



# **Молекулярное моделирование в применении к биомолекулам**

## **Лекция 8**

**Модификации молекулярной динамики и  
введение в методы Монте-Карло**

# Молекулярная динамика

Расчёт сумма сил действующих  
на атом

$$F_i = \sum_j F_{ij}$$

$\Delta t$

Расчет новых координат или  
интегрирование

$$\frac{\partial^2 r_i}{\partial t^2} = \frac{F_i}{m_i}$$

$$\frac{\partial r_i}{\partial t} = v_i; \quad \frac{\partial v_i}{\partial t} = \frac{F_i}{m_i}$$

# Интегратор leap-frog

Leap-frog самый быстрый вариант алгоритма Верле:

$$r(t + \delta t) = r(t) + \delta t v(t + \frac{1}{2} \delta t)$$

$$v(t + \frac{1}{2} \delta t) = v(t - \frac{1}{2} \delta t) + \delta t a(t)$$

*Тогда скорость в момент  $t$ :*

$$v(t) = \frac{1}{2} [v(t + \frac{1}{2} \delta t) + v(t - \frac{1}{2} \delta t)]$$

# Молекулярная динамика и неявный растворитель

Часто явно заданный растворитель, т. е. молекулы воды заменяют потенциалами.

- Методы на основе поверхности доступной растворителю.

$$\Delta G_{solv} = \sum_i \sigma_i ASA_i$$

- Метод Пуассона-Больцмана и его производные

# Молекулярная динамика и неявный растворитель

- Метод Пуассона-Больцмана точен, но технически очень сложен для счёта.

$$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} \Psi(\vec{r})] = -\rho^f(\vec{r}) - \sum_i c_i^\infty z_i q \lambda(\vec{r}) e^{\frac{-z_i q \Psi(\vec{r})}{kT}}$$

- Generalized Born и GBSA

$$G_s = \frac{1}{8\pi} \left( \frac{1}{\epsilon_0} - \frac{1}{\epsilon} \right) \sum_{i,j} \frac{q_i q_j}{f_{GB}} \quad f_{GB} = \sqrt{r_{ij}^2 + a_{ij}^2} e^{-D} \quad , \quad D = \left( \frac{r_{ij}}{2a_{ij}} \right)^2 \quad , \quad a_{ij} = \sqrt{a_i a_j}$$

# Молекулярная динамика и неявный растворитель

## Основные недостатки неявного растворителя:

- РВ и GB учитывают в основном электростатическую составляющую.
- Гидрофобный эффект не учитывается.
- Вязкость как результат столкновений и скоростей не рассчитывается и не учитывается
- Водородные связи воды с объектом интереса не могут быть учтены. Исчезает возможность учёта водных мостиков.

# Конформации в молекулярной динамике

Обычная молекулярная динамика это тепловое движение молекулы.

- Основной компонент такого движения это высокочастотные гармонические колебания атомов.
- Интерес представляют низкочастотные движения больших частей молекулы.
- Экстракция таких движений проводят с помощью Фурье преобразований.
- Низкочастотные колебания молекулы называют основной или существенной динамикой

# Моделирование амфифильных молекул

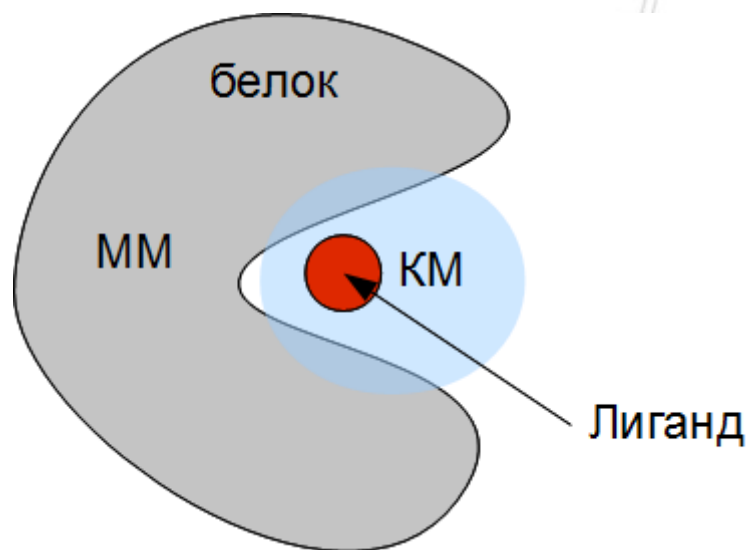
Амфифильные молекулы: фосфолипиды, гликолипиды и т.д.

