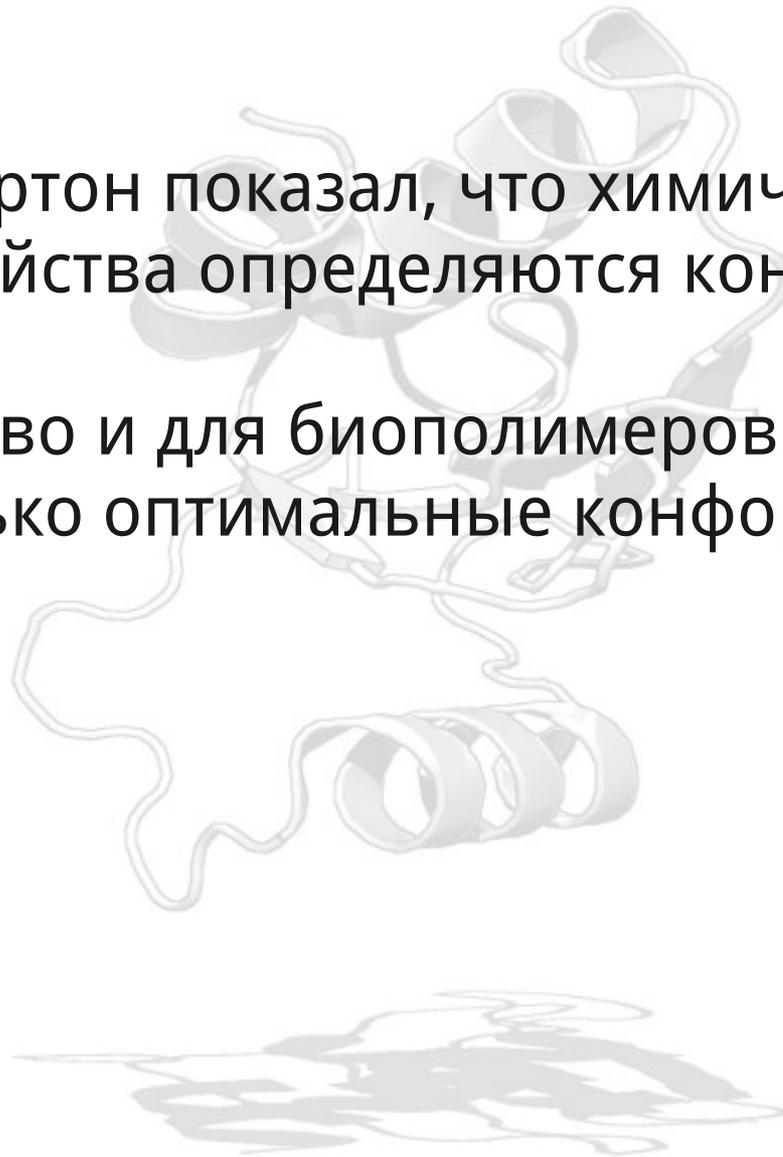


**Молекулярное моделирование в  
применении к биомолекулам**

**Лекция 9  
Анализ конформаций**

# Конформации

- В 1950 году Бартон показал, что химические и физические свойства определяются конформациями молекул
- Это справедливо и для биополимеров.
- Важны не только оптимальные конформации.



# Систематические методы исследования конформаций

## Поиск по сетке:

- Находим торсионные углы способный вращаться.
- Генерируем ротамеры по каждому углу с небольшим шагом.
- Минимизируем энергию для каждого ротамера .

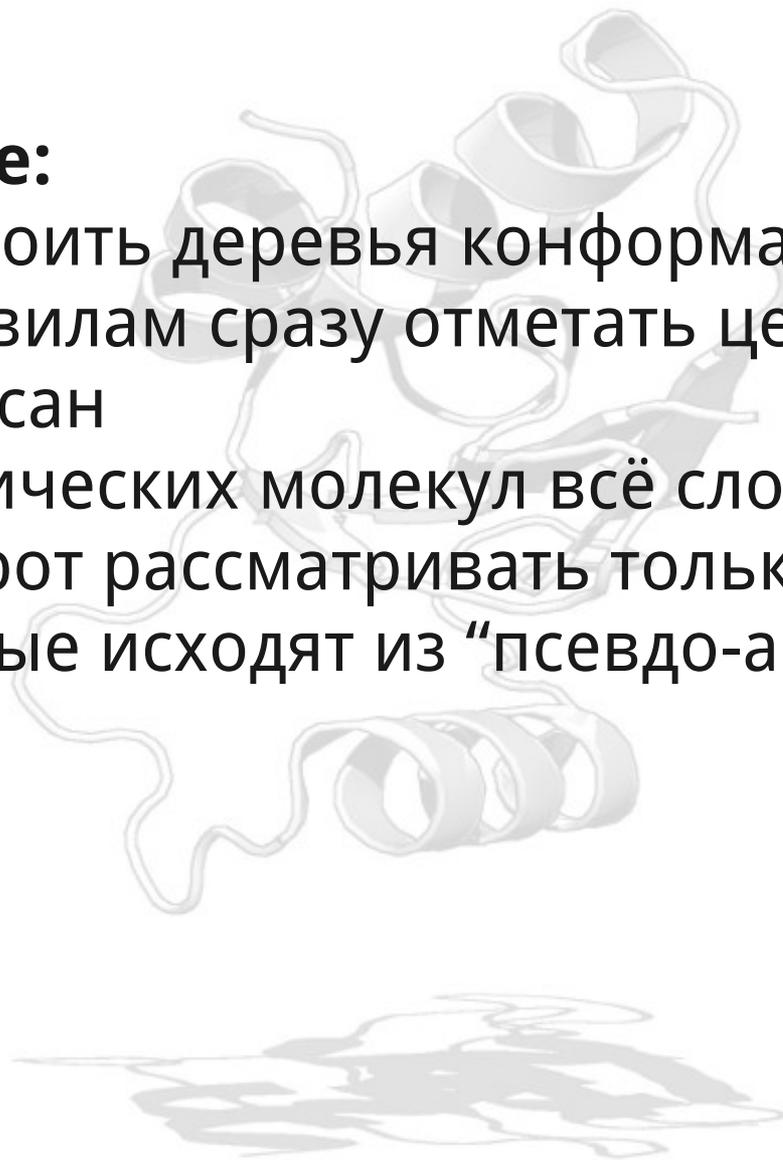
$$N_{conf} = \prod_{i=1}^N \frac{360}{\alpha_i}$$

