

# Моделирование макромолекул

Лекция

01.12.2021

## МОДЕЛИРОВАНИЕ БИОНАНОСТРУКТУР

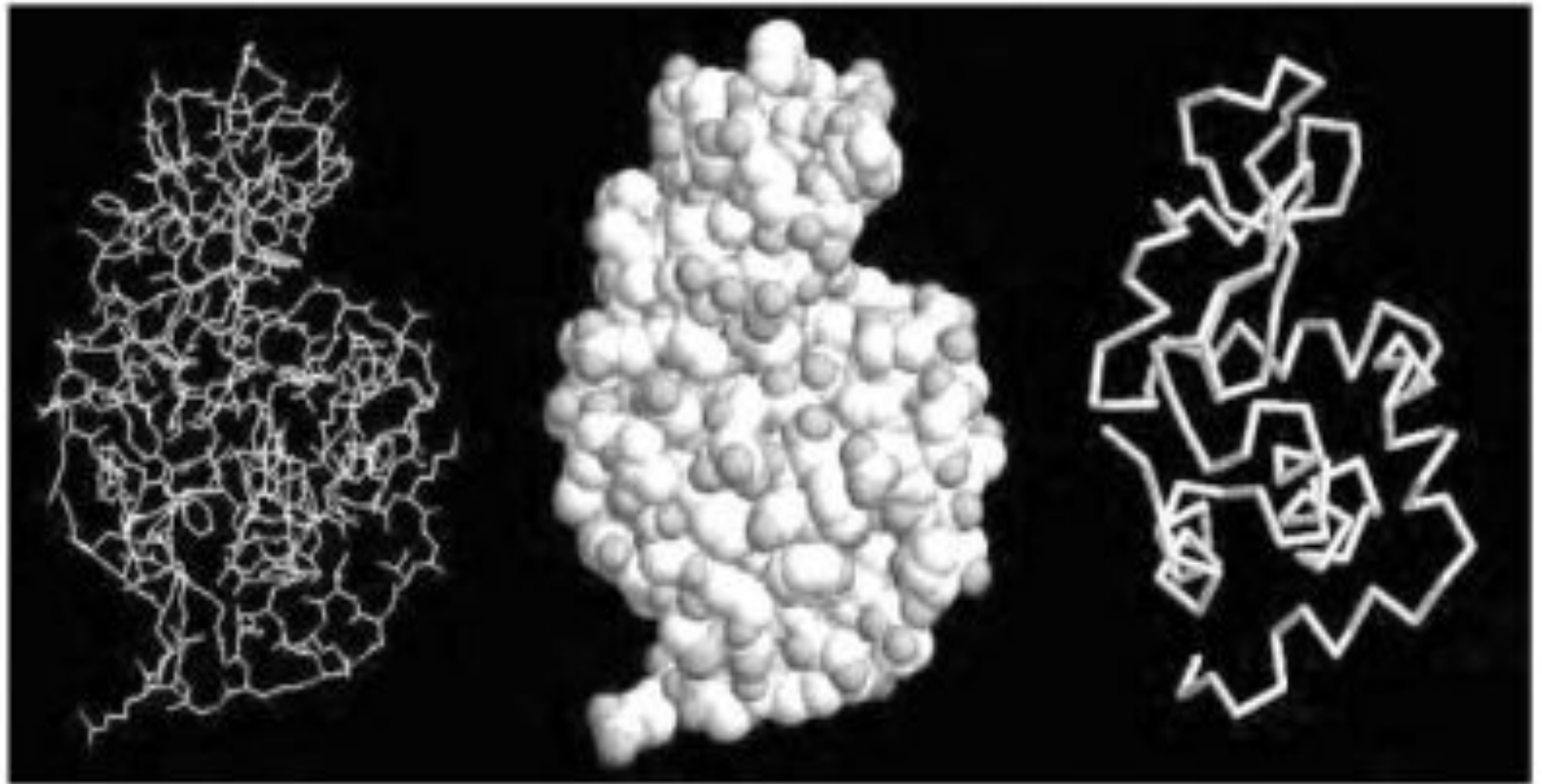
При конструировании новых бионаномашин компьютерное моделирование используется наравне с биологическим экспериментом, причем, зачастую, *итерационно*. Сначала молекулы *проектируются* на компьютере, затем лучшие из вариантов *синтезируются и тестируются*. Выводы, сделанные в результате моделирования и экспериментов учитываются в *следующем цикле* компьютерных разработок и т. д.

