

# Компьютерные методы в фармакологии



***В.Б.Сулимов***

***НИВЦ МГУ***

***Лекция № 12***

***Молекулярная динамика***

# Чтобы предсказать константу связывания молекулы-лиганда с белком

...мы должны рассчитать всего-навсего свободную энергию связывания комплекса белок-лиганд

$$\Delta G = \Delta H - T\Delta S$$

Изменение энтальпии  
при связывании  
лиганда

Изменение энтропии  
при связывании лиганда

**Термодинамика и статистическая физика**

# Расчет свободной энергии связывания лиганда с протеином

- **Методы молекулярной динамики:  $\Delta G = \Delta H - T\Delta S$** 
  - Вычисление траектории движения атомов системы белок+лиганд, находящейся в термостате
  - Описание движения атомов как классических частиц, взаимодействующих друг с другом с помощью классических потенциалов
  - Усреднение по траектории – усреднение по ансамблю
  - Время вычислений: сутки и более на одном процессоре
  - Обязательно нужны многопроцессорные вычисления на суперкомпьютерах

# Для каких задач применяется молекулярная динамика?

- Расчет **свободной энергии** связывания лиганд-белок
- Расчет относительной свободной энергии связывания лиганд-белок для близких по структуре молекул – оптимизация lead compound'a
- Наблюдение кинетики выхода лиганда из связанного состояния с белком – проверка результатов докинга
- Расчет растворения молекул – энергии сольватации
- Получение структуры белков – **protein folding**
- Моделирование наблюдаемых свойств конденсированных сред: свойства воды, свойства стекла, свойства новых материалов, получение аморфных структур из кристаллов.
- Поиск глобального минимума энергии на сложной поверхности

# Основные принципы

- Основывается на классических представлениях о движении атомов и молекул
- Численное решение уравнений движения (второй закон Ньютона) для системы из множества взаимодействующих объектов
- Взаимодействия между атомами основаны на межатомном потенциале  $U(r)$
- Квантовая молекулярная динамика – для каждого расположения ионных остатков решается уравнение Шредингера (обычно DFT) и находятся силы, действующие на ионные остатки, которые сдвигаются под действием этих сил
- Статистическая задача – движение ансамбля частиц при наличии термостата

# Эргодическая гипотеза

- Применение методов МД во многом основывается на эргодической гипотезе, полагая, что в среднем траектория не зависит от исходных условий
- Это позволяет использовать для анализа подсистему (например поведение одной молекулы белка вместо тысяч), получая и усредняя длинную траекторию
- На практике применимо далеко не всегда (например фолдинг, полимерные системы)
- Системы во многом чувствительны к исходным условиям, изначально близкие траектории могут быстро расходиться
- Это важно на длинных траекториях - микросекунды

# Программы для МД моделирования

## Классические атомы

### **CHARMM** (<http://www.charmm.org/>)

- Возможность использования силового поля MMFF94, другие силовые поля
- Эффективные модели растворителя
- Специальные опции для работы с макромолекулами (белки, фрагменты ДНК)

### **GROMACS** (<http://www.gromacs.org/>)

- Свободно распространяемая программа
- Высокая скорость работы
- Широкий набор инструментов для анализа траекторий
- Есть GPU возможности

# Программы для МД моделирования Классические атомы

**Amber** <http://ambermd.org/>

- Возможность использования универсального силового поля **Amber**
- Эффективные модели растворителя, белки, лиганды, вода

**LAMMPS** <http://lammps.sandia.gov/>

- Свободно распространяемая программа,
- Для неорганических твердых тел, кристаллов, аморфных

**DL\_POLY**

[http://www.cse.scitech.ac.uk/ccg/software/DL\\_POLY/index.shtml](http://www.cse.scitech.ac.uk/ccg/software/DL_POLY/index.shtml)

**DL\_POLY4 и DL\_POLY Classic**

**NAMD** – свободно распространяемая программа

<http://www.ks.uiuc.edu/Research/namd>

- Различные силовые поля, белки, ДНК, лиганды, вода



# Программы для МД моделирования

## Квантовая химия - DFT

**SIESTA** - <http://www.icmab.es/siesta/> - гауссовы базисы

**CASTEP** - <http://www.castep.org/>

**VASP** - <http://cms.mpi.univie.ac.at/vasp/> - базис плоские волны

**QUANTUM ESPRESSO** - <https://www.quantum-espresso.org> - FREE

**ABINIT** - <http://www.abinit.org/>

**CPMD** - <http://www.cpmd.org/>

# Параметры

- Экспериментальные данные и *ab initio* расчеты
- Для молекулярной биологии подходит только классическая МД
- Как правило каждый программный пакет имеет свой набор силовых полей (Charmm, AMBER, GROMOS, и др.).