Если у нас пять вращаемых связей и инкремент 30 то количество конформаций 248832

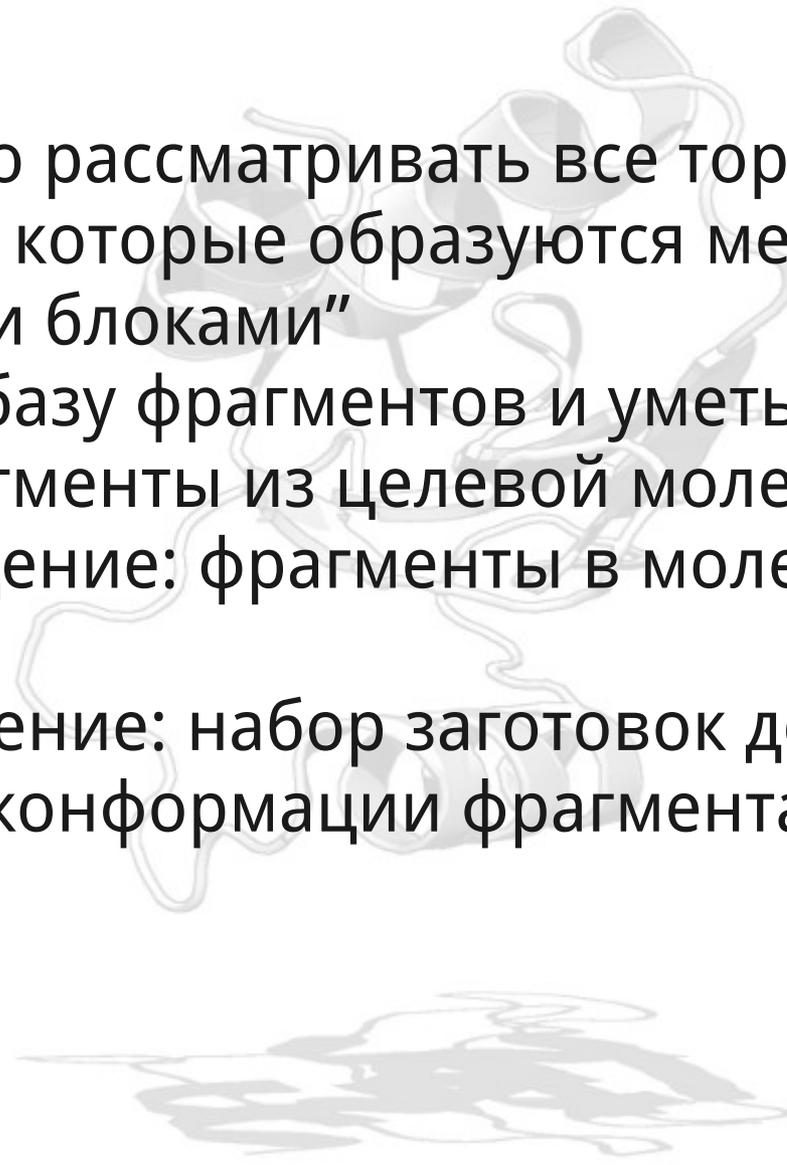
# Систематические методы исследования конформаций

## Поиск по сетке:

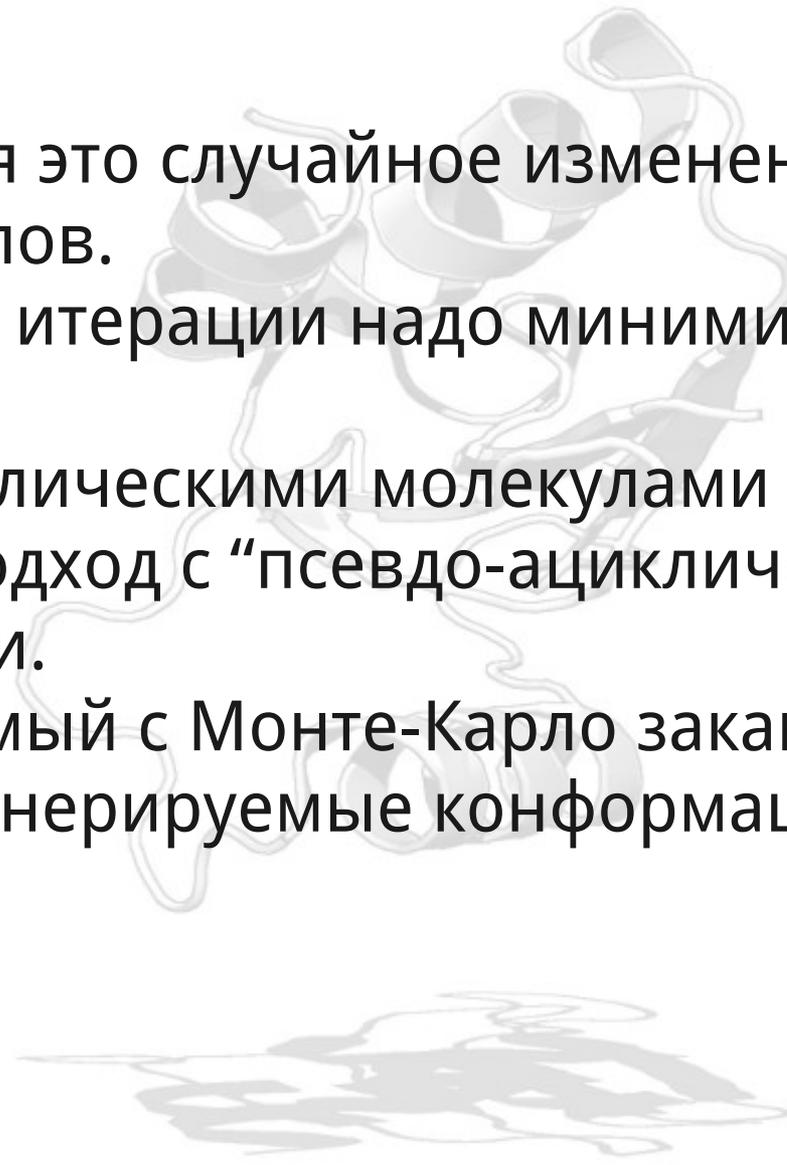
- Мы можем строить деревья конформаций и по некоторым правилам сразу отметить целые ветки
- Например: гексан
- В случае циклических молекул всё сложнее.
- Можно наоборот рассматривать только те ветки деревьев которые исходят из “псевдо-ациклической” конформации



