

Лекция 7. Модификации молекулярной динамики.

Курс: Молекулярное моделирование в применении к биомолекулам

Головин А.В. ¹

¹МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2020

Уравнение Шредингера

$$\left(-\frac{\hbar^2}{m} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] + V \right) \Psi(r, t) = i\hbar \frac{\partial \Psi(r, t)}{\partial t}$$

Или:

$$H\Psi = E\Psi; \quad H = \frac{-\hbar^2}{m} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r}$$

В молекулярной механике где аппроксимируем электронную плотность уравнениями классической физики.

$$F = m \frac{\partial^2 r}{\partial t^2}$$

Осталось придумать как следить за эволюцией системы во времени.



Простое уравнение силового поля (СП)

$$\begin{aligned}
 U = & \sum_{\text{bonds}} \frac{k_i}{2} (l_i - l_0)^2 + \sum_{\text{angles}} \frac{k_i}{2} (\phi_i - \phi_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned}$$

Молекулярная динамика

Монте-Карло

$$v(t + \frac{\Delta t}{2}) = v(t - \frac{\Delta t}{2}) + \frac{F(t)}{m} \Delta t$$

$$acc(o \rightarrow \eta) = \min \left(1, \exp \left\{ -\beta \left[U(r^N) - U(r^N) \right] \right\} \right)$$



Молекулярная динамика

Расчёт силы действующей на атом

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

Δt

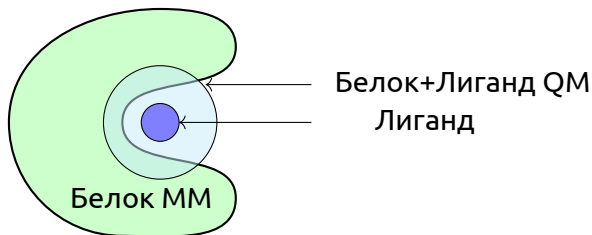
Расчёт новых координат

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



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.



Гибридное QM/MM моделирование

Простейший Гамильтониан для QM/MM системы:

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

К QM/MM части можно добавить и VdW составляющую:

$$H_{QM/MM} = - \sum_i^{elect} \sum_j^{MM_q} \frac{Q_j}{R_{ij}} + \sum_i^{nucl} \sum_j^{MM_q} \frac{Z_i Q_j}{R_{ij}} +$$

$$\sum_i^{nucl} \sum_j^{MM_{atoms}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$



Реализации описания QM/MM

Встраивание электронов

$$H^{QM/MM} = H_e^{QM} - \sum_i^n \sum_J^M \frac{e^2 Q_J}{4\pi\epsilon_0 r_{iJ}} + \sum_A^N \sum_J^M \frac{e^2 Z_A Q_J}{e\pi\epsilon_0 R_{AJ}}$$

- Ковалентные взаимодействия между QM и MM системами описываются соответствующими параметрами из MM
- Для QM системы в месте разрыва связи добавляют протон для восстановления системы до полного состояния.
- Сила действующая на этот протон "заглушку" в QM системе распределяется между атомами между которыми происходит раздел QM и MM системами



Реализации описания QM/MM

ONIOM

- Расчитываем энергию и градиент для QM системы с желаемым уровнем теории.
- Расчитываем энергию и градиент для MM системы с учетом ранее рассчитанных данных для QM системы.
- Рассчитываем MM энергию и градиенты для QM системы и вычитаем.

$$E_{tot} = E_I^{QM} + E_{I+II}^{MM} - E_I^{MM}$$

Этот подход можно использовать не только для двух уровней теории, но и больше.



Реализация QM/MM в Gromacs

- Добавляем атомы "заглушки":

```
[ virtual_sites2 ]  
LA QAtom MAtom 1 0.65
```

а связь описывается:

```
[ constraints ]  
QAtom MAtom 2 0.153
```

- Для атомов в QM системе надо поправить описание ковалентных связей:

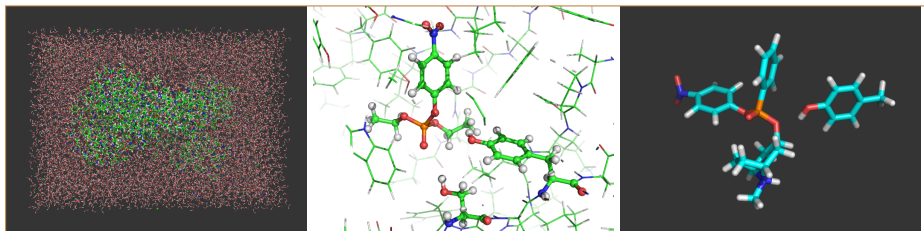
```
[ bonds ]  
QAtom1 QAtom2 5  
QAtom2 QAtom3 5
```

