

A faint, light gray background image of a protein structure, likely a ribbon diagram, is centered on the slide. The protein is shown in a complex, folded conformation with several alpha-helices and beta-strands. The text is overlaid on this image.

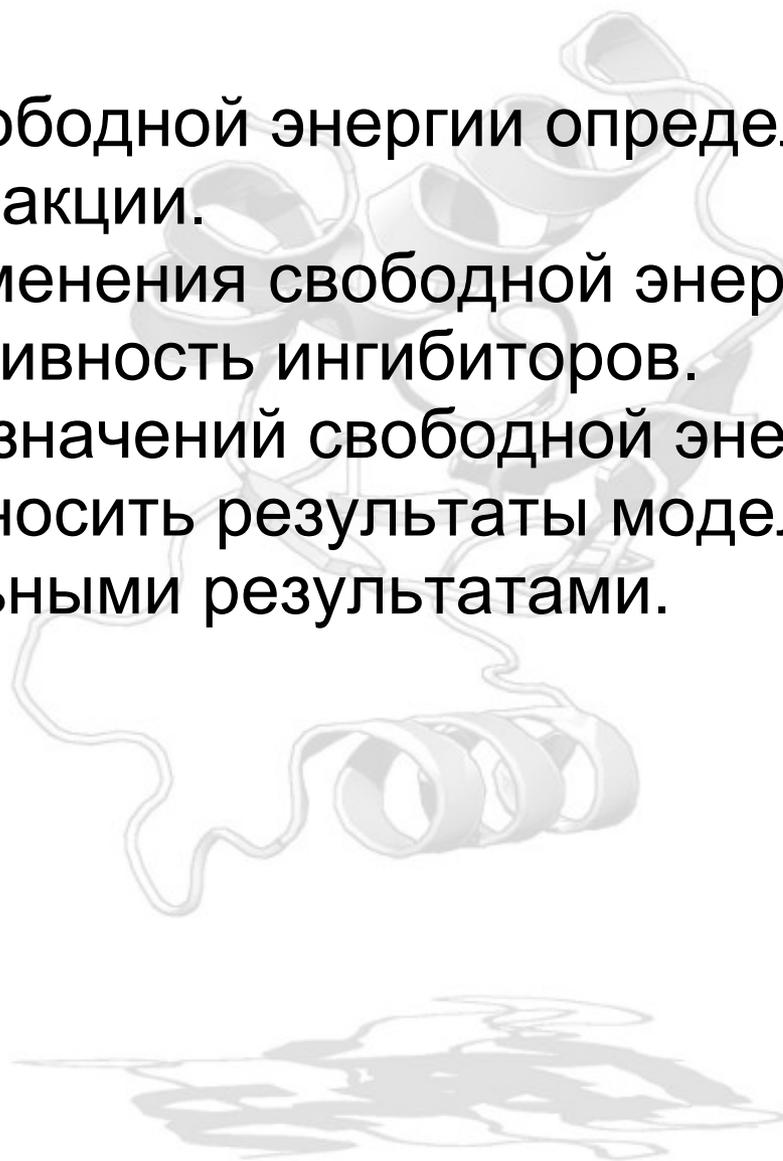
# **Молекулярное моделирование в применении к биомолекулам**

## **Лекция 10**

### **Расчёты свободной энергии**

# Свободная энергия

- Изменение свободной энергии определяет направление реакции.
- Сравнивая изменения свободной энергии можно изучать эффективность ингибиторов.
- Определение значений свободной энергии позволяет соотносить результаты моделирования с экспериментальными результатами.



