

СУПЕРКОМПЬЮТЕРЫ ДЛЯ РАЗРАБОТКИ ЛЕКАРСТВ

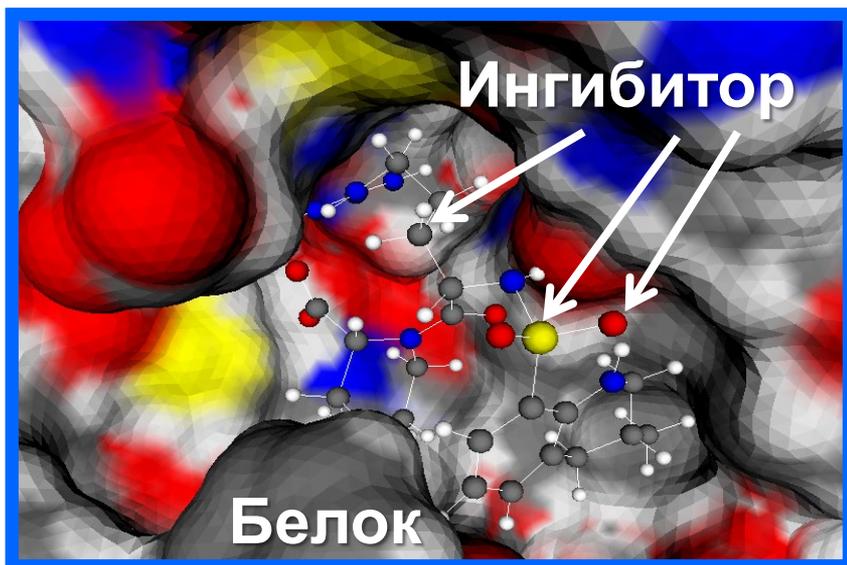
А.В. Сулимов, Д.К. Кутов, В.Б. Сулимов



Работа поддержана грантом РФФ, Соглашение №21-71-20031

Парадигма действия лекарства

- ▶ Для многих заболеваний известны белки, определяющие развитие патологии
- ▶ Такие белки – мишени для действия лекарств
- ▶ Молекулы-ингибиторы лекарства **связываются с активным центром** белка и блокируют его работу



**Блокирование
белка приводит к
излечению
заболевания**

Докинг – основной метод моделирования для разработки лекарств

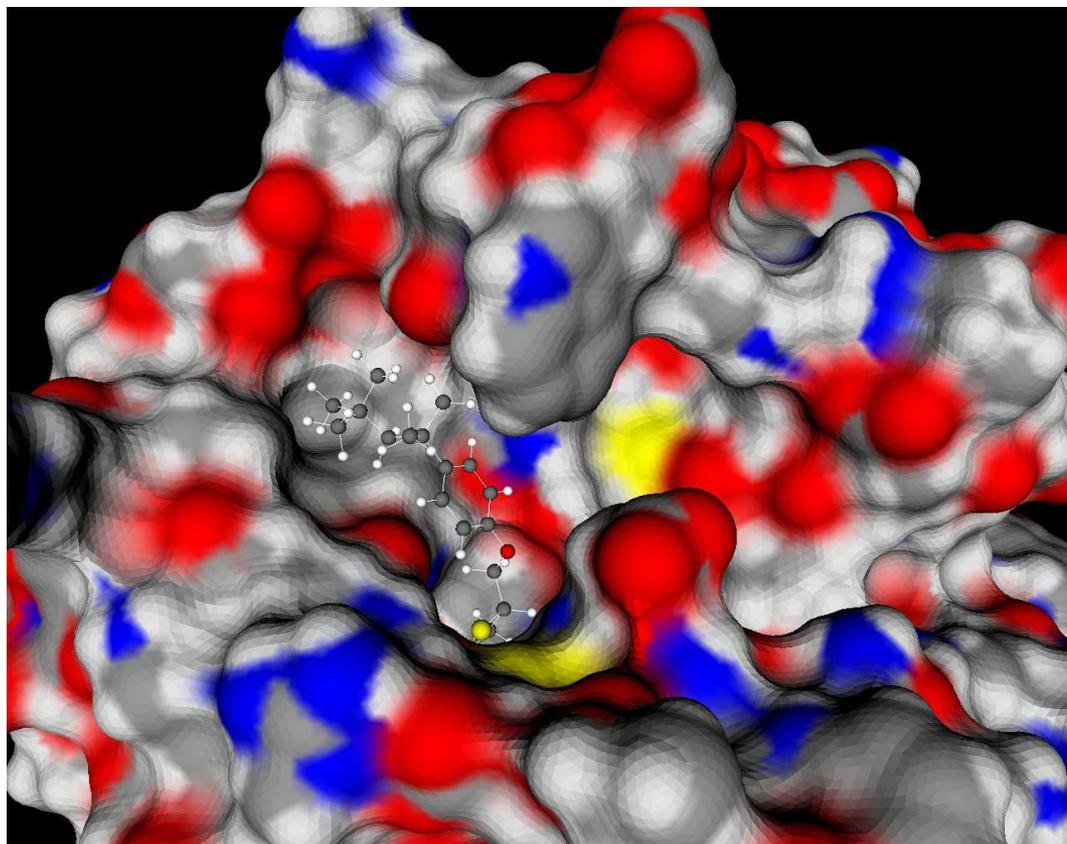
▶ Докинг:

- Positioning of the ligand in the active center of the protein
- Calculation of the binding energy of the protein-ligand ΔG_{bind}
- The higher the binding energy of the protein-ligand, the more likely inhibition in the experiment and the more effective the drug

▶ Accuracy of docking: **positioning**, ΔG_{bind}

Пандемия COVID-19 показала высокую востребованность докинга на начальном этапе поиска: за 4 года опубликованы многие сотни статей по использованию докинга для поиска ингибиторов белков-мишеней SARS-CoV-2

Парадигма докинга: лиганд связывается с белком вблизи **глобального минимума** энергии системы белок-лиганд



Докинг –
нахождение
глобального
минимума
энергии системы
белок-лиганд
на сложной
многомерной
энергетической
поверхности

Суперкомпьютеры стали широко применяться для поиска ингибиторов

- ▶ Поиск ингибиторов с помощью докинга – виртуальный скрининг баз данных органических веществ
- ▶ Виртуальный скрининг: от тысяч до десятков миллиардов молекул
- ▶ Поиск среди миллиардов молекул приводит к открытию **нескольких слабых ингибиторов** с экспериментальным подтверждением

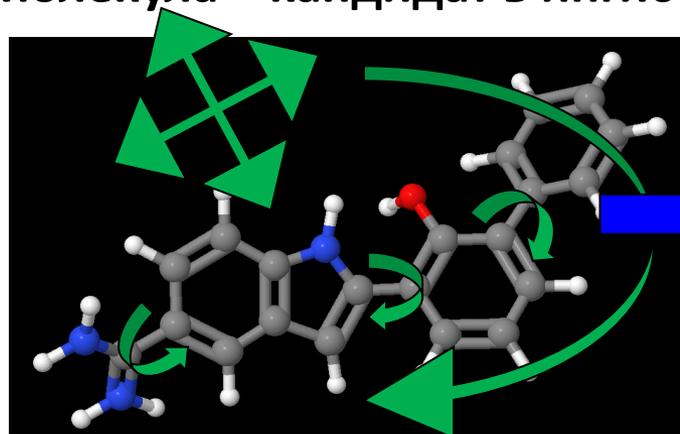
Причина низкой эффективности виртуального скрининга: низкая точность используемых программ докинга