- Ключевая особенность - образование разных фаз
- Возможно образование разных жидко-кристаллических фаз с выраженным порядком вдоль плоскости фазы.
- Параметр порядка :  $S = 1/2 \langle 3 \cos \phi_i \cos \phi_j - \delta_{ij} \rangle$
- Нередко используют контроль площади поверхности через поверхностное натяжение.
- К большим системам применяют крупно-зернистое моделирование.



# Гибридное QM/MM моделирование

- Основная идея: разделить большую систему на квантовую и молекулярную части.
- Электростатическое окружение из MM части чувствуется QM частью.
- MM часть принимает силы из QM части и соответственно адаптируется.



$$S = 1/2 \langle 3 \cos \phi_i \cos \phi_j - \delta_{ij} \rangle$$

# Гибридное QM/MM моделирование

- Простейший Гамильтониан для QM/MM системы:

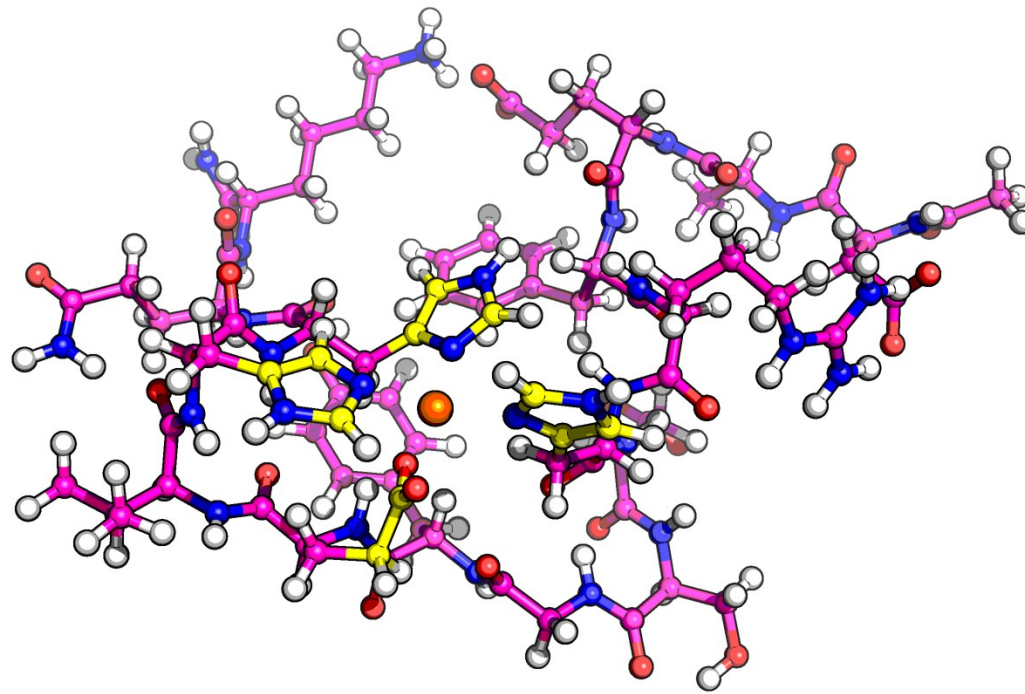
$$H = -1/2 \sum_i^{electr} \nabla^2 - \sum_i^{elect} \sum_j^{nucl} \frac{Z_j}{R_{ij}} + \sum_i^{elect} \sum_j^{elect} \frac{1}{r_{ij}} + \sum_i^{nucl} \sum_j^{nucl} \frac{Z_j Z_i}{R_{ij}} - \sum_i^{elect} \sum_j^{MM_{Q_j}} \frac{Q_j}{R_{ij}} + \sum_i^{nucl} \sum_j^{MM_{Q_j}} \frac{Z_i Q_j}{R_{ij}}$$

