

Компьютерные методы в фармакологии



Владимир Борисович Сулимов

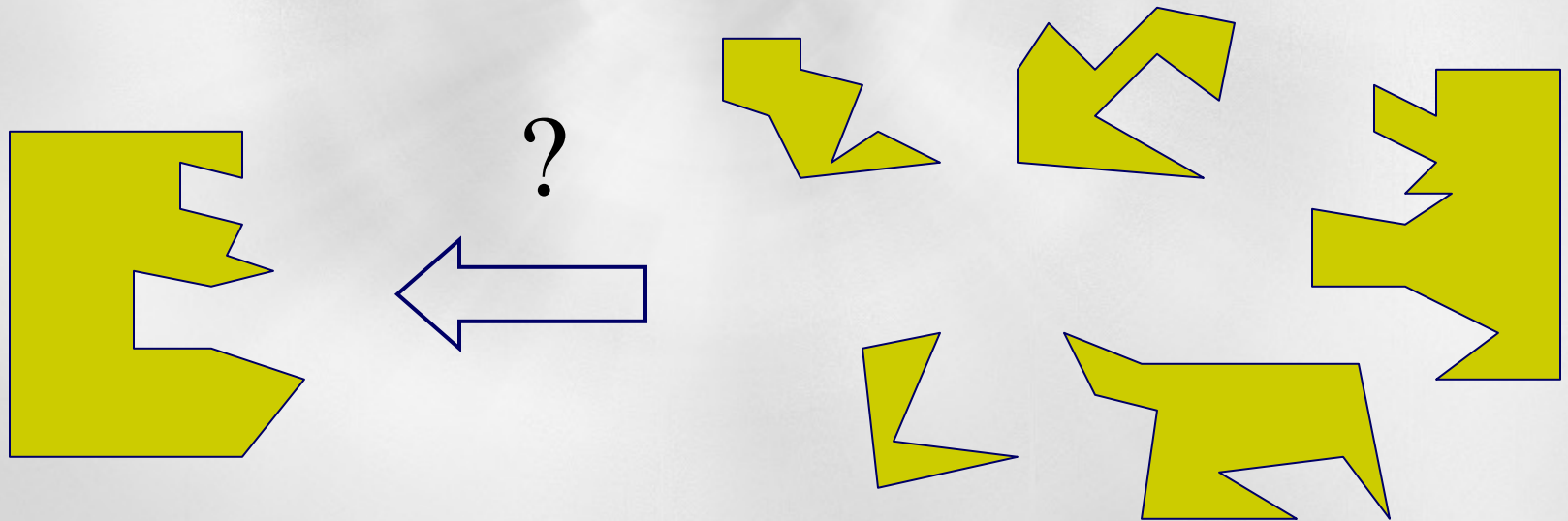
Научно-Исследовательский Вычислительный Центр МГУ

Лекция № 2

Методы молекулярного моделирования

Поиск ингибитора

- Как найти молекулу-ингибитор для заданного белка-мишени?
- Обычный путь: *in vitro* и *in vivo* экспериментальный очень затратный перебор различных химических соединений
- Новый путь: виртуальный перебор с помощью компьютеров - *in silico*!
- Молекула-кандидат в ингибиторы – лиганд (ligand) от латинского слова *ligare* = to bind: молекула, которая может связаться и сформировать комплекс с биомолекулой
- Компьютерный перебор баз данных лигандов



Докинг, скоринг и скрининг – основа конструирования лекарств

Молекулярное моделирование помогает быстрее и дешевле выполнять начальную стадию разработки

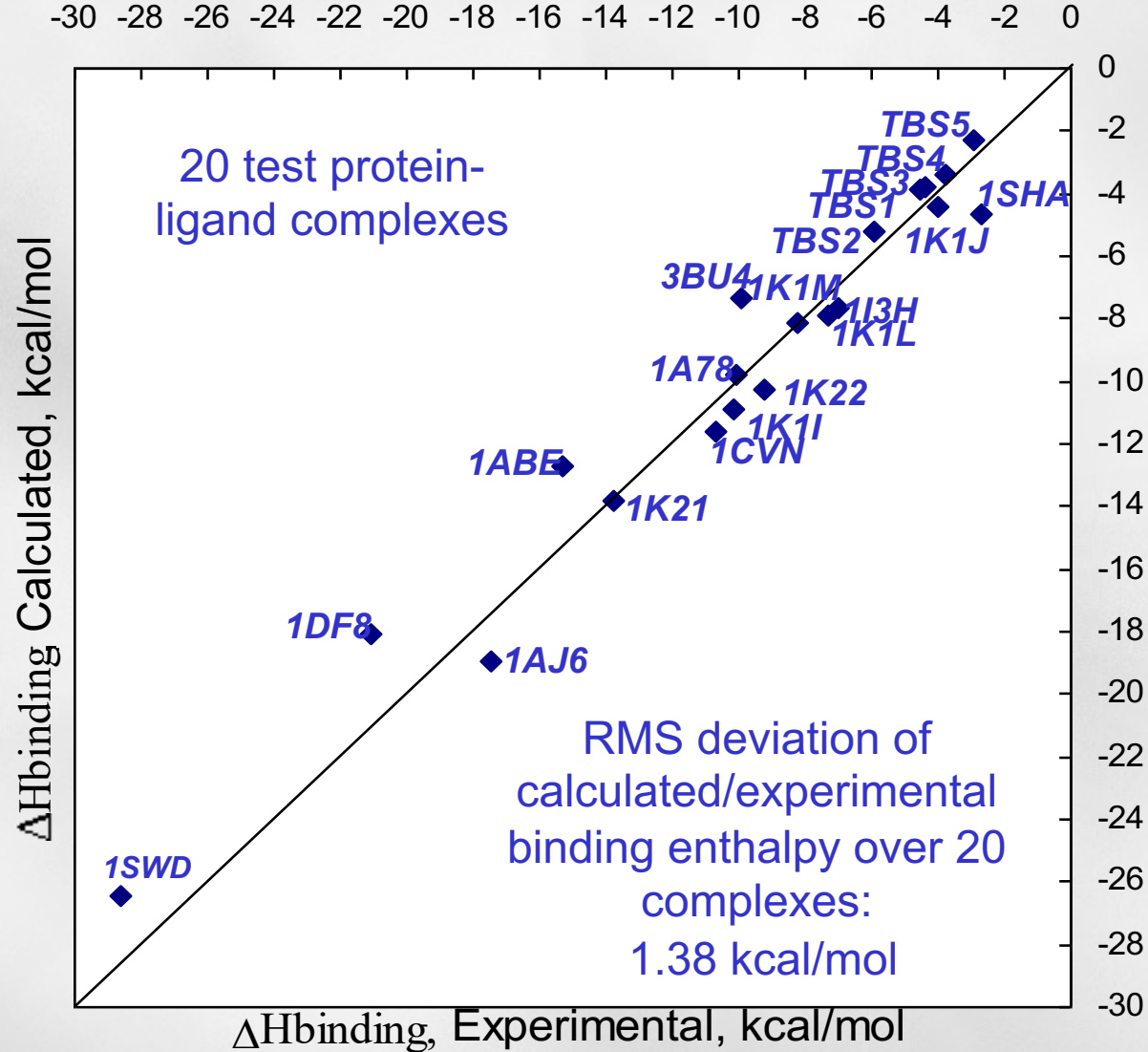
- ▶ **Докинг** – позиционирование лиганда в активном центре белка
- ▶ **Скоринг** – оценка энергии связывания лиганда с белком: чем сильнее лиганд связывается с белком, тем эффективнее лекарство – меньшая концентрация лекарства даст терапевтический эффект
- ▶ **Скрининг** – перебор больших баз данных молекул с целью поиска кандидатов в ингибиторы – **нужны суперкомпьютеры**