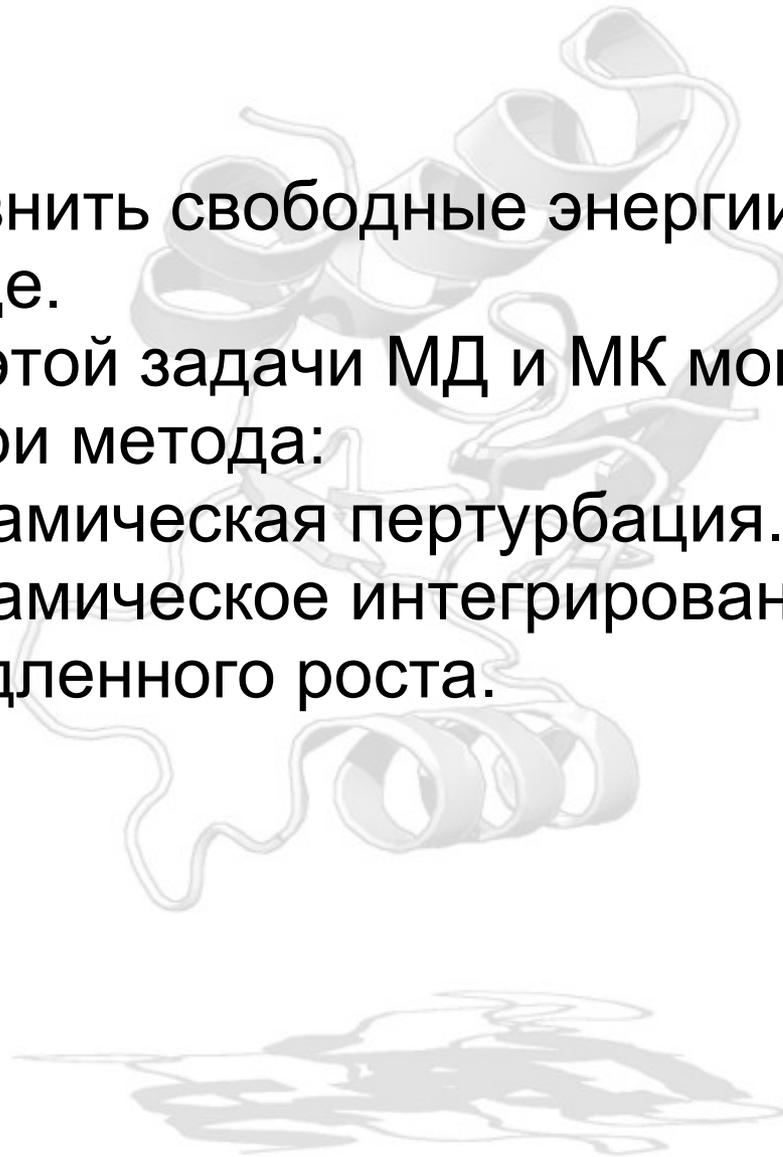
# Немного теории

- Так как большинство экспериментальных данных получено при постоянном давлении, то наиболее распространёно выражение свободной энергии как энергии Гиббса.
- МД и МК трудно использовать для расчёта энергии Гиббса, так как оба метода “не любят” те места фазового пространства где энергия не минимальна.

$$A = k_B T \ln \left( \int \int \partial p^N \partial r^N \exp \left( \frac{-H(p^N, r^N)}{k_B T} \right) \rho(p^N, r^N) \right)$$

# А что если:

- Нам надо сравнить свободные энергии этанола и этантиола в воде.
- Для решения этой задачи МД и МК могут подойти.  
Рассмотрим три метода:
  - Термодинамическая пертурбация.
  - Термодинамическое интегрирование.
  - Метод медленного роста.



# Термодинамическая пертурбация

X - это этанол в кубике воды

Y - это этантиол в кубике воды

$$\Delta A = A_Y - A_X = k_B T \ln \frac{Q_Y}{Q_X}$$

$$\Delta A = k_B T \left( \frac{\int \int \partial p^N \partial r^N \exp(-H_Y(p^N, r^N)/k_B T)}{\int \int \partial p^N \partial r^N \exp(-H_X(p^N, r^N)/k_B T)} \right)$$

$$\Delta A = k_B T \left( \frac{\int \int \partial p^N \partial r^N \exp(-H_Y(p^N, r^N)/k_B T) \exp(+H_X(p^N, r^N)/k_B T) \exp(-H_X(p^N, r^N)/k_B T)}{\int \int \partial p^N \partial r^N \exp(-H_X(p^N, r^N)/k_B T)} \right)$$

$$\Delta A = k_B T \langle \exp(H_Y(p^N, r^N)/k_B T - H_X(p^N, r^N)/k_B T) \rangle_0$$

Итак нам надо посчитать поведение этанола в воде и для найденных конформаций посчитать энергию этантиола. И наоборот.

# Термодинамическая пертурбация

- Предыдущий подход хорош если фазовые пространства молекул похожи или пересекаются.
- Если пространства не пересекаются давайте введём состояние 1 между состояниями X и Y. т. е.:

$$k_B T \ll |H_Y - H_X|$$

$$\Delta A = k_B T \langle \exp((H_Y - H_1)/k_B T) \rangle - k_B T \langle \exp((H_1 - H_X)/k_B T) \rangle$$

# Термодинамическая пертурбация, реализация

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

Задаём значение  $\lambda$  от 0 до 1 и делаем МД для каждого  $\lambda$

$$k_b(\lambda) = \lambda k_b(Y) + (1 - \lambda) k_b(X)$$

$$k_a(\lambda) = \lambda k_a(Y) + (1 - \lambda) k_a(X)$$

.....

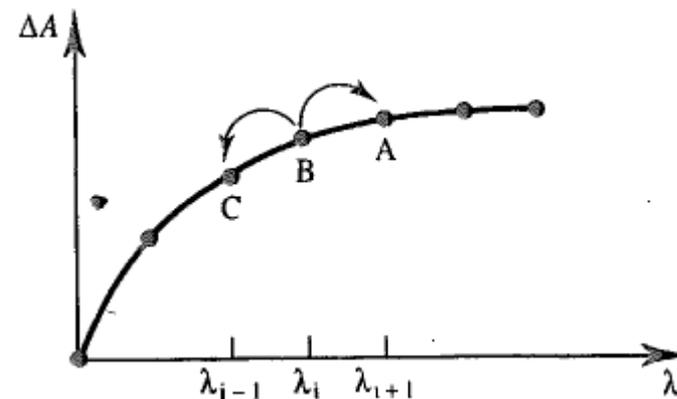
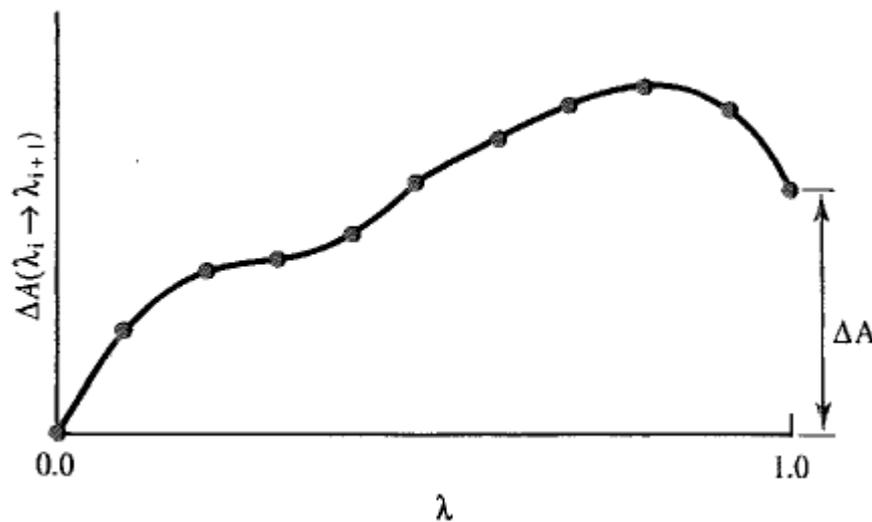
$$q(\lambda) = \lambda q(Y) + (1 - \lambda) q(X)$$