- К QM/MM части можно добавить и VdW составляющую

$$H_{QM/MM} = - \sum_i^{elect} \sum_j^{MM_{Q_j}} \frac{Q_j}{R_{ij}} + \sum_i^{nucl} \sum_j^{MM_{Q_j}} \frac{Z_i Q_j}{R_{ij}} + \sum_i^{nucl} \sum_j^{MM_{atoms}} 4 \epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right]$$

# Гибридное QM/MM моделирование

- Атомы связи
- Специальные орбитали



# Метод обмена репликами (REMD)

- Основная идея: запустить параллельно несколько счётов с разными температурами.
- Если разница между температурами не большая, то гистограммы распределения энергии должны пересекаться.
- Мы можем выбрать правило когда производить обмен конформациями.
- Если мы проводим обмен когда потенциальная энергия одной из реплик ниже чем других, то это похоже на моделирование отжига. Такой подход часто используется для моделирования самосборки.

# Метод Монте-Карло

- Основная идея предполагает поиск конформаций с низкой энергией на основе случайного изменения координат.
- Скорости у атомов не рассчитываются

Представим, что канонический ансамбль можно разбить:

$$Q_{NVT} = Q_{NVT}^{ideal} Q_{NVT}^{exess} \quad \text{где} \quad Q_{NVT}^{exess} = \frac{1}{V^N} \int dr^N \exp \left[ \frac{-U(r^N)}{k_B T} \right]$$

Энергия Гемгольца :

$$A = -k_B T \ln Q_{NVT} \rightarrow A = A^{ideal} + A^{exess}$$

# Расчёт термодинамических свойств

- Формально нам надо интегрировать по всему пространству конформаций, это невыполнимо.

Приближение:

1. Генерируем конформацию
2. Считаем потенциальную энергию  $\exp\left[\frac{-U(r^N)}{k_B T}\right]$
3. Считаем фактор Больцмана.
4. Суммируем известные факторы и находим среднее и переходим к пункту 1.

В результате средняя энергия определяется как:

$$\langle U(r^N) \rangle = \frac{\sum_1^{N_{trial}} U_i(r^N) \exp[U(r^N)/k_B T]}{\sum_1^{N_{trial}} \exp[U(r^N)/k_B T]}$$