Программа SOL – докинг 1 млн молекул

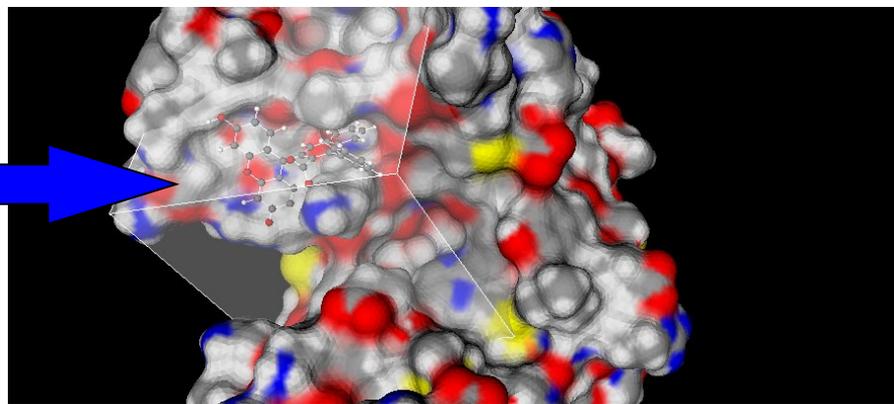
- ▶ Силовое поле MMFF94
- ▶ Предварительно рассчитанная сетка потенциалов для пробных атомов в поле белка (жесткий белок)
- ▶ Генетический алгоритм глобальной оптимизации
- ▶ Время расчетов на 1 CPU от 1 до 2 часов

Chinese National Compound Library \approx 1 000 000 organic molecules

Молекула – кандидат в ингибиторы



Активный центр белка-мишени



Для докинга 1 млн молекул необходимо 116 000 CPU*часов

Эффективность нахождения ингибиторов

- ▶ AutoDock Vina (1 лиганд < 1 минуты): несколько ингибиторов найдены в базах данных, содержащих более 1 миллиарда молекул
- ▶ Наша SOL (1 лиганд – 2 часа): ингибиторы найдены среди 20 000 молекул базы данных Воронежского ГУ: тромбина, факторов свертывания крови Xa, XIa и XIIa, урокиназы, главной протеазы коронавируса SARS-CoV-2
- ▶ Увеличение производительности суперкомпьютера МГУ в 1000 раз позволит с помощью SOL находить значительно больше ингибиторов в базах данных, содержащих многие миллионы и миллиарды молекул
- ▶ Существенно повысится эффективность применения докинга на начальном этапе разработки лекарств

Программа докинга SOL-P

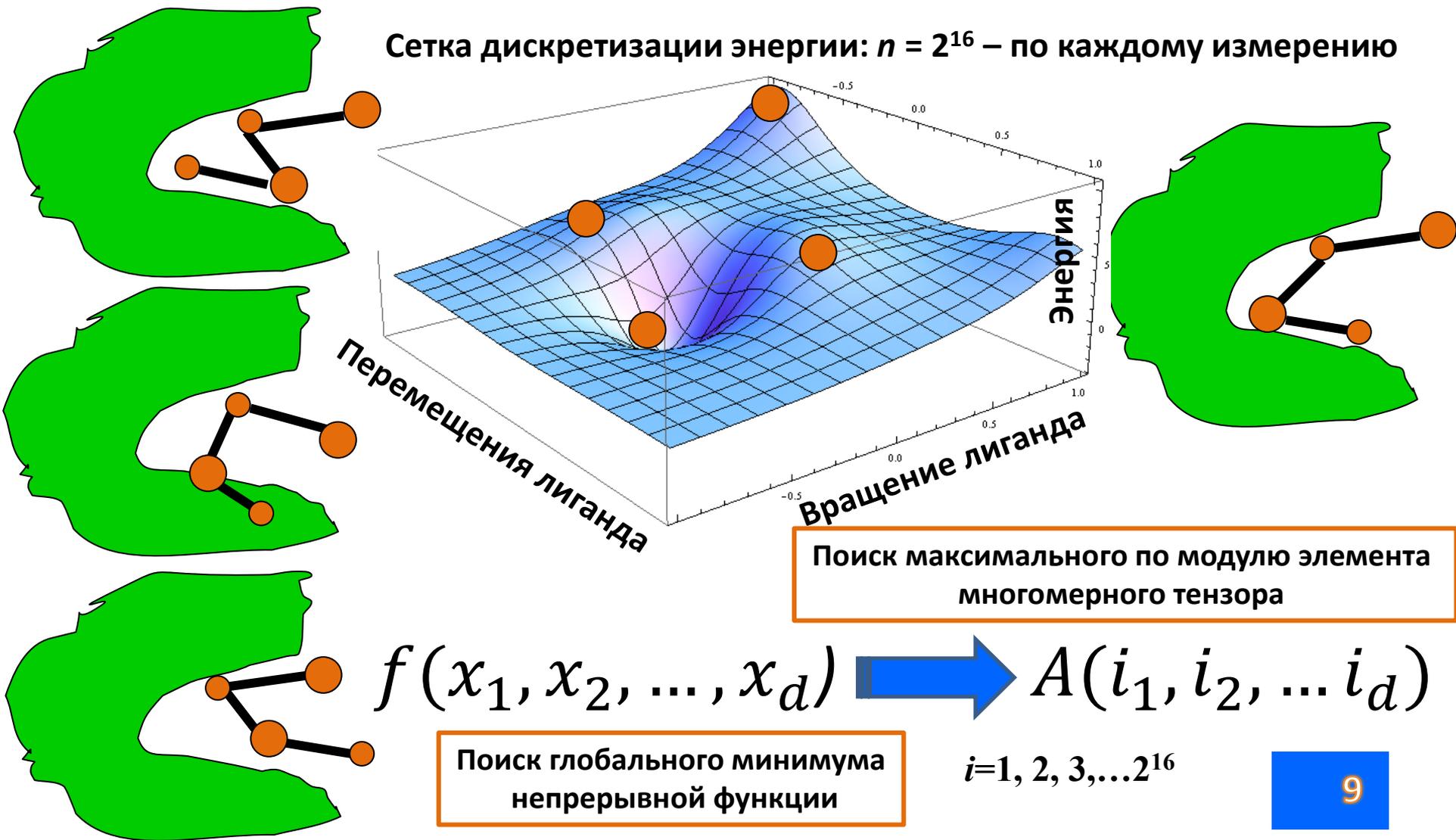
Новое поколение программ докинга – минимальное количество упрощений для повышения точности

