



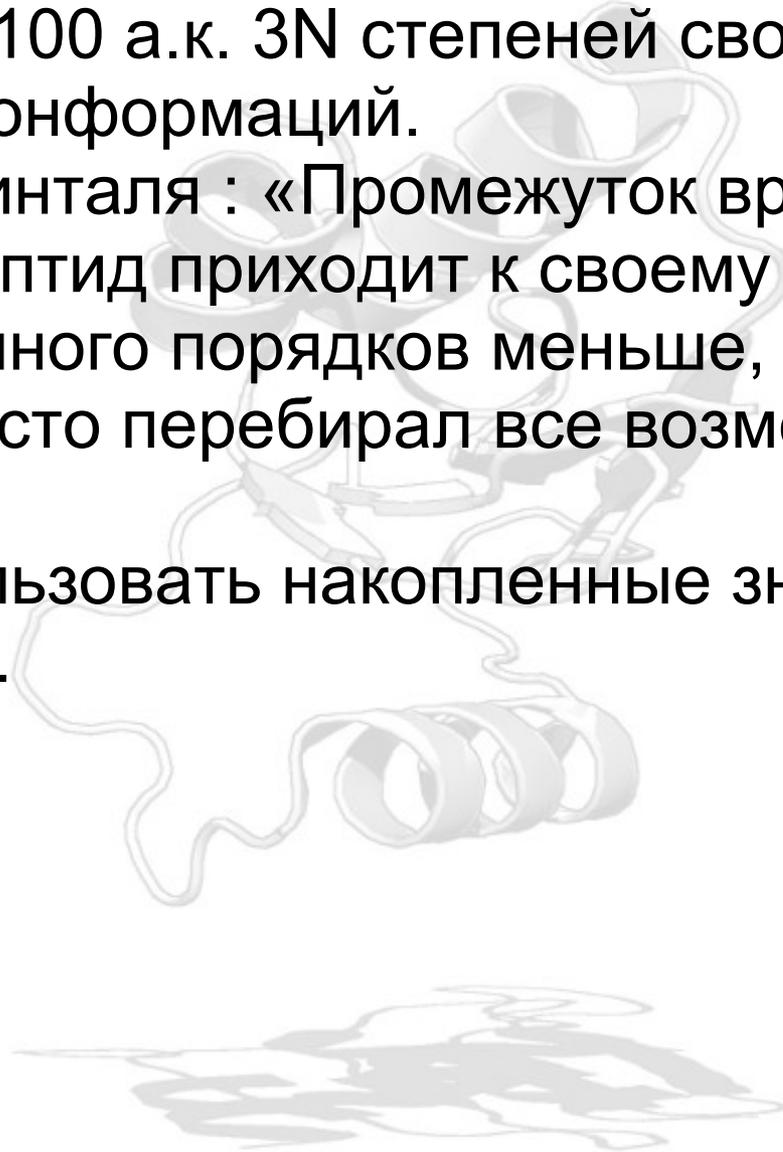
**Молекулярное моделирование в
применении к биомолекулам**

Лекция 10

Моделирование структуры белков

Основные проблемы

- Монте-Карло: 100 а.к. $3N$ степеней свободы, получаем 10^{48} конформаций.
- Парадокс Левинталя : «Промежуток времени, за который полипептид приходит к своему скрученному состоянию, на много порядков меньше, чем если бы полипептид просто перебирал все возможные конфигурации».
- Разумно использовать накопленные знания для моделирования.



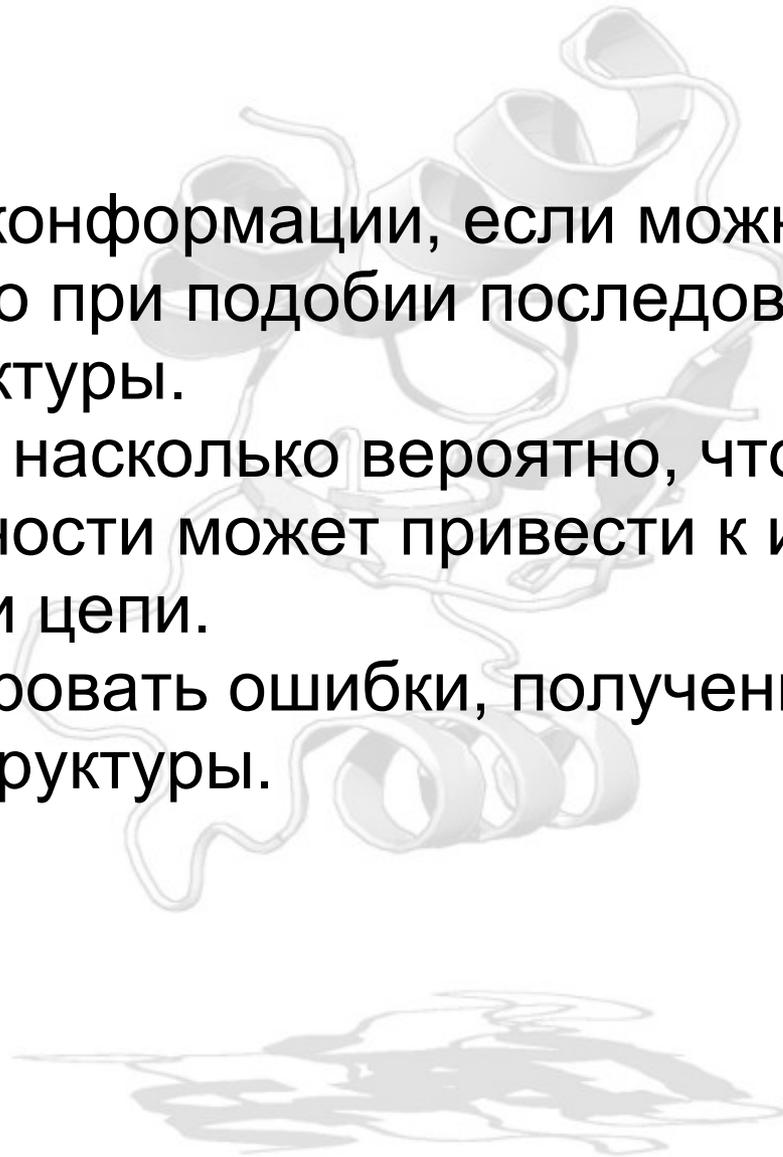
Последовательность-структура

Причины парадокса Левинталя:

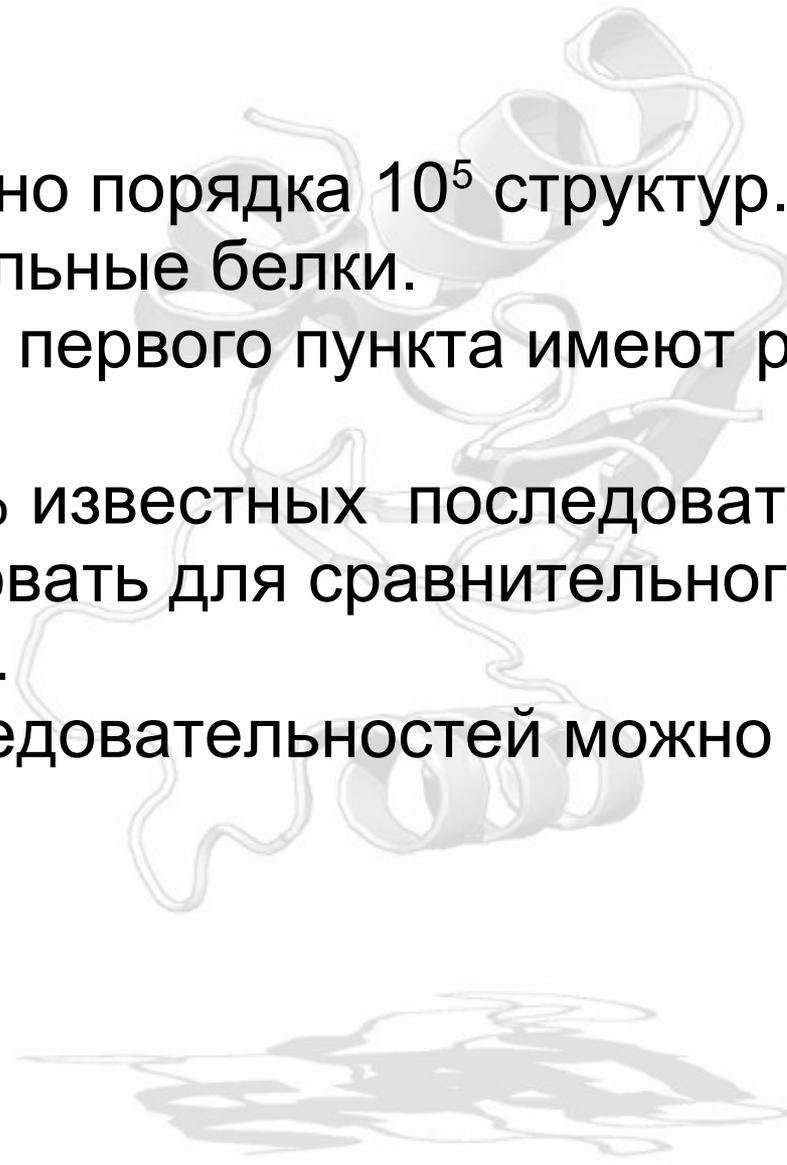
- Теоретические модели не соответствуют тому, что природа старается оптимизировать;
- в ходе эволюции были отобраны только те белки, которые легко сворачиваются;
- белки могут сворачиваться разными путями, не обязательно следуя глобально оптимальному пути.
- Считается, что структура определяется последовательностью, но иногда нужны другие факторы.
- Структура более консервативна, чем последовательность

