

ЛЕКЦИЯ

УДК 577.1
©Иванов

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ХИМИИ ДЛЯ МЕДИКО-БИОЛОГОВ

А.С.Иванов

ГУ НИИ биомедицинской химии им. В.Н. Ореховича РАМН
119121, Москва, Погодинская ул., д.10
факс: (095) 245-0857, эл. почта: ivanov@ibmh.msk.su

Кратко изложены основные принципы и подходы вычислительной химии, лежащие в основе методов молекулярного моделирования, используемых для решения задач в области биоинформатики. Рассмотрены основные положения, методы молекулярной и квантовой механики, комбинированные методы. Данная лекция входит в теоретический цикл "Биоинформатика и компьютерное конструирование лекарств" для студентов 4 курса Медико-биологического факультета РГМУ (специальности – биохимия, биофизика, медицинская кибернетика), может быть также рекомендована для других студентов и аспирантов медико-биологических специальностей.

Ключевые слова: лекция, вычислительная химия, молекулярное моделирование, молекулярная механика, квантовая механика

1. Основные положения

1.1. Определение. *Вычислительная химия* – это количественное математическое моделирование химических процессов на компьютере с использованием формальных правил теоретической химии.

1.2. Область применения вычислительной химии. Подходы вычислительной химии включают в себя разработку и применение компьютерных алгоритмов для предсказания различных свойств молекул и объяснения наблюдаемых феноменов. Методами вычислительной химии может быть осуществлен расчет и предсказание как количественных, так и не количественных параметров. Применимость подходов вычислительной химии для разных объектов моделирования приведена в таблице 1.

1.3. Компьютерное описание молекул. Существует два предельно разных подхода для компьютерного описания молекул.

- 1) Описание энергии молекул с помощью эмпирических уравнений (*молекулярная механика*);
- 2) Описание энергии молекул с помощью волновых функций (*квантовая механика*).

Рассмотрим кратко оба этих подхода.

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ХИМИИ

Таблица 1. Применимость методов вычислительной химии.

| <i>Моделируемые объекты</i> | <i>Рассчитываемые параметры</i> |
|---|---|
| Одиночные молекулы | <ul style="list-style-type: none"> - геометрия молекул; - длина химических связей и углы между ними; - энергетические барьеры при вращении вокруг связей; - колебательные частоты; - распределение электронов, электронные орбитали, дипольные моменты; - ионизационные потенциалы; - другие параметры (в принципе, любые физически наблюдаемые свойства). |
| Надмолекулярные системы (молекулярные ансамбли) | <ul style="list-style-type: none"> - межмолекулярные комплексы; - молекулярное узнавание; - молекулярная мимикрия; - догма молекулярной фармакологии. |
| Большое число молекул и их взаимодействие в течение определенного периода времени (моделирование молекулярной динамики) | <ul style="list-style-type: none"> - энтальпия; - свободная энергия; - теплоемкость; - константы равновесия; - коэффициенты распределения; - константы связывания. |
| Химические реакции | <ul style="list-style-type: none"> - переходные состояния; - ход реакции в течение короткого периода времени $\sim 10^{-10}$ с |

2. Молекулярная механика

Молекулярная механика представляет собой чисто эмпирический подход, основанный на применении принципов классической механики для расчета энергии молекул с использованием так называемых “силовых полей” (Force Fields). Хотя данный термин звучит очень солидно, в действительности силовое поле гораздо менее интересно. В сущности, оно представляет собой “всего лишь” набор чисел, являющихся параметрами потенциальных функций для описания внутри- и межмолекулярных взаимодействий.

В данном подходе молекула рассматривается как набор физических объектов (атомы), соединенных с помощью пружинок (химические связи) с различными параметрами эластичности (силовые константы). Силы, удерживающие атомы в молекуле вместе, могут быть описаны с помощью потенциальных силовых функций, зависящих от структурных параметров, таких как длина связи, валентный угол, различные нековалентные взаимодействия и т.д. Комбинация этих потенциальных энергетических функций и представляет собой силовое поле.

Молекулярная механика имеет ряд неоспоримых преимуществ:

- 1) метод не требователен к вычислительным ресурсам;
- 2) метод прост в использовании;
- 3) качество результатов (геометрия молекул) сравнимы с результатами, полученными с помощью сложных методов квантовой механики;
- 4) позволяет моделировать крупные молекулярные системы, состоящие из сотен и тысяч атомов (крайне важное преимущество, позволяющее моделировать биологические макромолекулы).

В “силовом поле” молекула описывается с помощью трех основных параметров:

- 1) типы атомов,
- 2) относительное расположение атомов в пространстве,
- 3) связи между атомами.

1) Типы атомов. Так как в основе молекулярной механики лежит механистическое представление молекул и атомов, то во всех без исключения версиях силовых полей различные варианты атомов (в зависимости от гибридизации и локального окружения, например, sp^2 , sp^3 и карбонильные атомы

углерода) просто представляются как различные объекты (разные типы атомов). Введение этих различий необходимо, так как параметры силового поля для каждого типа атома значительно различаются. Специализированные силовые поля, предназначенные для моделирования определенных видов молекул, как правило, используют большее число типов атомов, чем универсальные. Таким образом в молекулярной механике в списке типов атомов может присутствовать по несколько атомов C, N, O, S, P и др. В качестве объектов могут также использоваться так называемые “объединенные” типы атомов, которые описывают целые группы реальных атомов, например, метиленовая группа ($-\text{CH}_2-$), где два атома водорода объединены с углеродом в один объект (такое объединение возможно при условии, что атомы водорода не участвуют в формировании водородных связей).

2) Относительное расположение атомов в пространстве. В молекулярной механике позиция каждого атома в пространстве описывается стандартными трехмерными координатами.

3) Связи между атомами. Химические связи в молекулярной механике описываются соответствующими типами связей (число типов связей соответствует числу всех вариантов парных сочетаний всех типов атомов).

2.1. Геометрия (конформация) молекулы. Из химической теории известно, что существуют стандартные длины для всех типов химических связей и углов между ними. В большинстве случаев геометрия простых молекул соответствует этим величинам. Однако в более сложных молекулярных структурах, как правило, наблюдаются отклонения длин связей и углов от стандартных, что в соответствии с теорией должно приводить к повышению энергии системы. Объяснение этого факта кроется в том, что по законам термодинамики молекулярная система, стремясь к минимуму общей энергии, принимает для этого оптимальную конформацию. В сложных молекулярных структурах такая оптимальная конформация может быть достигнута за счет ряда локальных структурных напряжений и соответственно роста энергии отдельных связей и углов, если при этом происходит общий выигрыш в энергии системы.

В общем виде в расчеты молекулярной механики должны быть включены все виды сил (стерические, электростатические и т.д.). Вычисления молекулярной механики базируются на взаимосвязи структуры молекулы и ее энергии. Для нахождения правильной пространственной структуры молекулы необходимо исследовать энергию молекулярной системы и найти такую структуру, у которой энергия минимальна.

Для выполнения расчетов с помощью силового поля необходимо иметь:

- 1) уравнение для расчета энергии молекулярной системы как функции ее геометрии;
- 2) параметризацию силового поля (описание будет дано ниже);
- 3) алгоритм оптимизации структуры молекулы - итерационная процедура расчета новых координат атомов с расчетом энергии системы и поиском структуры с минимальной энергией (минимизация энергии).

