

Лекция 4

Модель конформационной энергии макромолекул.

Силовое поле. Трансферабельность параметров силового поля.

Способы расчета и оптимизации параметров силового поля.

Примеры силовых полей: AMBER94, OPLS, CHARMM, GROMOS96

Точность силовых полей.

Модель конформационной энергии макромолекул

Сумма энергий всех возможных групп валентно связанных атомов –

- деформация валентного остова

$$U = U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \Sigma \text{ валентные связи (12)} + \Sigma \text{ валентные углы (123)}$$

- внутреннего вращения

Σ квартеты (1234)

- энергия невалентных взаимодействий -
сумма по атомам разделенным по цепи валентных связей
+ Σ пары атомов разделенные 3 + связями

Аналитические модели потенциалов этих энергий получены ранее

$$U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{bond} \frac{1}{2} K_{bi} (b_i - b_{0i})^2 + \sum_{angle} \frac{1}{2} K_{\alpha i} (\alpha_i - \alpha_{0i})^2 +$$
$$+ \sum_{tors} \frac{V_{tk}}{2} [1 + \cos(n_k \phi + \delta_k)]$$
$$+ \sum_{i < j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right] + \sum_{i < j} \frac{q_i q_j}{r_{ij}}$$

Выражение для потенциальной энергии молекулы в методе молекулярной механики имеет большое количество параметров силового поля

Силовое поле

Параметры силового поля

- равновесная (основное состояние) структура идеального валентного остова
 \mathbf{b}_i – длины связей, α_i – величины валентных углов
- константы деформации длин связей и величин валентных углов
 K_{b_i} , K_{α_i}
- параметры торсионных потенциалов внутреннего вращения,
 V_i , n_i , δ_i ,
- параметры ван дер Ваальсовых взаимодействий пар атомов
 A_{ij} , B_{ij} или $(\epsilon_{ij}, R_{m_{ij}})$
- заряды на атомах q_i

Невозможно для каждой молекулы рассчитывать свои параметры – трудоемко.

Универсальное силовое поле - для классов молекул

- биополимеры
- классы лигандов

Универсальность силового поля основана на

- приближение атомов в молекуле
- гипотеза трансферабельности – (подобные атомы сохраняют свои свойства в разных молекулах)

Атомы в молекуле и трансферабельность

Молекула может быть разбита на атомы.

Максимальное количество сортов атомов –

- каждый атом в разных молекулах уникален.

Минимальное количество сортов атомов –

- число химических сортов H, C, ...

Разумный компромис –

- множество сортов атомов в классе молекул определяется с учетом уникальности и гипотезы трансферабельности

«атом в молекуле» =

- химический сорт + тип электронной гибридизации + тип ближайшего валентного окружения + тип химической группы

Например силовое поле amber94 для белков и нуклеиновых кислот
(*J. Am. Chem. Soc.* 1995, v.117, n19, p.5179-91)

имеет 13 атомов C, 7 атомов N, 5 атомов O, 12 атомов H

атомы в молекуле силовое поле **amber**

atom	Type “atom in molecule”	description
------	----------------------------	-------------

Carbon	CT	sp ³ с 4-вал связями
	C	sp ² карбонильный С
	CA	sp ² ароматический
	CM	sp ² атом с двойной связью
	CC	sp ² ароматический в 5-членном кольце с одним заместителем и связанный с азотом N
	CV	sp ² ароматический в 5-членном кольце связанный с атомом С и азотом N (3-ех вал) с неподеленной эл. парой :N<
	CW	
	CR	sp ² ароматический в 5-членном кольце связанный с атомом С и азотом NH
	CB	sp ² ароматический в 5-членном кольце связанный с двумя атомами N и N
	C*	sp ² ароматический в 5-членном кольце в месте сочленения с 6-членным кольцом
	CN	sp ² ароматический в 5-членном кольце связанный с двумя атомами С
	СК	sp ² ароматический при сочленении 5- и 6-ти членных колец и связанный с атомами СН и NH
	СQ	sp ² ароматический при сочленении 5- и 6-ти членных колец и связанный с атомами N и NR
азот	N	sp ² ароматический в 6-ти членном кольце и связанный с двумя атомами N и N с неподеленными эл. Парами
	NA	sp ² в амидах sp ² в ароматических кольцах связанный с Н