$$\epsilon(\lambda) = \lambda \epsilon(Y) + (1 - \lambda) \epsilon(X)$$

$$\sigma(\lambda) = \lambda \sigma(Y) + (1 - \lambda) \sigma(X)$$

# Термодинамическая пертурбация, реализация

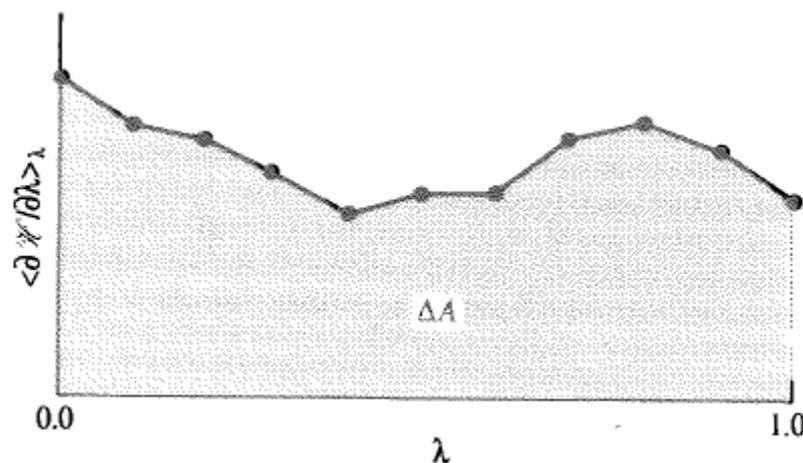
- На самом деле для каждого значения параметра сопряжения ( $\lambda$ ) надо сначала уравновесить систему и только после этого снимать значения энергий.
- Расчёт с  $\lambda$  от 0 до 1 это прямая выборка
- Расчёт с  $\lambda$  от 1 до 0 это обратная выборка
- Бывает двойная выборка



# Термодинамическое интегрирование

- Можно интегрировать изменение энергии по параметру сопряжения

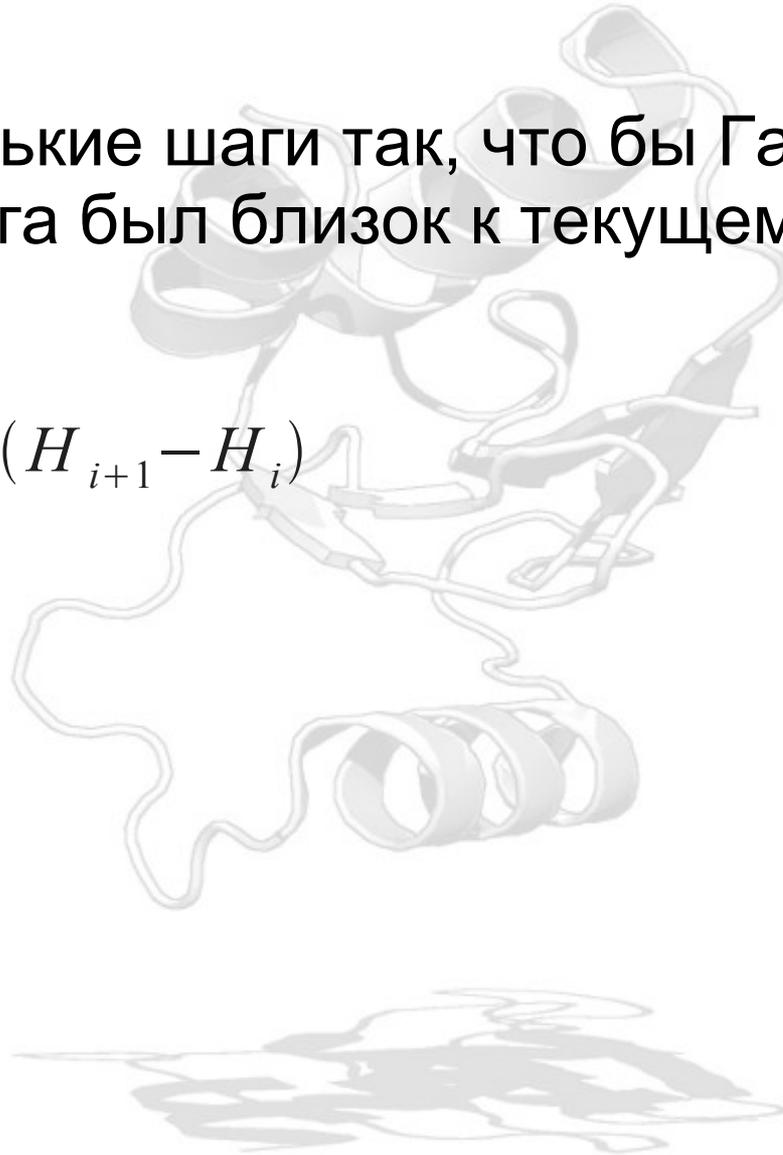
$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(p^N, r^N)}{\partial \lambda} \right\rangle_{\lambda} d\lambda; \quad \frac{\partial H}{\partial \lambda} \approx \frac{\Delta H}{\Delta \lambda}$$