Сравнительное моделирование

- Зачем искать конформации, если можно представить, что при подобии последовательностей подобны и структуры.
- Надо оценить, насколько вероятно, что отличие в последовательности может привести к изменению способа укладки цепи.
- Надо отфильтровать ошибки, полученные при определении структуры.

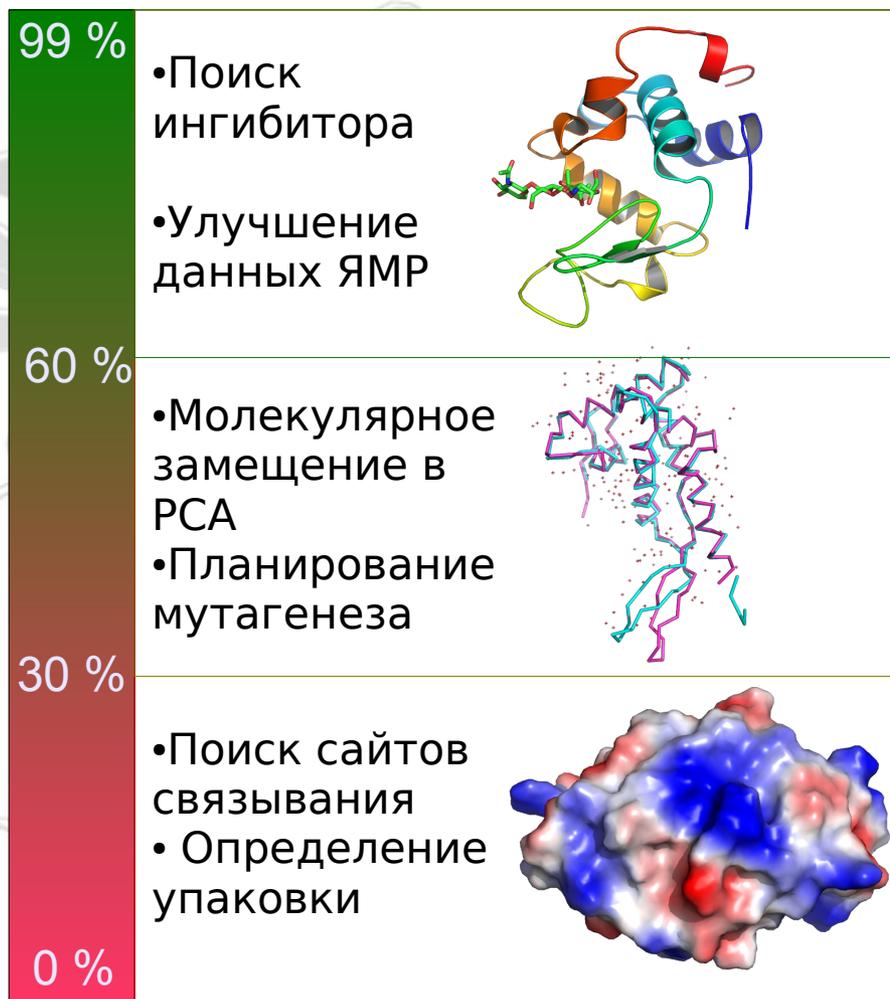


Известные структуры и последовательности

- Сейчас известно порядка 10^5 структур. Примерно 10% - это уникальные белки.
 - Только 30% из первого пункта имеют разрешение лучше 3.0 Å.
 - Примерно 25% известных последовательностей можно использовать для сравнительного моделирования.
 - Для 50% последовательностей можно предсказать способ укладки.
- 

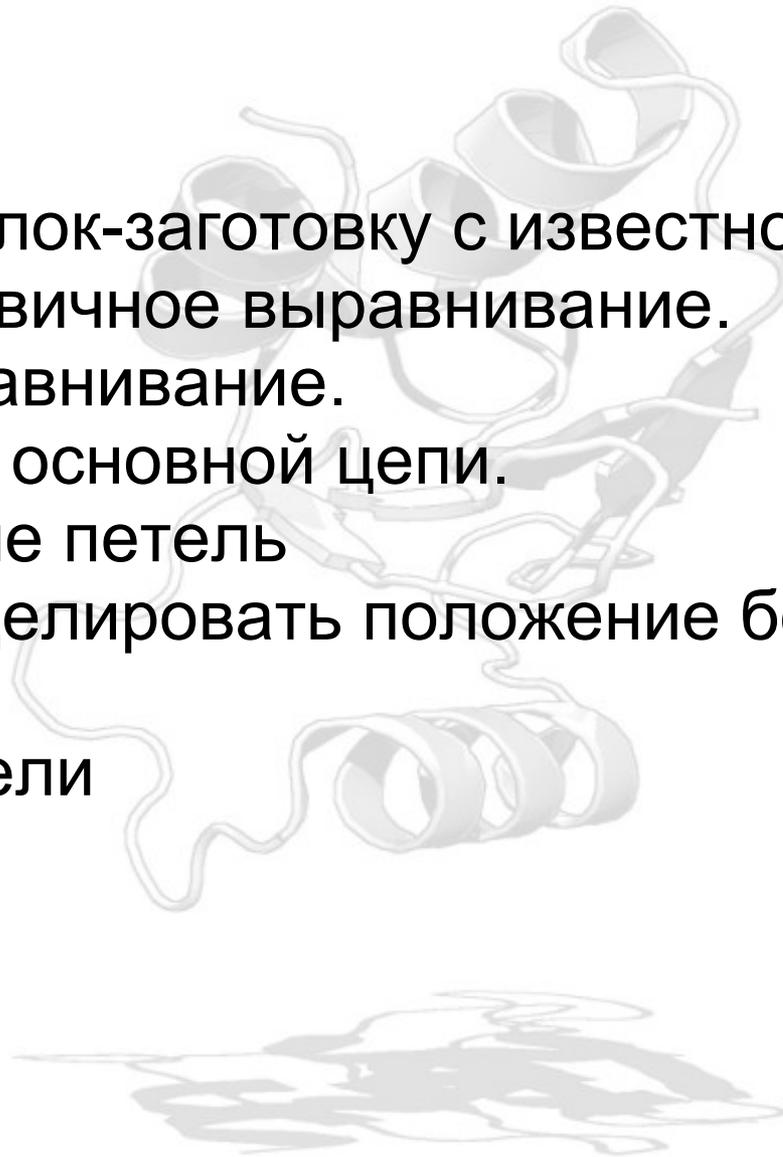
Степень идентичности и сравнительное моделирование

Sali, A. & Kuriyan, J. *Trends Biochem. Sci.* **22**, M20–M24 (1999)