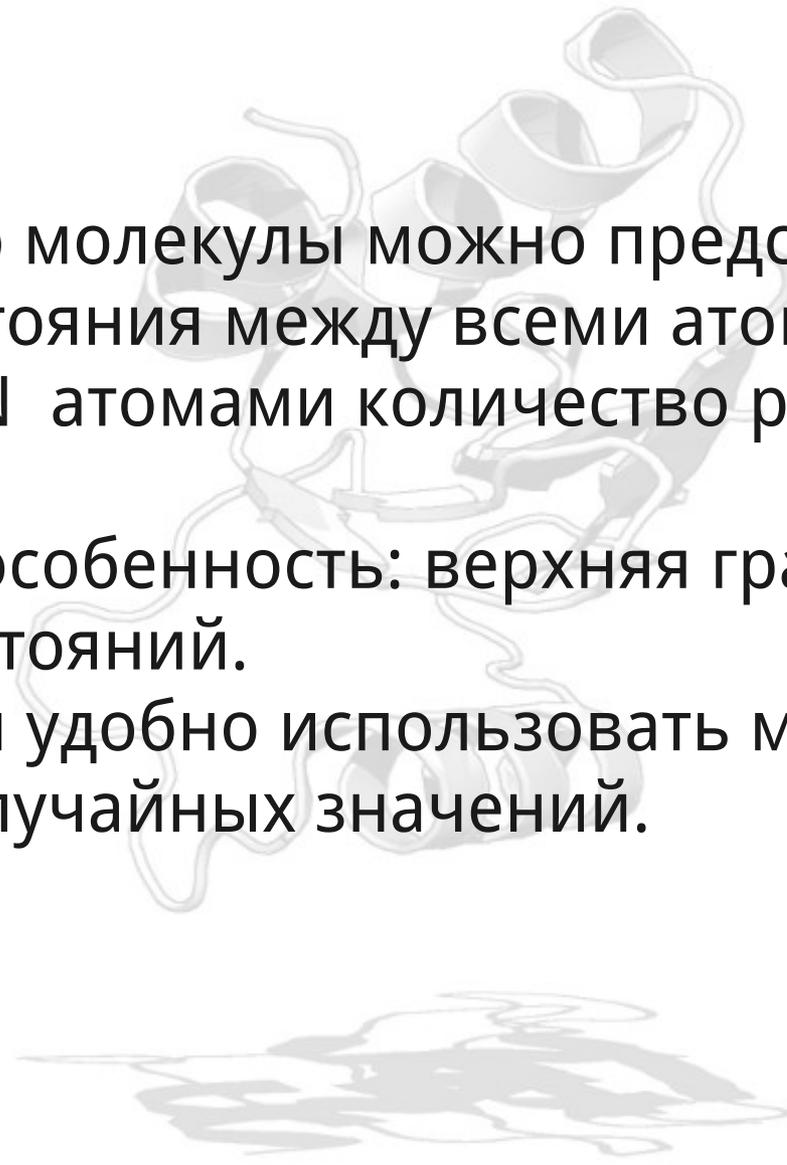
# Подходы по строению модели из блоков

- Не обязательно рассматривать все торсионные углы, только те, которые образуются между “строительными блоками”
  - Важно иметь базу фрагментов и уметь точно вычленять фрагменты из целевой молекулы.
  - Первое допущение: фрагменты в молекуле не зависимы
  - Второе допущение: набор заготовок должен покрывать все конформации фрагмента в составе молекулы
- 

# Методы случайного поиска

- Основная идея это случайное изменение значений торсионных углов.
  - В ходе каждой итерации надо минимизировать энергию.
  - В случае с циклическими молекулами снова применяется подход с “псевдо-ациклическими” конформациями.
  - Поиск сравнимый с Монте-Карло заканчивается при условии, что генерируемые конформации уже встречались.
- 

# Дистанционная геометрия

- Конформацию молекулы можно представить как попарные расстояния между всеми атомами.
  - В молекуле с  $N$  атомами количество расстояний  $N(N-1)/2$ .
  - Есть удобная особенность: верхняя граница для некоторых расстояний.
  - Тут становится удобно использовать матрицы для оптимизации случайных значений.
- 

# Дистанционная геометрия и ЯМР

- Прямое применение таких методов это поиск конформаций удовлетворяющих экспериментальным расстояниям из ЯМР.
- 2D-NOESY сигнал от близко расположенных атомов, убывает пропорционально 6ой степени от расстояния.
- 2D-COESY это информация об атомах разделённых 3мя связями.
- В итоге мы получаем несколько конформаций и если для некоторых частей у нас нет данных, то они могут иметь большую конформационную подвижность.

# Поиск конформаций на основе моделирования (МД или МК)

- В принципе оба метода могут быть использованы для поиска конформаций
- МК методы имеют ряд трудностей при работе с гибкими молекулами. Напоминанию, что МК методы не используют минимизацию энергии.
- Методы МД используют при высоких температурах. Конформации с минимальной энергией подвергают процедуре минимизации энергии.