Способы расчета и оптимизации параметров силового поля

Схема определения параметров силового поля

- Класс молекул для описания силовым полем – белки + нуклеиновые кислоты
- Репрезентативный набор молекул и молекулярных фрагментов для данного класса
- Разбиение на типы атомов в молекуле

Необходимы

- структурные данные молекул репрезентативного набора для определения параметров идеального валентного остова
- энергетические характеристики молекул репрезентативного набора

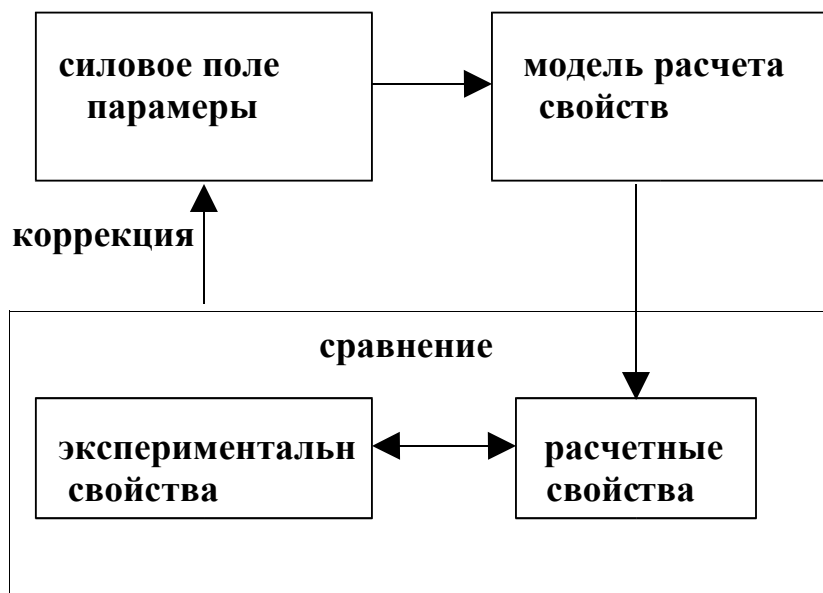
Источники данных

- экспериментальные данные
- высококачественные расчеты *ab initio* методами квантовой механики

Экспериментальные данные –

- атомные структуры высокого разрешения (высокой точности) метода рентгеноструктурного анализа монокристаллов прямо дают параметры идеальной валентной структуры
- спектроскопические данные – спектры ИК, видимой обл. и УФ источник частот колебаний связей и валентных углов
косвенно -> константы деформаций связей и углов
- спектры радиочастотные – заторможенные торсионные колебания
косвенно -> величины барьеров внутреннего вращения
- термофизические данные молекулярных кристаллов молекул репрезентативного набора – энергии сублимации
- термофизические данные для молекулярных жидкостей из молекул репрезентативного набора
косвенно -> энергии невалентных атом-атомных взаимодействий
- электрические поляризуемости молекул репрезентативного набора
- дипольные и квадрупольные моменты связей молекул репрезентативного набора
косвенно -> заряды на атомах, ван дер ваальсовы потенциалы атом-атомных взаимодействий

Решается оптимизационная задача
подбор параметров силового поля обеспечивающий **минимальное отклонение** расчетных свойств от наблюдаемых в эксперименте



Ab initio квантово-механические расчеты –

- равновесная структура молекул
- константы деформации валентных связей и углов
- величины барьеров внутреннего вращения
- заряды на атомах

По мере расширения возможностей кв.механических методов они будут все шире привлекаться к созданию силовых полей нового поколения.

Например, силовое поле MMFF94 впервые построено на большом массиве (более 2000) предельно качественных (HF/6-31G*, MP2/6-31G*) расчётов молекулярных характеристик квантово-механическими методами.

[Halgren, T.A. Merck molecular force field. 1. Basis, form, scope, parametrization and performance of MMFF94. *J. Comp. Chem.* **17**:490-519, (1996).]