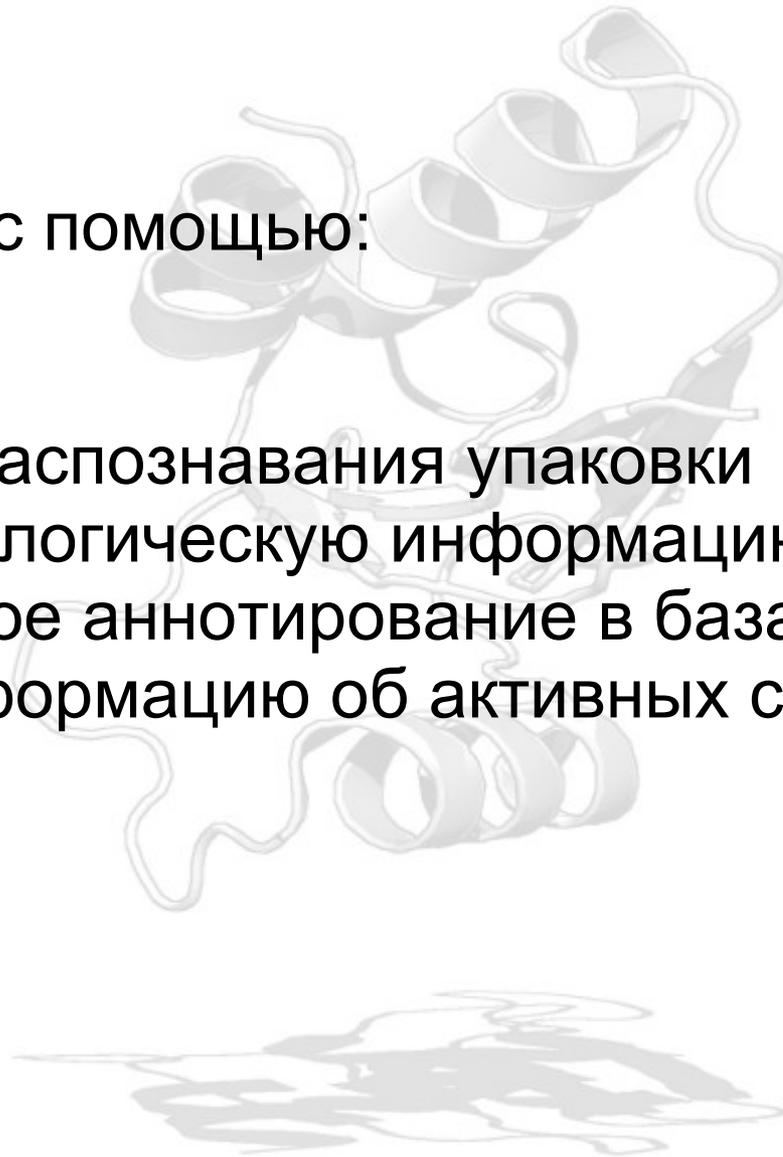
Как это реализовать?

- Надо найти белок-заготовку с известной структурой.
- Построить первичное выравнивание.
- Улучшить выравнивание.
- Построить ход основной цепи.
- Моделирование петель
- Достроить/моделировать положение боковых радикалов
- Проверка модели



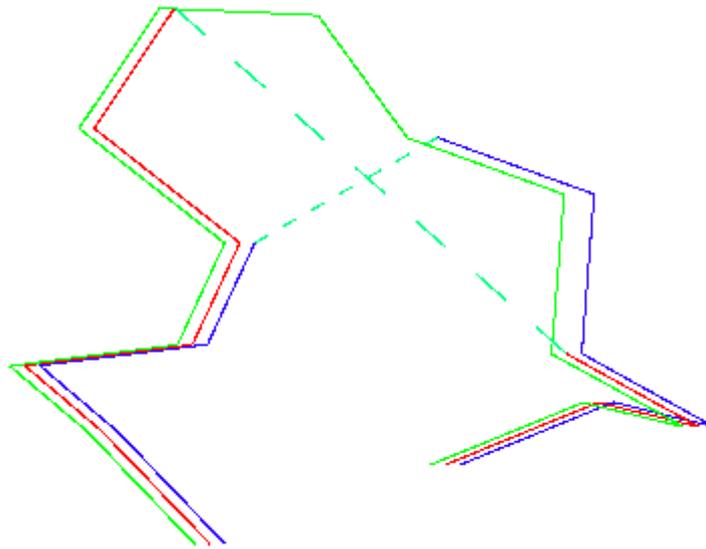
Поиск белка-заготовки

- Поиск по PDB с помощью:
 - Blast
 - Psi-Blast
 - Методов распознавания упаковки
- Используя биологическую информацию.
- Функциональное аннотирование в базах данных.
- Используя информацию об активных сайтах или мотивах.



Улучшение выравнивания

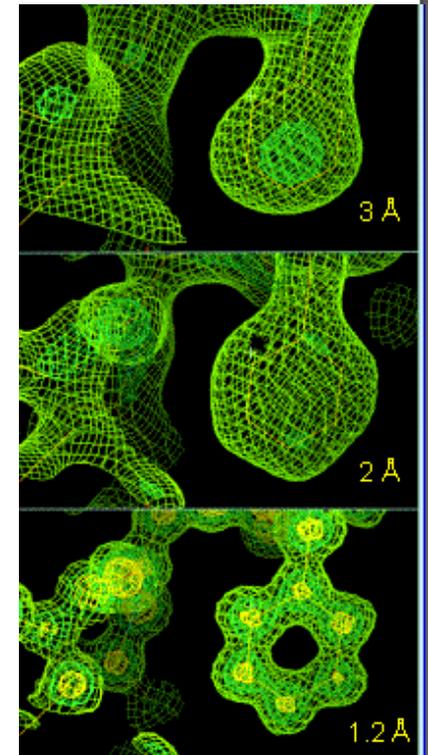
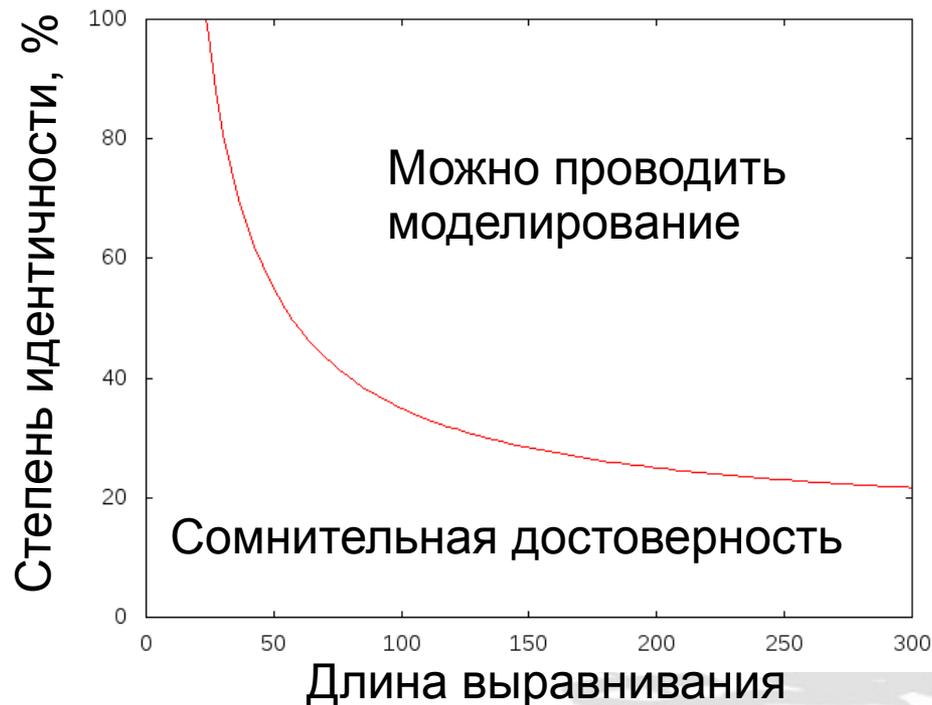
1	2	3	4	5	6	7	8	9	10	11	12	13	14
PHE	ASP	ILE	CYS	ARG	LEU	PRO	GLY	SER	ALA	GLU	ALA	VAL	CYS
PHE	ASN	VAL	CYS	ARG	THR	PRO	---	---	---	GLU	ALA	ILE	CYS
PHE	ASN	VAL	CYS	ARG	---	---	---	THR	PRO	GLU	ALA	ILE	CYS



Из книги "Professional Gambling" от Gert Vriend
<http://www.cmbi.kun.nl/gv/articles/text/gambling.html>

Качество белка-заготовки

- Выбор качественного белка-заготовки очень важен.
- Лучший вариант необязательно обладает лучшей степенью идентичности.
 - Белок 1: ID 93%, 3.5 Å разрешение. Хуже.
 - Белок 1: ID 90%, 1.5 Å разрешение. Лучше!