Основными целями компьютерного молекулярного моделирования могут быть:

- 1) предсказание равновесной молекулярной структуры и ее относительной стабильности в равновесном и неравновесном состоянии;
- 2) воспроизведение молекулярной подвижности в терминах фундаментальных внутри- и межмолекулярных сил.

Для достижения любой из этих целей необходимо знать функцию зависимости потенциальной энергии системы от ее молекулярной структуры и пространственных координат.

Накопленные знания и опыт говорят о том, что в молекулярных структурах существуют определенные закономерности между атомами, химическими связями, углами между ними, и т.д. Известны также параметры равновесной геометрии, частоты колебаний и энергии диссоциации.

Молекулярная механика подразумевает, что потенциальная энергия молекулярной системы, находящейся в равновесном состоянии, является простой

функцией межатомных расстояний, углов между связями, двугранных углов и ряда других параметров. Этот подход применим для описания различных молекул

Общая энергия молекулярной системы (E) описывается суммой энергий химических связей, углов между связями, торсионных (двугранных) углов, взаимодействий Ван-дер-Ваальса, электростатических взаимодействий и ряда других компонентов:

$$E = E_{св} + E_{угл} + E_{m.угл} + E_{vdw} + E_{эл.см} + \dots \quad (1)$$

Одной из проблем молекулярной механики является выбор "внутренних" координат для описания молекулярной энергии. Причина кроется в том, что число внутренних координат в молекуле обычно больше, чем число степеней свободы. Другими словами, многие параметры молекулярной системы являются избыточными, так как линейно зависимы.

Пример. В формальдегиде ($H_2C=O$) только 2 угла между связями из существующих 3 углов являются независимыми (так как сумма всех 3 всегда равна 360° для планарной молекулы). Таким образом, для описания энергии изменения углов в молекуле формальдегида достаточно всего 2-х из них, хотя нет каких либо определенных предпочтений в выборе любой пары углов из 3 возможных. Прямым следствием этого является тот факт, что силовые константы параметризации для данной молекулы являются не уникальными, так как все три константы линейно зависимы.

Явление линейной зависимости констант приводит к плохой переносимости параметров силового поля между различными молекулярными системами. Другими словами, не может быть одного универсального силового поля для оптимального описания любых молекул. Поэтому существует много различных вариантов силовых полей, оптимизированных для различных молекулярных систем.

Примером таких специализированных силовых полей могут быть поля, предназначенные для описания биологических молекулярных систем, такие как AMBER, CHARMM, TRIPOS, GROMOS, UFF, CVFF, ESFF и др.

2.2. Описание энергии молекулы. В силовом поле энергия молекулярной системы описывается набором потенциальных функций, которые условно можно разделить на две группы:

- 1) функции, относящиеся к химическим связям (bonded terms), описывают энергию связей, углов между ними и торсионных (двугранных) углов;
- 2) функции, не относящиеся к химическим связям (non-bonded terms), описывают внутри- и межмолекулярные взаимодействия (взаимодействия Ван-дер-Ваальса, электростатические взаимодействия и т.д.).

Как уже было сказано, энергия молекулы (E) описывается в силовом поле как сумма вкладов отдельных видов взаимодействий, каждое из которых описывается соответствующим уравнением. Все эти уравнения должны быть достаточно просты для быстрого их решения на современных компьютерах, чтобы можно было выполнять вычисления даже для очень больших молекул.

Крайне важно иметь в виду, что абсолютная величина E , вычисленная с помощью любого силового поля не имеет реального физического смысла (хотя и оценивается в единицах, обозначаемых как ккал/моль). Значения E , рассчитанные для одной и той же молекулярной системы с использованием разных силовых полей, могут сильно различаться. Такие же различия могут наблюдаться и при сравнении вычисленных значений и реальных величин энергии, определенных экспериментально. Однако разность вычисленных энергий для двух или более конформаций одной и той же молекулы имеет физический смысл как сравнение **относительных** конформационных энергий или различий конформационных энергий различных молекул.

Таким образом общая энергия молекулярной системы представляет собой сумму всех внутримолекулярных и межмолекулярных взаимодействий:

$$E = \sum E_{св} + \sum E_{угл} + \sum E_{m.угл} + \sum E_{vdw} + \sum E_{эл.см} + \dots \quad (2)$$

где:

- E - общая энергия системы,
- ΣE_{cv} - сумма энергий всех связей,
- $\Sigma E_{y\theta}$ - сумма энергий всех углов,
- $\Sigma E_{m.y\theta}$ - сумма энергий всех торсионных углов,
- ΣE_{vdw} - сумма энергий всех взаимодействий Ван-дер-Ваальса в системе,
- $\Sigma E_{эл.ст.}$ - сумма энергий всех электростатических взаимодействий в системе.

Пример. Расчет энергии в силовом поле TRIPOS включает все связи, углы, торсионные углы и нековалентные взаимодействия между атомами, не соединенными химическими связями (т.е. взаимодействия 1,4 и выше):

$$E = \Sigma E_{str} + \Sigma E_{bend} + \Sigma E_{oop} + \Sigma E_{tors} + \Sigma E_{vdw} + \dots \quad (3)$$

где:

- E_{str} - член, описывающий энергию изменения длины химической связи,
- E_{bend} - член, описывающий энергию изгиба угла между связями,
- E_{oop} - член, описывающий энергию изгиба химических связей у планарных атомов в сторону выхода атома из плоскости (*out of the plane*),
- E_{tors} - член, описывающий энергию вращения торсионного угла,
- E_{vdw} - член, описывающий энергию взаимодействия Ван-дер-Ваальса.

В расчет энергии может быть включен ряд дополнительных членов:

$$+ \Sigma E_{ele} + \Sigma E_{dist_c} + \Sigma E_{ang_c} + \Sigma E_{tor_c} + \Sigma E_{range_c} + \Sigma E_{multi} + \Sigma E_{fldfit} \quad (4)$$

где:

- E_{ele} - член, описывающий энергию электростатического взаимодействия,
- E_{dist_c} - член, описывающий энергию, обусловленную ограничениями по расстоянию,
- E_{ang_c} - член, описывающий энергию, обусловленную ограничениями по углу,
- E_{tor_c} - член, описывающий энергию, обусловленную ограничениями по торсионному углу,
- E_{range_c} - член, описывающий энергию, обусловленную ограничениями по интервалу,
- E_{multi} - член, описывающий энергию, обусловленную совмещением двух молекул (*multifit*),
- E_{fldfit} - член, описывающий энергию, обусловленную процедурой *field fit*.

2.2.1. Энергия химических связей (E_{cv}). В молекулярной механике химические связи рассматриваются как упругие элементы (пружинки) между атомами.

Для простой системы из двух атомов, соединенных ковалентной связью, зависимость энергии от длины связи d (расстояние между центрами атомов) описывается гармоническим приближением к равновесному значению d_0 . Наилучшее полное описание дает негармонический потенциал Морзе (рис. 1). Для молекулярных систем, содержащих более одной ковалентной связи, энергия всех химических связей описывается уравнением в виде суммы энергий всех связей:

$$E_{cv} = \Sigma E_{duc} \cdot [1 - e^{-\alpha(d-d_0)}]^2 \quad (5)$$

где:

- E_{duc} - энергия диссоциации связи,
- α - силовая константа,
- d - длина связи,
- d_0 - оптимальная длина связи.