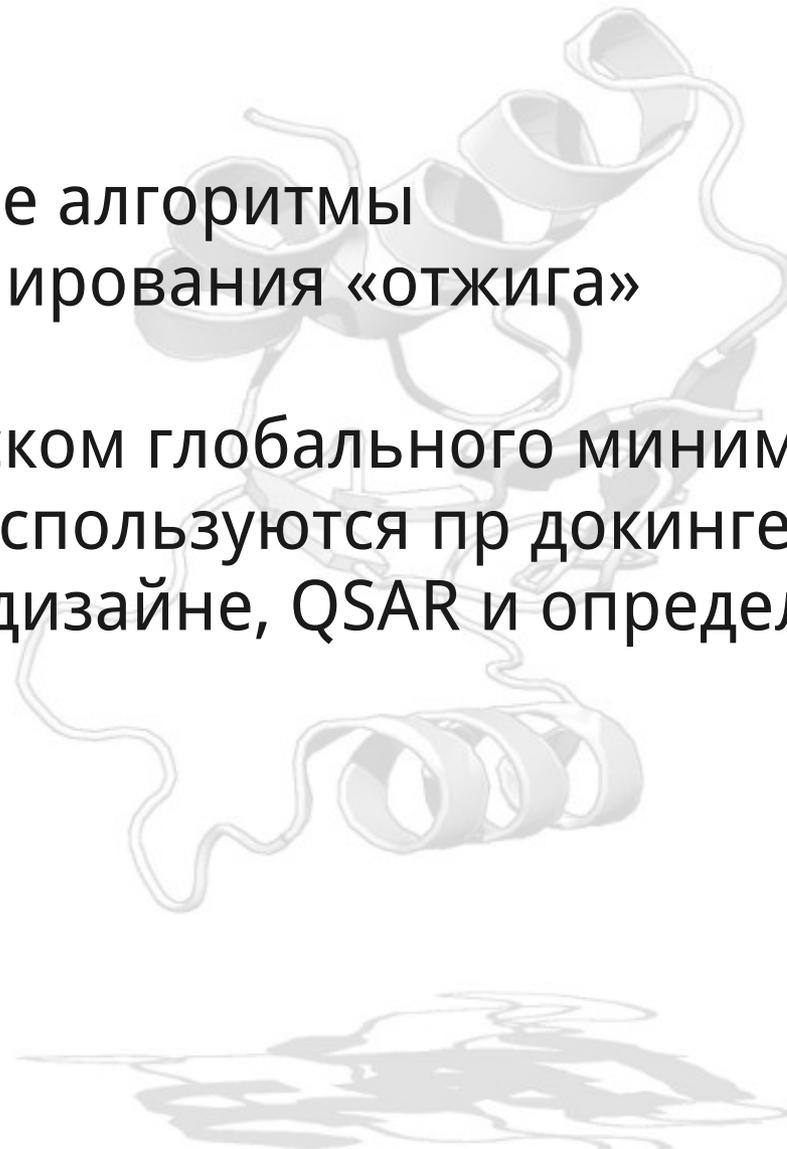
# Выбор метода для поиска конформаций

- Поиск конформеров циклогептадекана  $C_{17}H_{34}$

Метод	Количество конформеров
Систематический поиск	211
Случайный коор. поиск	222
Случайный торс. поиск	249
Дистанционные ограничения	176
МД	169

# Поиск глобального минимума

- Эволюционные алгоритмы
- Методы моделирования «отжига»
- На ряду с поиском глобального минимума эти методы также используются пр докинге, молекулярном дизайне, QSAR и определении фармакофора.



# Эволюционные алгоритмы

## Три основных метода

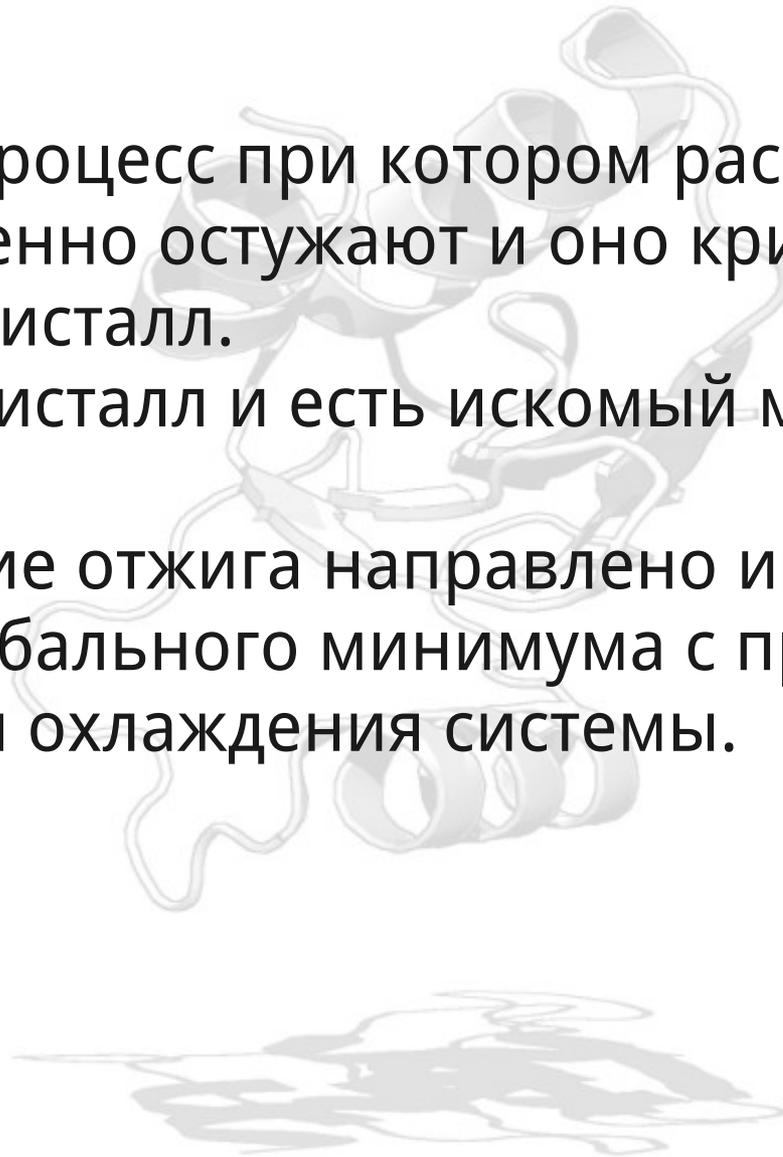
- Генетический алгоритм
- Эволюционное программирование
- Эволюционные стратегии

Общее: создаётся набор вероятных решений проблемы. Решения становятся хорошими или плохими и эволюционируют улучшаясь. Мера “хорошо” это внутренняя энергия молекулы определяемая из молекулярной механики.



# Моделирование отжига

- Отжиг — это процесс при котором расплавленное вещество медленно остужают и оно кристаллизуется в идеальный кристалл.
- Идеальный кристалл и есть искомый минимум энергии
- Моделирование отжига направлено именно на достижение глобального минимума с при моделировании охлаждения системы.