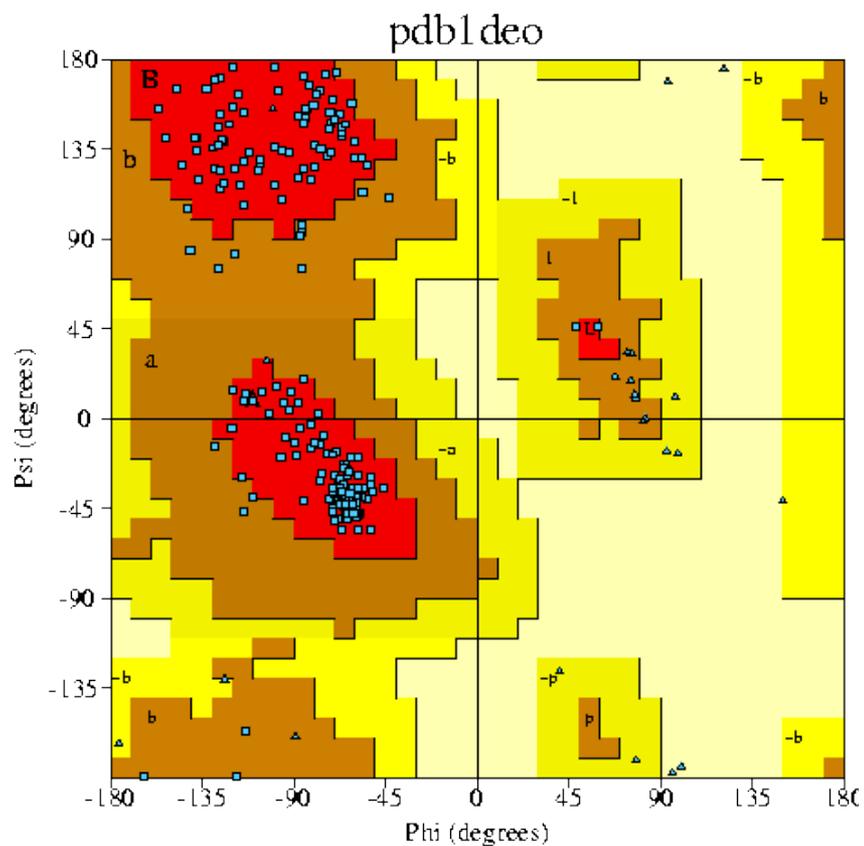


Если структура белка-заготовки получена методом ЯМР

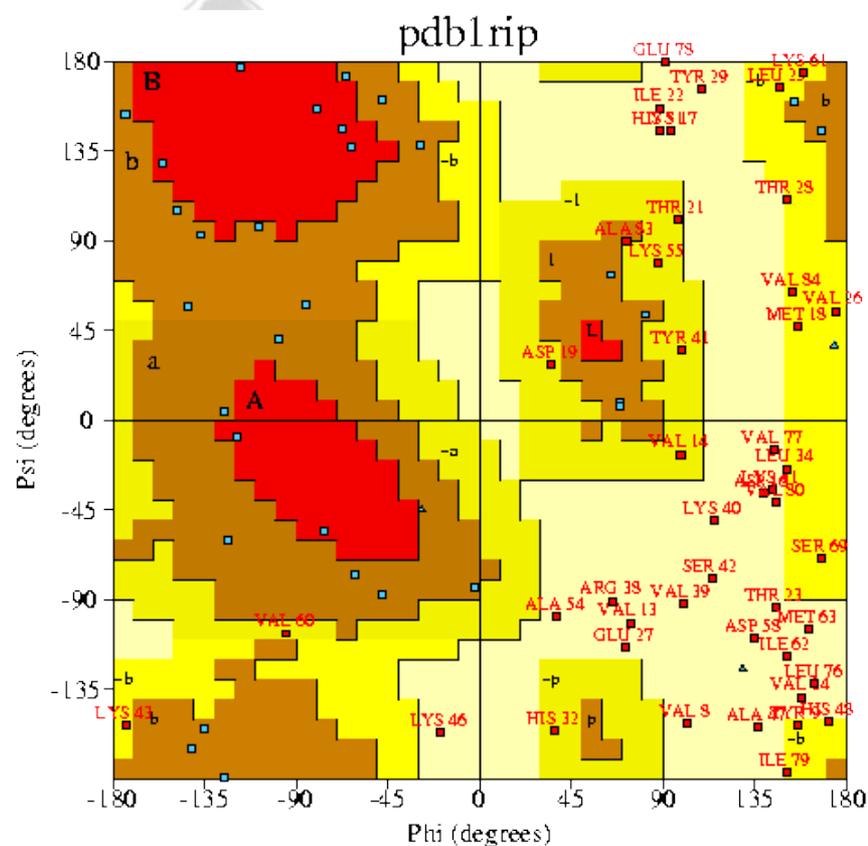
- Определимся, какие области определены лучше.
- Соотнесём с выравниванием.
- Если низкое значение идентичности выпадает на “подвижные” области, то структура подходит.



Качество заготовки, Рамачандран.



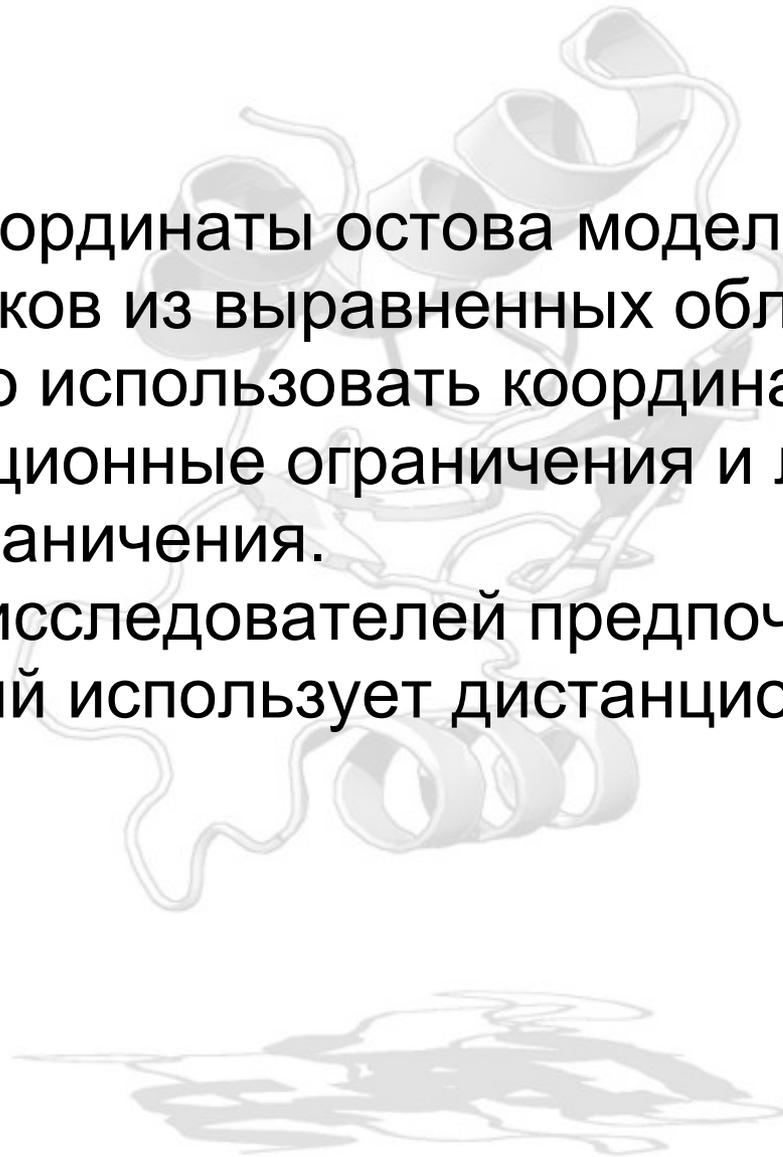
РСА, хорошая структура.