Такой подход, в частности, использовался при так называемом целенаправленном конструировании (или рациональном дизайне) молекул лекарств (*rational drug design*) наиболее впечатляющим из которых являются разрабатываемые лекарственные препараты против СПИДа. Компьютерное моделирование часто является единственно возможным методом изучения систем, недоступных экспериментатору, обеспечивая предсказания и направляя дальнейшие исследования и разработки.

## Моделирование макромолекул

Компьютерная графика *революционизировала* исследование бионаномолекул и сейчас она является незаменимым инструментом во всех молекулярных науках. Компьютерная графика позволяет визуализировать необычные формы, свойства и взаимодействия молекул в привычном для нас виде. Молекулярная графика является первым шагом в начале нового проекта, позволяющим исследователю представить себе и понять те молекулы, которые будут синтезированы или модифицированы.

Поскольку молекулы являются объектами, размеры которых меньше длины волны света видимого диапазона, то невозможно увидеть отдельную молекулу в микроскоп. Поэтому были разработаны различные способы представления структуры молекул (рисунок 102).



*a*

*б*

*в*

Рисунок 102 – Визуализация молекулярной структуры лизоцима: *a* – схема ковалентных связей; *б* – ван-дер-ваальсовая модель молекулы; *в* – "резиновая лента"

Можно изобразить пространственное расположение всех ковалентных связей в молекуле (рисунок 102(а)). Такое представление *наиболее полное*, но его сложно анализировать. Иногда удобно заполнить области, занимаемые каждым атомом, сферами соответствующего радиуса – *ван-дер-ваальсовая модель* (рисунок 102(б)). Такой способ визуализации даёт хорошее представление о *форме и размере* молекулы, но не видна внутренняя структура глобулы. Такое представление удобно для исследования межмолекулярного взаимодействия, но усложняет исследование процессов, связанных с особенностями вторичной структуры белка.

Для визуализации топологии белковой цепи используют схематическое изображение в виде "резиновой ленты", которая демонстрирует только последовательность пептидных связей в белке (рисунок 102(в)). Такое представление удобно для *исследования фолдинга биомолекул*.

Сегодня аппаратных средств компьютеров уже достаточно для интерактивного представления даже самых больших биомолекул. Программы визуализации доступны с сайта Базы данных белков (*Protein Data Base, PDB*) по адресу <http://sw-tools.pdb.org/index.html>.

### **3.6.2. Предсказание структуры и функций макромолекул**

Молекулярное моделирование позволяет исследователю строить любую желаемую молекулу на основе известной молекулярной геометрии составляющих атомов.

Затем к этой трехмерной структуре прикладывается поле сил, определяющих взаимодействие между каждыми двумя атомами молекулы. Такая система может быть использована для решения следующих задач.

1. *Оптимизация.* Черновой вариант молекулярной структуры может быть оптимизирован с тем, чтобы полученная структура лучше соответствовала наложенному силовому полю.

## Моделирование макромолекул

2. *Анализ нормальных мод.* Набор внутримолекулярных взаимодействий и атомных положений может быть проанализирован для специфических гармонических колебательных мод, представляющих основные движения изгиба и кручения целой молекулы.
3. *Молекулярная динамика.* Можно рассмотреть каким образом изменяется во времени движение компонентов молекулы по мере того как рассеивается энергия тепловых флуктуаций, изначально сообщённая молекуле.

---

4. *Возмущения свободной энергии.* Детально моделируется переход молекулярной системы из данного начального положения в различные конечные состояния с детализацией процессов образования и развития промежуточных состояний. Иногда

## Моделирование макромолекул

полезно моделировать физически *бессмысленный* путь реакции, который, тем не менее, термодинамически идентичен реальному переходу и компьютерное моделирование которого является разрешимой задачей, в то время как компьютерное моделирование реального перехода неразрешимо в принципе.