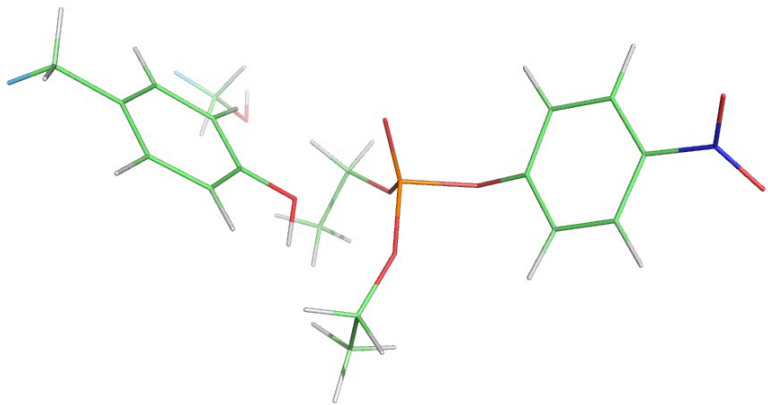
- В mdp файле описываем параметры для QM системы.

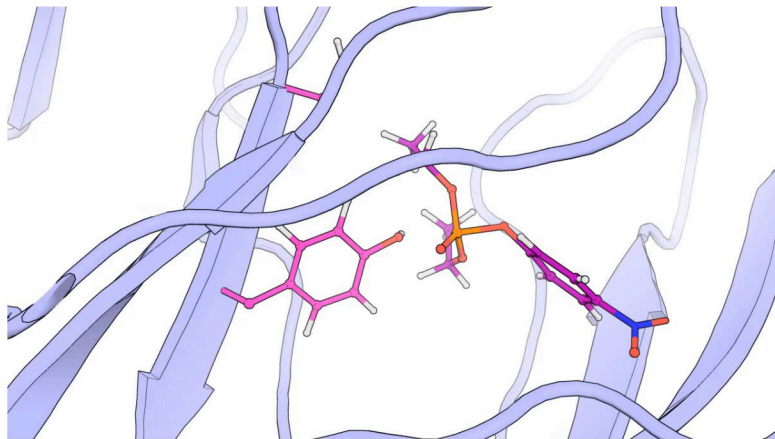
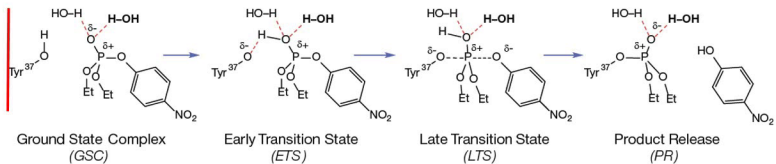


Гибридное QM/MM моделирование

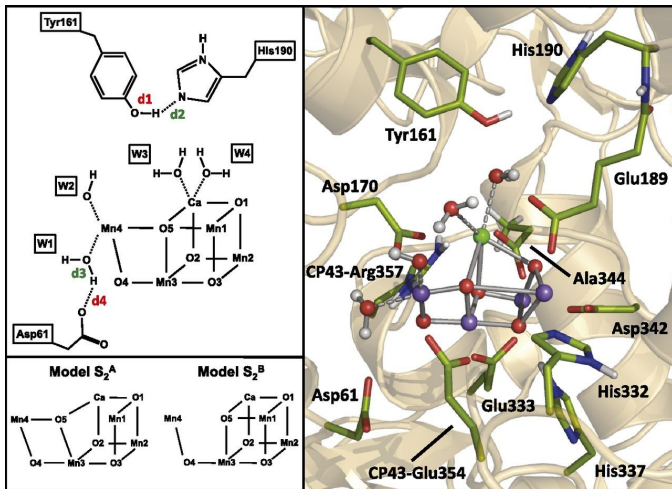
- Атомы связки
- Специальные орбитали







Ab initio QM/MM model of photosystem II. (Right) The QM region, consisting of 224 atoms, is shown in balls and sticks representation.