# Основные шаги

- **Создание структуры**
  - Экспериментальные данные (X-ray, ЯМР)
  - Моделирование (Modeller, редакторы)
  - Случайная генерация
- **Определение характера взаимодействий (силовое поле)**
- **Определение граничных условий**
  - Открытые границы системы
  - Фиксированные границы (ящик)
  - Периодические границы

# Основные шаги

- Определение внешних факторов
  - Температура, давление
  - Фиксированный объем
- Решение уравнений Ньютона
  - Дискретно по времени
  - Выбор алгоритма (Verlet, Leap Frog и т.д.)
  - Выбор интегратора
  - Длина траектории должна быть больше времени релаксации системы к равновесию

# Выбор начальных скоростей

- Случайное определение скоростей согласно распределению Максвелла-Больцмана

$$P(v)dv = 4\pi \left( \frac{m}{2\pi k_b T} \right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2k_b T}}$$

# Термостат

Основная идея – корректировать скорости частиц, чтобы получить заданную температуру системы

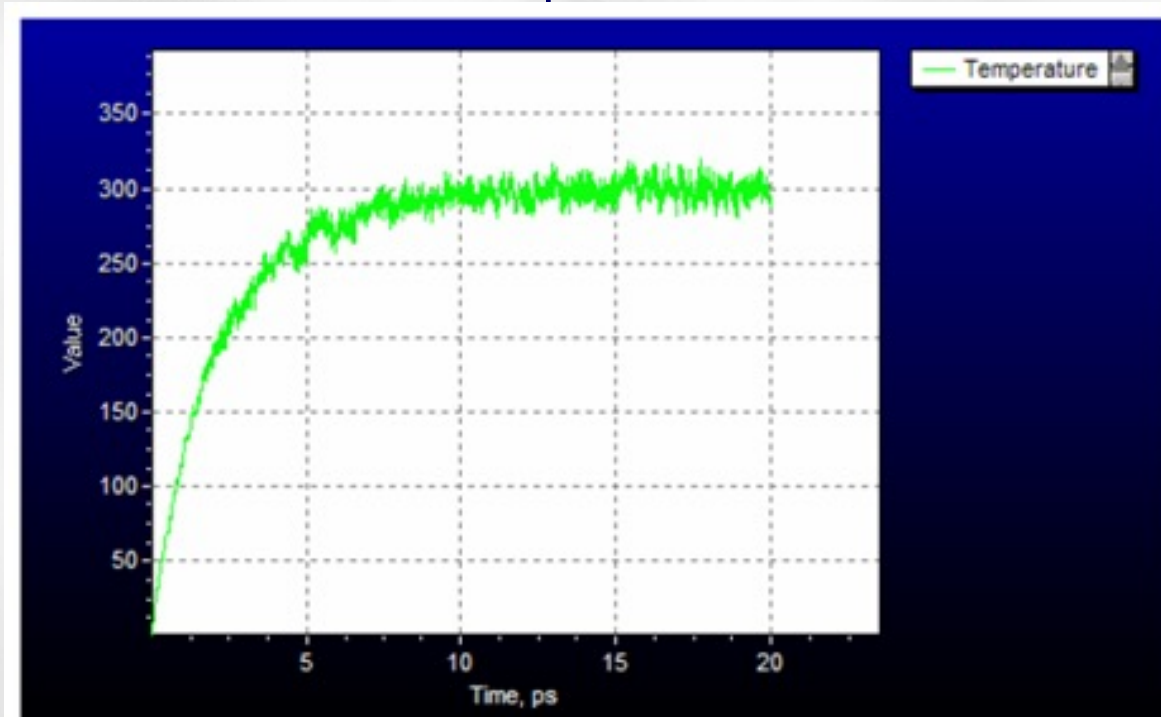
$$T \sim \langle v^2 \rangle$$

# Термостат Берендсена

Скорости всех частиц в системе одновременно умножаются на коэффициент:

$$\kappa^2 = 1 + \frac{\delta t}{\tau} \cdot \left( \frac{K_0}{K} - 1 \right),$$

$\delta t$  – шаг интегрирования,  $\tau$  – характерное время релаксации температуры системы,  $K_0$  - кинетическая энергия при заданной температуре,  $K$  – текущее значение кинетической энергии



# Langevin thermostat

$$\dot{r}_i = \frac{p_i}{m_i}, \quad \dot{p}_i = F_i - \gamma_i p_i + f_i,$$

$F_i$  – сила, действующая на атом  $i$  благодаря потенциалу со стороны других атомов,

$\gamma_i$  – коэффициент трения,

$f_i$  – случайная сила, имитирующая случайные удары, демпфирующие частицы между собой из-за трения.