Межмолекулярное взаимодействие

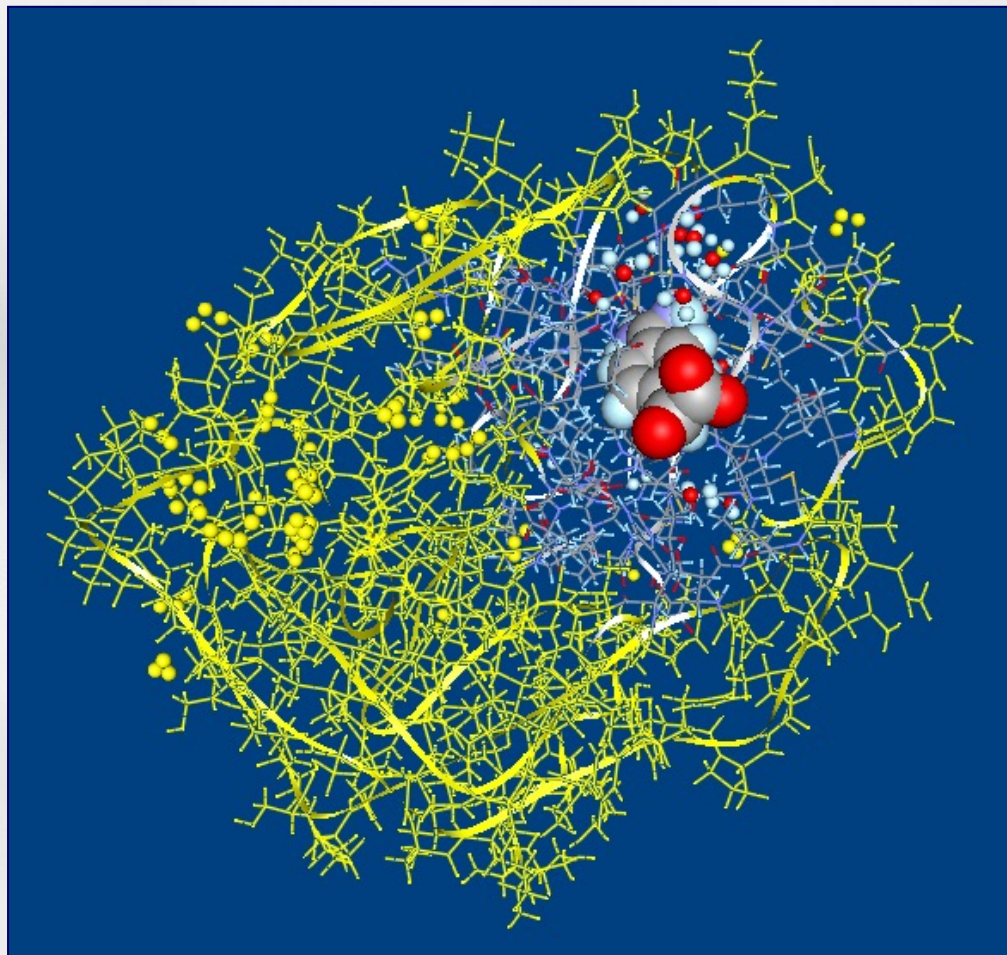
- Правильный учет сил между молекулами – основа успеха молекулярного моделирования лекарств
- Эти силы могут быть правильно рассчитаны только методами квантовой механики
- Силы между атомами и структура молекул правильно рассчитываются только методами квантовой механики
- Вывод: конструировать лекарства надо методами квантовой механики

Quantum calculation of binding enthalpy for protein-ligand complexes:

methodology validation



QM расчет ΔH связывания белок-лиганд



$$\Delta G = \Delta H - T \Delta S$$

Молекулярная механика

- Квантовая механика и квантовая химия
- Точные законы – очень трудоемкие вычисления – одна молекула из **нескольких десятков атомов** - **десятки и сотни часов**
- **Что делать?**
- Создать **силовое поле (force field)**– набор потенциалов, описывающих взаимодействие атомов и молекул, как **классических частиц**: расчеты – **десятки минут для тысяч атомов**

Силовые поля

- Knowledge-based: потенциалы из статистики взаимного расположения атомов
- Empirical: потенциалы из экспериментальных значений констант связывания
 - Обучающий набор комплексов протеин-лиганд
 - Тестовый набор комплексов протеин-лиганд
- Physical: силовые поля - потенциалы, полученные из физических законов, учет растворителя

Силовые поля – Force fields

- MM3
- Charmm
- Amber
- OPLS-2005
- OPLS3
- **MMFF94 – Merck – Большая Фармацевтическая компания**
- ReaxxFF

Силловые Поля – Force fields

- MMFF94 – *ab initio* force field
 - На основе **квантовой механики**
 - Валентные связи
 - Валентные углы
 - Торсионы – вращения вокруг связей
 - Межмолекулярное взаимодействие
 - Электростатические взаимодействия
 - Ван дер Ваальсовы взаимодействия

Молекулы

- Белки – протеины – тысячи атомов
 - Полипептидная цепь:



- Органические молекулы – десятки атомов:
миллиарды молекул: C, N, O, H, S, Cl, F, Br ...

Чтобы предсказать константу связывания молекулы-лиганда с белком

...мы должны рассчитать всего-навсего свободную энергию связывания комплекса белок-лиганд

$$\Delta G = \Delta H - T\Delta S$$

Изменение энтальпии
при связывании
лиганда

Изменение энтропии
при связывании лиганда

Термодинамика и статистическая
физика

Аккуратный расчет свободной энергии связывания лиганда с протеином

- Методы молекулярной динамики: $\Delta G = \Delta H - T\Delta S$
 - Вычисление траектории движения атомов системы белок+лиганд, находящейся в термостате
 - Описание движения атомов как классических частиц, взаимодействующих друг с другом с помощью классических потенциалов
 - Усреднение по траектории – усреднение по ансамблю
 - Время вычислений: сутки и более на одном процессоре
 - нужно использовать многопроцессорный режим
 - Реальный пример: Задача 60 тысяч атомов 100 нс, 224 процессора (16 узлов) считалась на суперкомпьютере Ломоносов-2 61 час – сейчас одна задача мах 20 часов.

Что измеряется в эксперименте?

С чем сравнивать расчет?