# Расчёт термодинамических свойств

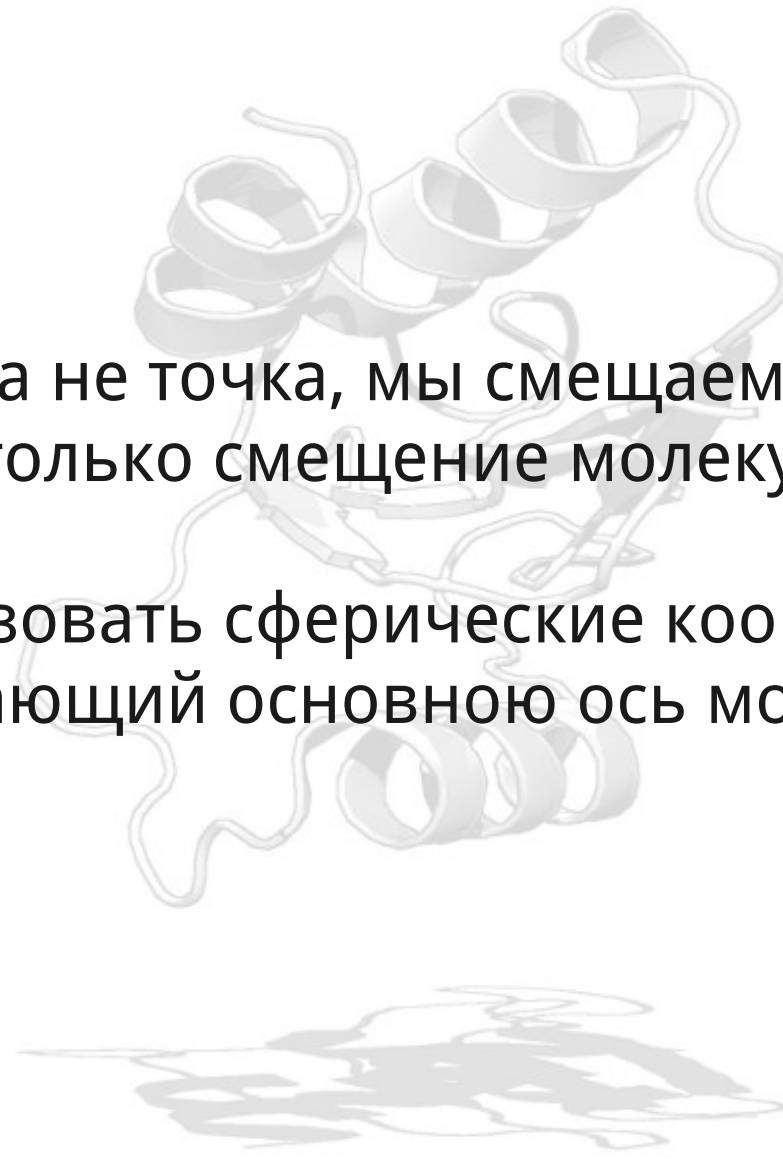
- Неудобно, что основной счёт приходит на ненужные состояния.
- Метрополис и соавторы предложили использовать цепи Маркова для генерации конформаций.

Пример:

1. генерируем  $x_{new} = x_{old} + (2\xi - 1) \delta r_{max}$
2. если новая энергия меньше старой то принимаем конформацию и используем дальше
3. если энергия выше то сравниваем фактор Больцмана со случайным числом 0:1 и если фактор Больцмана изменения энергии больше, то принимаем конформацию