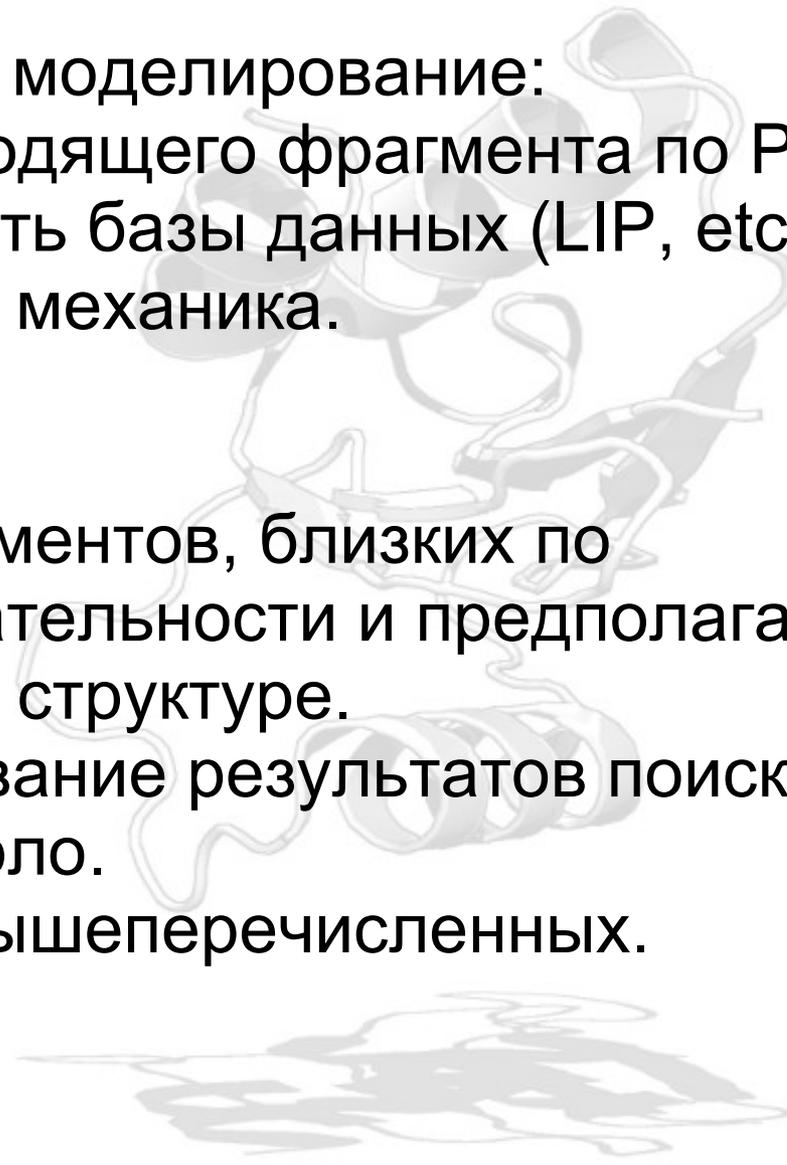
ЯМР, сомнительные данные.

Построение остова

- Генерируем координаты остова моделируемого белка для остатков из выравненных областей.
- Необязательно использовать координаты, могут подойти дистанционные ограничения и любые другие подходящие ограничения.
- Большинство исследователей предпочитают Modeller, который использует дистанционные ограничения.



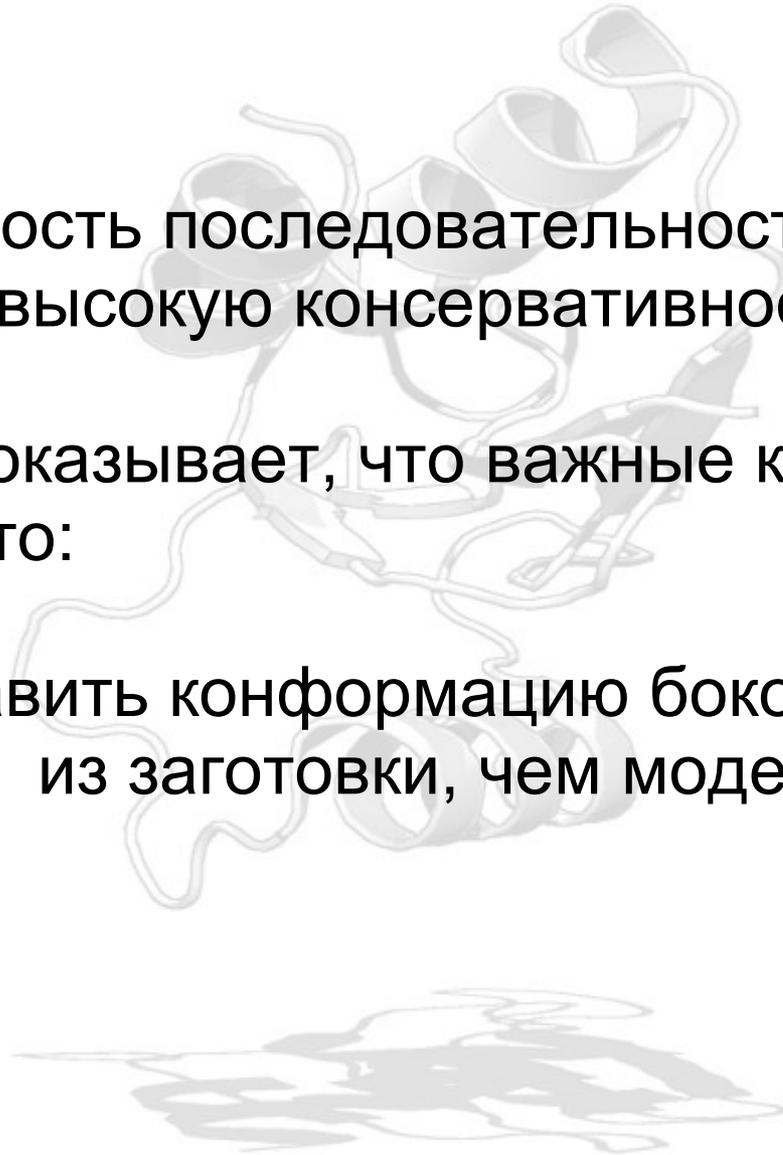
Моделирование петель

- Эмпирическое моделирование:
 - Поиск подходящего фрагмента по PDB
 - Использовать базы данных (LIP, etc..)
 - Молекулярная механика.
 - Монте-Карло.
 - Rosseta:
 - Поиск фрагментов, близких по последовательности и предполагаемой вторичной структуре.
 - Комбинирование результатов поиска с помощью Монте-Карло.
 - Комбинации вышеперечисленных.
- 

Моделирование боковых радикалов

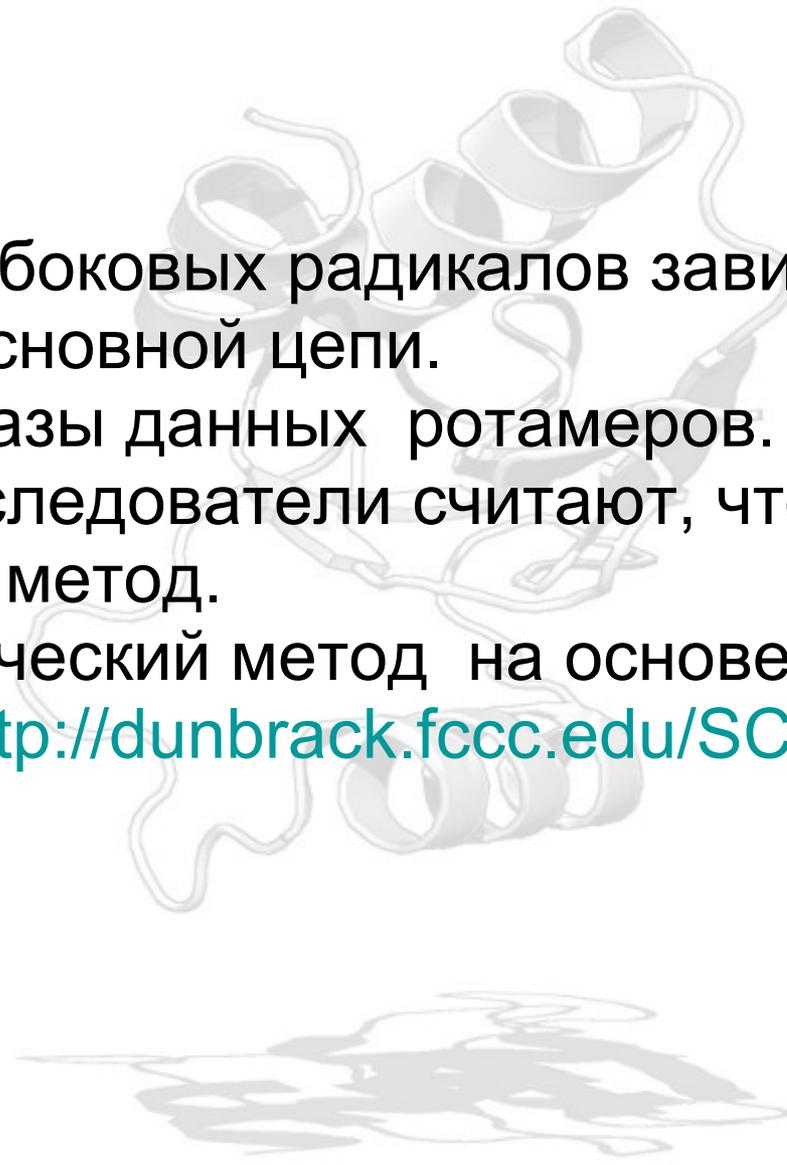
- Если идентичность последовательностей высока, то можно ожидать высокую консервативность третичных контактов.
- Если анализ показывает, что важные контакты консервативны то:

Лучше оставить конформацию боковых радикалов из заготовки, чем моделировать.