# Метод медленного роста

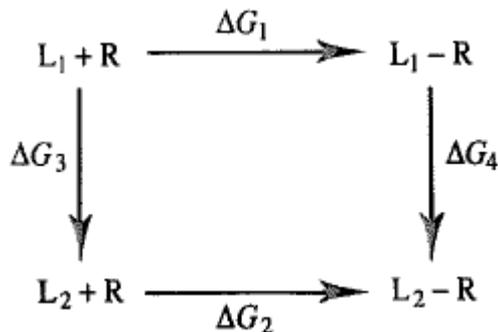
- Делаем маленькие шаги так, что бы Гамольтиан следующего шага был близок к текущему.

$$\Delta A = \sum_{i=0; \lambda=0}^{i=N; \lambda=1} (H_{i+1} - H_i)$$



# Термодинамические циклы

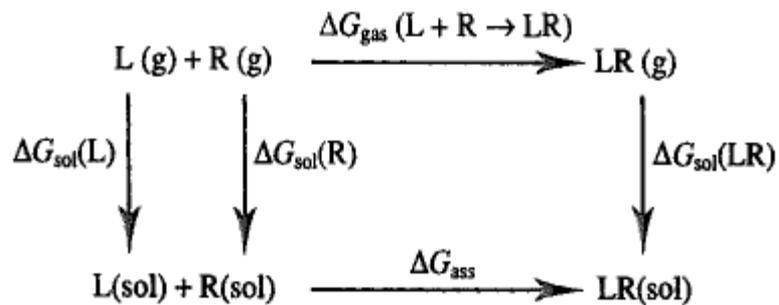
- Исследователей часто интересует энергия нековалентного связывания лиганда с рецептором.
- Допустим у нас есть два лиганда. Можно посчитать их  $\Delta\Delta G$  просто промоделировав процесс связывания, но это трудно исполнимо.
- Так как свободная энергия это функция состояния, то:



Надо всего-то посчитать переход из одного лиганда в другой, как в растворе, так и в белке.

# Расчёт абсолютного изменения свободной энергии

- Основная идея — это делать термодинамические циклы через состояния комплекса в растворе и газовой фазе.



$$\Delta G_{ass} = \Delta G_{gas}(R + L \rightarrow RL) + \Delta G_{sol}(RL) - \Delta G_{sol}(R) - \Delta G_{sol}(L)$$

подставляем  $\Delta G_{sol}(X) = \Delta G_{gas}(X \rightarrow 0) - \Delta G_{sol}(X \rightarrow 0)$  и *gas* сокращается

$$\Delta G_{ass} = \Delta G_{sol}(L \rightarrow 0) - \Delta G_{sol}(LR \rightarrow R)$$

# Расчёт изменения энтальпии и энтропии

- Изменение свободной энергии может быть рассчитано достаточно точно. В хороших случаях ошибка всего 1 ккал/моль.
- Изменение энтальпии можно было бы посчитать сравнив потенциальные энергии двух систем, но там достаточно большие значения с большими ошибками.
- На сегодняшний момент расчёт энтальпии и энтропии в моделирование даёт ошибки на порядок большие, чем расчёт свободной энергии.

# Компоненты свободной энергии

При использовании термодинамического интегрирования:

$$\Delta A = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(p^N, r^N)}{\partial \lambda} \right\rangle_{\lambda} d\lambda$$

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda} = \left\langle \frac{\partial H_{bond}}{\partial \lambda} + \frac{\partial H_{angle}}{\partial \lambda} + \dots \right\rangle_{\lambda}$$

$$\Delta A = \Delta A_{bonds} + \Delta A_{angles} \dots$$

- Необходимо отметить, что только сумма компонент является осмысленной.
- Интересное применение этого подхода получило исследование связывания биотина со стрептавидином.