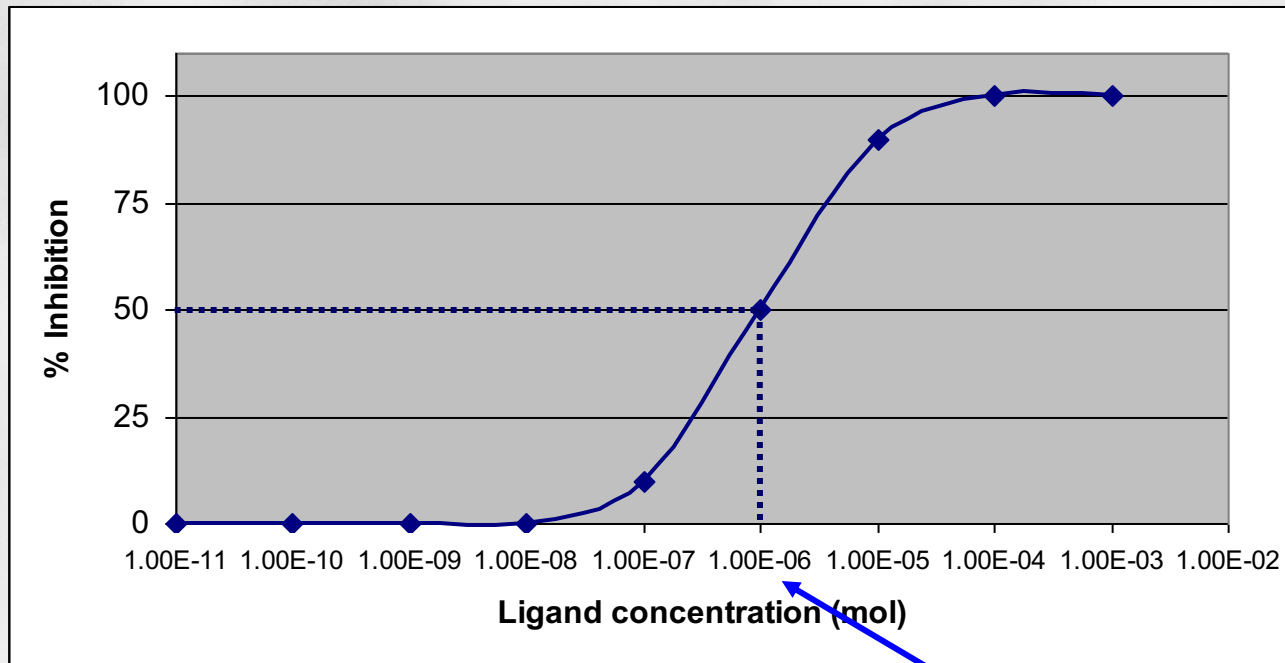
- ▶ Измерение IC50 – ингибирующая концентрация: концентрация ингибитора, при которой активность белка подавляется на 50%
- ▶ Чем меньше IC50 , тем выше ингибирующая активность вещества: IC50 = 1 наномоль лучше, чем IC50 =1 микромоль
- ▶ Чем больше энергия связывания белок-лиганд, тем меньше IC50 и лучше ингибитор
- ▶ Белок вместе с ингибитором кристаллизуются, применяется метод рентгеноструктурного анализа для определения координат атомов комплекса белок-лиганд и определения положения лиганда в активном центре белка

Что измеряется в эксперименте?
С чем сравнивать расчет?

- Измерение IC_{50} – ингибирующая концентрация
- IC_{50} – мера ингибирующей активности данного вещества
- IC_{50} – измеряется в экспериментах *in vitro*

Измерение ингибирующей концентрации IC_{50}

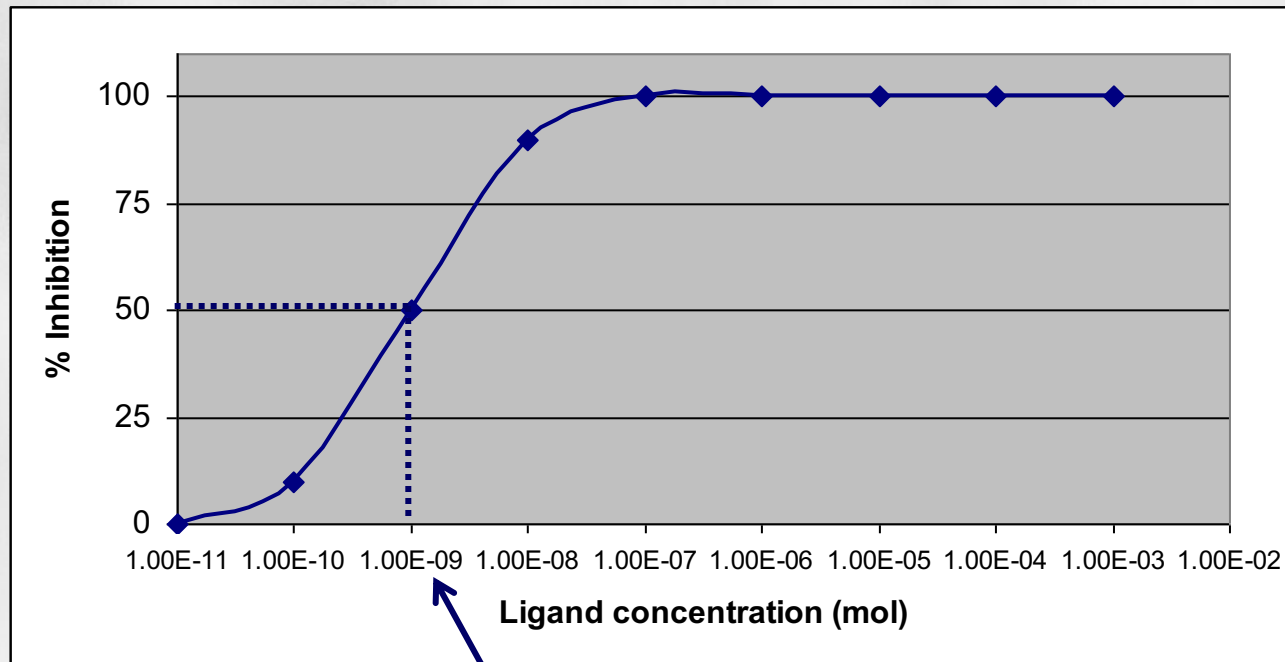
Лиганды со средней способностью связывания с белком