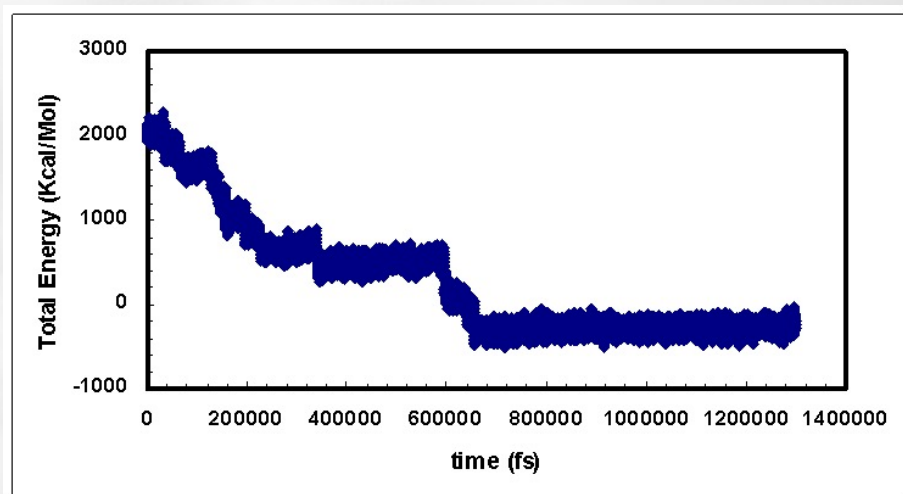
Случайные числа подчиняются распределению Гаусса с дисперсией:

$$\sigma_i^2 = 2m_i\gamma_i k_B T / \delta t \quad \delta t \text{ – шаг по времени}$$



# Достижение равновесия

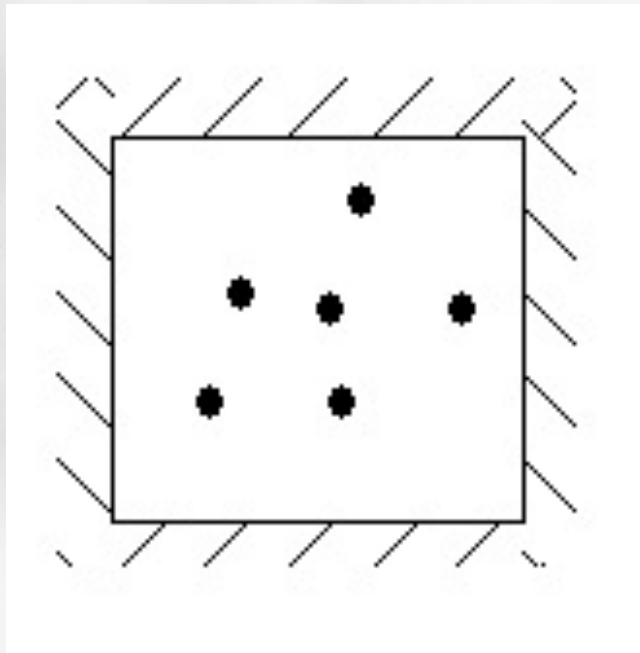
- При случайном начальном распределении скоростей и неточном определении координат система начинает траекторию далеко от равновесия



- При движении по траектории необходимо достичь равновесного значения энергии или температуры
- При этом система продолжит флуктуировать, но вокруг средних значений