Так как в районе d_0 форма кривой потенциала Морзе приближается к параболической, то в молекулярной механике для упрощения вычислений данная зависимость обычно заменяется гармоническим приближением в виде описания колебания упругого элемента (закон Гука, пунктирная линия на рис. 1) с силовой константой k , описывающей деформацию, и равновесным значением d_0 :

$$E_{cv} = \Sigma k(d - d_0)^2 \quad (6)$$

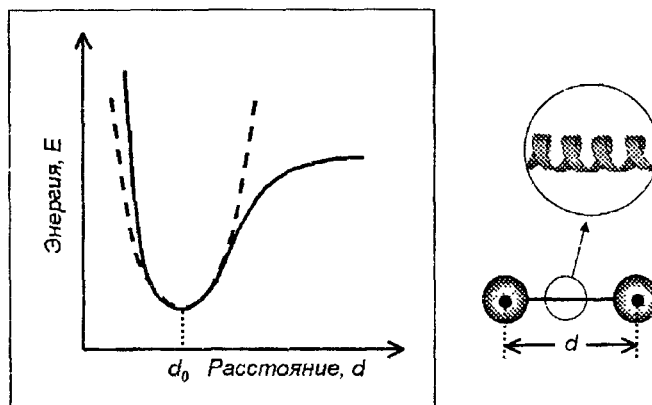


Рисунок 1.
Зависимость энергии от длины химической связи.
—— - потенциал Морзе; - - - - - закон Гука

где k - силовая константа, описывающая деформацию.

Для повышения точности описания на больших расстояниях может быть использовано кубическое уравнение:

$$E_{св.} = \sum k (d - d_0)^2 + k' (d - d_0)^3 \quad (7)$$

Необходимо отметить, что в уравнении 6 (или 7) должны быть использованы различные силовые константы (параметры k и d_0) для всех вариантов парных сочетаний атомов, как по их типам, так и по типам ковалентных связей. Полный набор этих силовых параметров описывает энергии связей всех возможных вариантов пар атомов, и является частью так называемого "силового поля", лежащего в основе метода молекулярной механики.

Так как данный подход является полностью эмпирическим, то существующие алгоритмы и программы молекулярной механики, хоть и основаны на общих принципах, отличаются друг от друга видом уравнений и наборами силовых параметров (силовыми полями). В качестве примера на рисунке 2. для трех пар атомов $[C(sp^3)-C(sp^3), C(sp^2)-O(sp^2)]$ и $N(sp^3)-H]$ приведены зависимости энергии химических связей от их длины и соответствующие силовые параметры из силового поля TRIPOS.

2.2.2. Энергия углов ($E_{угл.}$). Зависимость энергии системы, состоящей из 3-х атомов (рис. 3), от величины угла между двумя химическими связями также может быть описана гармонической функцией по типу закона Гука, как и в случае длины химической связи (см. рис. 1). Соответствующие силовые константы, как правило, в 5-7 раз меньше констант, описывающих энергию длины связи.

Как и в предыдущем случае, энергии всех углов между химическими связями в более сложной молекулярной системе описывается гармоническим потенциалом в виде суммы энергий для каждого угла:

$$E_{угл.} = \sum k (\Theta - \Theta_0)^2 \quad (8)$$

где:

- k - силовая константа, описывающая деформацию угла,
- Θ - угол между химическими связями,
- Θ_0 - равновесный угол.

Данное уравнение плохо описывает сильно напряженные системы (в этих случаях используют дополнительные члены с более высокими степенями и дополнительными силовыми параметрами, по типу уравнения 7 для химических связей).

На рисунке 4 в качестве примера из силового поля TRIPOS приведены зависимости энергии углов от их величины и соответствующие силовые параметры для следующих 3-х атомных систем $[C(sp^3)-C(sp^3)-C(sp^3), C(sp^3)-C(sp^2)-O(sp^2), C(sp^3)-C(sp^3)-N(sp^3)]$.

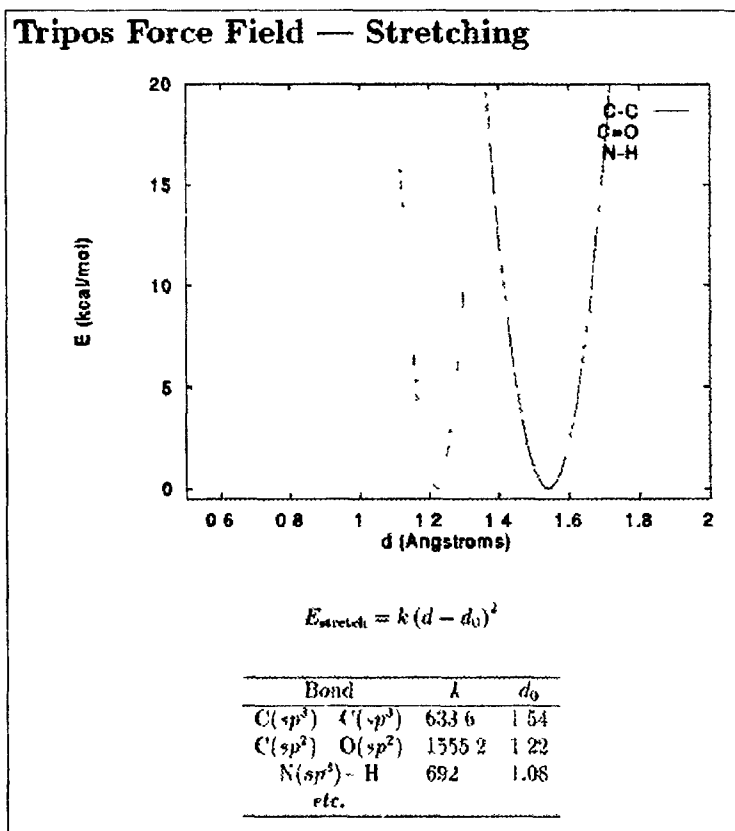


Рисунок 2

Пример из силового поля TRIPOS. Зависимость энергии ковалентных связей от длины и соответствующие силовые параметры для трех пар атомов C(sp³) - C(sp³), C(sp²) - O(sp²), N(sp³) - H

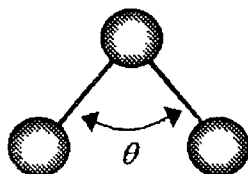


Рисунок 3

Система из 3-х атомов. Зависимость энергии молекулярной системы от величины угла θ может быть описана простым гармоническим колебанием подобно зависимости энергии химической связи от ее длины (см. п. 2.4.1.1).

Примеры некоторых валентных углов в структурах аминокислот приведены на рисунке 5.

2.2.3. Энергия торсионных углов. Торсионный угол, называемый также двугранным углом, является важным фактором формирования конформации макромолекул. Торсионный угол присутствует в любой системе из 4 последовательно соединенных атомов и описывает угол вращения (ω) вокруг средней связи (рис. 6А). Его значение соответствует двугранному углу между двумя плоскостями, проведенными через атомы 1-2-3 и 2-3-4 (рис. 6Б) и измеряется от 0 до $\pm 180^\circ$. Два крайних положения (0 и 180) соответствуют *цис* и *транс* конформациям (рис. 6В). Два промежуточных положения ($\pm 60^\circ$) соответствуют равновесным *гош* состояниям (рис. 6Г). На практике энергия *гош* состояний немного отличается от *транс* состояния в зависимости от атомов, составляющих торсионный угол.