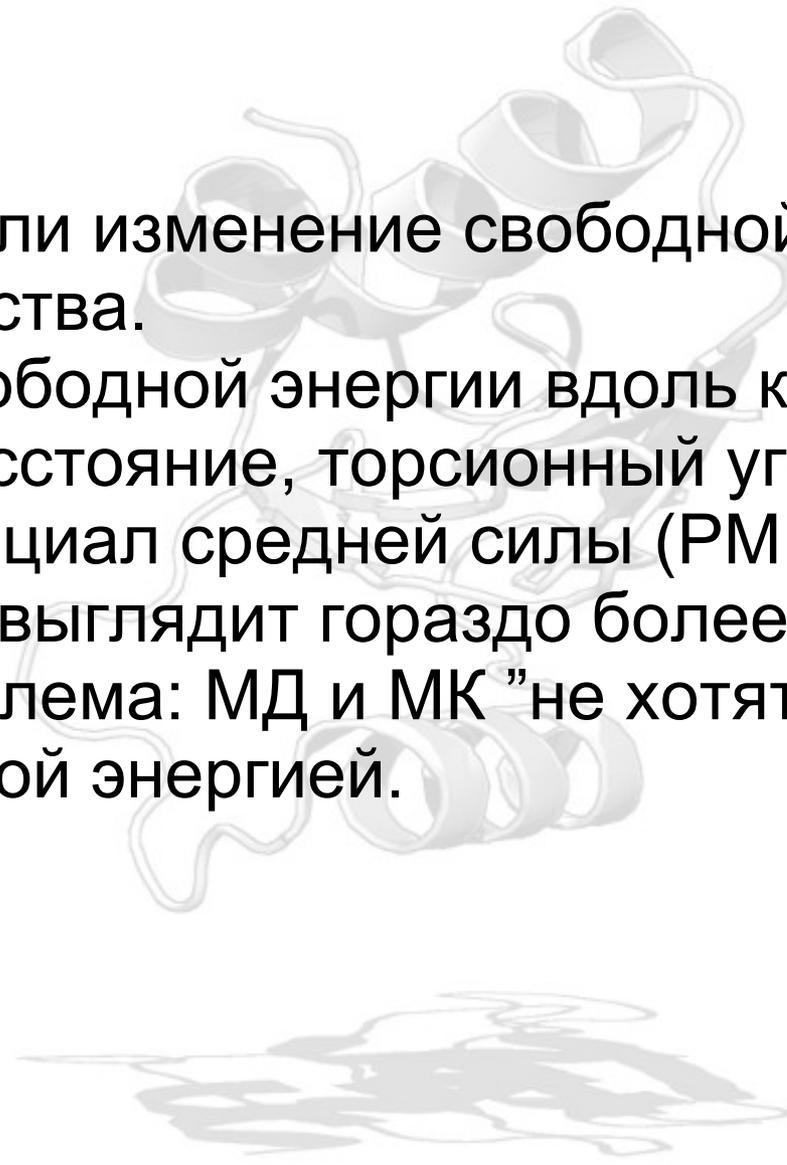
# Подводные камни

- Существует два основных источника ошибок:
  - Неточность расчёта гамильтиана, ошибки в силовом поле, неправильный расчёт взаимодействий .
  - Недостаточная выборка из фазового пространства.
- К сожалению нет рецепта для определения достаточности выборки.
- Сравнение прямой и обратной выборки может указывать на гистерезис. Если выборка маленькая гистерезис будет стремиться к 0. Это явный признак малой выборки.

# Особенности применения методов

- МК используют для малых жестких молекул.
- МД используют для крупных информационно подвижных молекул.
- Метод медленного роста почти не используют, так как считается, что молекула не успевает адаптироваться к изменению  $\lambda$ .
- Преимущество интегрирования и пертурбации это возможность уточнить некий диапазон  $\lambda$  без пересчёта остальных значений.
- Возможно динамическое изменение  $\lambda$ .
- Иногда используют модифицированные потенциалы.

# PMF

- Мы рассмотрели изменение свободной энергии при "мутации" вещества.
  - Изменение свободной энергии вдоль какой либо координаты (расстояние, торсионный угол и т. д.) называют потенциал средней силы (PMF)
  - Этот процесс выглядит гораздо более "физичным"
  - Опять же проблема: МД и МК "не хотят уходить" из областей с низкой энергией.
- 

# PMF, umbrella sampling

- Umbrella sampling (US) модифицирует потенциал, что позволяет эффективно исследовать области с высокой энергией.
- Модификация потенциала записывается как пертурбация:

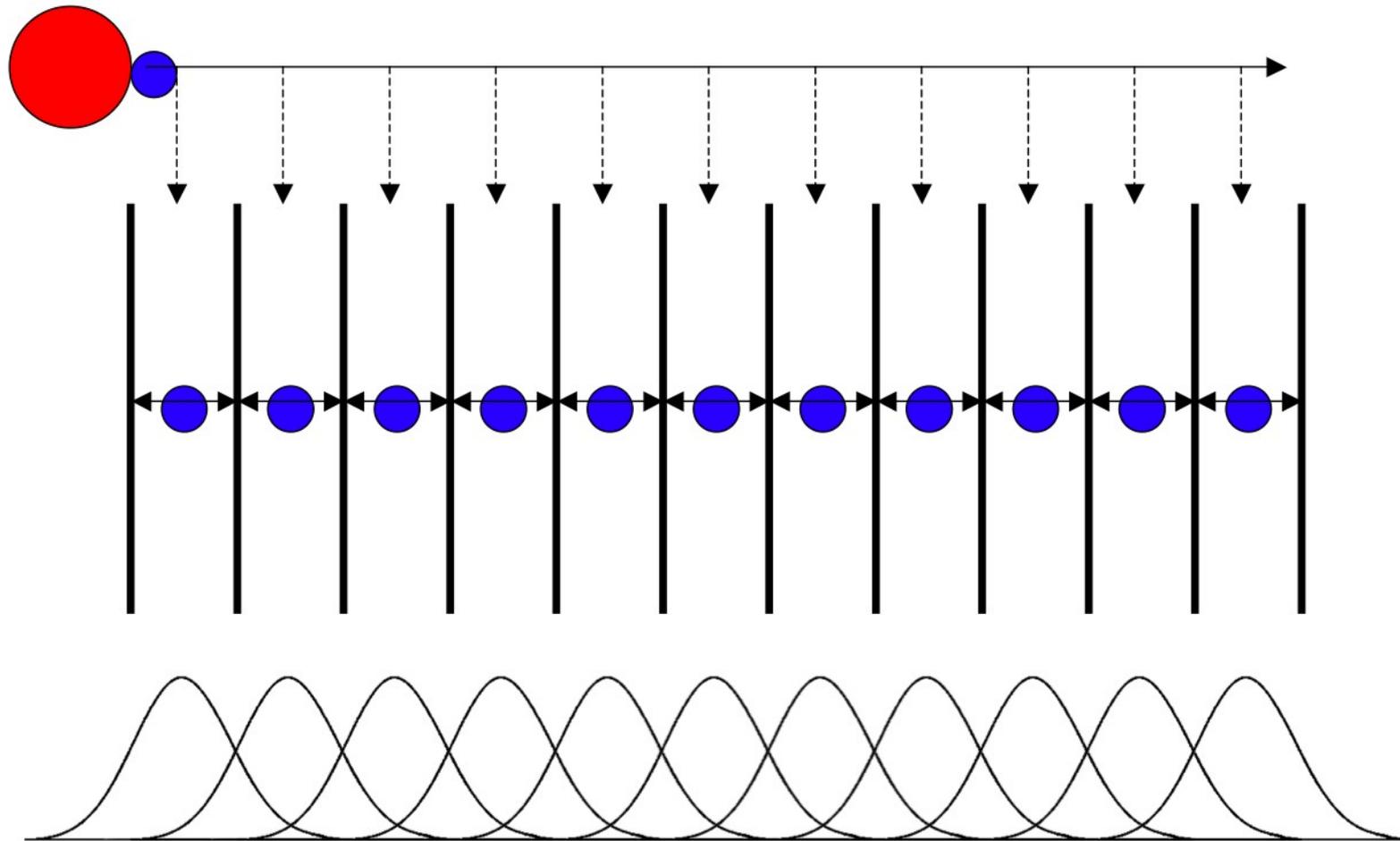
$$\dot{U}(r^N) = U(r^N) - W(r^N), \text{ часто } W(r^N) = k_w (r^N - r_0^N)^2$$

- Естественно, что распределение не Больцмановское. Но можно поправить:

$$\langle A \rangle = \frac{\langle A(r^N) \exp(+W(r^N)/k_B T) \rangle_W}{\langle \exp(+W(r^N)/k_B T) \rangle_W}$$

# Пример:

- Рассмотрим диссоциацию протофибрилы и одной молекулы аммилоидного пептида.



# Пример:

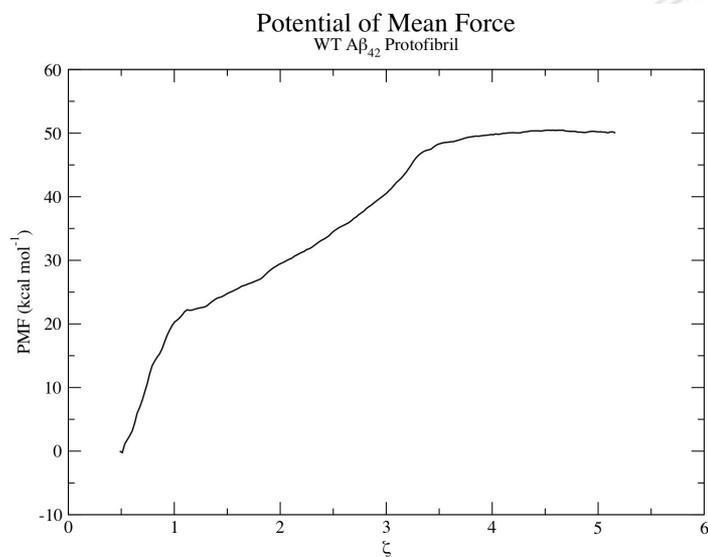
- **Запустим:**

```
grompp -f md_umbrella.mdp -c conf0.gro -p topol.top -o umbrella0.tpr  
...  
grompp -f md_umbrella.mdp -c conf450.gro -p topol.top -o umbrella22.tpr
```