Концентрация, при которой достигается 50% ингибирование (IC_{50}) = 10^{-6} моля

Измерение ингибирующей концентрации IC_{50}

Лиганды с **высокой** способностью связывания с белком

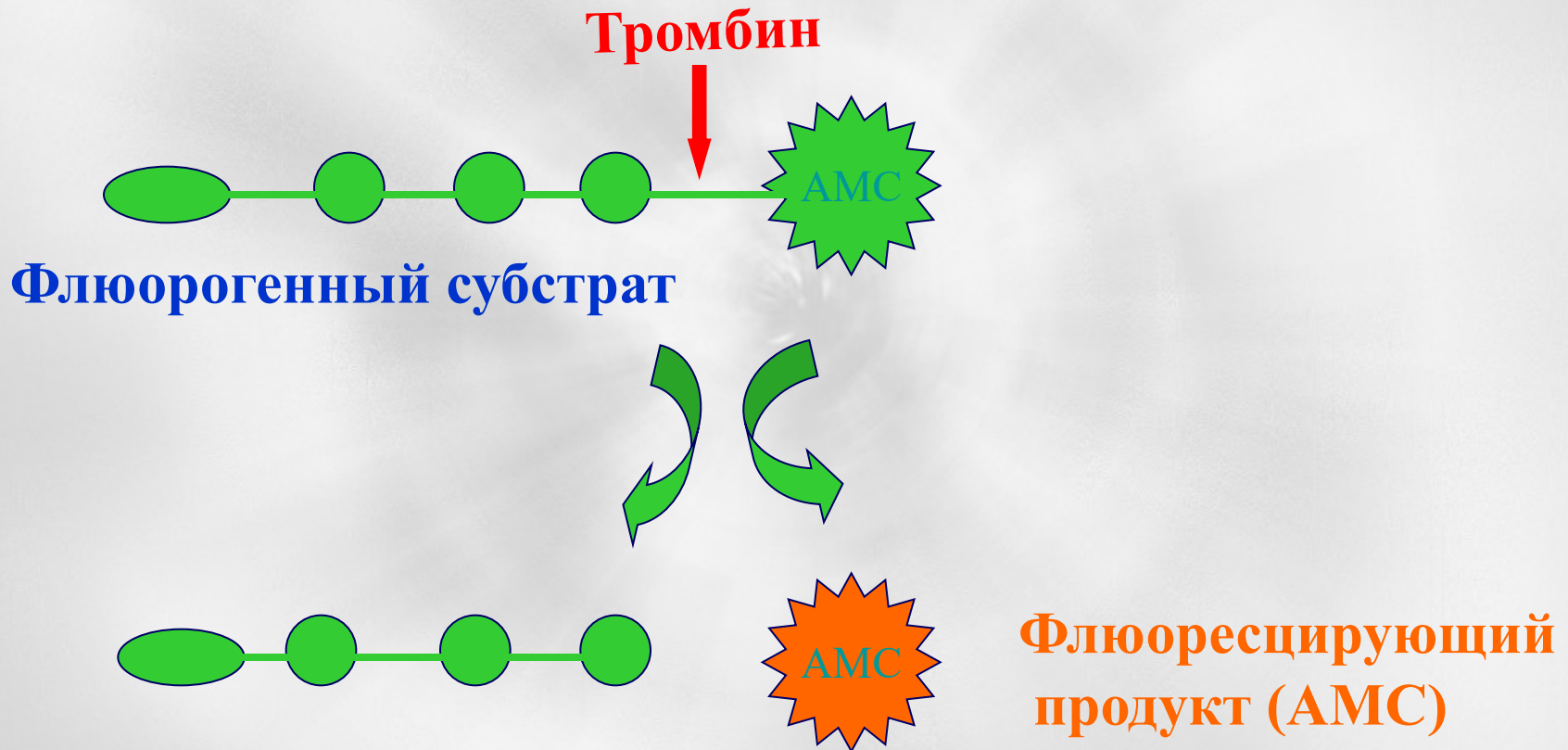


Концентрация, при которой достигается 50% ингибирование (IC_{50}) = 10^{-9} моля

Измерение IC_{50} при ингибировании тромбина

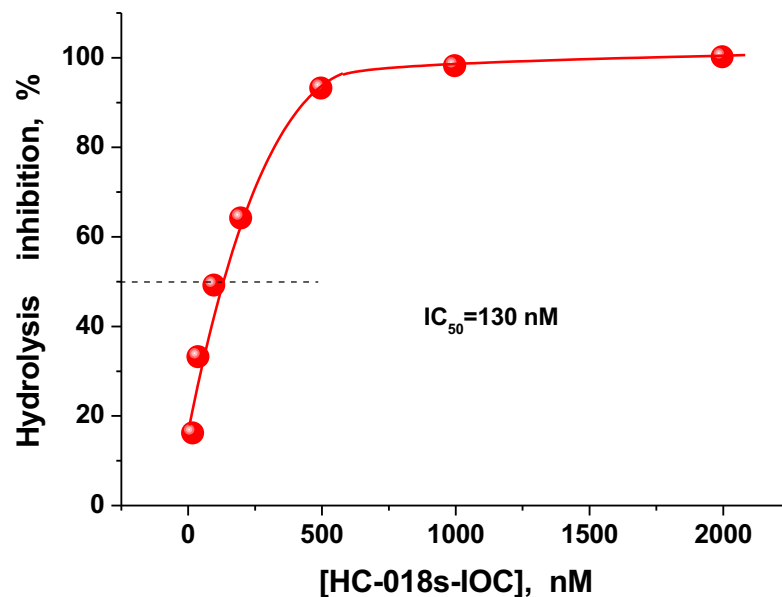
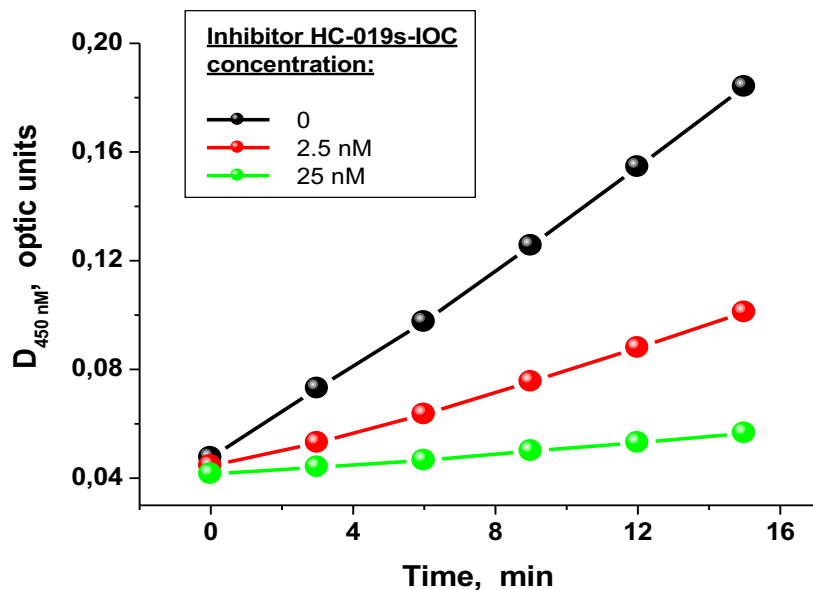
Схема гидролиза субстрата тромбином

такие методы измерений часто называются FRET
Fluorescence (Förster) Resonance Energy Transfer



Ингибитор тромбина уменьшает скорость гидролиза тромбином хромогенного (флуорогенного) субстрата:

IC_{50} – это концентрация ингибитора, при которой скорость гидролиза снижается до 50% от первоначальной величины



Комплексы Белок-Лиганд

- Protein Data Bank - PDB:
www.rcsb.org/pdb
 - X-ray эксперименты
 - 3D структуры закристаллизованных комплексов белок-лиганд
 - Положения только тяжелых атомов
 - Точность 1.5 Å - лучшие

Структуры молекул белков

- ▶ Белки кристаллизуются и методом рентгеноструктурного анализа экспериментально определяется пространственное расположение их атомов
- ▶ Определяются декартовы координаты «всех» атомов белка
- ▶ Декартовы координаты атомов комплексов белок-лиганд
- ▶ Информация о белках и комплексах белков с лигандами собрана в базе данных Protein Data Bank (PDB) <http://www.rcsb.org/>

Общее количество структур – 201 196 (13 февраля 2023); 69 511 в 2010

X-ray – 172 767 структур, ЯМР – 13 906 (13 февраля 2023),

Белки: 174 825 (13 февраля 2023)

Resolution характеризует точность определения координат атомов в структуре кристалла. Чем меньше этот параметр тем лучше. Можно работать, когда Resolution < 3.0 Å.