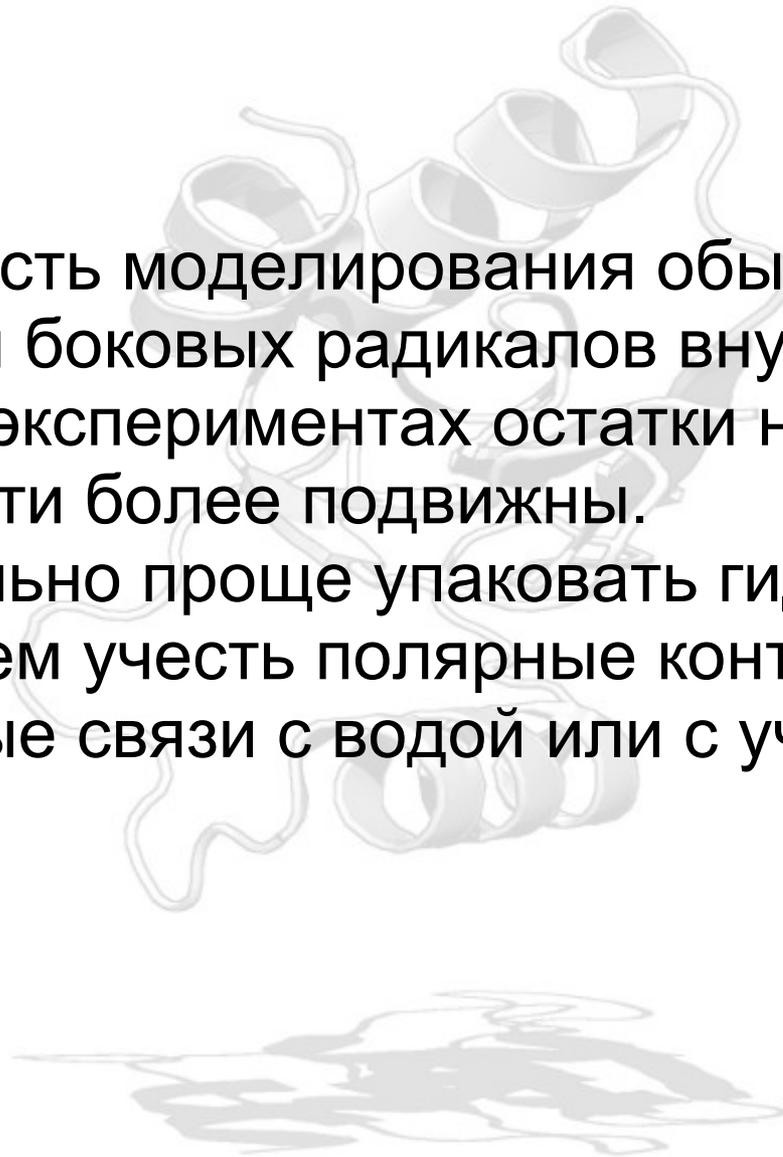
Моделирование боковых радикалов

- Конформация боковых радикалов зависит от конформации основной цепи.
- Существуют базы данных ротамеров.
- Некоторые исследователи считают, что SCWRL - самый удачный метод.
 - Это эмпирический метод на основе теории графов. <http://dunbrack.fccc.edu/SCWRL3.php>



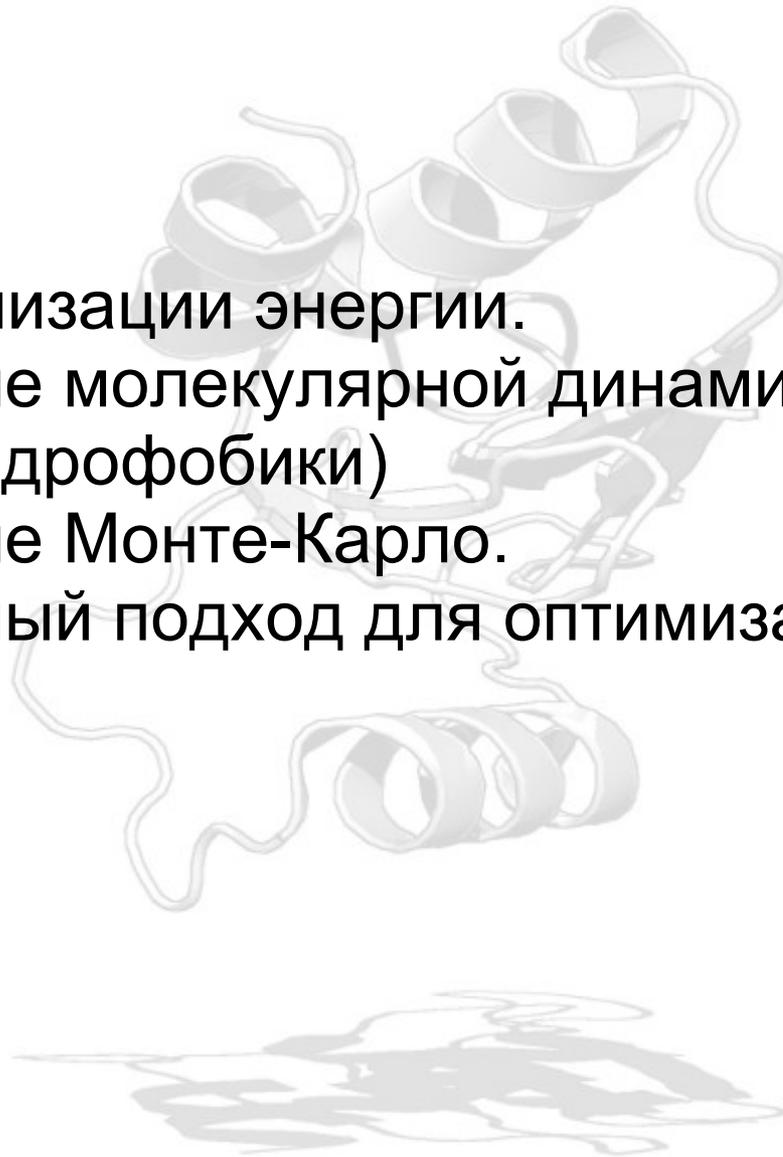
Точность моделирования боковых радикалов

- Высокая точность моделирования обычно достигается для боковых радикалов внутри глобулы.
 - Причина: в экспериментах остатки на поверхности более подвижны.
 - Вычислительно проще упаковать гидрофобные остатки, чем учесть полярные контакты и водородные связи с водой или с участием воды.