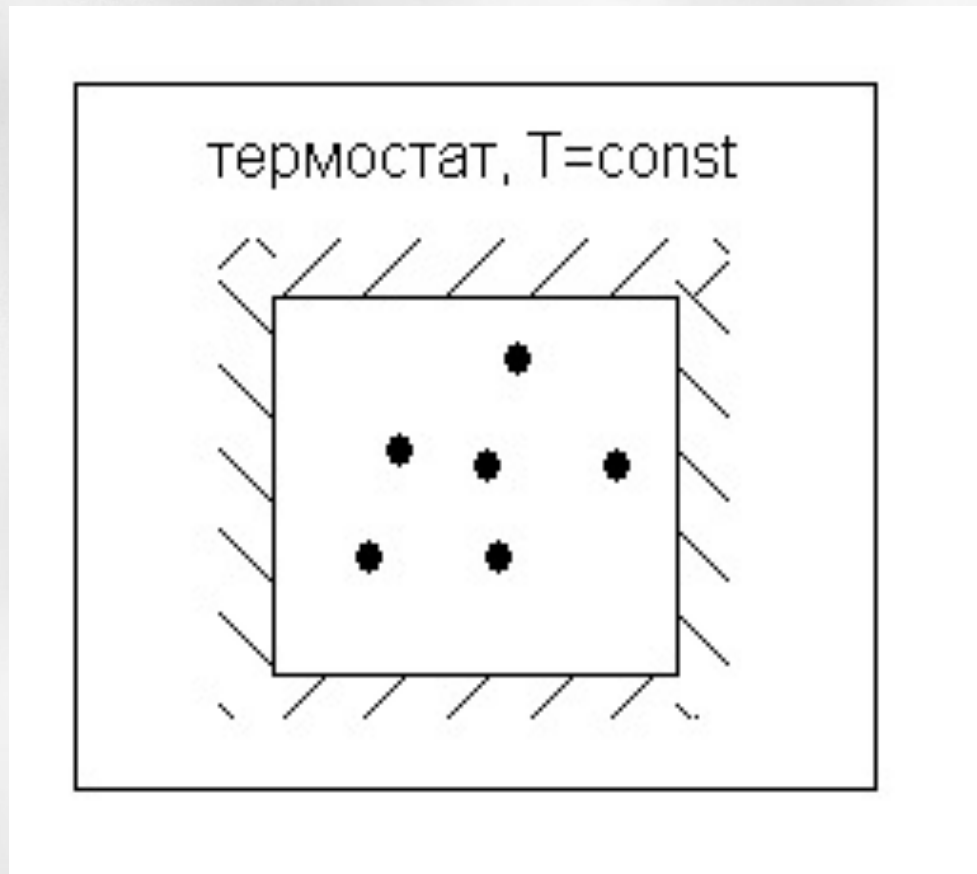
# Статистические ансамбли,

- $NVE$ -сохраняются число частиц, объем, энергия



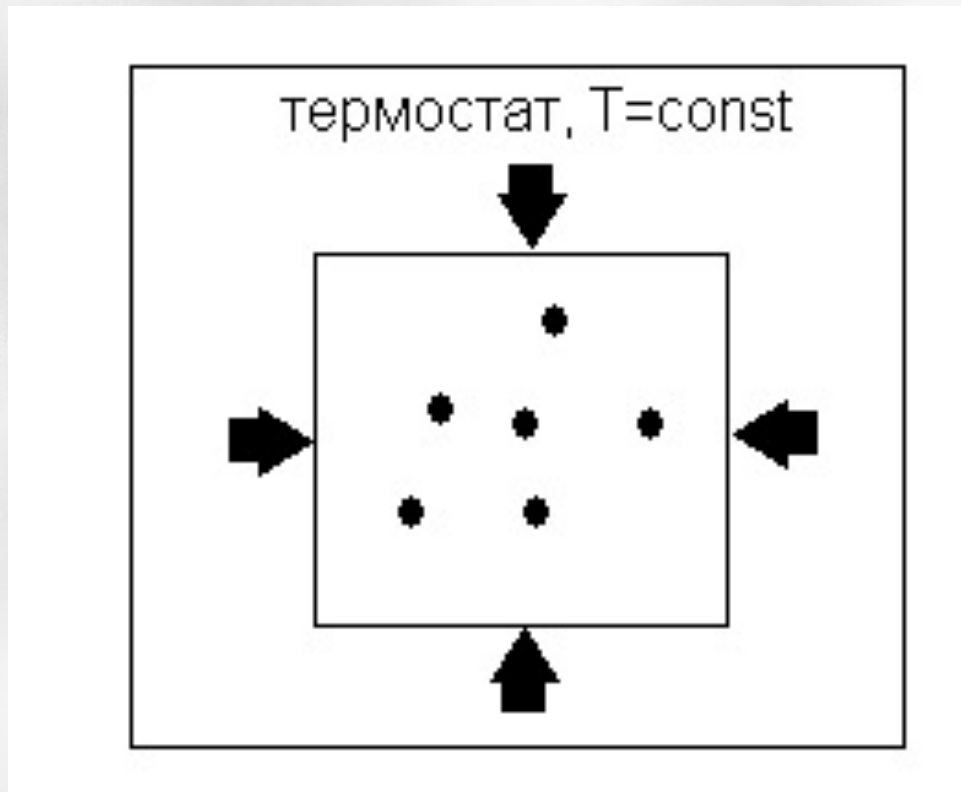
# Статистические ансамбли,

- $NVT$ -сохраняются число частиц, объем, температура



# Статистические ансамбли,

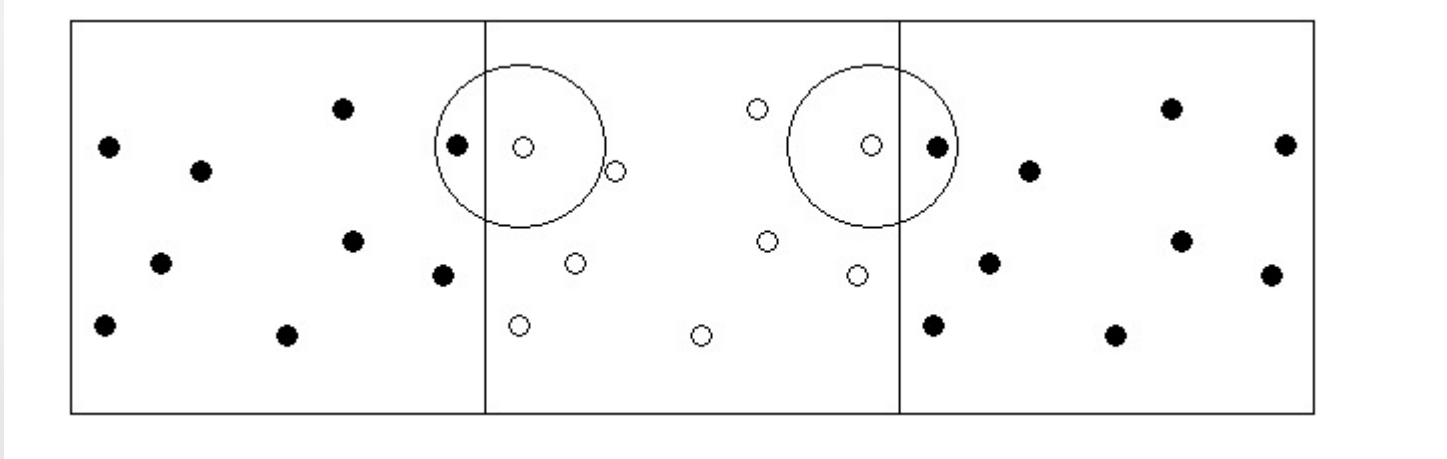
- NPT-сохраняются число частиц, давление, температура



# Граничные условия

- Открытые границы – система в «вакууме» -- неаккуратный подход
- Фиксированные границы – система в «коробке»
- Если система окружена молекулами растворителя, их массив можно делать любой формы: куб, сфера и т.д.
- Для снятия граничных эффектов применяют периодические граничные условия «размножение» системы

## Периодические граничные условия



Частица взаимодействует не только с остальными частицами ящика, но и с их отражениями

# Учет дальнедействующей части кулоновского потенциала

- Обрезание потенциала взаимодействия
- Метод реакционного поля
- Метод Эвальда
- **Метод PME**

# Обрезание потенциала взаимодействия

## Кулоновский потенциал

$$V = f \frac{q_1 q_2}{r}$$

Потенциал сдвигается так, чтобы на расстоянии, равном радиусу обрезания, потенциал и его производная равнялись нулю.

$$V = f \frac{q_1 q_2}{r} - f \frac{q_1 q_2}{r_c} + (r - r_c) \cdot f \frac{q_1 q_2}{r_c^2}, r \leq r_c$$

$$V = 0, r > r_c$$

При обрезании необходимо разбивать систему на нейтральные группы. Пример-молекула воды.