Докинг – начало компьютерной разработки новых ингибиторов

- Докинг - позиционирование молекулы-лиганда в активном центре белка-мишени
- Лиганд – от латинского слова *ligare* – *связываться* – это молекулярный объект, который может связываться с биомишенью

Программы докинга стали популярны. Десятки программ докинга востребованы при разработке лекарств. Каждая программа – уникальное сочетание моделей и приближений, скорости

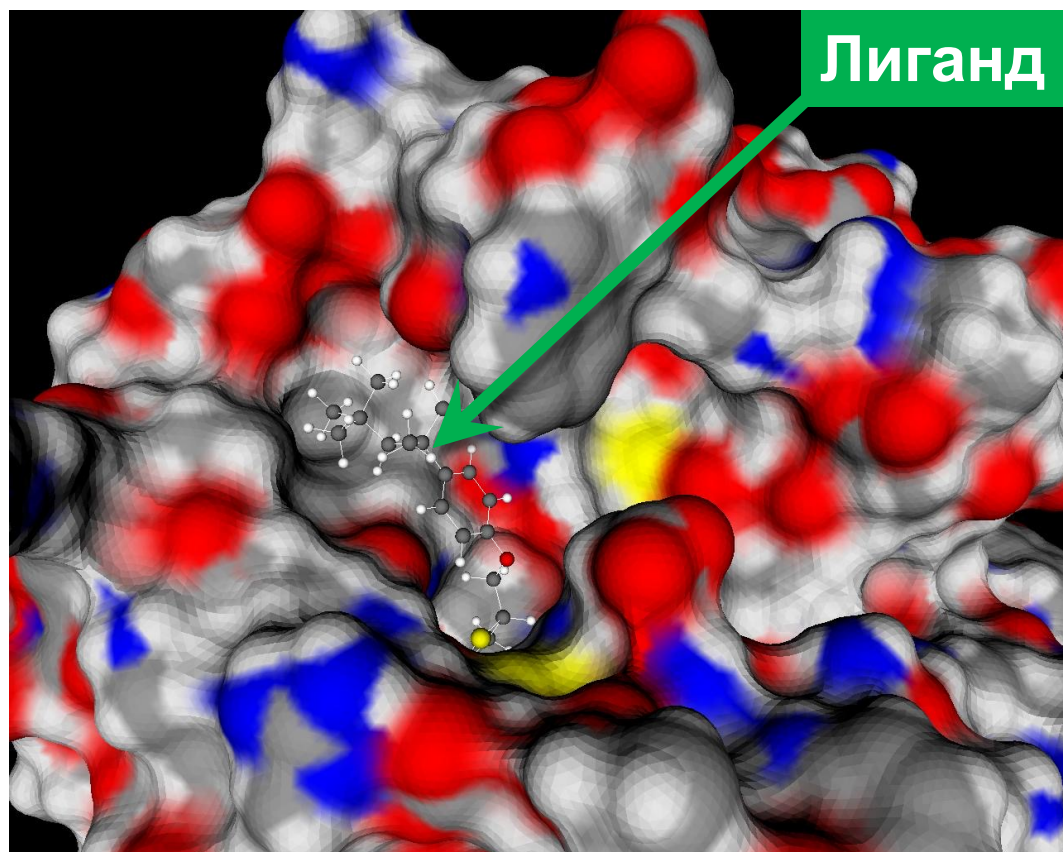
► Возможно ли увеличить точность докинга?

- Точность позиционирования – удовлетворительна
- Точность расчетов энергии связывания белок-лиганд ΔG_{bind} – плохая



Перебор больших баз данных органических молекул:
Поиск молекул с наибольшей энергией связывания с белком-мишенью

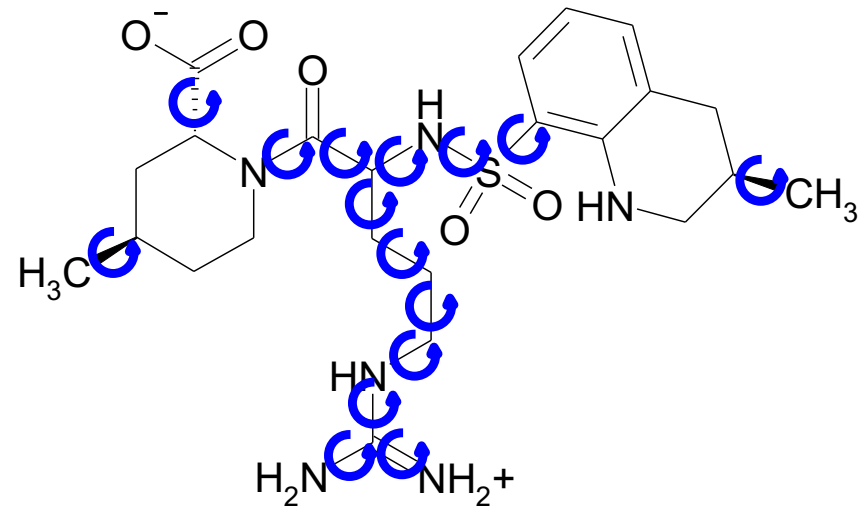
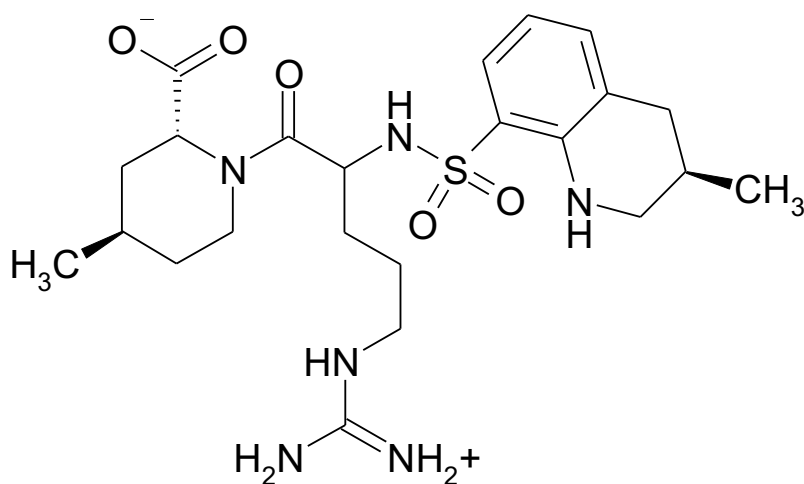
Парадигма докинга: в эксперименте
лиганд связывается вблизи **глобального**
минимума энергии системы белок-лиганд



Докинг –
нахождение
Глобального
минимума
энергии
системы белок-
лиганд

Основная проблема докинга: многомерность пространства глобальной оптимизации

Типичный лиганд – ингибитор тромбина, Аргатробан, антикоагулянт – препятствует свертываемости крови



Получаем задачу на нахождение глобального минимума в пространстве $15 + 6 = 21$ измерений!

Систематический перебор невозможен

Докинг

- ▶ Первые программы докинга появились в конце 1980х: Силовых полей тогда не было, не было и многопроцессорных компьютеров. Простейшие модели докинга – жесткий лиганд в жесткий белок: ключ (лиганд) соответствует замку (белок)
- ▶ Простейшие силовые поля, докинг на настольном компьютере, докинг на ноутбуке – нужно посчитать многие тысячи молекул.
Мантра всех программ докинга: кто быстрее!
- ▶ **Миллениум:** Созданы силовые поля, глобальная оптимизация на сложной многомерной энергетической поверхности