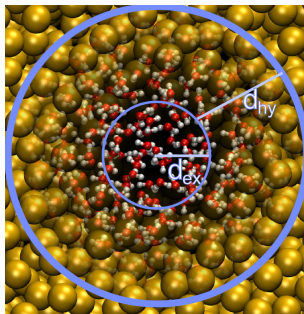


Daniele Narzi et al. PNAS 2014;111:8723-8728

Adaptive resolution scheme

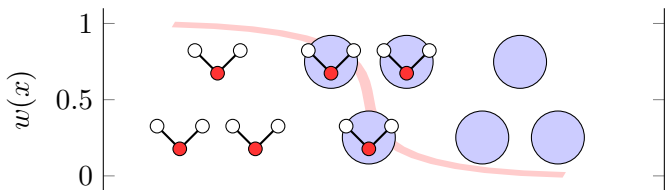
Описание взаимодействия между полноатомным и крупнозернистым описаниями системы

- В отличие от QM/MM уровень описания молекул системы может меняться "на лету"



Реализация AdResS

$$\vec{F}_{\alpha\beta} = w_{\alpha}w_{\beta}\vec{F}_{\alpha\beta}^{\text{ex,mol}} + [1 - w_{\alpha}w_{\beta}]\vec{F}_{\alpha\beta}^{\text{cg,mol}},$$



$$w(x) = \begin{cases} 0 & : & x > d_{\text{ex}} + d_{\text{hy}} \\ \cos^2\left(\frac{\pi}{2d_{\text{hy}}}(x - d_{\text{ex}})\right) & : & d_{\text{ex}} + d_{\text{hy}} > x > d_{\text{ex}} \\ 1 & : & d_{\text{ex}} > x \end{cases},$$



Метод обмена репликами (REMD)

- Основная идея: запустить параллельно несколько счётов с разными температурами.
- 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.



REMD

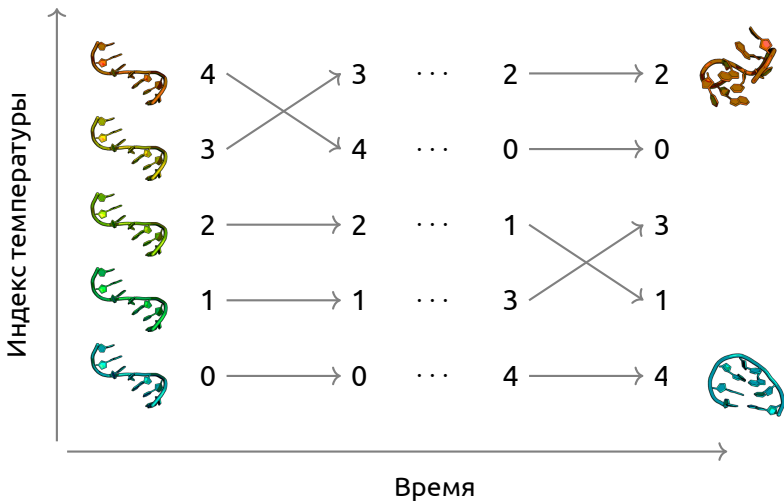
- Цель метода это ускорить сканирование (sampling) конформационного пространства.
- Применимо к переходам через значимые энергетические барьеры.
- В Gromacs обмен между репликами происходит случайно по условию:

$$P(1 \leftrightarrow 2) = \min \left(1, \exp \left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2} \right) (U_1 - U_2) \right] \right)$$

А скорости масштабируются: $(T_1/T_2)^{\pm 0.5}$



REMD



REMD: количество реплик

Разница между температурами (репликами):

$$U_1 - U_2 = N_{df} \frac{c}{2} k_B (T_1 - T_2)$$

где N_{df} это количество степеней свободы и "с" это величина от 1 до 2 для системы белок вода.

Если $T_2 = (1 + \epsilon)T_1$ тогда вероятность обмена:

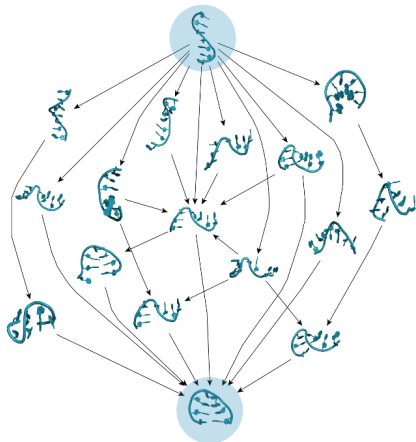
$$P(1 \leftrightarrow 2) = \exp\left(-\frac{\epsilon^2 c N_{df}}{2(1 + \epsilon)}\right) \approx \exp\left(-\epsilon^2 \frac{c}{2} N_{df}\right)$$

Таким образом для вероятности обмена $e^{-2} \approx 0.135$ получаем $\epsilon \approx 2/\sqrt{c N_{df}}$.