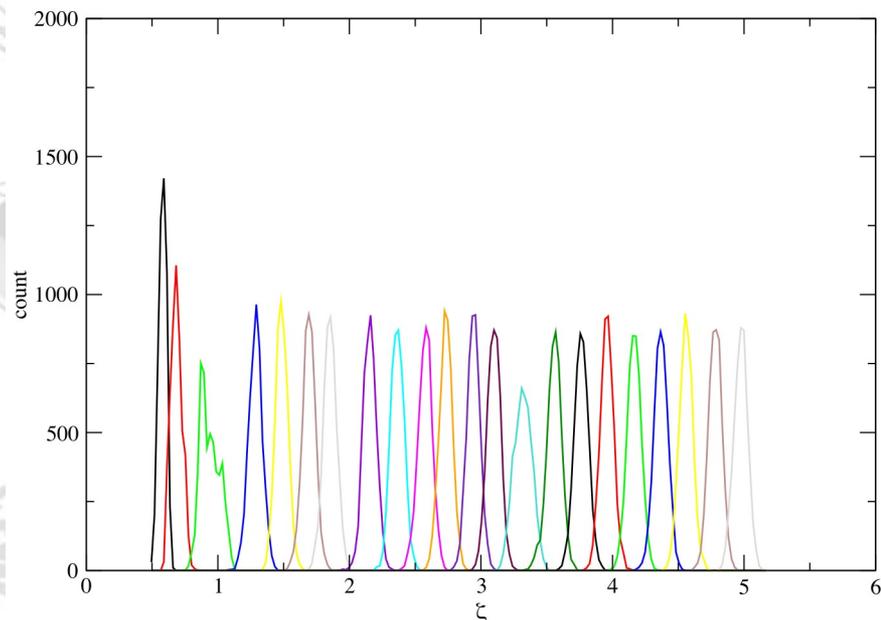
- **Обрабатываем:**

```
g_wham -it tpr-files.dat -if pullf-files.dat -o -hist -unit kCal
```

- **Итого:**

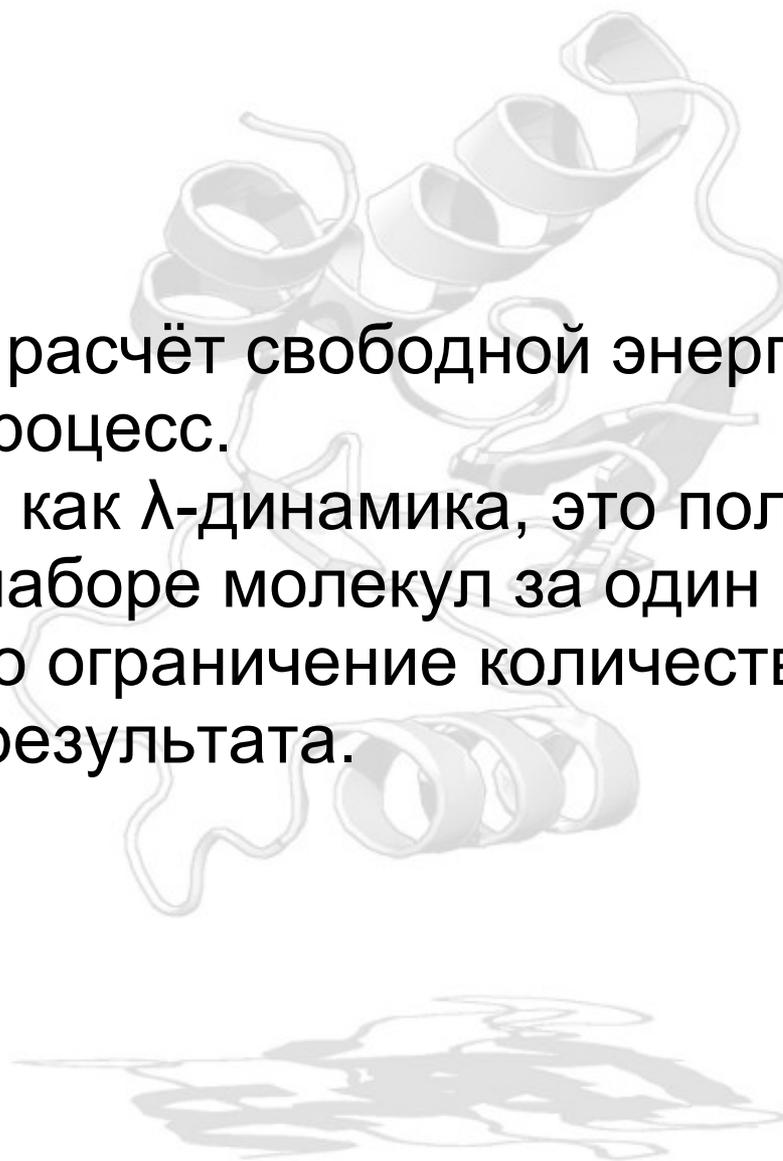


Umbrella histograms



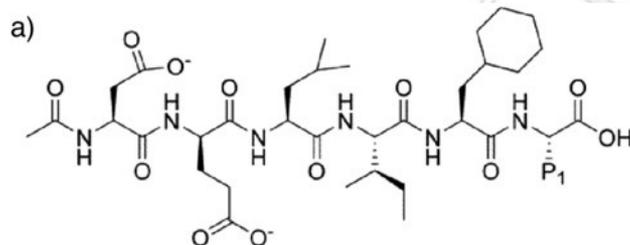
# «Быстрые» методы расчёта

- Очевидно, что расчёт свободной энергии это ресурсоёмкий процесс.
- Один из путей, как  $\lambda$ -динамика, это получение информации о наборе молекул за один расчёт.
- Другой путь это ограничение количества запусков для получения результата.