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ХИМИИ

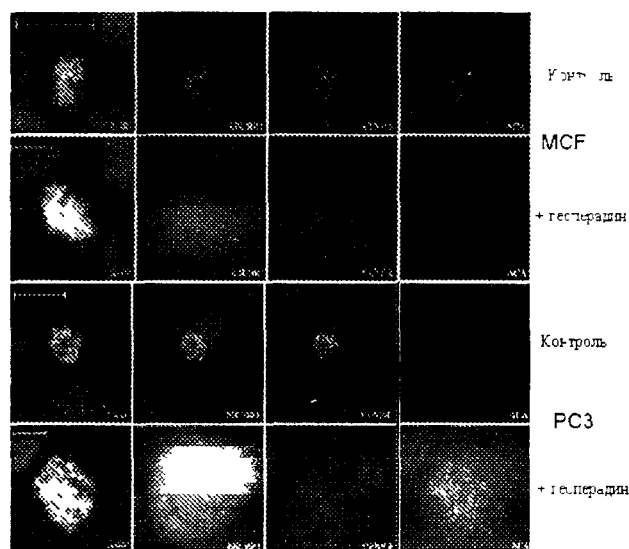


Рисунок 4

Пример из базисного пакета TRIPOS — зависимость энергии углов от величины θ и соответствующие силовые параметры для трех 3-х атомных систем: $C(sp) - C(sp) - C(sp)$, $C(sp) - C(sp) - O(sp)$, $C(sp) - C(sp) - N(sp)$.

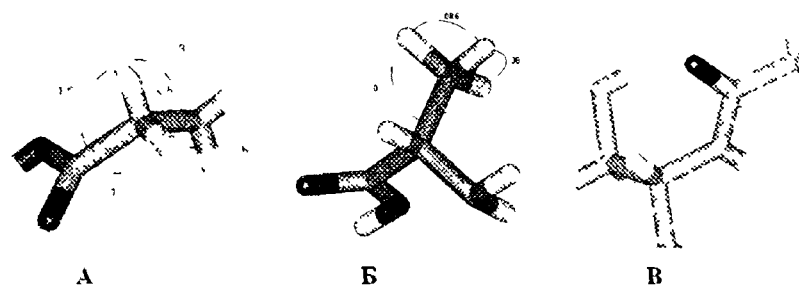


Рисунок 5

Примеры различных валентных углов между связями. В аминокислотах глицин (А) и аланин (Б) значения углов $\angle C_{\alpha} - C_{\beta} - N$ атомов равны примерно 109° , в то время как углы $\angle C_{\alpha} - C_{\beta} - N$ в основной пептидной цепи (В) равны примерно 120° , что указывает на планарность пептидной связи.

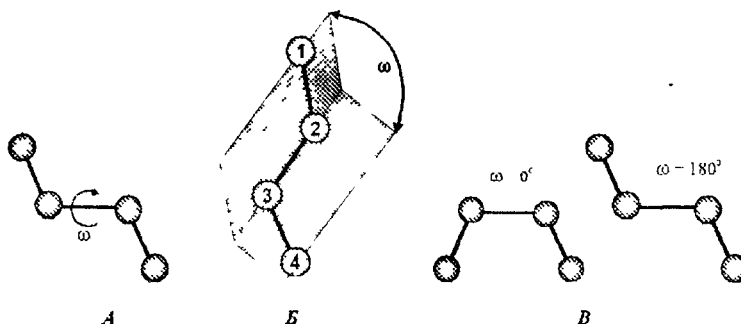


Рисунок 6

Торсионный угол и его равновесные конформации

Зависимость энергии молекулярной системы от величин торсионных углов описывается уравнением

$$E_{\text{торс}} = \sum_k [1 + \sum_{|S|} \cos(|S| \omega)] \quad (9)$$

где S — периодичность вращения, отражающая симметрию системы

Пример S равна 3 в случае этана и -2 в случае этена

Для дифференцирования энергий гош и транс конформаций в описании торсионного потенциала используются дополнительные параметры

Например в силовом поле MM2 уравнение выглядит так

$$E_{\text{торс}} = \Sigma \left[\frac{k_1}{2}(1 + \cos \omega) + \frac{k_2}{2}(1 + \cos 2\omega) + \frac{k_3}{2}(1 + \cos 3\omega) \right] \quad (10)$$

Это обеспечивает более точную параметризацию, но влечет значительное увеличение числа параметров даже для относительно небольших молекул

Внутримолекулярные движения, обусловленные изменением торсионных углов, в сотни раз более свободны по сравнению с движениями, обусловленными изменением длины химических связей. Причиной включения торсионного компонента во все силовые поля является необходимость гарантировать правильность оценки степени жесткости макромолекулярной цепи. В действительности она отражает стерические помехи, создаваемые соседними атомами и боковыми группами для вращения вокруг основной цепи.

Важность торсионного потенциала обусловлена тем, что многие силовые поля созданы для моделирования гибких макромолекул, где основные изменения конформации обусловлены вращением вокруг связей, т.е. изменением торсионных углов.

На рисунке 7 в качестве примера из силового поля TRIPOS приведены зависимости энергии торсионных углов от их величины и соответствующие силовые параметры для трех 4-х атомных систем $C(sp^3)-C(sp^3)-C(sp^3)-C(sp^3)$, $C(sp^3)-C(sp^2)-C(sp^2)-C(sp^3)$, $O(sp^2)-C(sp^2)-N(\text{amide})-H$.

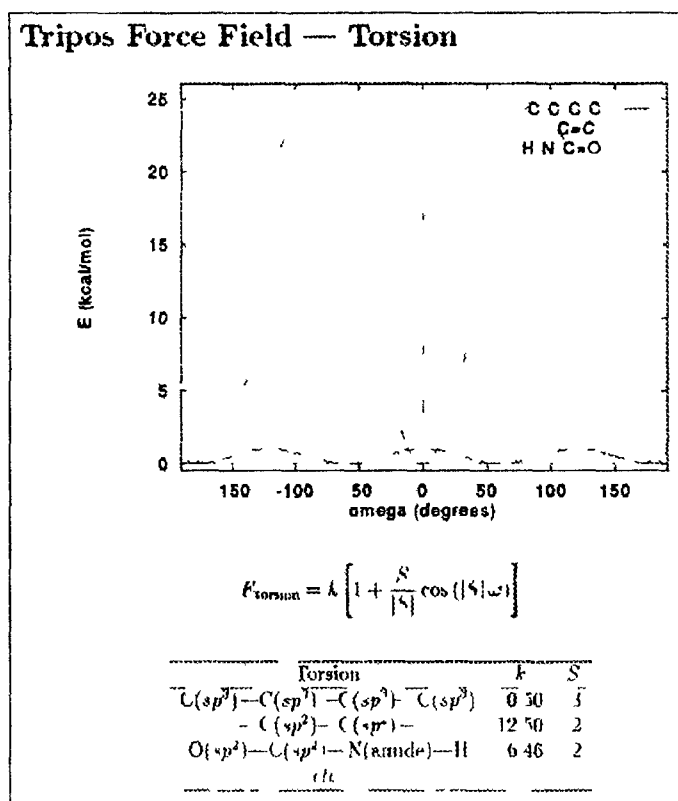


Рисунок 7

Пример из силового поля TRIPOS. Зависимость энергии торсионных углов от величины ω и соответствующие силовые параметры для трех 4-х атомных систем $C(sp^3)-C(sp^3)-C(sp^3)-C(sp^3)$, $C(sp^3)-C(sp^2)-C(sp^2)-C(sp^3)$, $O(sp^2)-C(sp^2)-N(\text{amide})-H$.

