанных молекулярных комплексов и кластеров.

Также необходимо помнить, что расчет движения атомов в ADMP, как и в AIMD в целом, производится на основе классических траекторий. Это приближение обосновано в случае достаточно тяжелых атомов. В случае таких атомов, как водород и гелий, большую роль начинают играть квантовые эффекты. Например, скорости многих реакций переноса протона, гидрирования, гидратации в большой степени определяются скоростью квантового туннелирования, которое не учитывается классической МД. Для учета квантового характера движения легких атомов в МД часто используют подход,

называемый *молекулярной динамикой на основе интегралов по траекториям* (*Path Integral Molecular Dynamics, PIMD*). Кроме того, большие усилия в настоящее время направлены на разработку т.н. *квантовой молекулярной динамики* (*Quantum Molecular Dynamics, QMD*), которую не следует путать с рассматриваемой здесь МД на основе квантовохимических потенциалов. В QMD движение атомов рассматривается как движение «волновых пакетов» или, как вариант, частиц, траектории которых учитывают размывание за счет квантового движения.

Расчеты методом AIMD позволяют эффективное и подробное описание процессов быстрого перераспределения энергии и динамики на сложных ППЭ. Достоинствами МД моделирования является то, что при исследовании сложных процессов нет необходимости искать локальные минимумы ППЭ и переходные состояния между ними и можно описывать процессы, протекающие по нескольким различным каналам с очень невыраженными или разрыхленными переходными состояниями.

Контрольные вопросы

1. Что такое молекулярная динамика? Для моделирования каких систем и процессов она используется?

2. Чем отличается молекулярная динамика с эмпирическими потенциалами и неэмпирическая молекулярная динамика (молекулярная динамика с квантовохимическими потенциалами)?

3. Какие разновидности неэмпирической молекулярной динамики вы знаете? Чем они отличаются?

4. Вам необходимо изучить конформационное состояние глобулы белка в воде. Какой подход целесообразно использовать для решения этой задачи?

5. Вам необходимо изучить времена жизни промежуточных состояний молекулярного комплекса $\text{SO}_3 \cdot (\text{H}_2\text{O})_2$ и типичные времена химических превращений в нем. Какой подход целесообразно использовать для этого?

Список литературы

1. Bagaturyants A., Vener M. Multiscale modeling in nanophotonics: Materials and simulations. – Jenny Stanford Publishing, 2017. – 291 p.
2. Gallup G.A. Valence bond methods. Theory and applications. – Cambridge University Press, 2002. – 238 p.
3. Бейдер Р. Атомы в молекулах. Квантовая теория. – М.: Мир, 2001. – 532 с.
4. Буренин А.В. Симметрия квантовой внутримолекулярной динамики. – Н. Новгород: ИПФ РАН, 2012. – 416 с.
5. Квасников И.А. Термодинамика и статистическая физика. Т. 2. Теория равновесных систем: статистическая физика: Учебное пособие. – М.: Едиториал УРСС, 2002. – 432 с.
6. Павинский П.П. Введение в термодинамику и статистическую физику. – Л.: ЛГУ, 1984. – 236 с.
7. Головина Л.И. Линейная алгебра и некоторые ее приложения: Учебное пособие для вузов. – М.: Наука, 1985. – 392 с.
8. Уилкинсон Дж.Х. Алгебраическая проблема собственных значений. – М.: Наука, 1970. – 564 с.
9. Domcke W., Yarkony D.R. Role of conical intersections in molecular spectroscopy and photoinduced chemical dynamics // Annu. Rev. Phys. Chem. 2012. – V. 63. – P. 325–352.
10. Абаренков И.В., Братцев В.Ф., Тулуб А.В. Начала квантовой химии. – М.: Высшая школа, 1989. – 303 с.
11. Фудзинага С. Метод молекулярных орбиталей. – М.: Мир, 1983. – 461 с.
12. Уилсон С. Электронные корреляции в молекулах. – М.: Мир, 1987. – 304 с.
13. Минкин В.И., Симкин Б.Я., Миняев Р.М. Теория строения молекул. – Ростов-на-Дону: Феникс, 1997. – 558 с.
14. Fitts D.D. Principles of quantum mechanics: as applied to chemistry and chemical physics. – Cambridge University Press, 2002. – 351 p.
15. Cook D.B. Handbook of computational quantum chemistry. – Oxford University Press, 1998. – 743 p.
16. Jensen F. Introduction to computational chemistry. – Wiley, 2001. – 429 p.
17. Lewars E. Computational chemistry. Introduction to the theory and applications of molecular and quantum mechanics. – NY: Kluwer Academic publishers, 2003. – 471 p.
18. Warshel A. Computer modeling of chemical reactions in enzymes and solutions. – Wiley-VCH, 1997. – 236 p.
19. Koch W., Holthausen M.C. A chemist's guide to the density functional theory. – Wiley-VCH, 2001. – 293 p.
20. Parr R., Yang W. Density functional theory of atoms and molecules. – NY: Oxford University Press, 1989. – 333 p.
21. Leach A.R. Molecular Modelling. Principles and applications. – London: Pearson Education Ltd., 2001. – 744 p.