Типичные методы позволяют моделировать биомолекулы, состоящие из нескольких тысяч атомов, а типичные расчеты молекулярной динамики охватывают интервалы времени в несколько наносекунд.

Этого достаточно, чтобы рассмотреть такие "*быстрые*" процессы как процесс катализа или локальных структурных изменений, но недостаточно для рассмотрения более продолжительных во времени процессов таких, как белковый фолдинг или ассоциация биомолекул.



## Предсказание белкового фолдинга

Прежде чем бионанотехнология действительно станет распространенной технологией необходимо решить одну очень важную задачу. Мы должны быть способны *предсказать глобулярную структуру белка, исходя из его аминокислотной последовательности*. Без такого умения мы будем просто имитировать эволюцию, случайным образом модифицируя существующие белки до тех пор, пока они не изменятся в желаемую нам сторону. Белковый фолдинг является серьезной проблемой по двум причинам.

1. Прежде всего, это *масштаб (абсолютная величина) проблемы*. Средний белок имеет несколько сотен аминокислот. Каждая соединена с соседями двумя гибкими связями, которые имеют целый набор устойчивых конформаций. Кроме того, каждая из аминокислот имеет гибкий боковой радикал, который может быть тоже может принимать много стабильных конформаций. Все вместе эти многочисленные торсионные степени свободы определяют невообразимо большое пространство конформаций, оперировать которым не в состоянии даже самые современные суперкомпьютеры.

2. Вторая проблема лежит в *методе*, который используется для оценки стабильности каждой пробной конформации в ходе компьютерного итерационного эксперимента.

Свёрнутый белок имеет тысячи внутренних контактов, каждый из которых вносит *ничтожный* вклад в стабилизацию общей структуры.

Множество *молекул воды* высвобождается в ходе фолдинга белка, когда белковые цепи прячут свои гидрофобные участки внутрь глобулы. Это высвобождение молекул воды является значительной силой, заставляющей белки приобретать глобулярную структуру.

С другой стороны, *энтропия "сопротивляется"* образованию внутрибелковых связей и отрыву молекул воды от белковой нити. С точки зрения снижения энтропии жёсткая белковая глобула, имеющая единственную конформацию, энергетически менее выгодна, чем гибкая расплетённая белковая нить, имеющая огромное множество конформаций.

Энергия, которая выделяется при образовании внутрибелковых контактов и высвобождении молекул связанной воды расходуется на сворачивание цепи в компактную форму.

В целом для системы *эти два противоположных вклада практически уравнивают друг друга.*

Именно эту очень малую разницу, которая и представляет собой энергию стабилизации, мы и должны предсказать, когда пытаемся решить проблему белкового фолдинга, выбирая ту единственную третичную структуру, которая имеет наибольшую энергию стабилизации.

Однако величина этой энергии вычисляется как *разность между двумя большими величинами*, каждая из которых получена в результате суммирования огромного количества индивидуальных вкладов атомов молекулы белка. Даже малая погрешность определения величины вклада взаимодействий каждого из атомов белка в результате дает *суммарную (кумулятивную) ошибку*, которая больше, чем величина энергии стабилизации.

*Оба фактора – и (1) огромное пространство конформаций, и (2) кумулятивные ошибки в целевых функциях, – в сумме сводят на нет предсказания белкового фолдинга.*

Наиболее успешные приближения используют *упрощённые* модели, часто аппроксимирующие белковую цепь некоторым подобием кристаллической решётки, чтобы уменьшить число координат в пространстве конформаций. Однако, эти приближения все ещё далеки от предсказания трёхмерных структур при конструировании биомолекул.

**Гомологическое моделирование.** В настоящее время *самые хорошие предсказания* структуры белков получают *гомологическим моделированием*. При этом белки моделируются на основе известной структуры *подобных* белков. Проанализировав структуру всех белков Базы данных белков, оказалось, что белки, имеющие 30% *идентичных участков* первичной структуры, имеют *гомологичные структуры*. В этих структурах фолдинг и топология белковой цепи подобны, но локальные детали конкретных петель могут различаться. Гомологическое моделирование использует преимущества этого наблюдения. Структура нового белка может быть смоделирована на основе структуры известного белка, имеющего подобную аминокислотную последовательность (если, конечно, такая информация существует).