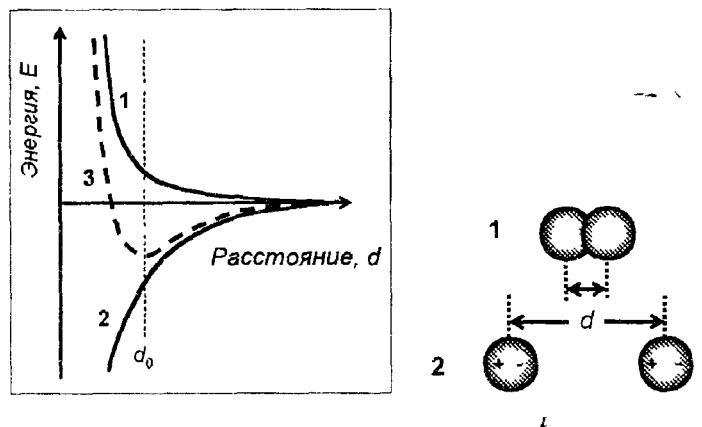


Рисунок 8

Зависимость энергии взаимодействия Ван-дер-Ваальса от расстояния между атомами i и j
 1 - электронное отталкивание, 2 - дисперсионное взаимодействие (притяжение между двумя индуцированными диполями - взаимодействие Лондона), 3 - результирующая кривая зависимости энергии взаимодействия Ван-дер-Ваальса от расстояния между двумя атомами

2.2.4. Энергия Ван-дер-Ваальса. Взаимодействие Ван-дер-Ваальса (VDW) происходит между атомами, находящимися на близком расстоянии но не связанными химическими связями между собой (взаимодействие между атомами i и j , разделенными двумя или более связями) Оно является одним из основных факторов, обуславливающих плотную упаковку атомов в молекулярной системе

Взаимодействие обусловлено двумя противоположно действующими эффектами (рис 8), зависящими от размера атомов, описываемых так называемыми радиусами Ван-дер-Ваальса, и от расстояния между центрами атомов (d)

1) электронное отталкивание, относящееся к близкодействующим эффектам, так как пропорционально расстоянию между атомами (d) в -12 степени (рис 8, кривая 1)

$$E = \varepsilon \left(\frac{d_0}{d}\right)^{12} \quad (11)$$

где ε - константа (~ 0.2 ккал/моль),

d_0 - оптимальное расстояние, равное примерно сумме радиусов VDW (порядка 3,5 Å)

2) дисперсионное взаимодействие, известное также как взаимодействие Лондона, - притяжение между двумя индуцированными диполями и пропорциональное расстоянию между атомами в -6 степени, (рис 8, кривая 2)

$$E = -4\varepsilon \left(\frac{d_0}{d}\right)^6 \quad (12)$$

Результирующая энергия E_{vdw} является суммой двух эффектов (рис 8, кривая 3) и для одной пары атомов она равна $\sim 0,1$ ккал/моль E_{vdw} описывается уравнением Леонарда-Джонсона, часто называемым также потенциалом "6-12" (название произошло от степеней двух членов уравнения)

$$E_{vdw} = \Sigma \varepsilon \left[\left(\frac{r_i + r_j}{d}\right)^{12} - 2 \left(\frac{r_i + r_j}{d}\right)^6 \right] \quad (13)$$

где ε - константа,

r_i и r_j - радиусы VDW атомов i и j ,

d - расстояние между центрами атомов

Довольно часто член "-12" заменяют экспоненциальной функцией, как, например, в потенциале Букингема

$$E_{vdw} = \Sigma A e^{B/d} - C d^{-6} \quad (14)$$

где A , B и C - константы

Tripos Force Field — Van der Waals

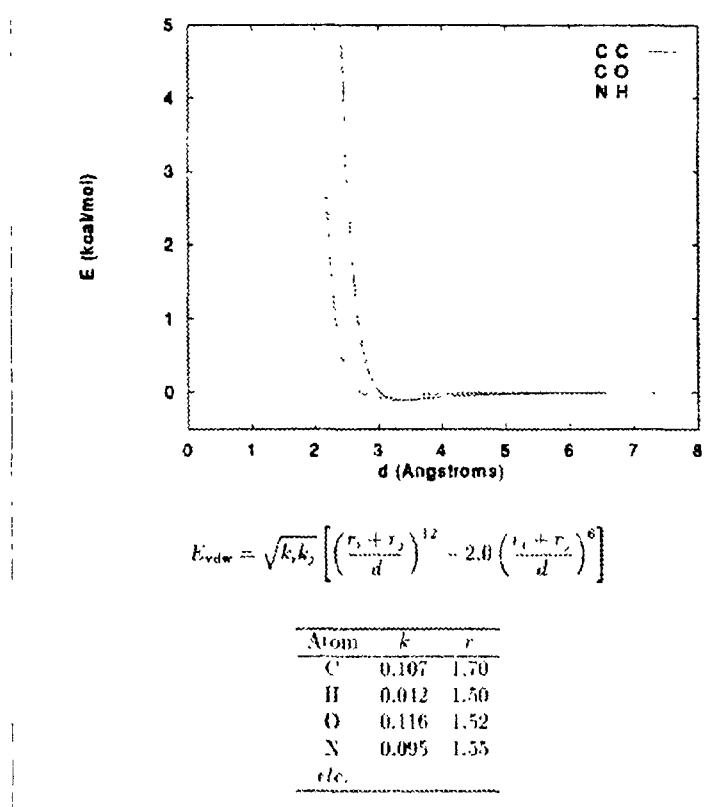


Рис 9.

Пример из силового поля TRIPOS. Зависимость энергии VDW от расстояния между атомами и соответствующие значения радиусов VDW и константы для атомов C, H, O, N.

На рисунке 9 в качестве примера из силового поля TRIPOS приведены зависимости энергии VDW от расстояния между атомами, и соответствующие значения радиусов VDW и константы для атомов C, H, O, N.