# Моделирование отжига

- Отжиг — это процесс при котором расплавленное вещество медленно остужают и оно кристаллизуется в идеальный кристалл.
- Идеальный кристалл и есть искомый минимум энергии
- Моделирование отжига направлено именно на достижение глобального минимума с при моделировании охлаждения системы.
- Важно тщательно контролировать температуру
- Не факт что метод может привести к глобальному минимуму, но если несколько запусков привели к одной структуре, то можно это предположить.

# Определение структуры белка ограниченной МД

- Суть идеи состоит в том, что в молекулярную динамику вводятся потенциалы, которые увеличивают энергию если конформация не соответствует экспериментальным данным.
- Например: проверяем в ходе оптимизации геометрии R-фактор структуры и если он отличается от экспериментального, то увеличиваем энергию.

$$E_{tot} = U_{pot} + E_{sf} \quad E_{sf} = S \sum \| |F_{obs}| - |F_{calc}| \|^2$$

# Использование МД при оптимизации данных ЯМР

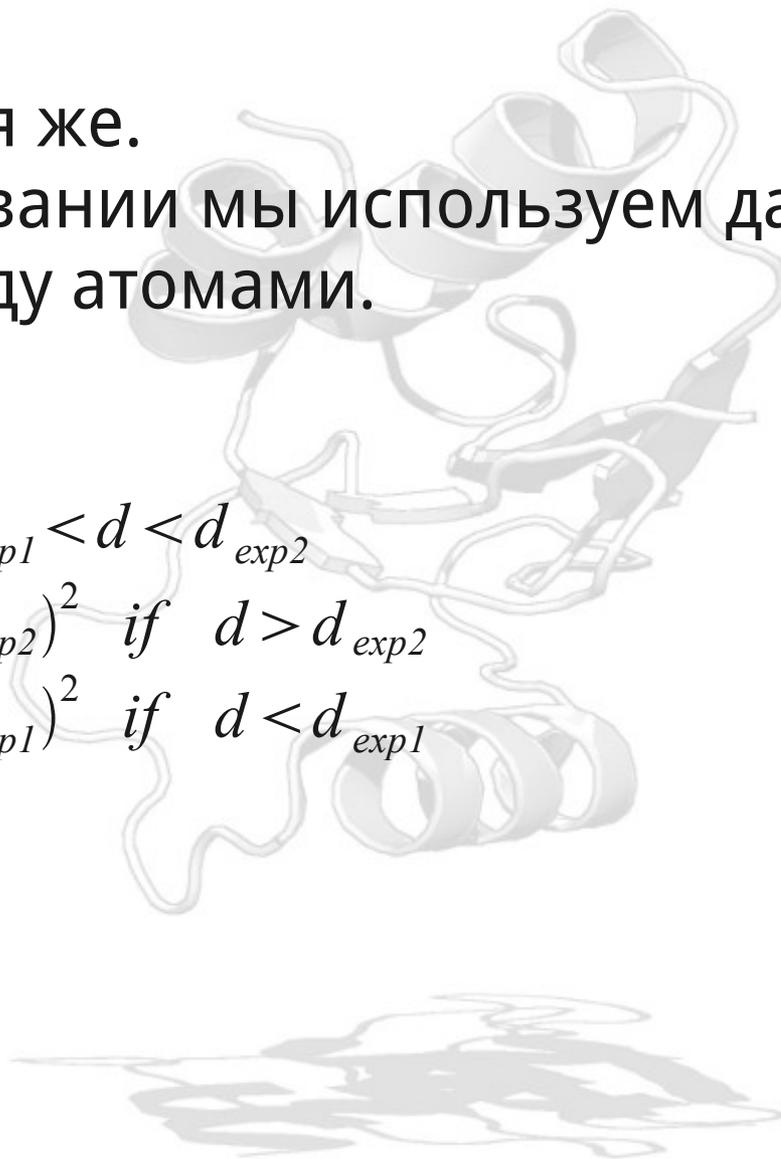
•Суть идеи такая же.

При моделировании мы используем данные о дистанции между атомами.

$$E(d) = 0 \quad \text{if} \quad d_{exp1} < d < d_{exp2}$$

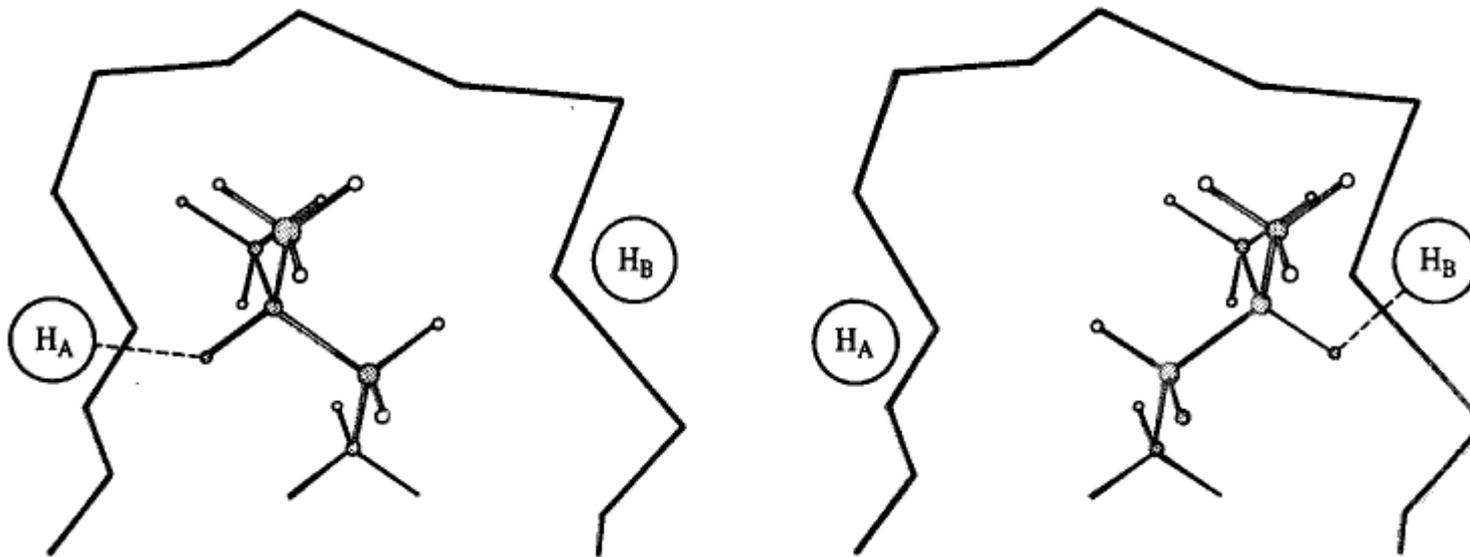
$$E(d) = k(d - d_{exp2})^2 \quad \text{if} \quad d > d_{exp2}$$