22. Применение теории графов в химии. / Под ред. Н. С. Зефирова, С. И. Кучанова. — Новосибирск: Наука, 1988. — 306 с.
23. Химические приложения топологии и теории графов: Пер. с англ. / Под ред. Р. Кинга. — М.: Мир, 1987. — 560 с.

Предметный указатель

Z-матрица	29	Полуэмпирические методы	14, 15
Алгоритм BFGS	42	Постреакционный комплекс	71
Базисные наборы	19	Предреакционный комплекс	66
Базисные функции	16, 17	Приближение МО ЛКАО	6
Вариационная теория переходного состояния, VTST	69	Принцип Aufbau	7
Внутренние координаты	27	Программа Avogadro	39
Внутренняя координата реакции, IRC	55	ChemCraft	39
Времязависимая теория функционала плотности, TD-DFT	14	CRYSTAL	24
Детерминант Слейтера	6	GAMESS	23
Дробные координаты	33	Gaussian	23
Избыточные внутренние координаты	41	GaussView	39
Константа скорости реакции	68	HyperChem	24
Локальные минимумы ППЭ	41	Moltran	40
Матрица Гессе	53	NWchem	24
плотности	7	VASP	24
Метод бракнеровских пар, BD	11	VMD	39
противовеса	52	Процедура CCP	8
самосогласованного реакционного поля, SCRF	73	Псевдоатомы	30
Методы <i>ab initio</i>	10	Псевдопотенциалы	20, 21
AIMD, BOMD,	85	Расчет при фиксированной геометрии	47
CPMD, ADMP	86	Расчет термодинамических функций	58, 59
конфигурационного взаимодействия, CI	11	Стандартная теория переходного состояния	68
Модель поляризуемого континуума, PCM	74	Стационарная точка ППЭ	64
Молекулярная динамика	83	Суперячейка	33
Неограниченный метод Хартри-Фока, UHF	9	Теория возмущений	
Ограниченный метод Хартри-Фока, RHF	9	Меллера-Плессета, MP2	10
Оптимизация геометрии	41	Теория связанных кластеров, CCSD	10
Переходное состояние	66	Теория функционала плотности, DFT	12
Поверхность потенциальной энергии, ППЭ	41	Уравнение Шредингера	5
		Уравнения Рутана	7
		Хартри-Фока	6
		Функционалы DFT	13
		Экстраполяционные методы	
		G3, G4, CBS, W1BD	62
		Энергия нулевых колебаний, ZPE	59

Приложение

Соотношения между различными единицами измерения, используемыми в квантовохимических расчетах

$$1 \text{ а.е. энергии (Хартри)} = 27.21 \text{ эВ}$$

$$1 \text{ эВ} = 23.06 \text{ ккал/моль}$$

$$1 \text{ ккал/моль} = 4.184 \text{ кДж/моль}$$

$$1 \text{ а.е. длины (Бор)} = 0.529177 \text{ \AA}$$

$$1 \text{ \AA} = 10^{-10} \text{ м}$$

$$1 \text{ а.е.м. (масса электрона)} = 9.11 \cdot 10^{-31} \text{ кг}$$

$$1 \text{ у.а.е.м. (Дальтон)} = 1.66 \cdot 10^{-27} \text{ кг}$$

$$1 \text{ Дебай} = 3.34 \cdot 10^{-30} \text{ Кл}\cdot\text{м} = 0.208 \text{ e}\cdot\text{\AA}$$

$$1 \text{ атм} = 101325 \text{ Па} = 760 \text{ мм рт.ст.}$$

$$1 \text{ бар} = 100000 \text{ Па}$$

$$1 \text{ Торр} = 133.32 \text{ Па} = 1 \text{ мм рт.ст.}$$

Станислав Константинович **Игнатов**

**КВАНТОВОХИМИЧЕСКОЕ МОДЕЛИРОВАНИЕ
МОЛЕКУЛЯРНЫХ СИСТЕМ**

Учебное пособие

Федеральное государственное автономное образовательное учреждение
высшего образования «Национальный исследовательский Нижегородский
государственный университет им. Н.И. Лобачевского».
603022, Нижний Новгород, пр. Гагарина, 23.

Подписано в печать 10.12.2024 г. Формат 60×84 1/16.
Печать цифровая. Бумага цифровая. Гарнитура Таймс.
Усл. печ. л. 6,2. Заказ № 1071. Тираж 50 экз.

Отпечатано в типографии
Нижегородского госуниверситета им. Н.И. Лобачевского.
603000, г. Нижний Новгород, ул. Большая Покровская, 37.