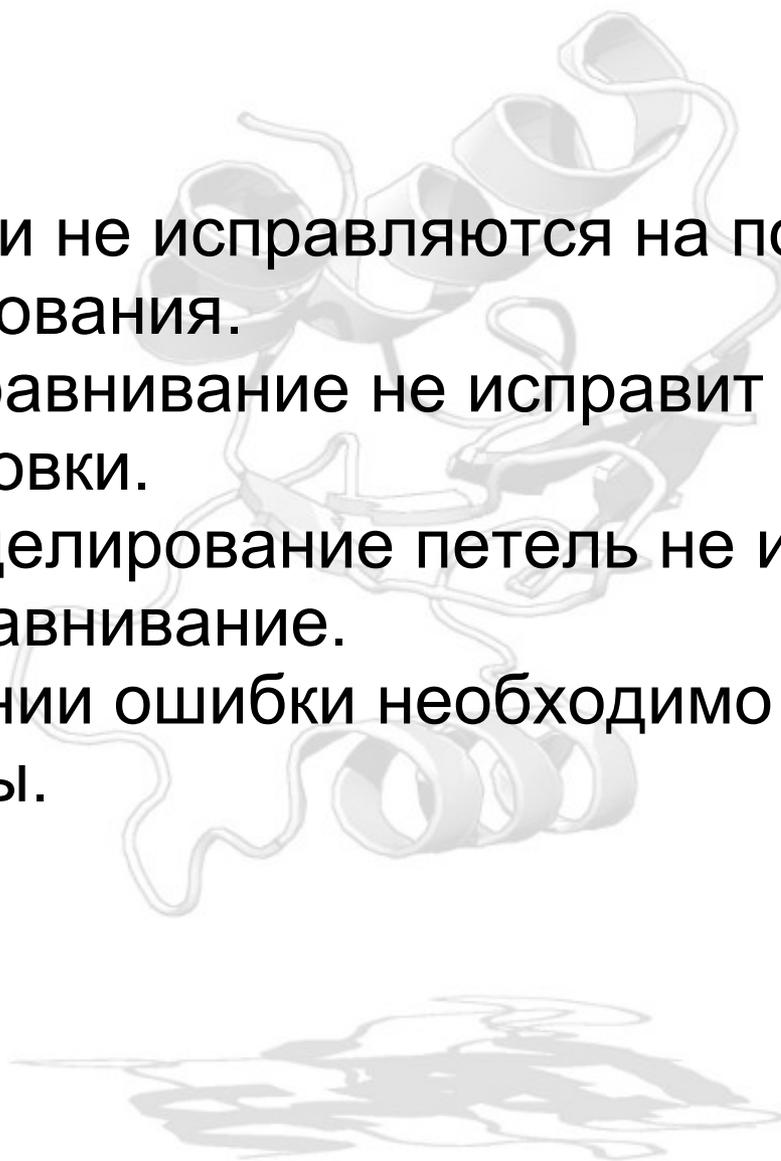
Улучшение модели

- Методы минимизации энергии.
- Моделирование молекулярной динамики (оптимизация гидрофобики)
- Моделирование Монте-Карло.
- Любой известный подход для оптимизации структуры.



Ошибки

- Обычно ошибки не исправляются на последующих этапах моделирования.
 - Хорошее выравнивание не исправит плохой выбор белка-заготовки.
 - Хорошее моделирование петель не исправит плохое выравнивание.
- При обнаружении ошибки необходимо повторять некоторые этапы.



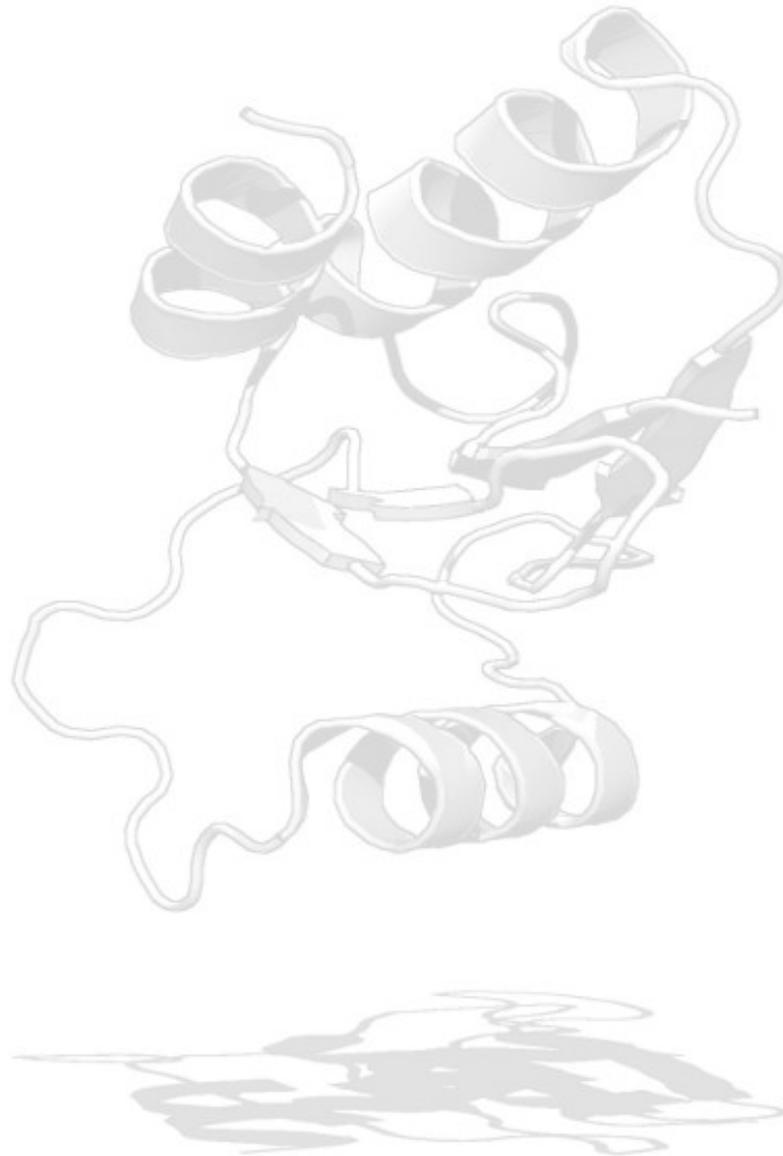
Проверка

- Большинство программ для моделирования по гомологии выдают правильные значения для связей и валентных углов.
- Карта Рамачандрана в большинстве случаев для модели выглядит так же, как для белка-заготовки
- Проверка на ориентацию или положение заряженных остатков может быть полезна.
- Использование любых экспериментальных данных:
 - Остатки активного центра.
 - Места модификаций.
 - Места контактов.

ProQ сервер оптимизирован на поиск правильной модели, а не нативной структуры.

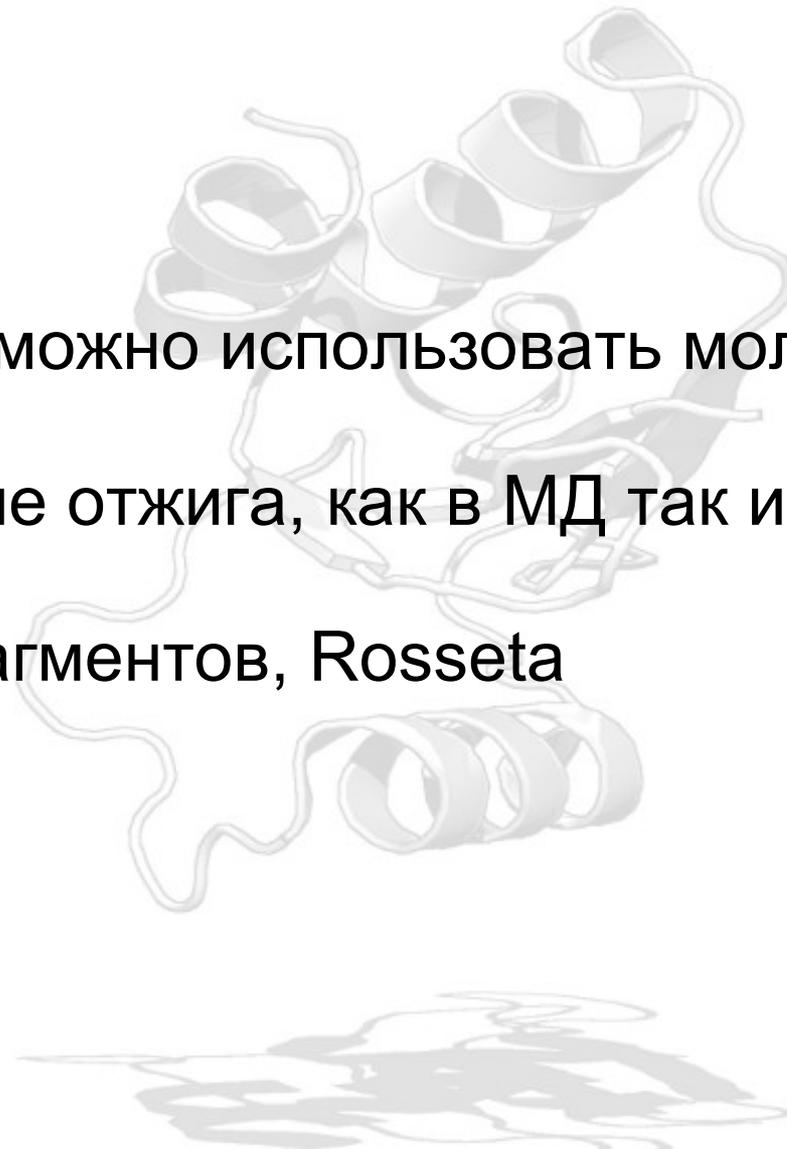
Ресурсы для гомологичного моделирования

- Modeller
 - SwissModel
 - Eva-CM
 - Nest
- и т. д.



