

Лекция 7

Метод молекулярной динамики.

Реализация метода молекулярной динамики.

Проблема моделирования макромолекулы в водном растворе. Основные модели. Конечные системы - водная капля. Модель бесконечной конденсированной системы – периодические граничные условия. Влияние водного раствора на конформационную динамику макромолекулы.

Расчет термических и динамических характеристик макромолекулы методом молекулярной динамики. Термодинамическая стабильность. Существенные конформационные движения.

Области применения методов статистического моделирования в биоинформатике.

Метод Молекулярной Динамики и конформационный анализ биополимеров

- Макромолекула биополимера с точки зрения физики – система многих взаимодействующих частиц (атомов) – статистическая система.

- Экспериментальное доказательство – обратимая термическая денатурация белков и нуклеиновых кислот – разрушение и восстановление уникальной структуры

- Свойства системы N частиц определяются Поверхностью Потенциальной Энергии – $U(r_1, \dots, r_N)$

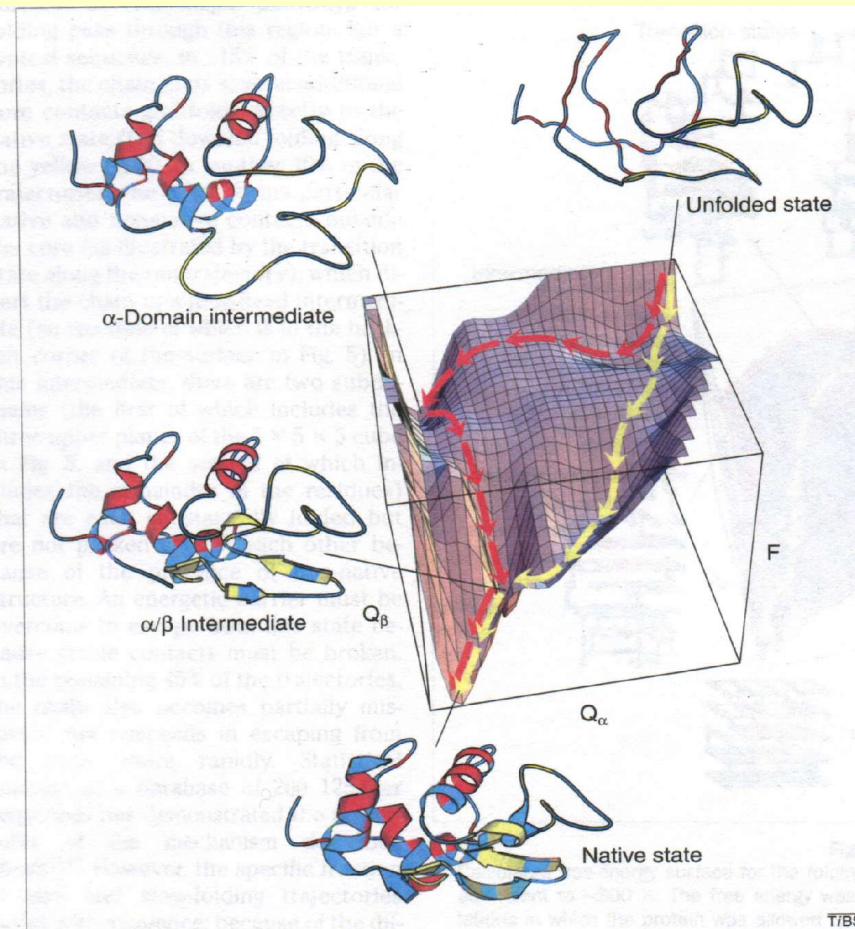
$$W(r_1, \dots, r_N) = (1/Z) \exp(-U(r_1, \dots, r_N)/kT)$$

- Минимумы ППЭ – метастабильные конфигурации (конформации)

- Глобальный минимум – стабильная (нативная) структура

- конформационный анализ – изучение свойств ППЭ

Пример ППЭ белка



Схематическое представление Пов.Своб.Энергии лизоцима (белок 129 остатков). **Желтые** стрелки быстрый путь, небольшая концентрация интермедиата. **Красный** путь-медленная самосборка с большой концентрацией интермедиатов домена

Роль компьютерных экспериментов в науке

•Огромна – срединная между **ЭКСПЕРИМЕНТОМ** и **АНАЛИТИЧЕСКОЙ** теорией в физике конденсированного состояния,

К.Э. расширяет возможности теоретического исследования свойств вещества путем **симбиоза моделей взаимодействия между частицами и прямым расчетом ансамбля наивероятных состояний** системы

-твердое тело (кристалл, аморфное)