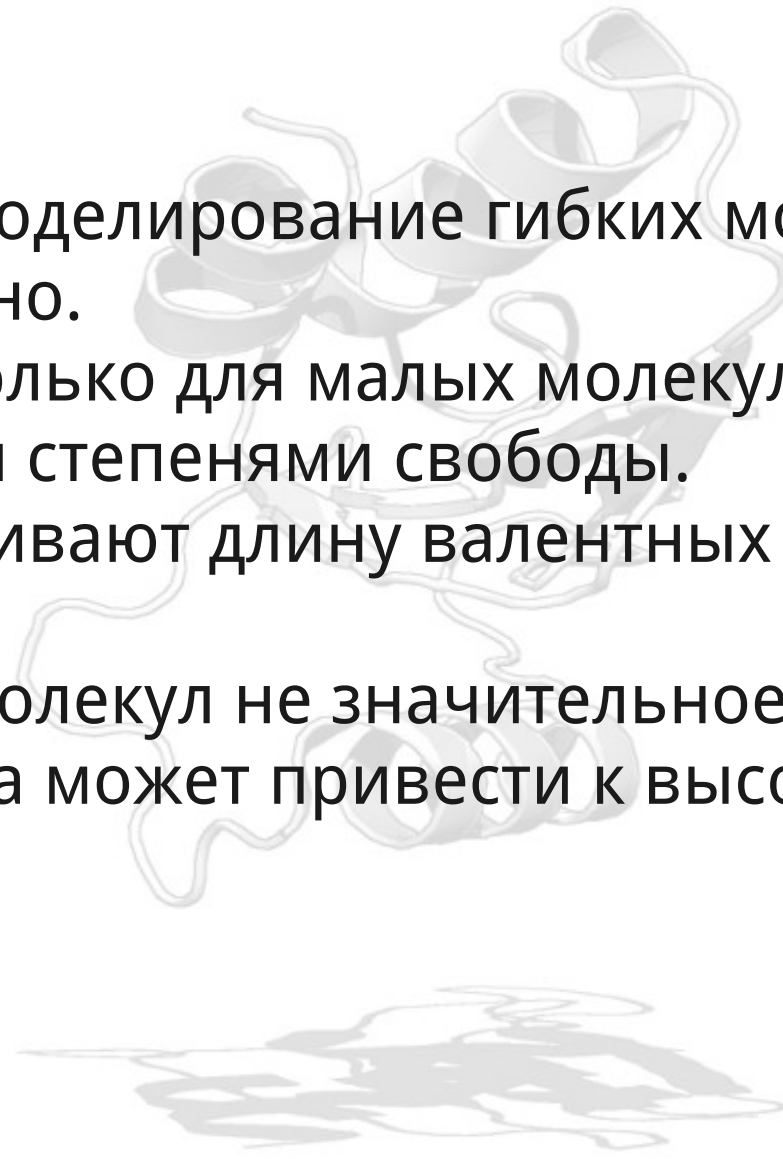
# Моделирование жестких молекул (rigid body)

- Так как молекула не точка, мы смещаем центр масс.
- Нам важно не только смещение молекулы, но и её ориентация.
- Удобно использовать сферические координаты и вектор отображающий основную ось молекулы.





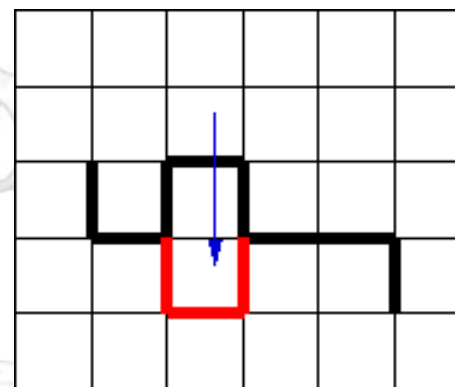
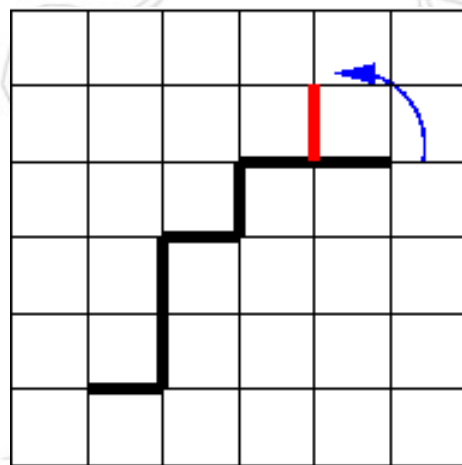
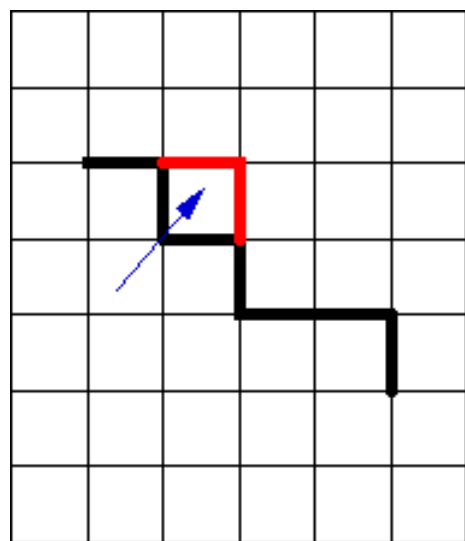
# Моделирование гибких молекул