Докинг

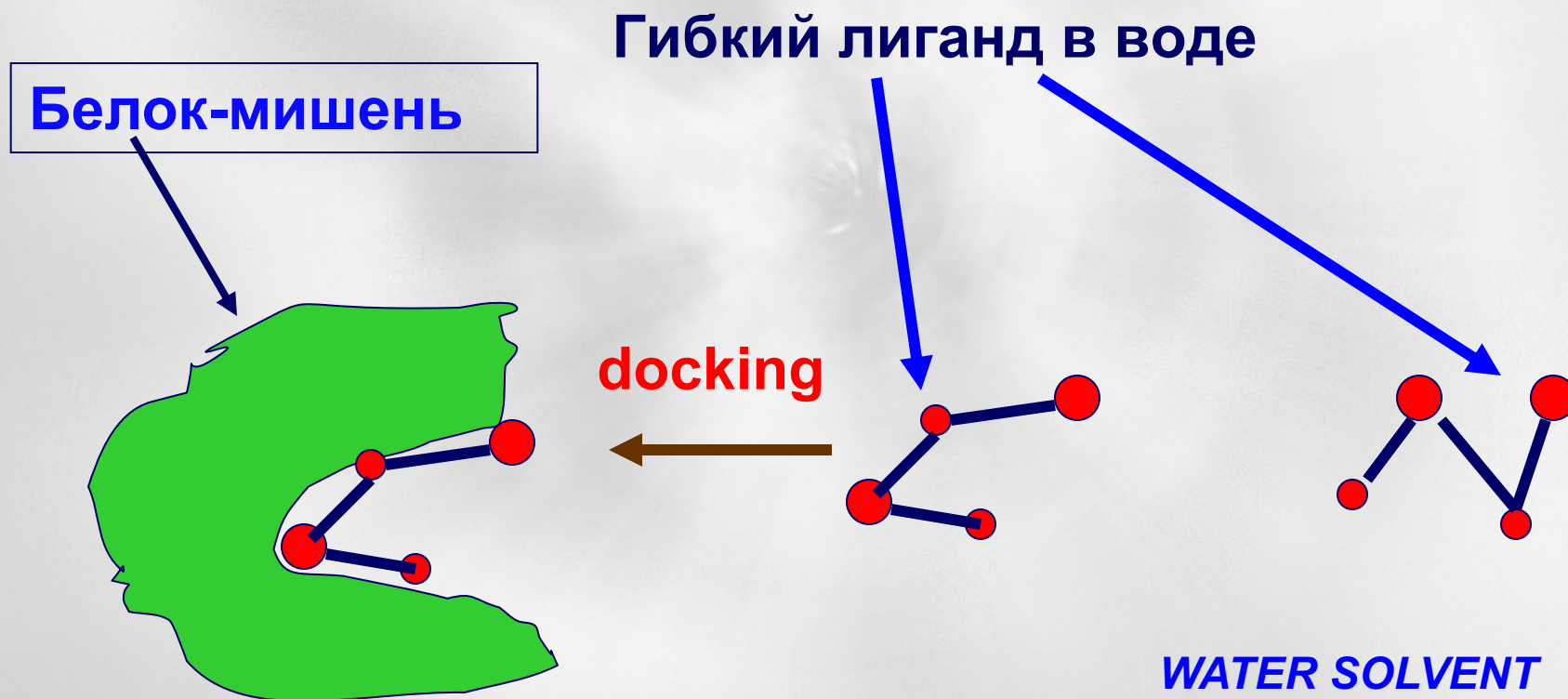
- ▶ 2005-2010 - Суперкомпьютеры, **мантра: кто быстрее, кто больше**, виртуальный скрининг многих десятков тысяч молекул, сейчас – миллионы и миллиарды молекул
 - ▶ Суперкомпьютеры: **кто точнее?** Многопроцессорный докинг 1 лиганда, минимум упрощений и без подгоночных параметров
 - ▶ Точность определяет эффективность применения компьютерного моделирования для разработки лекарств, эффективной всего конвейера разработки лекарств.
-
- ▶ **Квантовый докинг – энергия с помощью квантовой механики**
 - ▶ **2х ступенчатый виртуальный скрининг:**
 - Скрининг больших баз данных молекул – классический докинг SOL
 - Квантовый докинг лучших молекул, отобранных SOL

Докинг

- ▶ Начальный этап разработки лекарства: виртуальный скрининг многих десятков тысяч молекул, сейчас – миллионы и миллиарды молекул – это существующие, лежащие на полке молекулы
 - ▶ Зачем это нужно?
 - ▶ 1. найти новые молекулы-ингибиторы и подтвердить их ингибирующую способность в экспериментах *in vitro*.
 - ▶ Далее, отталкиваясь от этих молекул, синтезировать новые более сильные ингибиторы, и сделать основу лекарства
-
- ▶ 2. Среди существующих лекарств найти новые ингибиторы для новых белков-мишеней. Такие соединения очень быстро могут пройти все испытания и быть зарегистрированы в качестве новых лекарств – новое применение.