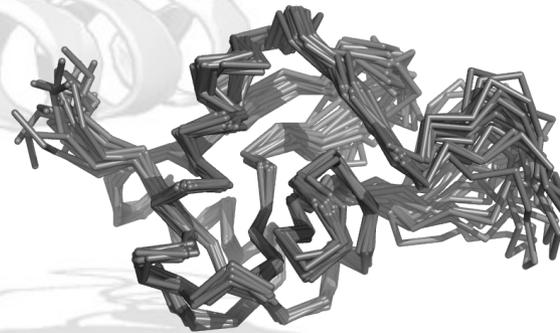
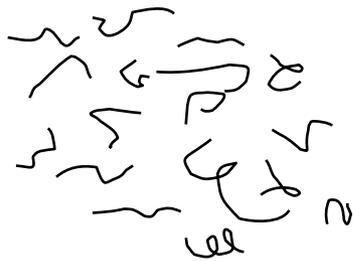
Предсказание структуры белка *Abinitio*

- Теоретически можно использовать молекулярную динамику.
- Моделирование отжига, как в МД так и в Монте-Карло.
- На основе фрагментов, Rosseta



Abinitio, Rosseta

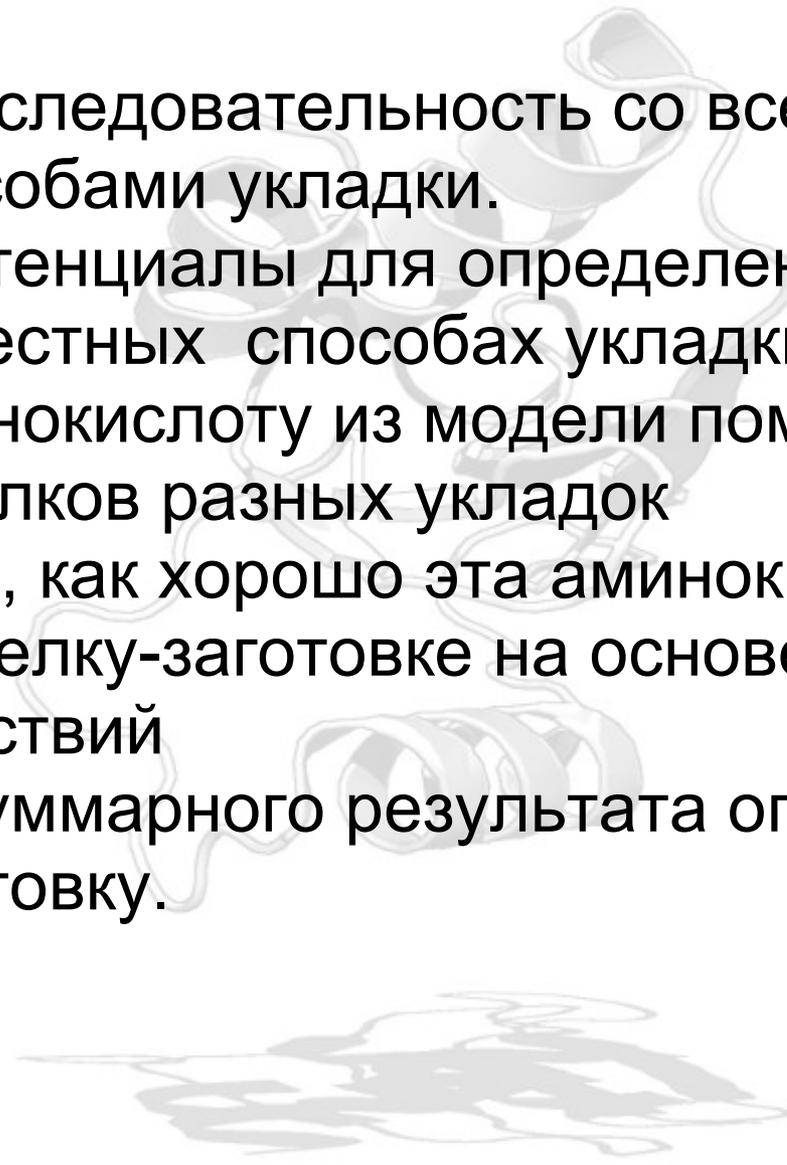
- Метод использует информацию о предсказании вторичной структуры
- Сравниваем фрагменты от 3 до 9 остатков с библиотекой известных структур. Строим эти фрагменты.
- Соединяем эти фрагменты и используем Монте-Карло для оптимизации третичной структуры.



Abinitio, Rosseta

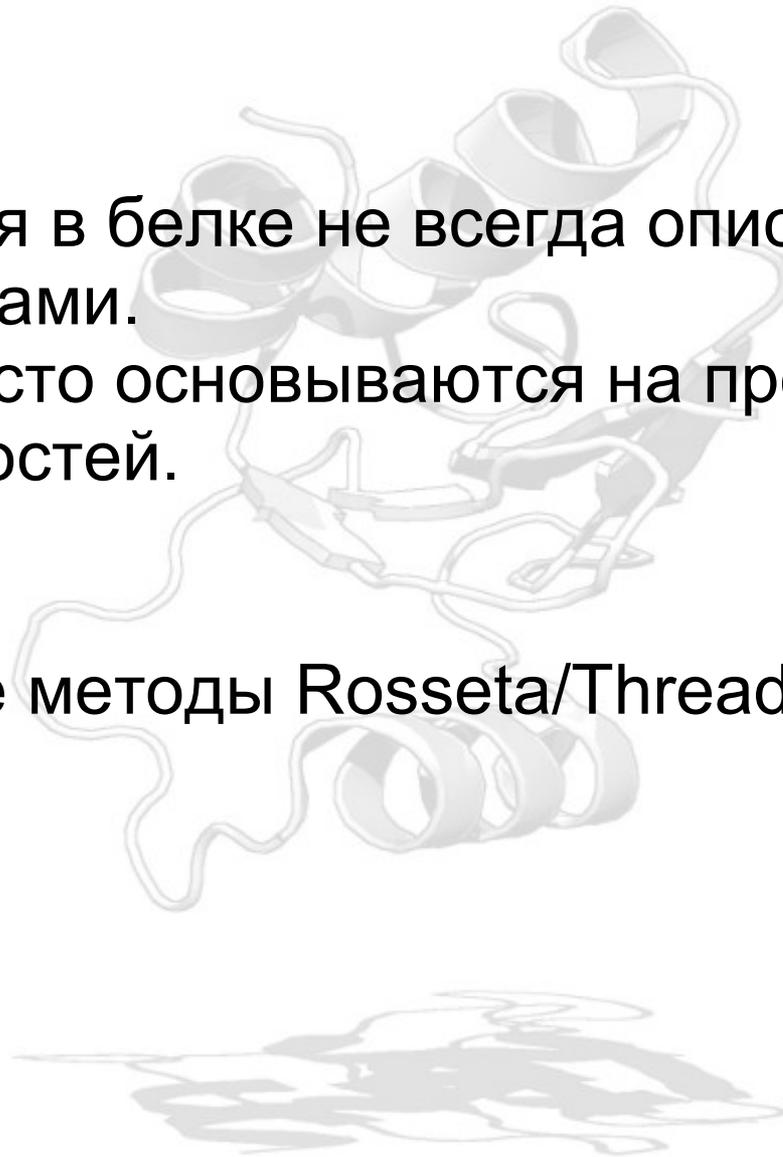
- Для определения хорошей конформации используют специальные потенциалы, которые делают модель похожей на нативную
- Что можно использовать:
 - Потенциалы для третичных контактов
 - Гидрофобные потенциалы
 - Потенциал для уменьшения радиуса вращения молекулы
 - Водородные связи и т.д.
- Можно добавить знание о дисульфидных мостиках, местах связывания катионов металлов и т. д.

Threading — протягивание нити

- Сравниваем последовательность со всеми известными способами укладки.
 - Используем потенциалы для определения тенденций в известных способах укладки.
 - Каждую аминокислоту из модели помещаем в позиции белков разных упадок
 - Определяем, как хорошо эта аминокислота подходит белку-заготовке на основе парных взаимодействий
 - На основе суммарного результата определяем белок-заготовку.
- 

Threading — недостатки

- Взаимодействия в белке не всегда описываются парными контактами.
- Потенциалы часто основываются на профилях последовательностей.
- Есть гибридные методы Rosseta/Threading: I-Tasser



Мета-серверы

- Сравнение разных методов.
- Большинство методов предсказывают правильную укладку в первых 10-20 результатах.
- Удаление структур с высоким значением параметров модели, но с единственной укладкой.
- Суперпозиция результатов, взвешивание.
- Часто выдают только позиции атомов остова.



Заключение

- Суть современного моделирования белков - эмпирическая
 - Чем больше известной информации используется при моделировании, тем точнее модель.
 - Каждый метод имеет недостатки.
 - Критический анализ модели позволяет выявить ошибки и улучшить модель.
- 