- ▶ Классическое силовое поле MMFF94, без упрощений, без подгоночных параметров
- ▶ Нет сетки потенциалов взаимодействия белок-лиганд
- ▶ Многопроцессорные вычисления: несколько сотен вычислительных ядер на докинг 1 молекулы
- ▶ ТТ-алгоритм докинга: Метод глобальной оптимизации с помощью разложение тензоров в тензорные поезда – Tensor Trains или просто ТТ, разработан в Институте вычислительной математики РАН под руководством академика Е.Е. Тыртышникова
- ▶ Возможен докинг очень гибких лигандов и докинг с учетом подвижности атомов белка

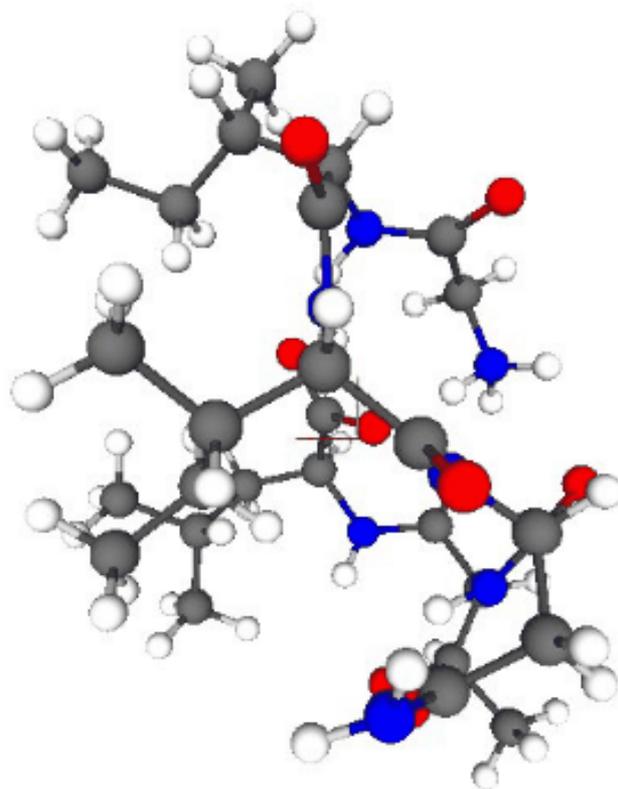
Преобразование непрерывной функции энергии в многомерный тензор