- Монте-Карло моделирование гибких молекул проводить сложно.
  - Реалистично только для малых молекул и молекул с замороженными степенями свободы.
  - Часто замораживают длину валентных связей и валентные углы.
  - Для больших молекул незначительное изменение торсионного угла может привести к высоким энергиям.
- 

# Моделирование полимеров

## Решёточные модели полимеров.

- Представим что один мономер занимает один кубик (аналог крупно-зернистого моделирования)



# Методы выборочного поиска (Biased Monte-Carlo)

**Суть идеи:** сканировать не все пространство конформаций, а только объект интереса.

**Пример:** Нас интересует поведение молекулы в воде. Поведение воды вдали от молекулы не важно.

Существуют:

- Силовое предпочтение — направляет движение молекулы вдоль действия силы.
- Умное (Smart) предпочтение — тоже использует силы для определения движения. В случае силового предпочтения есть ограничения на движение. В большинстве случаев оба метода очень похожи.

# Проблема квази-эргодичности

**Суть проблемы:** При наличии высокоэнергичных барьеров метод Метрополиса оказывается в ловушке локальных минимумов энергии. Происходит накопление данных для весьма ограниченной области фазового пространства.

Существуют два подхода для решения этой проблемы:

- Подход “прыжок-ход” (Jump-walking или J-walking).
- Мультиканонический подход.

# J-walking

**Суть идеи:** При движении в пространстве при заданной температуре, разрешены переходы в пространство с более высокой температурой.

Простейшее решение: проводить 2 и более моделирования параллельно при разных температурах. Моделирование при высокой температуре делают с предпочтением в фазовое пространство при низкой температуре.

# Мультиканонический ансамбль

При каноническом ансамбле вероятность (  $P$  ) оказаться в точке с энергией  $E$ , пропорциональна фактору Больцмана умноженному на плотность состояний между точками  $E_1$  и  $E_2$ .

Фактор Больцмана:  $w_B = \exp \left[ \frac{-E}{k_B T} \right]$

$$P_{\text{canon}}(T, E) \sim n(E) w_B(E)$$

Этот вариант предполагает существенные скачки энергии при переходах.

# Мультиканонический ансамбль

В мультиканоническом ансамбле вероятность не зависит от энергии в определённых диапазонах энергии. Это равносильно замене фактора Больцмана на некий мультиканонический фактор  $w$ .

$$P_{mu}(E) \sim n(E) w_{mu}(E) = constant$$

Определение этого фактора не тривиальная задача. Сначала делают поиск при высокой температуре. На основе результатов определяют состояния близкие к наименьшей энергии и наоборот.

$$w_{mu}(E) = \exp[-\beta(E) E - \alpha(E)]$$

# Большой канонический ансамбль

Большой канонический ансамбль предполагает постоянство химического потенциала.

$$\mu = k_B T \ln \Lambda^3 z \quad \text{где} \quad \Lambda = \sqrt{h^2 / 2\pi m k_B T}$$

Ключевые особенности большого канонического ансамбля:

- Смещение частицы происходит методом Метрополиса
  - Частица может быть уничтожена
  - Частица может появиться в случайном положении
- Необходимо аккуратно подобрать частоту уничтожений/появлений.



# Что выбрать молекулярную динамику или Монте-Карло?

**Молекулярная динамика:** незаменима когда нужно найти свойство базирующиеся на эволюции системы во времени.

**Монте-Карло:** гораздо удобнее для работы с заданным ансамблем. Работать с большими молекулами не удобно, но возможность «нефизических» переходов позволяет проводить эффективный поиск в фазовом пространстве.

**Методы комплиментарны:** не удивительно, что существуют подходы с использованием обоих методов одновременно.