2.2.5. Энергия электростатических взаимодействий. Для описания электростатических взаимодействий в молекулярной механике используется классический закон Кулона

$$E_{ij} = \sum \frac{q_i \cdot q_j}{D \cdot d} \quad (15)$$

где q_i и q_j - порционные заряды атомов i и j ,

D - диэлектрическая константа,

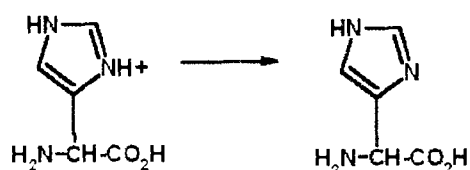
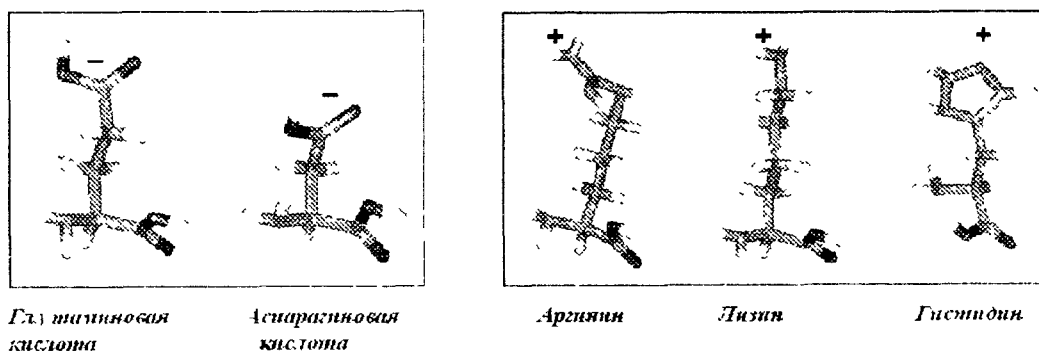
d - межатомное расстояние

Как известно, в белках встречаются аминокислоты, несущие как отрицательные, так и положительные заряды (рис. 10). Гистидин при нейтральных pH имеет положительный заряд, в тоже время при щелочных pH он теряет протон и становится нейтральным. Заряды присутствуют также на концах пептидной цепи и у отдельных аминокислот в растворе.

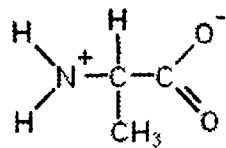
Положительно заряженные остатки лизина и аргинина могут вступать в сильные электростатические взаимодействия с отрицательно заряженными группами аспарагиновой и глутаминовой кислот. Эти взаимодействия в белках носят название солевых мостиков (рис. 11).

Электростатическое взаимодействие играет также важную роль при связывании белками неорганических ионов, которые могут играть ключевую роль в механизмах функционирования белков (рис. 12).

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ХИМИИ



Потеря положительного заряда на гистидине при переходе от нейтральных значений pH к щелочным



Изолированная аминокислота аланин в растворе.

Рисунок 10
Заряженные аминокислоты

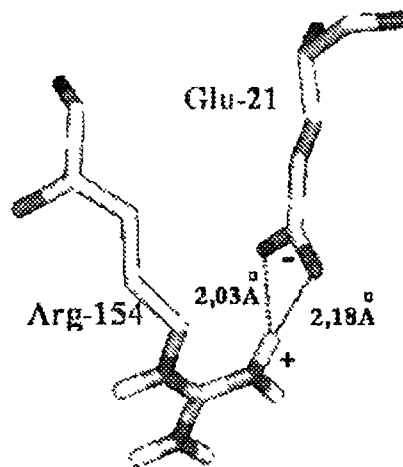


Рисунок 11

Солевой мостик в структуре химотрипсина. Электростатическое взаимодействие отрицательно заряженного Glu-21 с положительно заряженным Arg-154

2.2.6. Другие виды взаимодействий. 1) Водородные связи. Энергия водородного взаимодействия (E_H) описывается уравнением

$$E_H = \Sigma \left(\frac{X}{d^{1/2}} \right) - \left(\frac{Y}{d^{1/2}} \right) \quad (16)$$

где X и Y – константы,
 d – расстояние

В белках в водородном взаимодействии могут участвовать различные группы как основной, так и боковых цепей (рис 13)

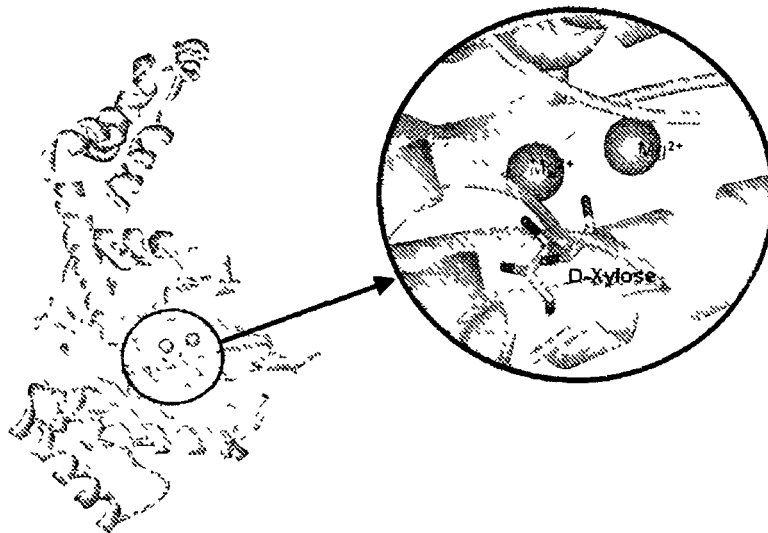
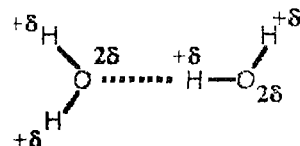


Рисунок 12

Пример связывания неорганических ионов с белком. Ксилоизомераза содержит два иона Mg^{2+} в активном центре для координации и поляризации субстрата, а также для стабилизации переходного состояния в ходе катализируемой реакции



Водородное взаимодействие между молекулами воды

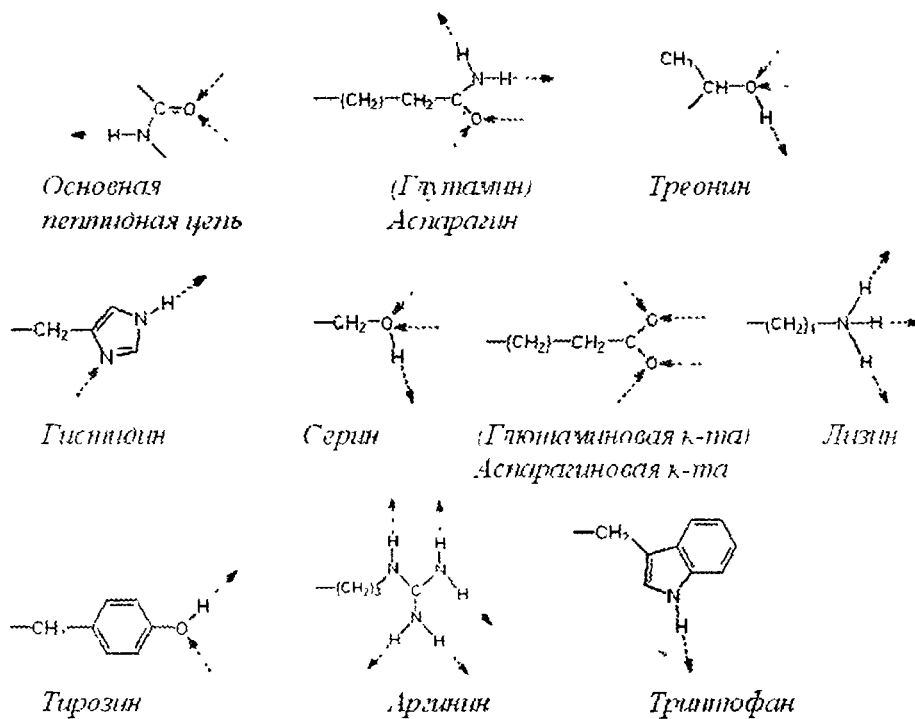


Рисунок 13

Доноры и акцепторы водородных связей в белках. Стрелками показаны возможные направления водородных связей для доноров и акцепторов