-поверхность тв.тела, жидкости

-жидкое состояние, раствор

-биополимер в растворе

Методы генерации ансамблей наивероятных состояний

- Тотальное сканирование фазового пространства – возможно для малых молекул
макромолекулы (нет), число конформеров 3^{100}
- метод Монте-Карло $\{r\} \rightarrow \{r'\}$, $\Delta U = U(r') - U(r)$,
 $p_{12} \sim \exp(-\Delta U/kT)$ – метод Метрополиса,
- Методы псевдо-глобальной оптимизации {генетический алгоритм, метод отжига(simulated annealing)}
- метод молекулярной динамики – достаточно эффективный и информативный

МЕТОД МОЛЕКУЛЯРНОЙ ДИНАМИКИ

Принципы:

- атомы биополимера и молекул растворителя = система взаимодействующих частиц
 - классические уравнения движения для каждого атома в поле сил остальных частиц системы
- ⇒ траектории атомов $\{\mathbf{r}_i(t)\}$ от времени

$$m_i d^2 \mathbf{r}_i / dt^2 = \mathbf{f}_i - \text{полная сила}$$

Возможности:

- Конформационная динамика –
 - основные МОДЫ существенных конформационных движений
 - равновесные величины конформационных параметров
 - термические флуктуации конформационных параметров
- Термодинамические свойства –
 - энергии конформаций
 - энергии связывания лиганда с макромолекулой – drug design

История метода МД

1957 - Alder and Wainwright – расчет фазовой диаграммы жидкости из твердых шаров;

1960 - J. В. Gibson, - расчет радиационных дефектов в т.т.

1967- L. Verlet свойства жидкого аргона,

1976 – динамика биополимера в вакууме

1982 – динамика небольшого белка в водном растворе

1987 – тРНК в водном растворе

Сегодняшняя роль МД – тв.тело, жидкости, биополимеры, конструирование лекарственных препаратов

Ограничения метода МД

- классическая механика - $\lambda = \sqrt{\frac{2\pi\hbar^2}{mkT}}$ Длина волны Де Бройля

$\lambda/d(\text{at-atom}) \ll 1$, классическая система, ~ 0.1 для атомов C,N,O при комнатных T

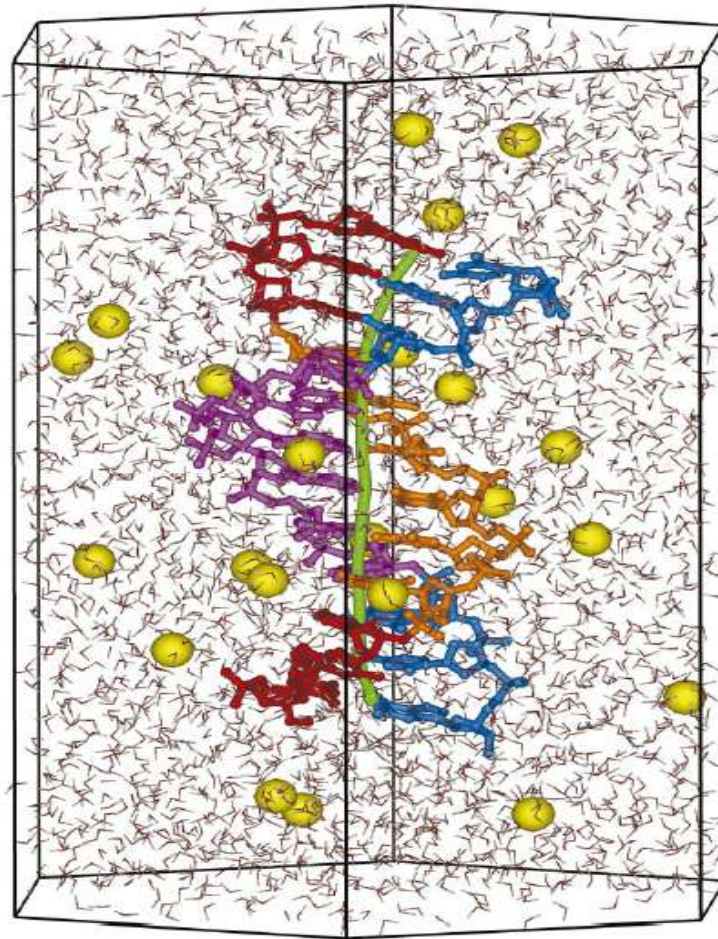
- - размер системы (зависимость от размера)
- - расчет достаточного ансамбля состояний, репрезентативная выборка
- - множественность состояний (конфигураций, конформаций) – достижимость состояний, время достижения
- биополимеры – реалистичные МД модели - 1996