Ключевые программы для дизайна лекарств

Docking – правильное позиционирование лиганда в активном центре белка-мишени

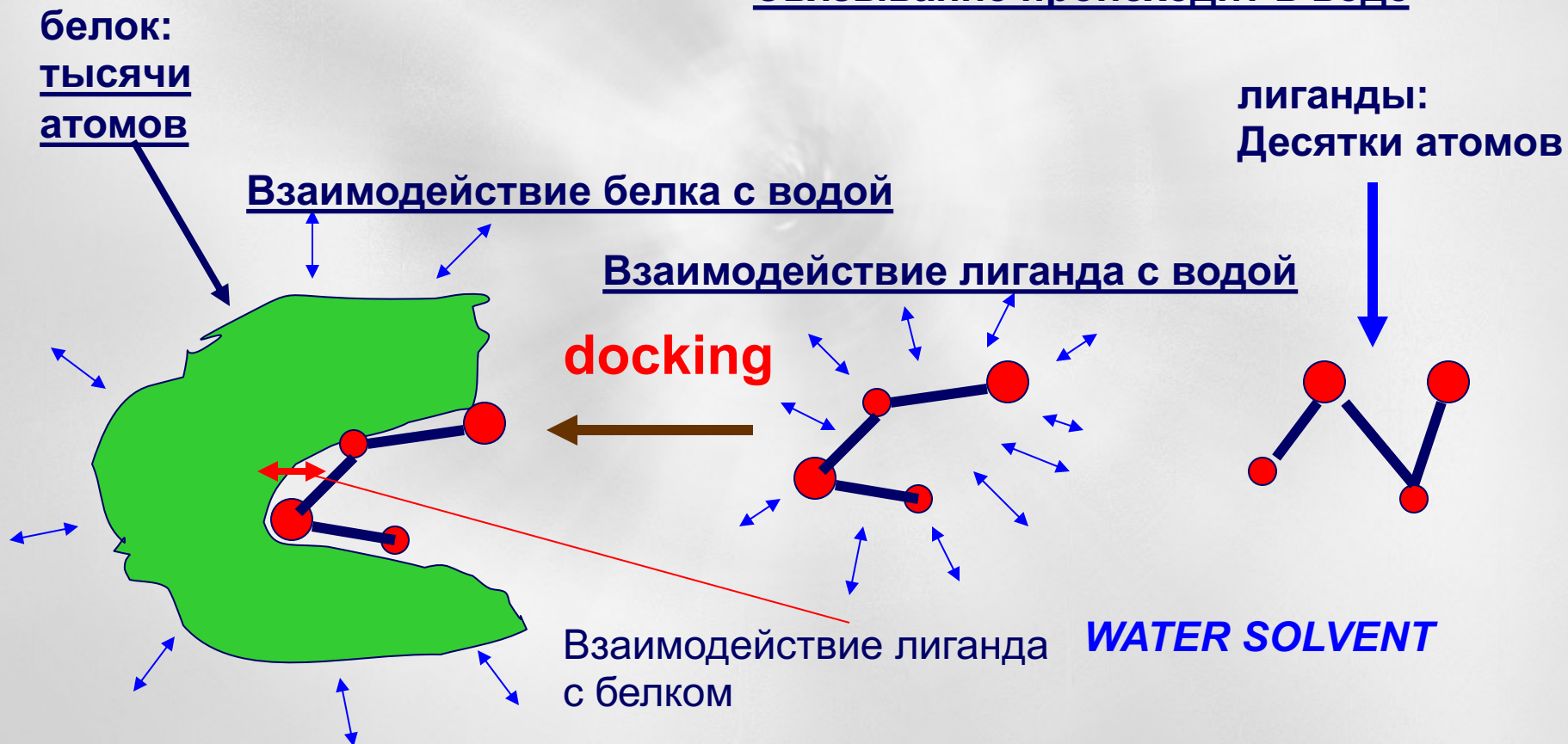


Scoring – правильная оценка энергии связывания белок-лиганд

Слабые
межмолекулярные
взаимодействия

Высокая точность: ~ 1 kcal/mol ~ 0.05 eV

Связывание происходит в воде



Энергия связывания белок-лиганд

- Взаимодействия между атомами белка и лиганда: **электростатика** и **силы Ван-дер-Ваальса**
- Энергия десольватации:
 - Подсчет баланса взаимодействий: лиганда с растворителем, белка с растворителем, комплекса белок-лиганд с растворителем
- Энергия внутренних напряжений лиганда
- Изменение числа вращательных степеней свободы лиганда при связывании – **неправильно! Надо честно считать изменение энтропии**

Энергия десольватации ΔG_{desolv} :

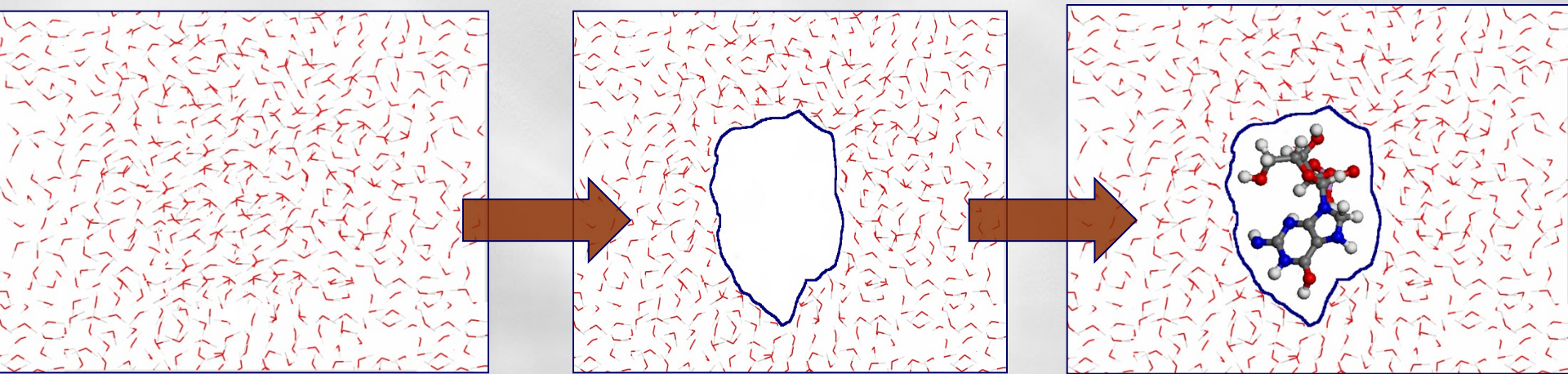
$$\Delta G_{desolv} = G_{solv}(\text{белок-лиганд}) \\ - G_{solv}(\text{белок}) - G_{solv}(\text{лиганд}),$$

где $G_{solv}(X)$ – энергия сольватации молекулы X ,

В величину $G_{solv}(X)$ вносят вклад следующие взаимодействия:

- ❖ Электростатическое взаимодействие молекулы с растворителем
- ❖ Ван-дер-Ваальсово взаимодействие молекулы с растворителем
- ❖ Кавитация: статистическое образование полости в растворителе, имеющей форму растворяемой молекулы

Процесс растворения молекулы (solute) в растворителе (solvent)



$$\Delta G_{\text{растворения}} = \Delta G_{\text{кавитации}} + \Delta G_{\text{взаимодействий}}$$

$$\Delta G_{\text{взаимодействий}} = \Delta G_{\text{vdW}} + \Delta G_{\text{электростатические}}$$