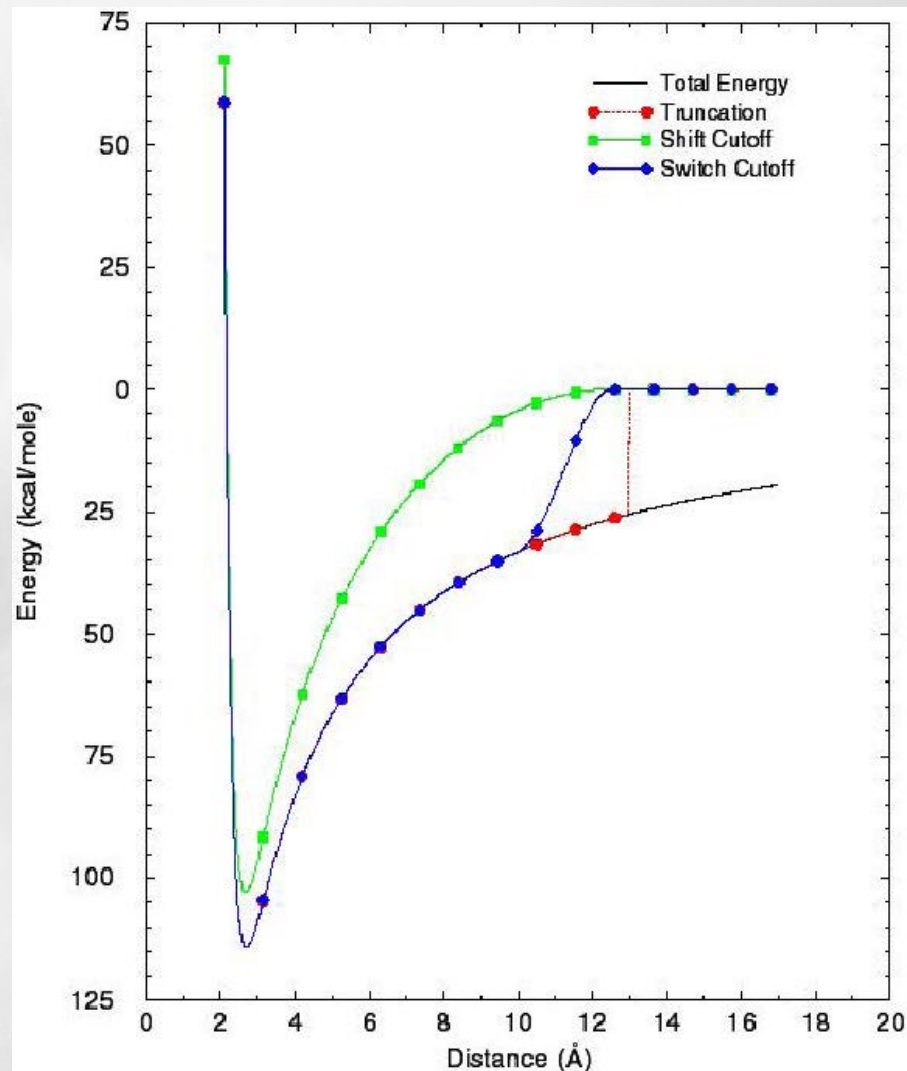


# Обрезание взаимодействий

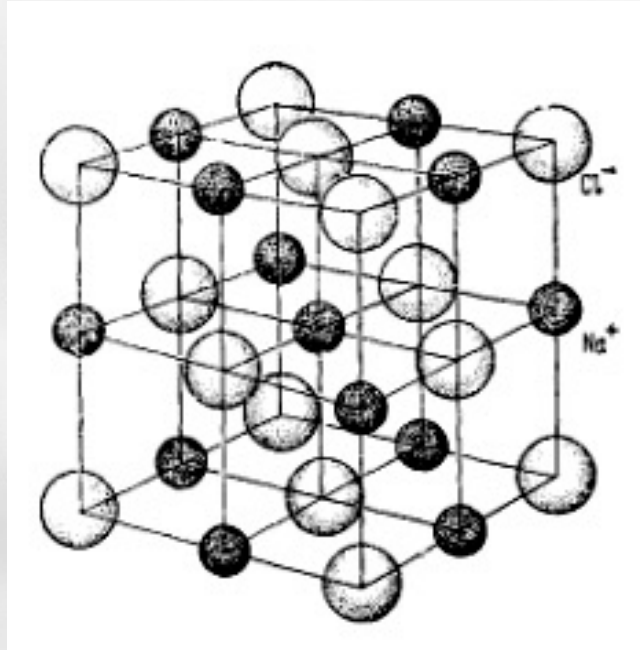
- В идеале необходимо учитывать взаимодействие каждого атома с каждым
- На практике это слишком дорого – пренебрегаем взаимодействиями на больших расстояниях
- Обычно используют расстояние  $\sim 15\text{\AA}$
- Электростатические и дисперсионные взаимодействия обрезаются на разных расстояниях

# Методы

- Простое обрезание – начиная с заданного расстояния взаимодействия не учитываются
- Shift – функция изменяется так, чтобы на заданном расстоянии выходила на ноль
- Switch – приближаясь к расстоянию обрезания функция смещается к нулю



# Метод Эвальда для кристаллов



Энергия взаимодействия иона с остальными ионами  $V = Q \sum \frac{q_i}{r_{ij}}$

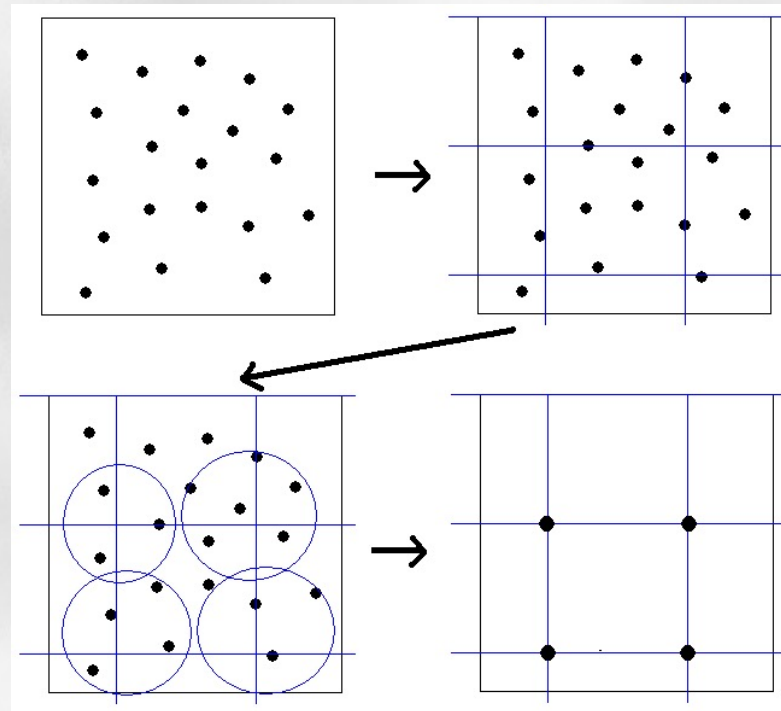
Идея метода:

Суммирование ведется по электрически нейтральным слоям

Это улучшает сходимость суммы

# Метод PME

Метод Эвальда работает медленно. В практических расчетах распространение получила его модификация – метод PME (Particle-Mesh Ewald).