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ХИМИИ

Водородные связи играют большую роль в стабилизации вторичных структур в белках (рис 14)

2) Нарушение планарности При искажении планарных как правило ароматических, структур возникают напряжения и энергия молекулярной системы возрастает (рис 15)

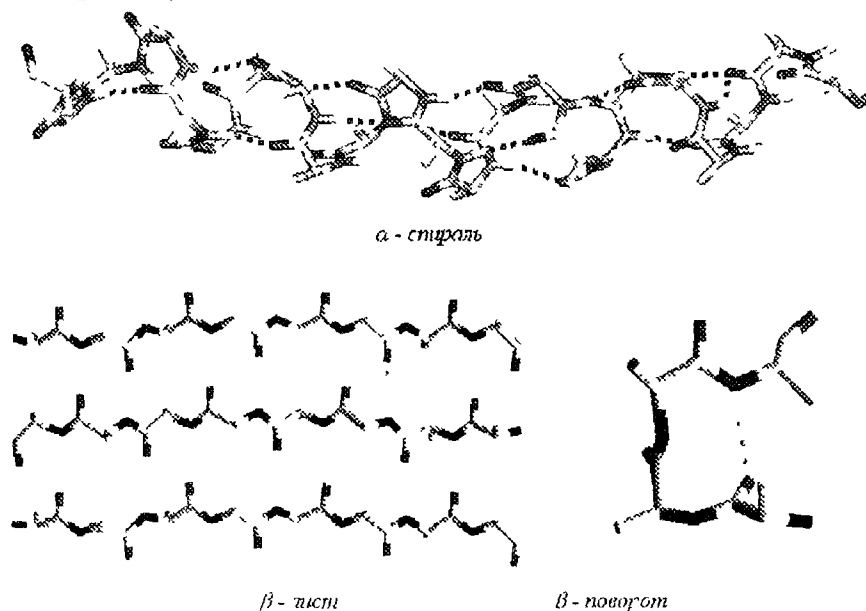


Рисунок 14

Примеры вторичных структур, стабилизированных водородными связями

Для учета этого энергетического компонента вводится дополнительный член (E_{oop} , *out of plane*)

$$E_{oop} = \sum k\chi^2 \quad (17)$$

где k - константа,

χ - отклонение C атома от плоскости, образованной атомами 1, 2 и O (см рис 15)

2.2.7. Комбинированные энергетические члены В молекулярной механике все энергетические члены рассматриваются как независимые, т.е. изменение длины химической связи не влияет на углы и т.д. Однако в реальных молекулярных системах это не так. На рисунке 16 в качестве примера приведены две структуры бутана в энергетически минимальных цис и транс конформациях, где различие величин углов и длины химической связи свидетельствуют о их взаимозависимости.

Для учета этих явлений в силовых полях вводят так называемые комбинированные энергетические члены (cross-terms). Так для приведенного выше примера о взаимосвязи длины химической связи и углов в молекуле бутана уравнение будет выглядеть так

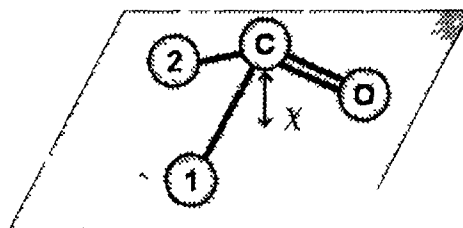


Рисунок 15

Пример нарушения планарности в системе из 4 атомов (карбонил), которые в равновесном состоянии должны быть в одной плоскости

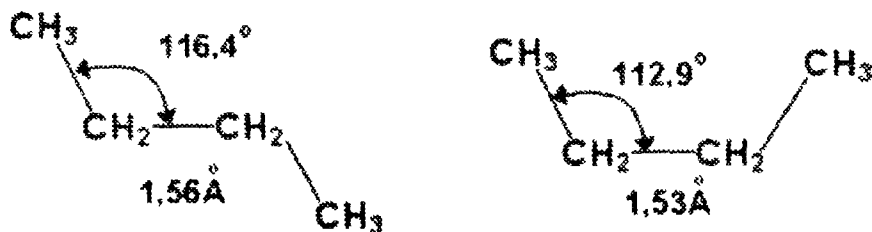


Рисунок 16

Две структуры бутана в энергетически минимальных *цис* и *транс* конформациях демонстрируют взаимозависимость величин углов и длины химической связи

$$E_{\text{св-зв}} = k(d - d_0)(\Theta - \Theta_0) \quad (18)$$

2.2.8. Относительная энергетическая оценка взаимодействий. В таблице 2 приведена сравнительная шкала (по порядку величины энергии) для различных видов взаимодействий и преодоления энергетических барьеров. Можно видеть, что кардинальное изменение структуры молекулы за счет поворота вокруг двойной связи (например, переход *цис-транс*) или разрыва ковалентной связи требует огромных энергетических затрат, на порядки превосходящие изменение энергии молекулярной системы при небольших конформационных перестройках.

2.3. Параметризация. Основной проблемой молекулярной механики являются сотни силовых параметров, используемых в эмпирических уравнениях расчета энергетических составляющих молекулярной системы. Общее число необходимых параметров значительно превышает число атомов в системе.

Как уже упоминалось выше, атомы классифицируются в соответствии с типом гибридизации, химическим окружением и типом связей. Для расчетов молекулярной механики нужны параметры для каждого типа атома, связанного с каждым другим типом атомов, и для всех типов связей (одиночная, двойная, тройная, ароматическая, амидная, и т.д.). Ситуация становится еще более сложной при рассмотрении торсионных параметров, комбинированных параметров и т.д.

Откуда берутся все эти многочисленные параметры?

Основными источниками параметров для химических связей и углов являются

- 1) рентгеноструктурный анализ высокого разрешения небольших молекул;
- 2) спектроскопические данные;
- 3) теоретические расчеты силовых констант методами квантовой механики (*ab initio*).

Первоначально выбранные параметры затем оптимизируются с помощью итерационных вычислений для воспроизведения известных экспериментальных данных, таких как геометрия молекул, теплота образования, частоты колебаний, торсионные барьеры, параметры кристаллической решетки, энергия сублимации, геометрия и энергия водородных связей, свойства жидкостей (плотность, теплоты испарения, и т.д.) и свободные энергии сольватации.

Таблица 2 Сравнительная шкала взаимодействий

| Вид взаимодействия | Порядок величины энергии (ккал/моль) |
|---|--------------------------------------|
| Взаимодействие Ван-дер-Ваальса в воде для 2-х атомов | - 0,1 |
| То же в вакууме | - 0,3 |
| Водородная связь в воде | - 1,0 |
| Водородная связь в вакууме | - 5,0 |
| Изменение угла между связями на 10° | 2,0 |
| Изменение длины связи на 0,1 Å | 2,5 |
| Барьер при вращении торсионного угла вокруг одиночной связи | 3,0 |
| Барьер при вращении торсионного угла вокруг двойной связи | 20 |
| Барьер разрыва одиночной связи | 100 |

2.4. Оптимизация структуры молекулы (минимизация энергии). Так как энергия молекулярной системы является функцией координат атомов, то программы молекулярной механики пытаются сгенерировать координаты, соответствующие структуре с минимальной энергией, что достигается с помощью процедуры минимизации. Все используемые для этой цели методы минимизации называются методами последовательного спуска. Они представляют собой итерационные вычислительные процедуры, в которых координаты атомов изменяются при каждой итерации с целью последовательного снижения величины энергии.