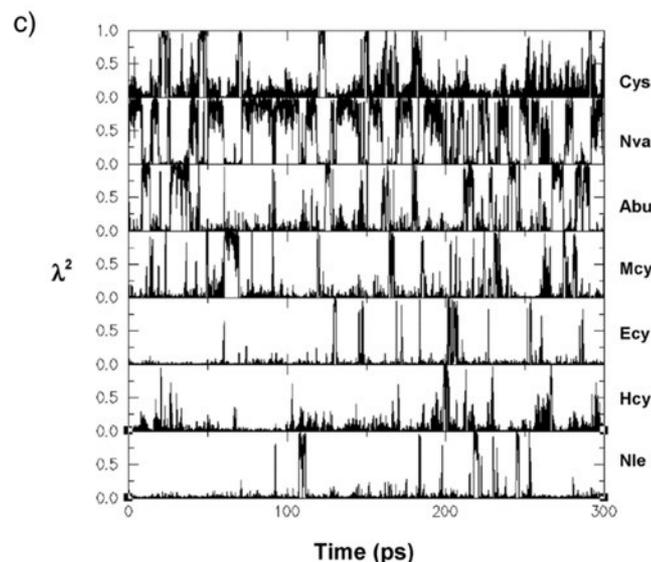
# $\lambda$ -динамика

- Основная суть — это изменение  $\lambda$  в ходе моделирования, причём изменение не только от одного вещества к другому, но к множеству других.
- Например исследуем как меняется  $\lambda$  при исследовании многих заместителей.



b)

Name	P1	Ki ( $\mu\text{M}$ )
Cys	$\text{CH}_2\text{SH}$	0.12
Nva	$(\text{CH}_2)_2\text{CH}_3$	0.13
Mcy	$\text{CH}_2\text{SCH}_3$	0.5
Abu	$\text{CH}_2\text{CH}_3$	0.6
Nle	$(\text{CH}_2)_3\text{CH}_3$	2.0
Ecy	$\text{CH}_2\text{SCH}_2\text{CH}_3$	5.0
Hcy	$(\text{CH}_2)_2\text{SH}$	80.0



# Linear response (LR)

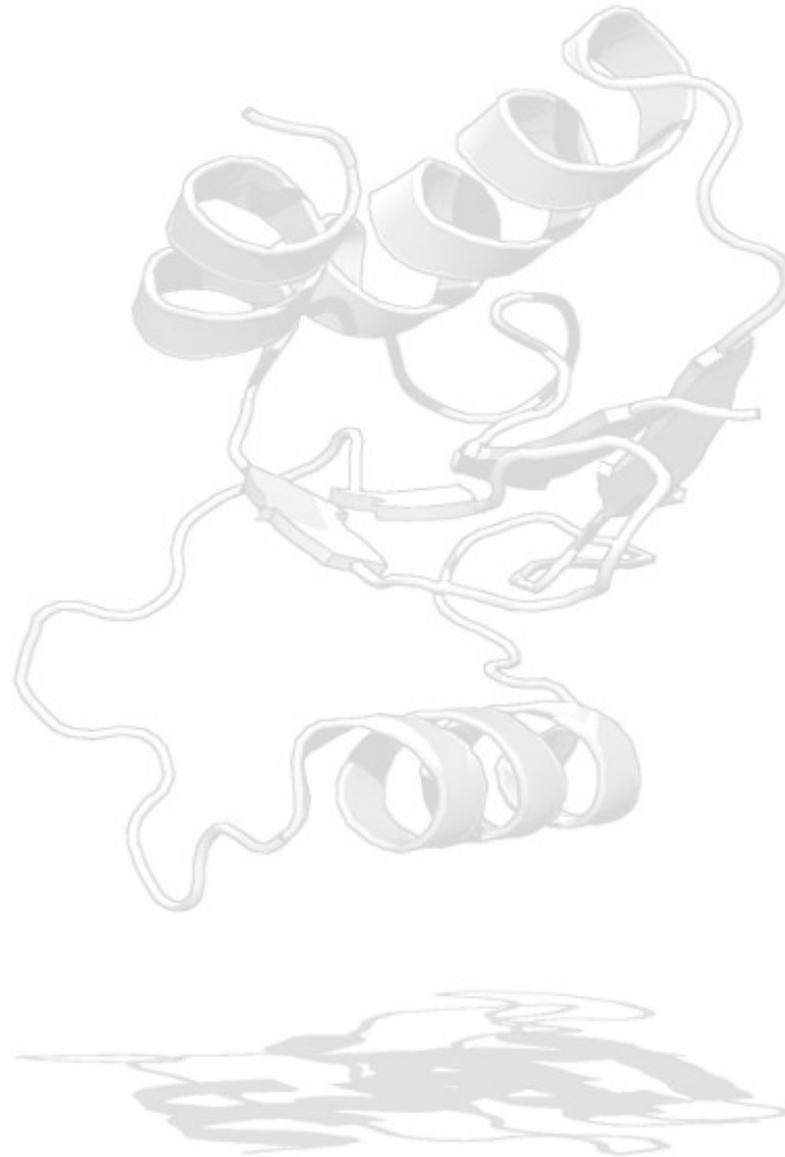
- Суть идеи состоит в проведении всего двух запусков комплекса и лиганда в воде.

$$\Delta G = \beta (\langle U_{LR}^{el} \rangle - \langle U_{LS}^{el} \rangle) + \alpha (\langle U_{LR}^{VdW} \rangle - \langle U_{LS}^{VdW} \rangle)$$

$$\Delta G_{hydr} = \beta \langle U_{hydr}^{el} \rangle + \alpha \langle U_{hydr}^{VdW} \rangle + \gamma SASA$$

- Где коэффициенты находятся либо аналитически, либо, как во втором случае эмпирически, подгонкой.

# Вопросы?



...