$$\Delta G_{\text{SOLV}} = \Delta G_{\text{NES}} + \Delta G_{\text{ES}}$$

Энергетические единицы

$$1 \text{ эВ} \approx 23 \text{ kcal/mol}$$

$$1 \text{ эВ} = 1.602176...10^{-19} \text{ Дж}$$

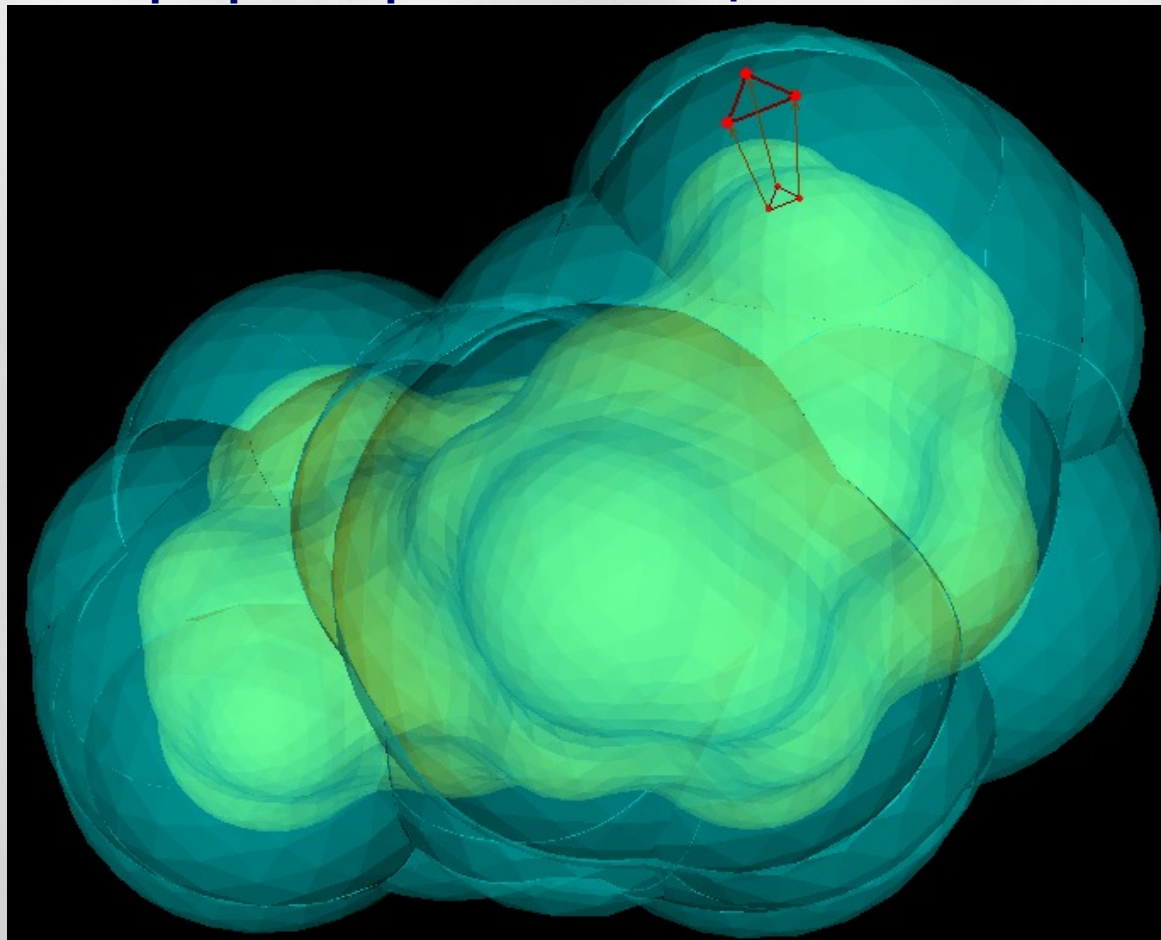
$$1 \text{ cal} = 4.1868 \text{ Дж}$$

$$N_A = 6.022141...10^{23}$$

$$1 \text{ эВ} = 1.602...10^{-19} \times N_A \times 10^{-3} / 4.1868$$

$$1 \text{ эВ} = 23.045121 \text{ kcal/mol}$$

Автоматическое построение гладкой триангулированной поверхности вокруг молекулы любой формы и любого размера: программа построения такой поверхности **TAGSS** разработана в Лаборатории Вычислительных Систем и Прикладных Технологий Программирования НИВЦ МГУ

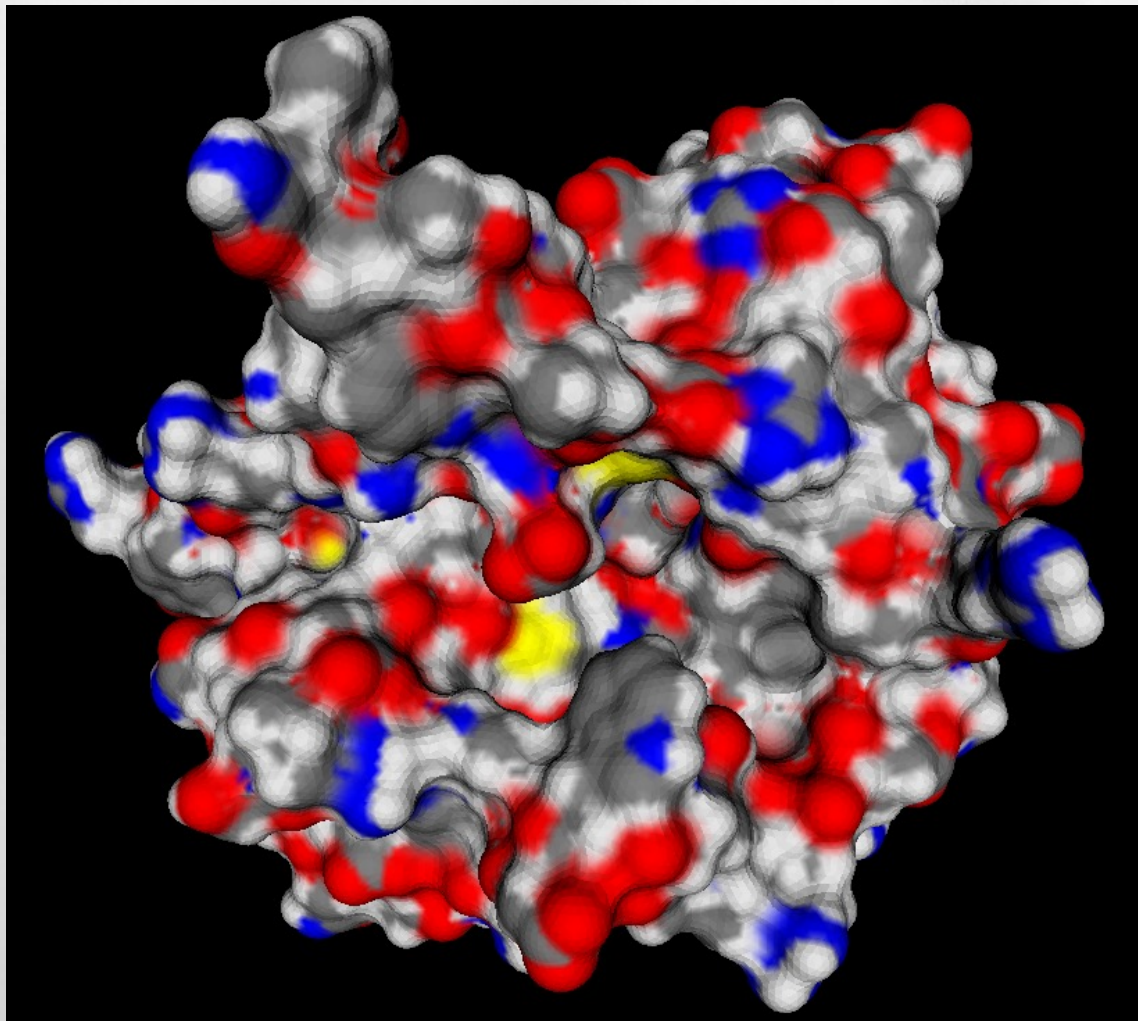


Желтая поверхность –
для подсчета
электростатических
взаимодействий
молекулы с
растворителем

Прозрачная поверхность –
для подсчета VdW
взаимодействий молекулы
с растворителем

Прозрачная поверхность
получается путем
соответствующего
преобразования из желтой

Поверхность доступная растворителю вокруг белка тромбина



Здесь различные цвета различных областей поверхности соответствуют разным типам атомов, находящихся вблизи этих областей:

- **Красный** – кислород
- **Синий** – азот
- **Серый** – углерод
- **Белый** – водород
- **Желтый** – сера

Программы докинга

- **AutoDock** – The Scripps Research Institute – **USA**
- **ICM** – Molsoft, LLC (Abagyan) San Diego, CA, **USA**
- **DOCK6** – University of California, San Francisco, **USA**
- **GOLD** - The Cambridge Crystallographic Data Centre - **UK**
- FlexX – BioSolveIT - **Germany**
- Sybyl – Tripos - **USA**
- Discovery Studio – Accelrys - **USA**
- Glide – Schrödinger - **USA**
- Lead-Finder – Moltech, Ltd – **Москва, Россия**
- SOL – НИВЦ МГУ - **Москва, Россия**
- FLM – НИВЦ МГУ - **Москва, Россия**
- SOL-T, SOL-P - **Москва, Россия**

Программы докинга против COVID-19

- ▶ В настоящее время известно около 50 программ докинга и с десяток сайтов для выполнения докинга online
- ▶ Каждая программа уникальна совокупностью используемых моделей, аппроксимаций и подгоночных параметров
- ▶ Большинство программ рассчитаны на работу на ноутбуке (1 CPU) и быстрый докинг – меньше минуты на 1 лиганд
- ▶ Пандемия высветила чрезвычайную востребованность докинга, особенно докинга больших баз данных, имеющих сотни тысяч и даже миллионы молекул
- ▶ Цель – быстрая разработка противовирусных препаратов
- ▶ Несколько белков коронавируса SARS-CoV-2 считаются терапевтическими мишенями, и их атомистическая пространственная структура измерена – можно проводить **Computer Aided Structural Based Drug Design**

Спасибо за внимание

- *...Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils...*
- *...Каждый медицинский метод есть инновация; а кто не хочет применять новые средства, должен ждать новых бед...*

Sir Francis Bacon (1561-1626)



OF INNOVATIONS