Физический смысл метода PME заключается в том, что система из  $N$  заряженных частиц заменяется сеткой, заряды в узлах которой вычисляются по определенному алгоритму

# Вычисление $\Delta G$

$$F(\lambda = 1) - F(\lambda = 0) = \int_{\lambda=0}^{\lambda=1} d\lambda \left\langle \frac{\partial \mathcal{U}(\lambda)}{\partial \lambda} \right\rangle_{\lambda}.$$

- Free energy perturbation и термодинамическое интегрирование
- Переход из одного состояние в другое путем замены параметров атомов – дискретно за 10-20 шагов с уравниванием системы в промежутке (FEP) или непрерывно на каждом шаге(ТИ)
- Наиболее надежный метод расчета энергии взаимодействия белка с лигандом

# Классическая книга по Молекулярной Динамике

## Computer Simulation of Liquids

M. P. ALLEN

*H. H. Wills Physics Laboratory  
University of Bristol*

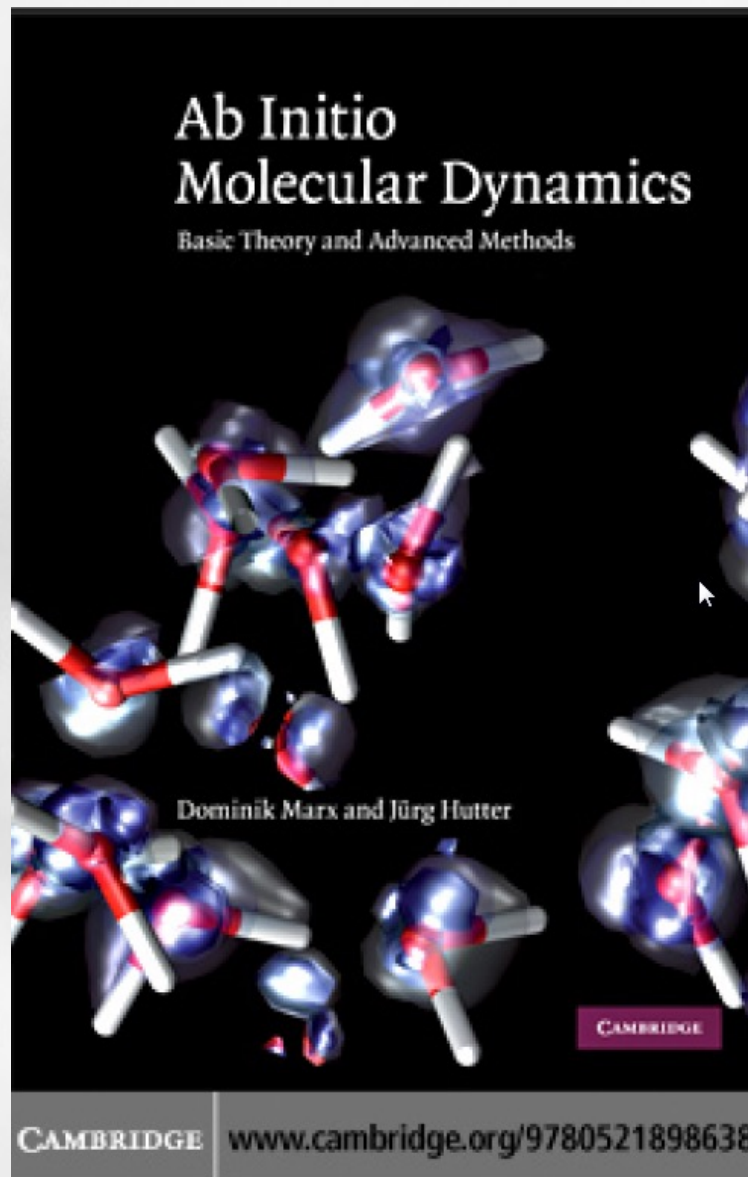
and

D. J. TILDESLEY

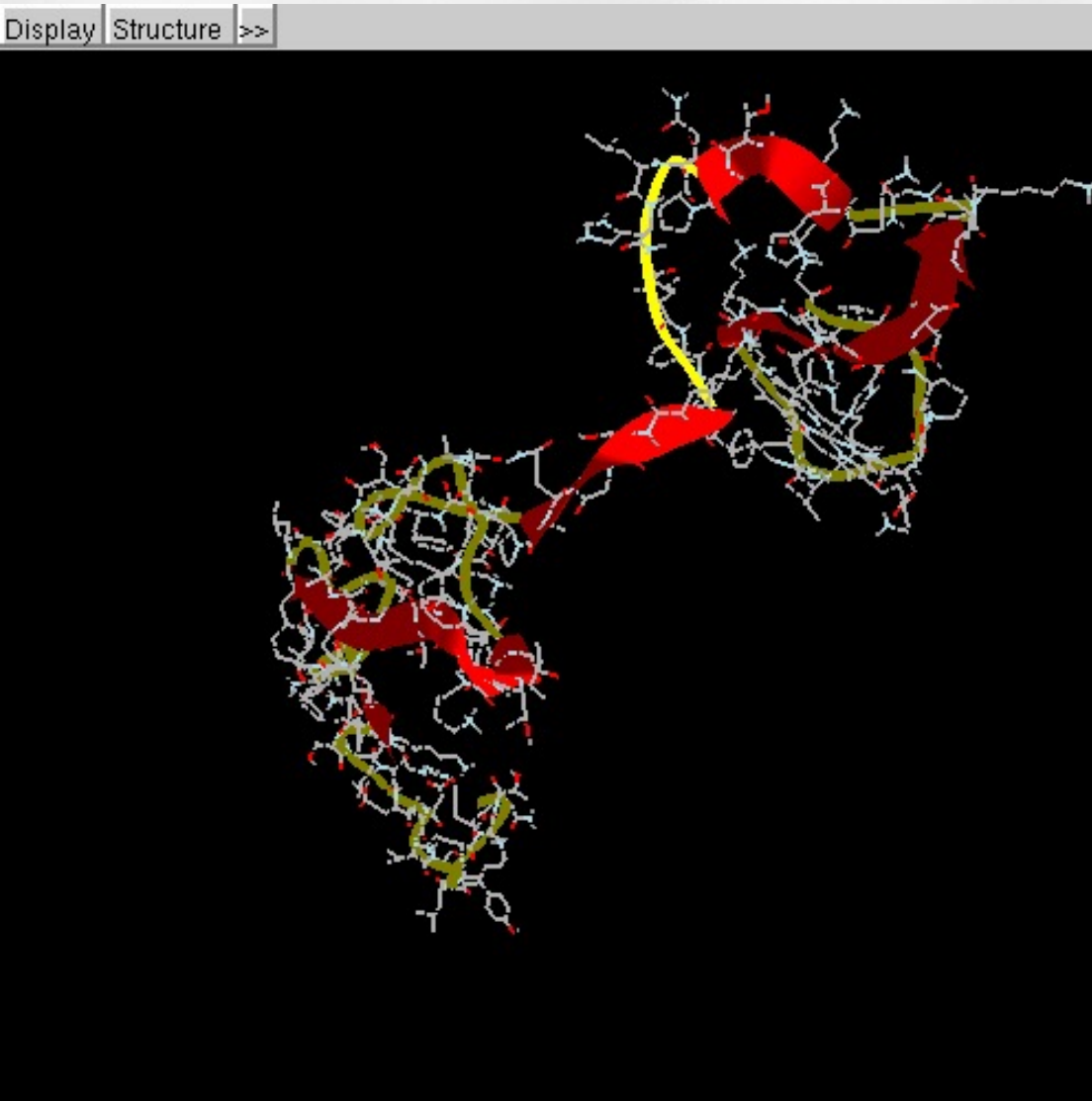
*Department of Chemistry  
The University, Southampton*

**Clarendon Press    Oxford**

# Dominik Marx & Jürg Hutter



# Фолдинг цитохрома С



Цитохромы – белки, активный центр которых связан с молекулой гема. Аморфная структура с двумя доменами постепенно сворачивается в компактную глобулу. Желтым показаны неструктурированные фрагменты, которые сворачиваясь в  $\alpha$ -спираль или  $\beta$ -лист становятся зелеными и красными лентами



Parallel programming lead to the new generation of docking and postprocessing programs:

- No grid: direct ligand-to-protein docking
- Quantum chemical docking
- High accuracy of protein-ligand binding free energy calculations:

**Free energy  $\Delta G = \Delta H - T\Delta S$**

**Entropy**

How to calculate?

**Molecular dynamic is too time consuming!**

# Спасибо за внимание

- *...Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils...*
- *...Каждое лекарство есть инновация; а кто не хочет применять новые средства, должен ждать новых бед...*

*Sir Francis Bacon (1561-1626)*



*OF INNOVATIONS*