$$E(d) = k(d - d_{exp1})^2 \quad \text{if} \quad d < d_{exp1}$$



# Усреднённое по времени ЯМР

• Проблема:



Leach, 1996



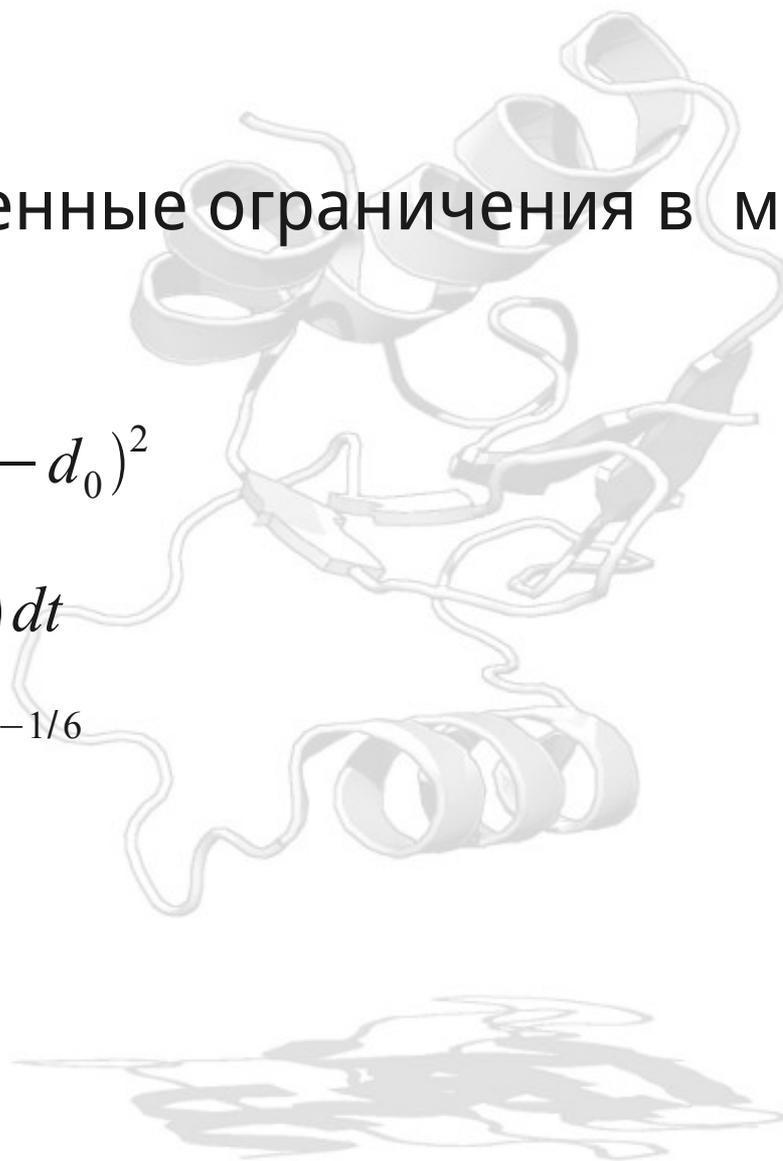
# Усреднённое по времени ЯМР

- Время-усредненные ограничения в молекулярной динамике:

$$E(d) = k(\langle d(t) \rangle - d_0)^2$$

$$\langle d(t) \rangle = \frac{1}{t} \int_0^t d(t) dt$$

$$d_{NOESY} = \langle d(t)^{-6} \rangle^{-1/6}$$

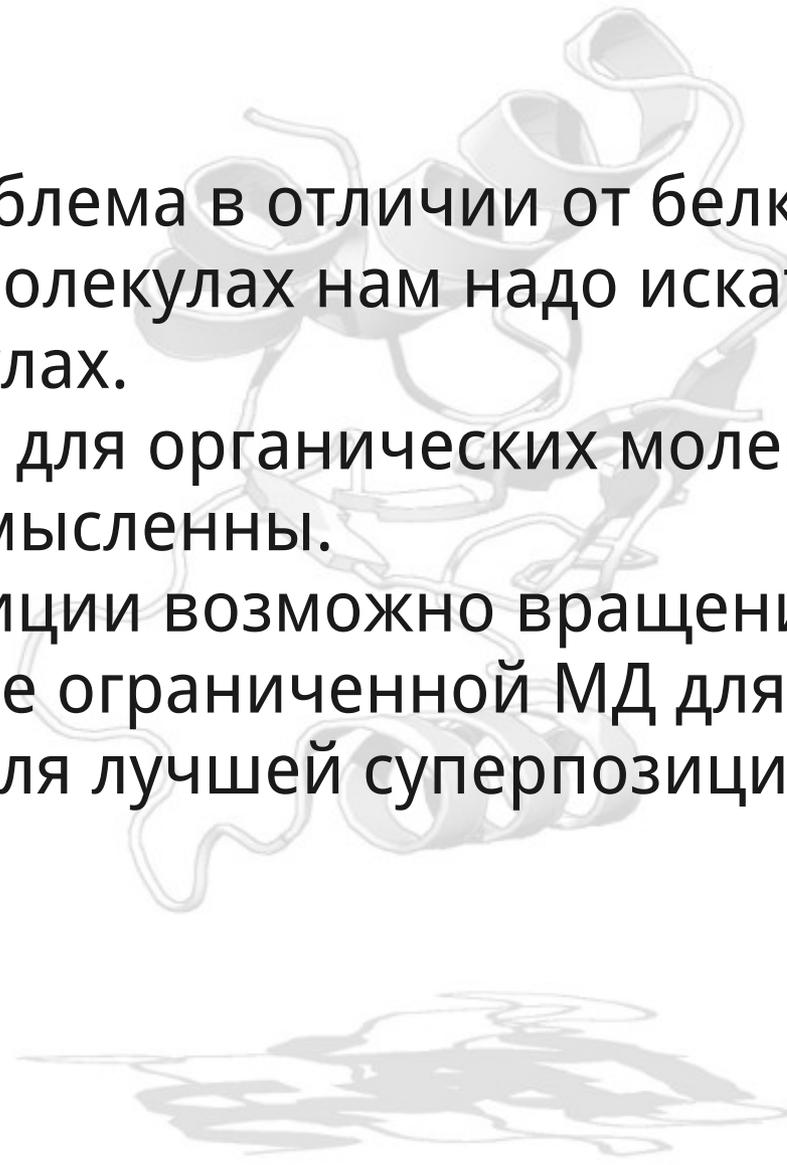


# Базы данных структур

- PDB
- PubChem
- Cambridge structural database
- Inorganic structural database



# Суперпозиция молекул

- Основная проблема в отличии от белков, в органических молекулах нам надо искать подобные атомы в молекулах.
  - Имена атомов для органических молекул в основном бессмысленны.
  - При суперпозиции возможно вращение связей.
  - Использование ограниченной МД для поиска конформации для лучшей суперпозиции.
- 
- A faint, grayscale illustration of a protein structure, likely a ribbon diagram, is visible in the background of the slide. It shows a complex, folded chain of atoms, with various loops and helices. The image is semi-transparent and serves as a decorative element related to the topic of molecular superposition.

# Анализ принципиальных компонент (РСА)

- Молекулу можно описать многими параметрами.
- Значения параметров часто коррелируют.
- Минимизация количества параметров позволяет достоверно описывать систему.

Пример:

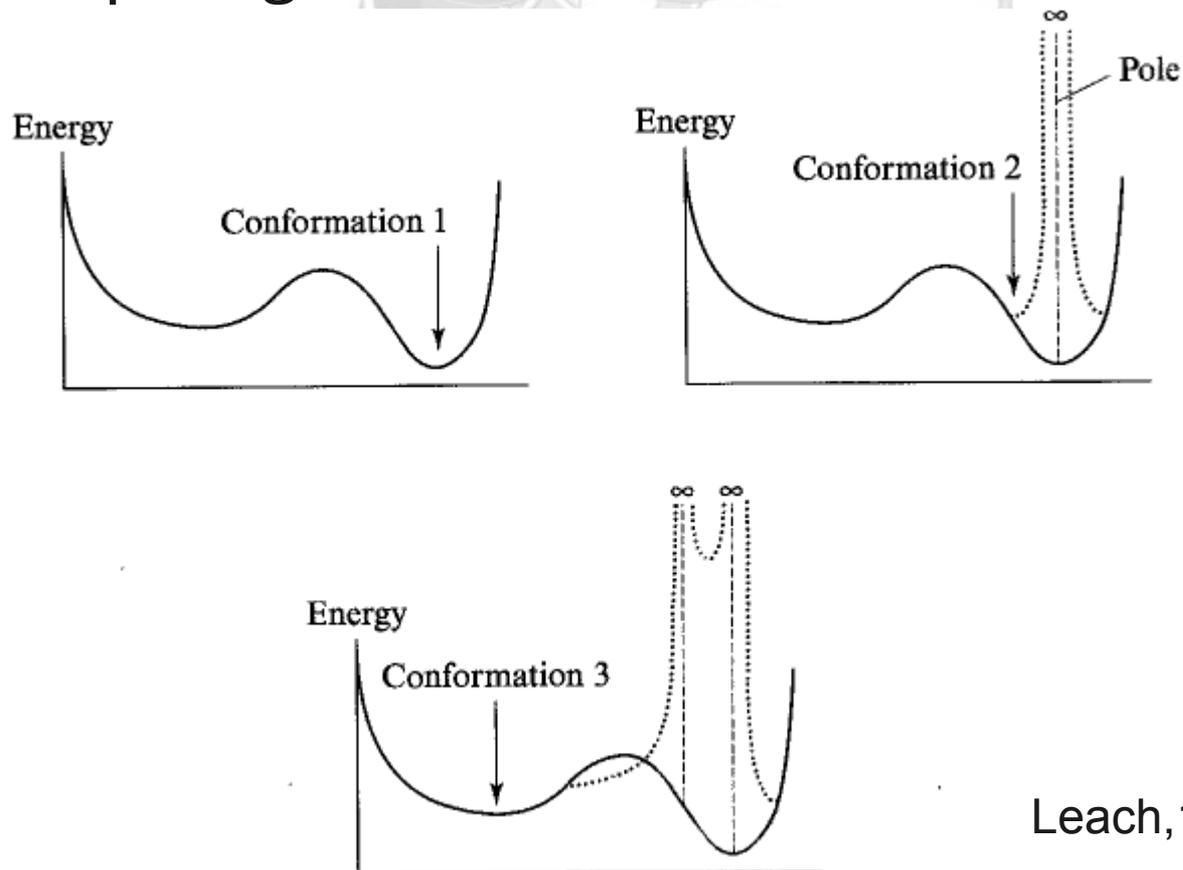
*пусть коррелируют  $x$  и  $y$  введём  $p = x + y$*

$$p_i = \sum_j^v c_{i,j} x_j$$

Расчёт  $p_i$  проводят с помощью традиционного ковариационного матричного анализа