Силовые ММ поля

$$E_{tot} = \sum_{bonds} K_b (l - l_0)^2 + \sum_{angles} K_a (\vartheta - \vartheta_0)^2 + \sum_{torsions} K_t [1 + \cos(n\phi + \gamma)] + \sum_{i,j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}} \right]$$

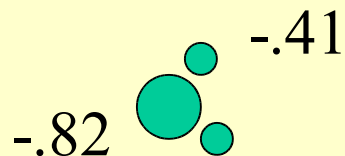
- гипотеза **трансферабельности** – подобные хим. группы → одинаковые взаимодействия.
- атом ММ = (хим. атом + свойства хим. фрагмента)
- Параметры – описание свойств молекулярных органических жидкостей и кристаллов (малые молекулы); Кв. мех. Расчеты.
- AMBER, OPLS, GROMOS, CHARMM
- **РЕАЛИЗМ** атом-атомных взаимодействий !?

Проблема растворителя



- Минимизация влияния границ растворитель – вакуум (среда)
- Периодические граничные условия-кристалл из основных боксов – искусственная, не характерная периодичность

Молекулярные модели воды



SPC, TIP3



Правильное описание структуры ближнего порядка жидкой воды и структуры льда,

Энергия испарения ~ 10 kcal/mol

Протокол МД эксперимента

- начальные координаты атомов $\{r_1, \dots, r_N\}$ – PDB / сборка молекулярным конструктором + локальная оптимизация (исключение напряженных контактов – релаксация)
- сольватация биополимера – водный бокс + PBC
- релаксация структуры воды вокруг биополимера
- медленный разогрев системы: $T_1 \sim 10\text{K}$, $T_2, \dots, T \sim 300\text{K}$
- начальные скорости атомов $\{v_1, \dots, v_N\}$ – случайные с Максвелловским распределением $\sim \exp(-mv^2/kT)$ при заданной T
- приведение системы в равновесие при (T, P)
- симуляция для расчета равновесных характеристик

Контроль МД эксперимента

Контроль температуры : $3/2NkT = m_i v_i^2 / 2$ протокол масштабирования скоростей $v_i \rightarrow k_T(T/T_0) v_i$

Контроль давления растворителя – масштабирование размеров основного бокса и координат молекул растворителя $r_i \rightarrow k_p(T/T_0) r_i$

$$PV = Nk_B T - \frac{1}{D} \left\langle \sum_i \sum_{j>i} r_{ij} \frac{d\phi}{dr} \Big|_{r_{ij}} \right\rangle. \quad \text{Вириальная формула P}$$

Возможные МД ансамбли: NEV ; NTV;
NTP

Интегрирование МД уравнений движения

- МД → реалистичные траектории - сохранение полной энергии, ограничения на временной шаг $\Delta t \sim 1-2 * 10^{-15}$ sec
- многошаговые схемы интегрирования – быстро и медленно меняющиеся силы, $f(t) = f_{fast}(t) + f_{slow}(t)$

Иерархия времен релаксации системы:

- метастабильная конформация ~ 100 ps ($t_{MD} \sim 1$ ns)
- конформационные переходы $\sim 10^{-9} - 10^{-6}$ s ($t_{MD} \sim 10-1000$ ns)
- самоорганизация белка $\sim 10^{-3} - 1$ s ($t_{MD} \sim 1$ s)

Свойства системы из МД данных

- средняя структура биополимера в наблюдаемой конформации

$$\langle A(t) \rangle = (1/N_t) \sum A(t_i)$$

- термические флуктуации структурных параметров

$$\langle A(t) \rangle^2 = (1/N_t) \sum [A(t_i) - \langle A \rangle]^2 - \text{степень конф.}$$

подвижности

- средняя потенциальная энергия
- свободная энергия

спектральные частотные характеристики колебаний

Оценка свободной энергии методом МД

Как рассчитать F или G :