И если мы контролируем длину связей, то: $N_{df} \approx 2 N_{atoms}$ и при $c = 2$ надо использовать: $\epsilon = 1/\sqrt{N_{atoms}}$.



Результат REMD



Коллективные переменные (CV)

- Положение всех атомов в пространстве явно избыточная информация для описания некоторых простых процессов
- Коллективные переменные это некоторые значения которые наиболее значительно изменяются в ходе процесса, который нас интересует
- Трудно предположить заранее все важные коллективные переменные
- Недостаток описания приводит к гистерезису



Примеры коллективных переменных

ALPHABETA
ANTIBETARMSD
CONSTANT
DHENERGY
DISTANCE
NOE
PATH
RDC

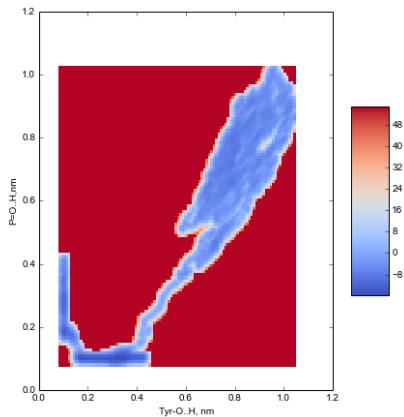
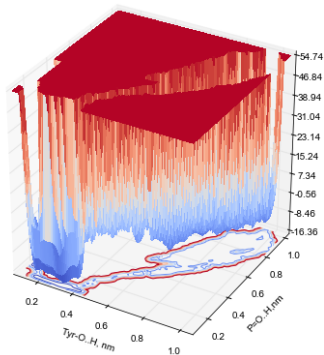
ALPHARMSD
CELL
CONTACTMAP
DIHCOR
ENERGY
PARABETARMSD
POSITION
TORSION

ANGLE
CH3SHIFTS
COORDINATION
DIPOLE
GYRATION
PATHMSD
PROPERTYMAP
VOLUME



Пример анализа на основе CV

$$E \approx -\frac{1}{\beta} \log(P)$$



Использование CV для влияния на МД, ABMD

adiabatic biased MD

$$V(\rho(t)) = \begin{cases} \frac{K}{2} (\rho(t) - \rho_m(t))^2, & \rho(t) > \rho_m(t) \\ 0, & \rho(t) \leq \rho_m(t), \end{cases}$$

$$\rho(t) = (CV(t) - TO)^2$$

Суть метода состоит в движении системы к заданным значениям CV используя гармонический потенциал, который изменяет термические флуктуации если система не движется к заданным значениям CV.



Использование CV для влияния на МД, steered MD

Позволяет добавлять временно зависимый гармонический потенциал на одну или несколько переменных.

$$V(\vec{s}, t) = \frac{1}{2} \kappa(t) (\vec{s} - \vec{s}_0(t))^2$$

или

$$\begin{aligned} H_\lambda(X, t) &= H(X) + U_\lambda(X, t) \\ &= H(X) + \frac{k(t)}{2} (s(X) - \lambda(t))^2 \\ &= H(X) + \frac{k(t)}{2} (s(X) - s_0 - vt)^2. \end{aligned}$$



Использование CV для влияния на МД, WALLS

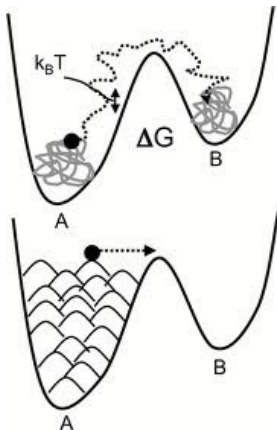
Суть очень проста выставление "стен" в виде штрафа за пределами значения CV

$$\sum_i k_i ((x_i - a_i + o_i) / s_i)_i^e$$



Метадинамика

$$V(\vec{s}, t) = \sum_{k\tau < t} W(k\tau) \exp\left(-\sum_{i=1}^d \frac{(s_i - s_i^{(0)}(k\tau))^2}{2\sigma_i^2}\right).$$



Well tempered metadynamics

$$V(s, t) = \sum_{t'=0, \tau_G, 2\tau_G, \dots}^{t' < t} W e^{-V(s(q(t')), t') / \Delta T} \exp \left(- \sum_{i=1}^d \frac{(s_i(q) - s_i(q(t')))^2}{2\sigma_i^2} \right)$$

Этот вариант метадинамики позволяет добавлять энергию более аккуратно, что приводит к тому, что добавленное количество энергии начинает стремиться к энергии перехода между состояниями.