# Poling

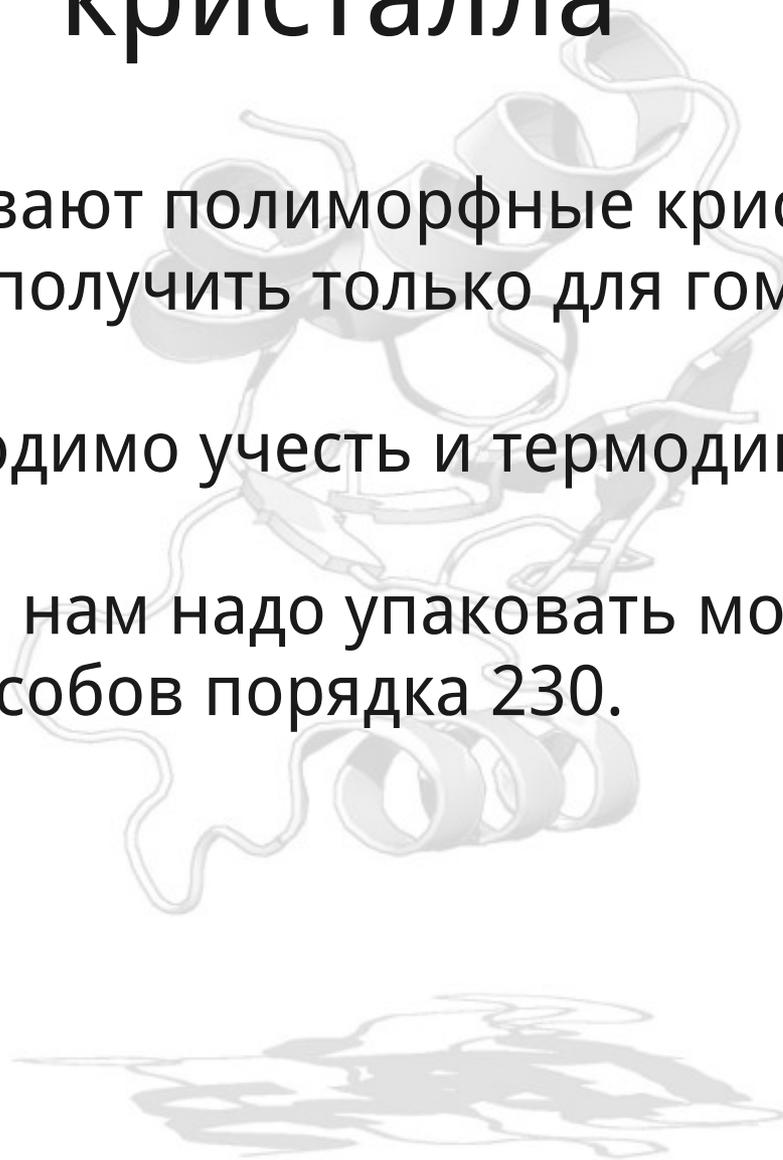
- Очевидно, что одна конформация, представляющая глобальный минимум малоинформативна.
- Для поиска конформаций близких к минимуму используют poling.



Leach, 1996

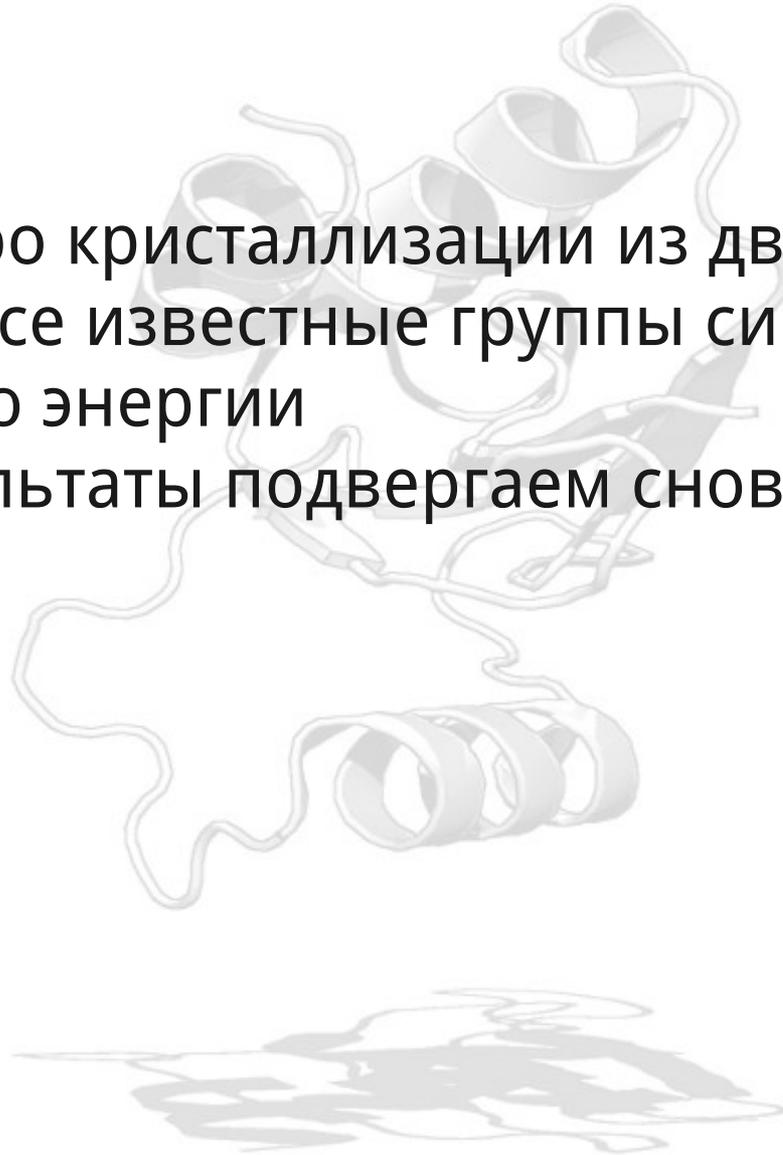
# Классическая задача, моделирование структуры кристалла

- Проблема: бывают полиморфные кристаллы. Но данные можно получить только для гомогенных кристаллов.
- Бывает необходимо учесть и термодинамику и кинетику.
- Неприятность: нам надо упаковать молекулы в кристалл, а способов порядка 230.



# Метод PROMET

1. Создадим ядро кристаллизации из двух молекул .
2. Применяем все известные группы симметрии и кластеризуем по энергии
3. Лучшие результаты подвергаем снова шагу 2 и кластеризуем.
4. И т.д.



# Альтернативный подход

1. Разрешим все возможные варианты симметрии.
2. Размещаем молекулы и разрешаем движение.  
Либо случайным образом либо систематическим.
3. Лучшие результаты кластеризуем.
4. Проводим очень точную оптимизацию геометрии.