$$F_A = -kT \ln w_A$$

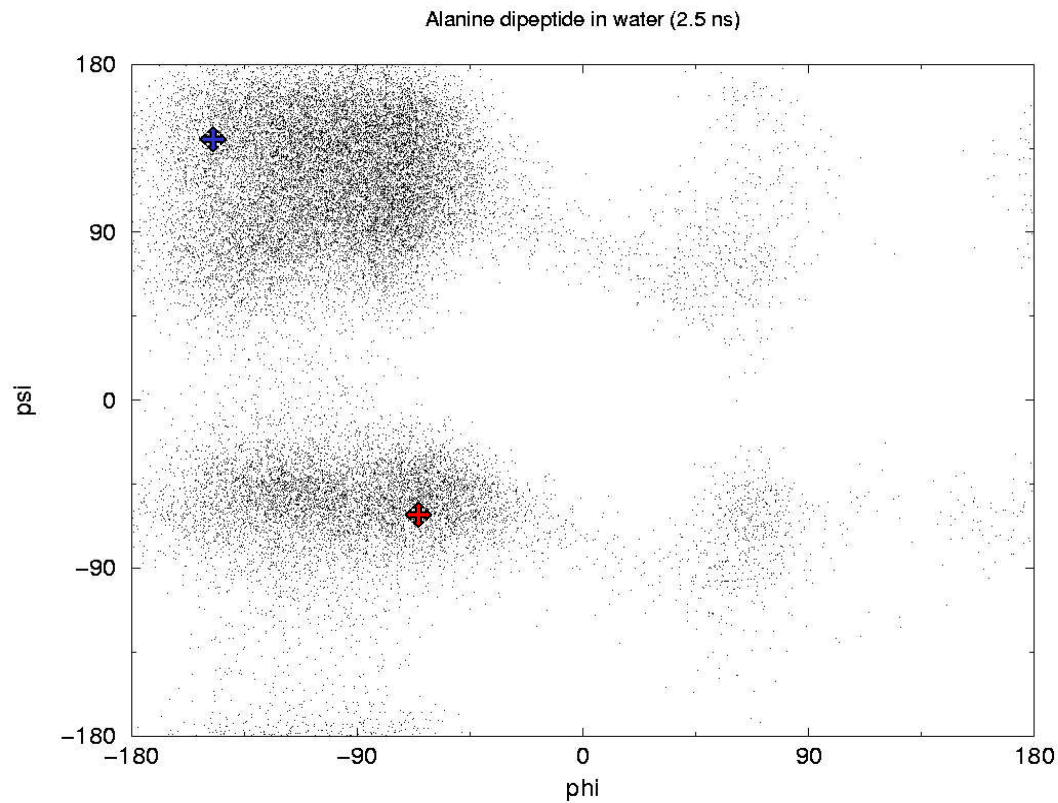
- Свободная МД дает распределение w_A по микро конформациям
- Разбить фазовое пространство на **ячейки**, анализировать МД мгновенные **состояния** $\{r_i(t)\}$ с равномерным интервалом t ,
- получить **гистограмму** вероятности нахождения МД системы в ячейках фазового пространства

$$w_i \approx -kT \ln (n_i / \sum n_j)$$

[n_i – число МД реализаций в ячейке]

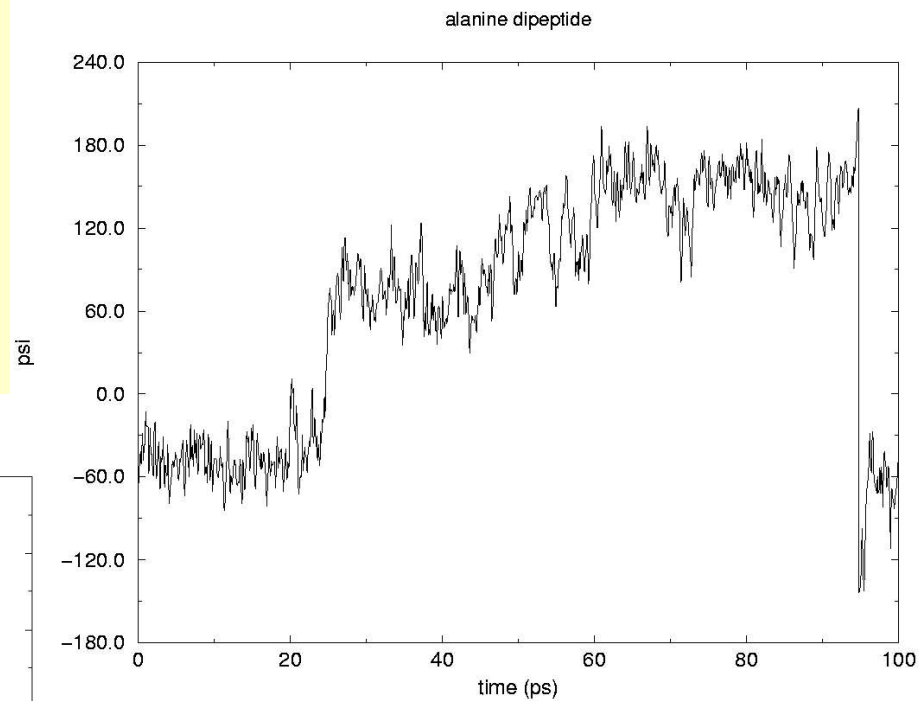
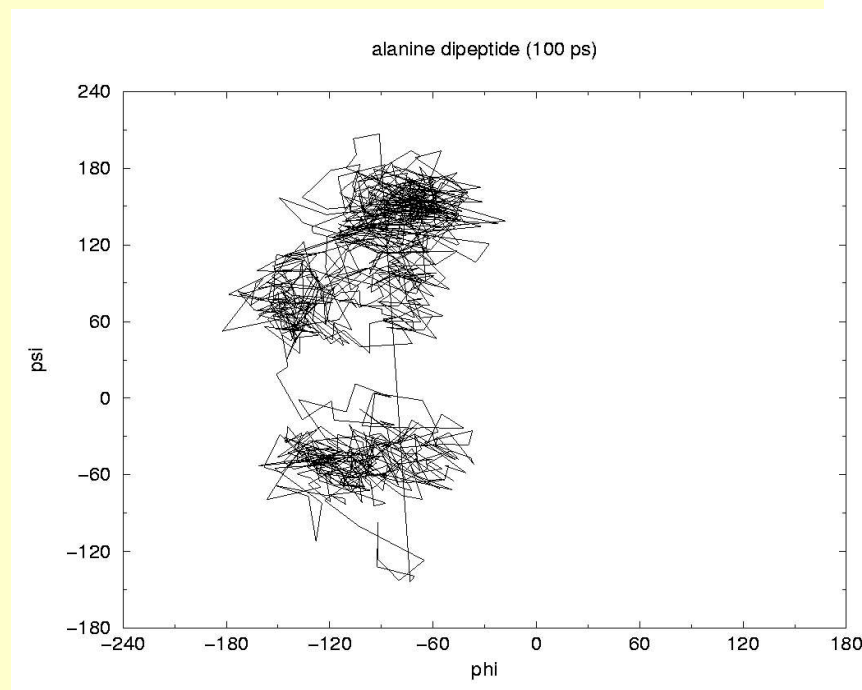
Свободная энергия