*Компьютерное моделирование*, в таком случае, используется для предсказания *структуры свободных петель* и для *определения координат специфических аминокислот*, которыми отличается новый белок от уже изученного.

Для белков, имеющих 60% и более гомологичности в первичной структуре, такие модели могут быть *очень точными*. В диапазоне гомологичности 30–60% первичной структуры такие модели могут быть полезны для предсказания *общих свойств* белковой структуры, таких как идентификация аминокислотных остатков, расположенных на поверхности, или могут быть использованы в предсказании общей формы глобулы.

**Предсказание локальной структуры белка на основе его аминокислотной последовательности.** Современные алгоритмы для предсказания локальной структуры белков на основе их последовательности также достаточно совершенны. Эти методы откалиброваны с исполь-

зованием известной структуры многих белков. При сканировании структуры нового белка эти алгоритмы классифицируют каждый участок как  $\alpha$ -спираль,  $\beta$ -структуру или другую структуру, или классифицируют участки как поверхностные или внутриглобулярные.

Эти методики дают корректные предсказания, если первичная структура белка известна не менее чем на 70%. В таком случае возможно корректное предсказание фолдинга белка. Подобные методики, в частности, являются успешными при идентификации *трансмембранных участков мембранных белков*. Поскольку такие участки имеют необычные свойства (они взаимодействуют с мембраной, а не с водным окружением), то в таком случае алгоритмы, предсказывающие белковую структуру, достигают 95% точности.

Однако отметим, что *оба* подхода, и (1) гомологическое моделирование и (2) предсказание вторичной структуры могут иметь лишь ограниченное применение в бионанотехнологии. Оба метода основаны на факте, что рассматриваемые аминокислотные последовательности являются развитием белковой структуры вследствие эволюции. Исследователь *изначально полагает*, что первичная белковая структура укладывается в стабильную, функциональную, свернутую глобулу. В то же время существует множество примеров, когда замена *единственной* аминокислоты полностью изменяет (разрушает) известную структуру и вышеперечисленные методы предсказания не в силах идентифицировать такого типа "катастрофу" вследствие единственной замены.

Однако успешное гомологическое моделирование может быть рассмотрено в качестве хорошей *стартовой точки* в бионанотехнологии. Используя его можно начинать разрабатывать модифицированные белки, начиная с существующих стабильных функциональных белков, и, модифицируя их по шагово, стремиться получить искомую новую функциональность.

## Моделирование докинга молекул

Специфическое взаимодействие между молекулами (молекулярное узнавание) является основой большинства биомолекулярных процессов.

*Ферменты* распознают форму (геометрическую и химическую) субстратов и создают такое окружение, которое благоприятствует протеканию необходимой химической реакции. *Антитела* имеют центры связывания, которые идеально соответствуют эпитопам антигенов. *Белки* взаимодействуют друг с другом и управляют функционированием друг друга через специфические места связывания. Для бионанотехнологии важно уметь предсказывать такие взаимодействия.

В настоящее время разработаны точные, самосогласованные методы для предсказания связывания молекул – лигандов или ингибиторов – с биомолекулами. На рисунке 103 показан пример компьютерной симуляции такого "докинга".