Более подробно вопрос минимизации энергии молекулярных систем рассматривается в отдельной лекции.

2.5. Известные силовые поля. Существует довольно много различных силовых полей. Ряд из них относится к универсальным полям общего назначения, которые имеют мало параметров и относительно низкую точность (например, силовые поля DREIDING, UFF). Другие поля являются узко специализированными с большим числом параметров и высокой точностью (например, силовые поля MM2, MMX и MM3 для небольших углеводородных молекул; поля CVFF, PCFF, специализированные для фторсодержащих углеводородов). Для решения молекулярно-биологических задач наибольшее распространение получили силовые поля TRIPOS, AMBER, GROMOS, CHARMM и некоторые другие.

3. Квантовая механика

Данный подход основан на разработанном Эрвином Шредингером списании движения электрона в атоме с помощью трехмерных волновых функций (вероятность распределения электронов в атоме описывается волновым процессом подобно звуковым волнам).

Общий вид уравнения Шредингера:

$$H\psi = E\psi \quad (19)$$

где: H – гамильтониан,

ψ – волновая функция,

E – энергетический оператор.

Если мы знаем все волновые функции системы, то, в принципе, мы можем вычислить любую наблюдаемую физическую характеристику этой системы.

Волновая функция атомов описывается в виде комбинации атомных орбиталей (χ_i):

$$\psi = \chi_1 \chi_2 \chi_3 \dots \chi_n \quad (20)$$

В свою очередь волновая функция молекул описывается комбинацией молекулярных орбиталей (ϕ_i):

$$\psi = \phi_1 \phi_2 \phi_3 \dots \phi_n \quad (21)$$

где ϕ – линейная комбинация атомных орбиталей:

$$\phi_i = \sum C_{ik} \chi_k \quad (22)$$

где C_{ik} – орбитальный коэффициент.

Существует два подхода в вычислениях методами квантовой механики:

- 1) вычисления *ab initio*,
- 2) полуэмпирические вычисления.

С помощью данных подходов могут быть вычислены самые различные параметры молекулярной системы: молекулярные энергии, геометрия, заряды, электронные плотности, высшие заполненные молекулярные орбитали (HOMO), нижние свободные молекулярные орбитали (LUMO) и т.д.

3.1. Расчеты *ab initio*. Все атомные орбитали описываются экспоненциальными функциями или функциями гауссиана, приведенными для простоты вычислений к экспоненциальной форме.

Число функций определяет качество базового набора и, следовательно, точность любого вычисления.

Наиболее известные методы расчета *ab initio*: STO-3G, STO-6G, 3-21G, 6-31G, MIDI, MINI и DZP.

Основная проблема, ограничивающая применимость данного подхода, состоит в том, что число вычислений связано с числом базовых функций в четвертой степени (N^4). Т.е. даже для малых молекул необходимо выполнять миллионы вычислений.

3.2. Полуэмпирические расчеты. Данный подход базируется на использовании ряда приближений для уменьшения числа вычислений. В основном это игнорирование электронов внутренних орбиталей, которые не вовлекаются в образование межатомных связей. Соответствующие интегралы заменяются параметрами, полученными из экспериментальных данных.

Наиболее известные полуэмпирические методы: CNDO, MNDO, AM1 и PM3.

4. Комбинированные подходы

Разработка точных и надежных методов представления химических систем с помощью комбинации методов квантовой и молекулярной механики представляет одну из главных задач вычислительной химии в последнее десятилетие. Такое представление позволяет осуществлять точное квантово-механическое моделирование небольших, ограниченных участков больших молекулярных систем (например, моделирование реакционно-способных групп в функционально важных областях), в то время как основная часть системы моделируется молекулярной механикой. Таким образом удастся увеличить точность и надежность моделирования функционально важных участков в сложных молекулярных системах без значительного увеличения объема вычислений.

Ключевой проблемой в реализации комбинированного подхода является создание точной и аккуратной границы между областями системы, описываемыми разными методами. Данная проблема очень сложна и несмотря на существующие значительные успехи до настоящего времени полностью пока не решена.

Одним из наиболее эффективных подходов является метод Хартри-Фока (Hartree-Fock). Он начинается с установления областей молекулы, в которых действуют квантовая и молекулярная механики. Граница между этими областями описывается так называемыми "замороженными орбиталями", которые ограничиваются определением одиночных связей. Замороженные орбитали получают из специальных расчетов малых модельных молекул. Между двумя областями, описываемыми разными методами, моделируются электростатические взаимодействия и взаимодействия Ван-дер-Ваальса.

5. Рекомендуемая литература и источники в сети Интернет

1. В.Г. Дашевский (1982) Конформационный анализ органических молекул, Химия, М.
2. В.Г. Дашевский (1982) Конформационный анализ биополимеров, Химия, М.
3. У. Буркерт, Н. Эллинджер Молекулярная механика (пер. с англ) (1984), Мир, М.
4. Т. Кларк (1990) Компьютерная химия (пер. с англ), Мир, М.
5. J.-P. Doucet, J. Weber (1996) Computer-Aided Molecular Design: Theory and Applications, Academic Press, London.
6. K. Anthony K. Rappe, Carla J. Casewit, John Choi (1997) Molecular Mechanics Across Chemistry, University Science Books.
7. K.B. Lipkowitz, D.B. Boyd, (eds) Reviews in Computational Chemistry (Volumes 1- 14). A book series of reviews and tutorials. Dep. Chemistry IUPUI, Indiana 46202-3274, USA.
8. Murphy R. B., Philipp D. M., Friesner R. A. (2000) J. Comp Chem., **21**(16), 1442-1457.
9. Oliver S. Smart Course "Principles of Protein Structure Using the Internet" <http://www.cryst.bbk.ac.uk/PPS2/course/>

10. Training Course: Molecular Design and Computer Assisted Combinatorial Chemistry <http://www.ics.trieste.it/chemistry/combinatorial/tc-moleculardesign2002.htm>

Поступила 01.06.2004

**BASIC PRINCIPLES OF COMPUTATIONAL CHEMISTRY
FOR MEDICO-BIOLOGISTS**

A.S. Ivanov

V.N.Orekhovich Institute of Biomedical Chemistry RAMS,
Pogodinskaya str 10, Moscow. 119121 Russia
fax: (095) 245-0857, e-mail: ivanov@ibinh.msk.su

The lecture describes the basic principles of computational chemistry underlying the methods of molecular modelling used in bioinformatics area. The basic positions, methods of molecular and quantum mechanics and combined approaches are considered. This lecture is from theoretical cycle "Bioinformatics and Computer-Aided Drug Design" for fourth year students of Medico-Biological Department of Russian State Medical University (specialty - biochemistry, biophysics and medical cybernetics). It can also be advised for all students and post-graduate students of medico-biological specialties.

Key words: computational chemistry, molecular modeling, molecular mechanics, quantum mechanics