Alanine dipeptide 2.5 ns



β α_L
 α_R β^{inv}

Свободная энергия



Два перехода через / барьер

Равновесность – множество переходов

Метод термодинамического интегрирования

Принцип:

Если, молекулярная система определяется функцией потенциальной энергии координат: $U_{\text{pot}}(\mathbf{X})$

тогда: функция потенциальной энергии определяет молекулярную систему.

пусть: U_A описывает систему А, а U_B систему В

тогда: гибридная функция $U^*(\mathbf{X}, \lambda) = \lambda U_B + (1 - \lambda) U_A$ описывает систему В при $\lambda = 1$ и А $\lambda = 0$

при $0 < \lambda < 1$ имеем химерную систему

Симуляция химерной системы при изменении $0 < \lambda < 1$ переводит систему из состояния А в В

Из стат.механики (термодинамической теории возмущений):

$$\frac{\partial U^*(\mathbf{X}, \lambda)}{\partial \lambda} = U_A(\mathbf{X}) - U_B(\mathbf{X})$$
$$\Delta G_{A \rightarrow B} = \int_0^1 d\lambda \left\langle \frac{\partial U^*(\mathbf{X}, \lambda)}{\partial \lambda} \right\rangle_\lambda$$

- Метод МД дает возможность симуляции химерной системы и расчета средних по ансамблю $\langle \rangle$ состояний химерной системы
- Метод ТТВ имеет разнообразные применения:
 - оценка влияния аминокмслотных замен на **стабильность** молекулы белка,
 - зависимость **энергии связывания** лиганда от строения лиганда

Стабильность белков

Термодинамический цикл для оценки стабильности

expt.

Native: Folded (1) \Leftrightarrow Unfolded (2)

$\Downarrow\Uparrow$

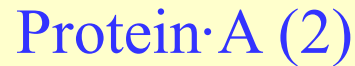
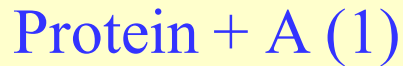
$\Downarrow\Uparrow$ MD simulation

Mutant: Folded (3) \Leftrightarrow Unfolded (4)

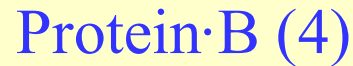
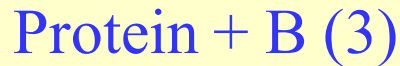
$$\Delta G^\circ_{12} + \Delta G^\circ_{24} = \Delta G^\circ_{13} + \Delta G^\circ_{34}$$

$$\Delta G^\circ_{12} - \Delta G^\circ_{34} = \Delta G^\circ_{13} - \Delta G^\circ_{24} = \Delta\Delta G$$