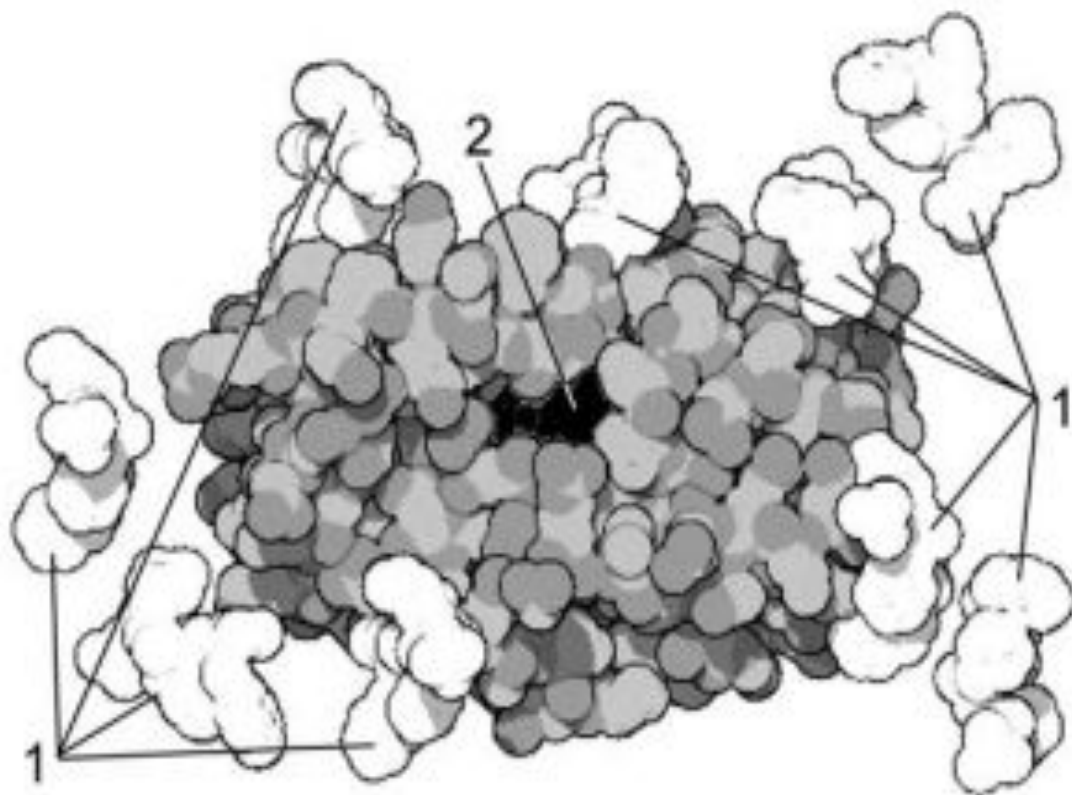


Рисунок 103 – Моделирование молекулярного докинга молекул лекарственного препарата саквинавир (*saquinavir*) к ВИЧ-протеазе: 1 – различные конформации молекул саквинавира; 2 – наиболее "удачная" конформация саквинавира, связанная в активном центре протеазы

Компьютерные программы проводят *поиск наилучших* центров связывания молекул лекарств с бионаномашинами. Если моделирование предсказывает *достаточно сильное* взаимодействие между молекулами, то тогда данное лекарственное средство синтезируют и проводят клинические испытания.

Наиболее совершенные компьютерные программы соединяют в себе *два* алгоритма.

Сначала они используют быстрый алгоритм для поиска множества способов, с помощью которых молекулы могут быть подогнаны друг к другу геометрически.

Затем к геометрическому подобию подключают *систему межатомных взаимодействий*, оптимизация которых (поиск конфигурации с минимальной энергией) позволяет точно предсказать энергию взаимодействия. *Компьютер перебирает множество возможных способов* связывания молекул, используя энергетический критерий для определения наилучшего из них.

Безусловно, для того, чтобы такие расчетные задачи на современных компьютерах возможно было *решить в разумное время*, используются различные *упрощения*. Как правило, предполагается, что белковая молекула является абсолютно (или в значительной степени) жёсткой – тем самым снижается количество возможных конформаций, которые следует рассчитать. Но это упрощение может создавать значительные трудности, например, когда белок "смыкается" вокруг малой молекулы при её связывании. Современные методики, использующие стандартные методы, являются успешными в половине случаев. В остальных случаях необходимо внимательно учитывать любые движения аминокислотных петель внутри белка, которые могут оказать существенное влияние на результат.

Что же касается моделирования *взаимодействия между большими биомолекулами*, особенно между белками, то эта задача все ещё слишком сложна для более-менее точного решения. В настоящее время не существует "готового к употреблению" метода для её решения, однако многие лаборатории уже тестируют такие методы.