Обсуждение

Преимущества

- Ускорение событий путём выталкивания системы из известной области
- Знание результата не нужно, возможно все, что может быть в этой системе
- Возможность восстановить профиль поверхности потенциальной энергии

Недостатки

- Динамика и температура изменены человеком.
- Данные о кинетике процессов не доступны (а может и нет : [arXiv:1309.5323](https://arxiv.org/abs/1309.5323), Tiwari & Parrinello)



Выбор CV

Правильный выбор CV это критический момент, вот некоторые требования:

- CV должны описывать процесс и интересы
- Включать все медленно изменяющиеся степени свободы
- Количество CV должно быть не большим
- Используйте химическую/физическую интуицию
- Подход проб и ошибок



Метадинамика с обменом между CV

Используя правило Метрополиса:

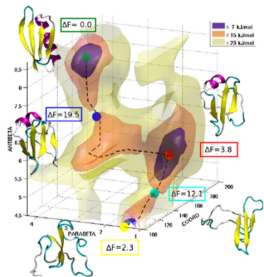
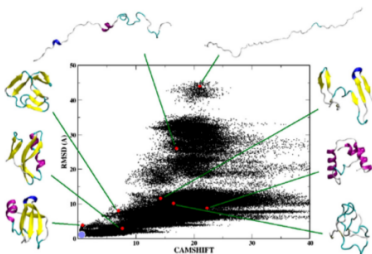
$$\min \left(1, \exp \left[\beta (V_G^a(x^a, t) + V_G^b(x^b, t) - V_G^a(x^b, t) - V_G^b(x^a, t)) \right] \right)$$

Запускаются несколько параллельных траекторий с разными CV, которые обмениваются конформациями как в REMD



Примеры с использованием метадинмики

Образование структуры белка на основе знания химических сдвигов и реконструкция FES белка :

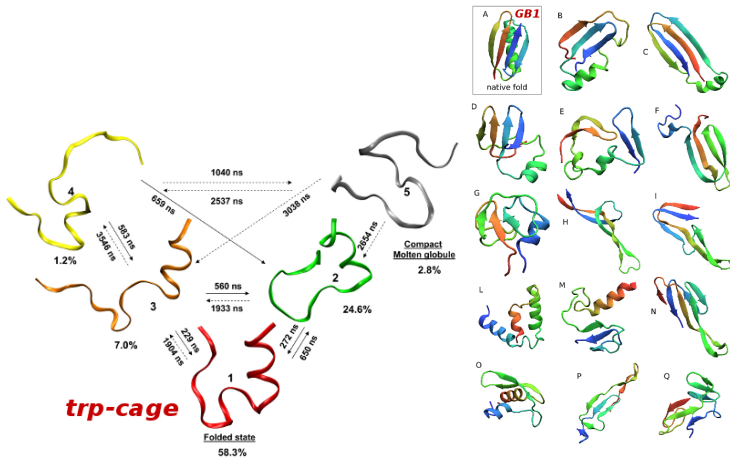


doi: 10.1073/pnas.1218350110.



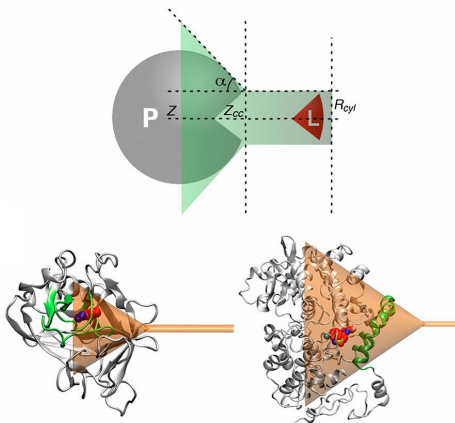
Примеры с использованием метадинимики

Изучение процесса фолдинга белка:



Примеры с использованием метадинамики

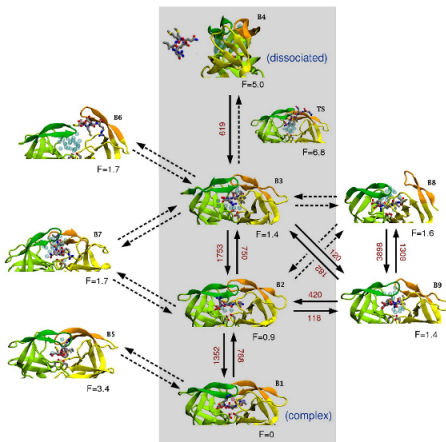
Определение пути лиганда к сайту связывания:



doi: 10.1073/pnas.1303186110

Примеры с использованием метадинмики

Аккуратное определение механизма связывания лиганда:



JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, 131, 11811 (2009)