Связывание лигандов



МД симуляция ТТВ



$$\Delta G^\circ_{12} + \Delta G^\circ_{24} - \Delta G^\circ_{13} - \Delta G^\circ_{34} = 0 \quad (1)$$

$$\Delta G^\circ_{12} - \Delta G^\circ_{34} = \Delta G^\circ_{13} - \Delta G^\circ_{24} = \Delta \Delta G^\circ \quad (2)$$

$\Delta G^\circ_{12} - \Delta G^\circ_{34}$ **разность энергий связи лигандов А и В**

ESIS метод

- **ESIS** – Explicit Solvent Implicit Solvent model

$$F_A = \langle U_{\text{mol}}(\mathbf{x}) \rangle_A - TS_A(\mathbf{x}) + \langle W(\mathbf{x}) \rangle_A$$

усреднение по микросостояниям принадлежащим наблюдаемой макроскопической конформации A

IMPLICIT SOLVENT MODEL

- Свободная энергия сольватации

$$W(\mathbf{x}) = \int_0^1 d\lambda \frac{\int U_{m,s}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \exp\{-\beta[\lambda U_{m,s}(\mathbf{x}, \mathbf{y}) + U_{s,s}(\mathbf{y})]\}}{\int d\mathbf{y} \exp\{-\beta[\lambda U_{m,s}(\mathbf{x}, \mathbf{y}) + U_{s,s}(\mathbf{y})]\}}$$

$$W(\mathbf{x}) = G_{\text{cav}} + G_{\text{s,vdw}} + G_{\text{pol}}$$

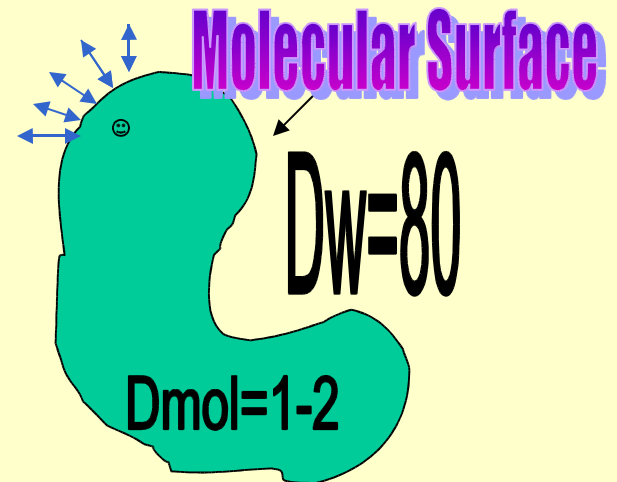
$$G_{\text{cav}} = \gamma S_{\text{mol}}$$

$$G_{\text{s,vdw}} = \langle U_{\text{s,mol}} \rangle$$

$$G_{\text{pol}} = \int_0^1 d\lambda \sum_i [q_i \langle V_{\text{solv}}(\mathbf{r}_i) \rangle_\lambda]$$

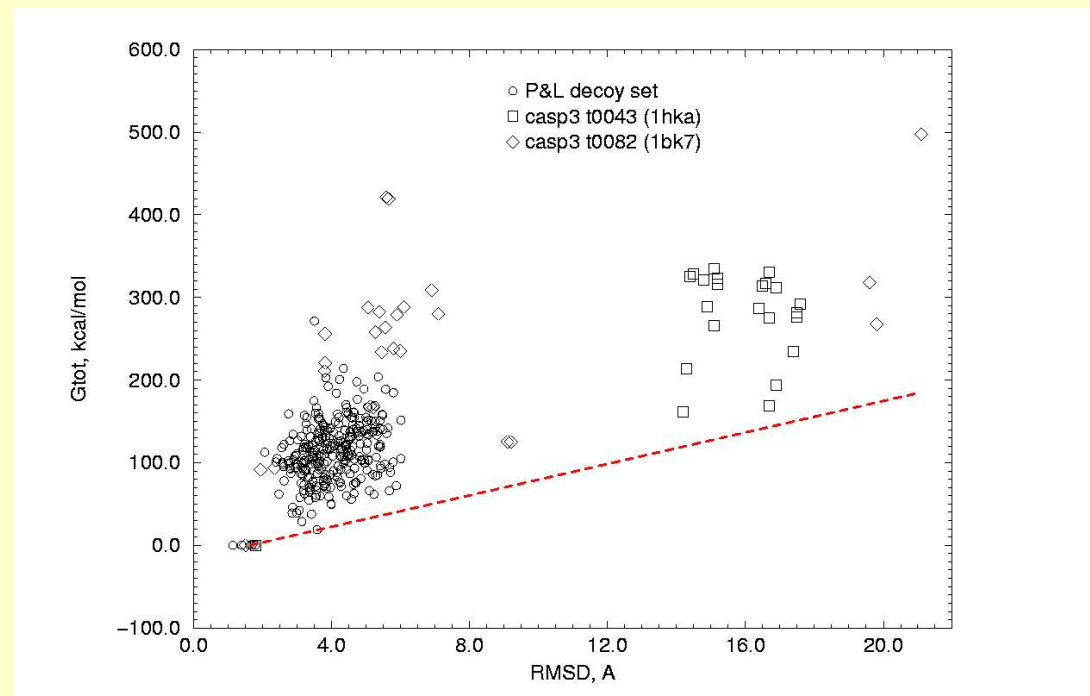
$$\langle V_{\text{solv}}(\mathbf{r}_i) \rangle \cong \text{Poisson Eq.}$$