Проблема объединения макромолекул, вследствие большого размера обеих молекул, намного более сложная, чем предсказание взаимодействия белка с малыми молекулами. Сейчас работа ведется в *двух* направлениях.

1. В первом реальные биомолекулы упрощаются, и вместо них используется модель их поверхностей с соответствующими химическими группами на них. Возможные комплексы молекул получают сначала простым перебором всех возможностей топологического докинга, а затем исследуют наиболее подходящие конфигурации более детально.
2. Второй подход, в котором "в лоб" моделируется атомная структура и атом-атомные взаимодействия, *основан на успехах геномики и протеомики*, в которых структура и функции белков определены с использованием сравнения структур генов [10].

Этот подход только недавно стал применяться и уже показал своё преимущество и эффективность в предсказании конформаций для эффективного межмолекулярного связывания и предсказания энергии связывания таких молекулярных комплексов.

## Моделирование новых функций молекул

Значительное количество успешных решений в бионанотехнологии было достигнуто благодаря видоизменению природных биомолекул с целью придания им новых функций.

Конечно же, чем лучше мы знаем структуру макромолекулы, тем больше уверенности в успехе наших модификаций этой макромолекулы. Простое молекулярное моделирование является основным инструментом позволяющим проектировать изменения в структуре молекул, глядя на экран монитора. Это позволяет исследователю конструировать новые структуры, оптимизировать их, следя за тем, насколько хорошо эти изменения вписываются в общую структуру макромолекулы. Такой подход в сочетании с молекулярной динамикой обеспечил большинство успехов молекулярной нанотехнологии сегодня. Большое число наноразмерных моделей было создано с использованием жёсткого атомного скелета, используя необходимую геометрию межатомных связей. На практике молекулярное моделирование интенсивно используется для разработки технологий точечных мутаций (*site-directed*

*mutations*) с целью повышения стабильности и изменения функциональности белков. Однако поскольку в компьютере можно спроектировать *все что угодно* и число таких вариантов бесконечно, то, для того, чтобы такое моделирование давало практический результат, исследователь должен иметь определённый опыт и обладать креативными качествами.

Многие лаборатории разработали *специализированные методы, помогающие в молекулярном дизайне*. Например, многие методы были созданы для того, чтобы минимизировать ручную рутинную работу. Эти способы позволяют проводить автоматизированную подгонку вместо интуитивного поиска вручную методом проб и ошибок (*hit-or-miss approach*, метод "тыка"), при моделировании они эффективнее учитывают изменение свободной энергии (возможность самопроизвольного протекания процесса). Многие из таких методик были созданы в области компьютеризированной разработки лекарств. Их целью является дизайн молекулы лекарства, которая идеально вписывается в активный центр бионаномашины, блокируя нормальное функционирование макромолекулы.

Некоторые программы начинают с попыток *докинга тысяч малых молекулярных фрагментов*, каждый из которых состоит из 5–10 атомов. Затем, чтобы спроектировать молекулу лекарства, те фрагменты, которые наилучшим образом вписались в активный центр, соединяются с теми фрагментами, которые хорошо подходят к соседним с активным центром участкам белка.

Другой подход к решению той же проблемы состоит в том, что *дизайн начинают с "базисной" молекулы*, которая связывается в глубине активного центра. Конструирование же молекулы лекарства осуществляется *добавлением* атомов по одному к этой базисной молекуле до тех пор, пока активный центр не будет полностью заполнен. Такими методами могут быть сконструированы превосходные кандидаты на новые лекарства (лиды), но есть и опасность, когда торопливый и нетерпеливый разработчик может спроектировать настолько экзотическую молекулу, которую ни один химик не в силах будет синтезировать.

# Моделирование макромолекул

## Вопросы для самоконтроля

1. Какова роль компьютерного моделирования в бионанотехнологии?
2. Назовите две причины, не позволяющие до настоящего времени осуществить компьютерное моделирование белкового фолдинга.
3. В чём заключается гомологическое моделирование структуры белков?
4. Для каких целей используют компьютерное моделирование молекулярного докинга?