Современные силовые поля

Amber

CHARMM

Gromos96

Gromacs

OPLS

Имеются пакеты программ с соответствующими названиями которые выполняют симулирование основных классов биополимеров в соответствующих силовых полях

Точность перечисленных силовых полей примерно одинакова с небольшим преимуществом **AMBER**

Ошибки воспроизведения

- длин связей $\sim 0.015 \text{ \AA}$,
- валентных углов $\sim 1.2^\circ$,
- барьеров вращения ~ 0.39 ккал/моль,
- воспроизведение энергий конформаций на аланин-трипептиде простой молекуле $\sim 1.2-1.4$ ккал/моль относительно высококачественных кв.расчетов
- на фрагментах нуклеиновых кислот расчет стабильности 26 пар нуклеиновых оснований и энергии стекинг взаимодействий 10 типов минидуплексов **AMBER** наиболее точно воспроизводит результаты квантово-механических расчётов, средним абсолютным отклонением 0.9 ккал/моль

Основные источники погрешности современных силовых полей второго поколения

- пренебрежение электростатической поляризуемостью атомов локальным электростатическим полем
- грубая форма ван дер ваальсова потенциала и комбинационные правила для параметров
- пренебрежение трехчастичными ван дер ваальсовыми взаимодействиями
- ограниченность монополярной аппроксимации электростатических взаимодействий

каждый из отмеченных моментов может вносить погрешности $\sim 10\%$ в конформационной энергии макромолекулы

Резюме:

достигнутая точность силовых полей вполне позволяет применять методы моделирования для решения широкого круга задач структурной биологии

Точность современных силовых полей

Точность описания экспериментальных данных силовыми полями может быть оценена на примере силового поля MMFF94, которое имеет ошибки воспроизведения длины связей $\sim 0.014 \text{ \AA}$, валентных углов $\sim 1.2^\circ$, частот колебаний $\sim 61 \text{ см}^{-1}$ ($< 10\%$), конформационных энергий $\sim 0.38 \text{ ккал/моль}$ и барьеров вращения $\sim 0.39 \text{ ккал/моль}$ на большом наборе малых молекул. Однако, при описании конденсированных систем, малые ошибки в потенциалах межмолекулярных взаимодействий могут транслироваться в большие ошибки в рассчитанных термодинамических параметрах. Тесты на конденсированных системах показывают, что силовое поле MMFF94 описывает термодинамические характеристики жидких метана, этана, н-бутана, н-пентана и н-октана с неприемлемой ошибкой ($> 10\%$), тогда как поле OPLS-AA оказывается наилучшим. С другой стороны, тесты на малой биомолекуле, аланин-трипептиде, показывают, что силовые поля MMFF94, AMBER [83] и OPLS-AA воспроизводят энергии 10 конформаций со среднеквадратичной ошибкой 1.2-1.4 ккал/моль, причем «системой отсчета» служили высококачественные квантово-механические расчёты. Найдено также, что аппроксимация электростатических взаимодействий зарядами на атомах недостаточно точна для описания водородных связей. Силовые поля AMBER, CFF95, CVFF, CHARMM, OPLS, POLTEV были протестированы на фрагментах нуклеиновых кислот. Были рассчитаны стабильности 26 пар нуклеиновых оснований и энергии стекинг взаимодействий 10 типов минидуплексов в вакууме, и сравнены с высококачественными квантово-механическими расчётами в расширенном базисе. Найдено, что силовое поле AMBER наиболее точно воспроизводит результаты квантово-механических расчётов, с коэффициентом корреляции 0.98, коэффициентом линейной регрессии 1.10 и средним абсолютным отклонением 0.9 ккал/моль.

Основные источники погрешности

силовых полей второго поколения являются следующие моменты: 1) пренебрежение не аддитивными взаимодействиями, именно, электростатической поляризуемостью атомов локальным электростатическим полем и трехчастичными ван дер ваальсовыми взаимодействиями; 2) форма ван дер ваальсова потенциала и комбинационные правила для параметров; 3) ограниченность монополярной аппроксимации электростатических взаимодействий. Эти погрешности могут не проявляться при оптимизации параметров силового поля на конденсированных гомогенных системах, однако, каждый из отмеченных моментов может вносить погрешности $\sim 10\%$ в гетерогенной системе - макромолекула в водном растворе. Силовое поле следующего поколения должно быть простым по количеству параметров, но правильно отражать реалии физических взаимодействий.