$\sim 10^3$ faster



Protein conformation Stability

- Protein decoy set ~ 1000 of compact structures + NATIVE



Анализ существенной динамики

МД траектория $\{X_i\}(t) \rightarrow C_{ij}$ – матрица позиционных флуктуаций

$$C_{ij} = \langle (x_i - \langle x_i \rangle) (x_j - \langle x_j \rangle) \rangle$$

Диагонализация C_{ij} определяет набор обобщенных координат, которые независимы (ортогональны)

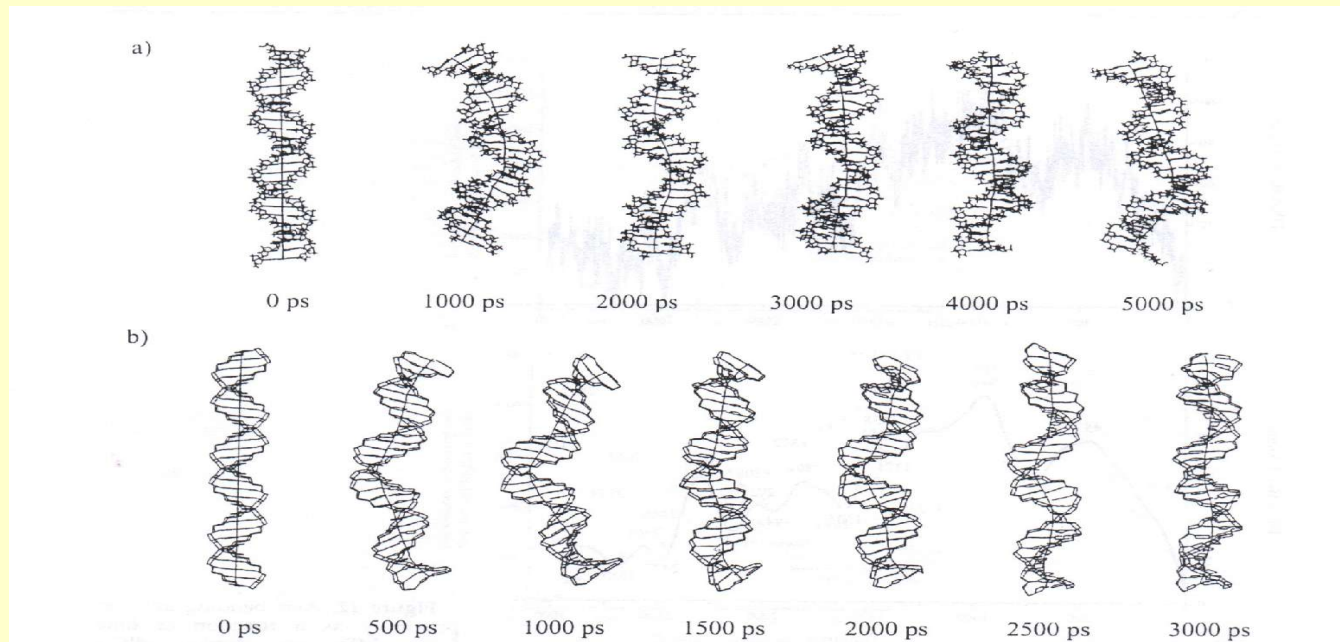
-**собственные векторы** = обобщенные координаты (ОК)

-**собственные значения** = амплитуда флуктуаций ОК

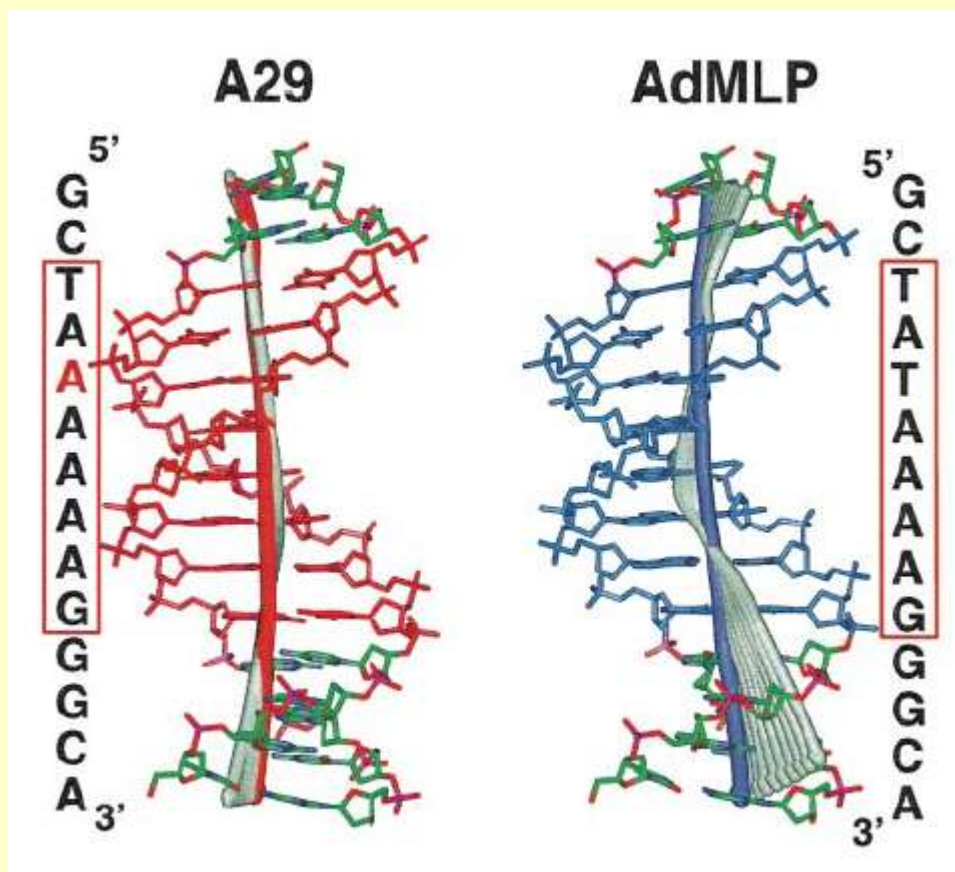
- основная часть движений (80%) сосредоточена в малом числе ОК ~ 10 движений

► Example of long time DNA MD simulation

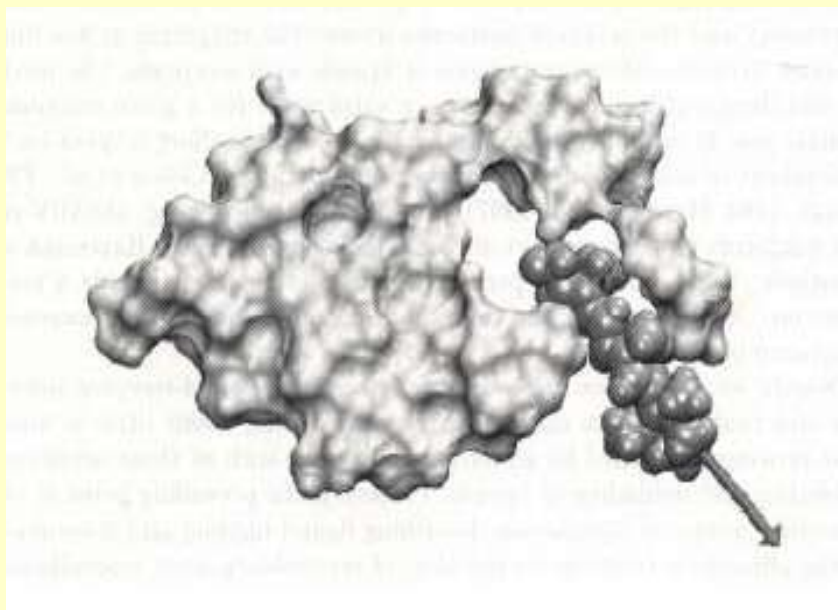
- DNA MD trajectory



Анализ существенных движений DNA изгибы



Механистическое понимание процессов взаимодействия —
Экстракция лиганда **biotine** из связывающего центра
белка **avidine**



МД - липидный двойной слой в воде