Число измерений = Число степеней свободы системы

Сетка дискретизации энергии: $n = 2^{16}$ – по каждому измерению

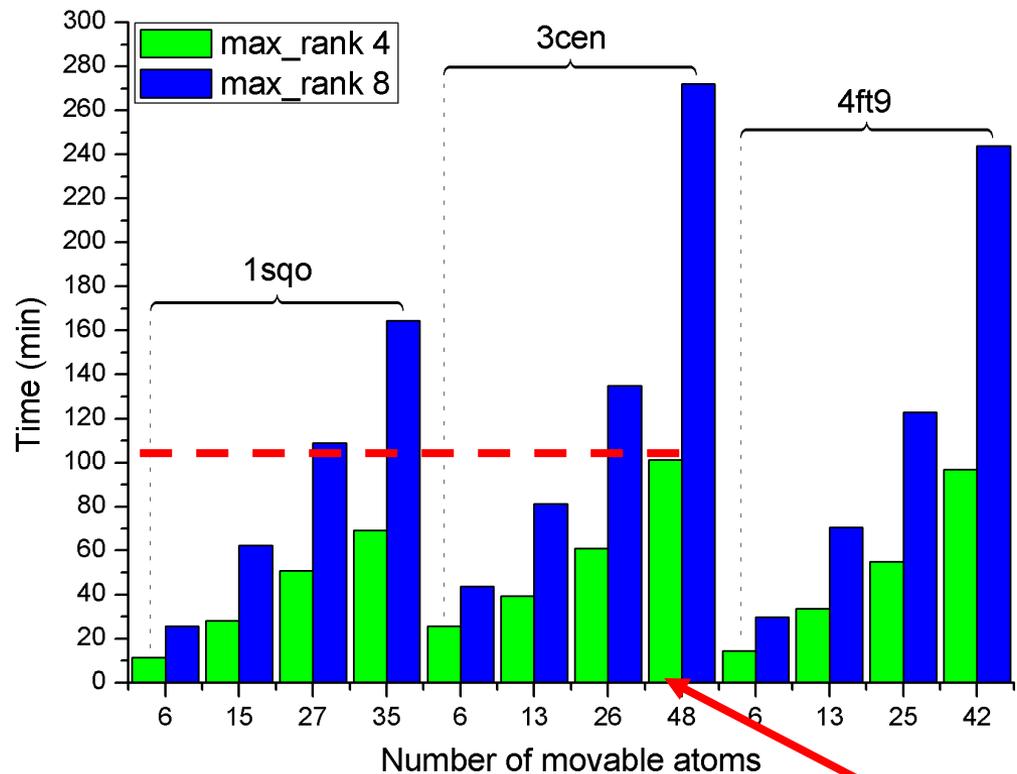


SOL-P: Гибкий олигопептид 25 внутренних вращательных степеней свободы



ТТ-алгоритм докинга существенно превосходит по своей эффективности докинг, основанный на генетическом алгоритме глобальной оптимизации

SOL-P: времена докинга 1 лиганда на 256 cores



SOL-P точнее SOL,
так как нет многих упрощений, используемых в SOL. Сейчас SOL-P используется для докинга только отдельных молекул

Докинг гибкого лиганда с 7 торсионными углами в белок с 48 подвижными атомами

Увеличение ресурсов суперкомпьютера МГУ на 3 порядка позволит проводить виртуальный скрининг тысяч молекул с помощью SOL-P

Путь к квантовому докинг

- ▶ **Квантовый квазидокнинг.** Гипотеза: Рельеф энергетической поверхности белок-лиганд (положения минимумов и барьеров между ними) слабо зависит от способа вычисления энергии
- ▶ 1-ый шаг квазидокнинга: нахождение широкого спектра низкоэнергетических минимумов (несколько тысяч) системы белок-лиганд в силовом поле MMFF94: около 10 000 CPU*часов на 1 лиганд
- ▶ 2-ой шаг: пересчет энергий этих минимумов квантово-химическим методом PM7 с континуальной моделью COSMO и определение глобального минимума энергии; свободная программа MOPAC
- ▶ Разработана программа FLM: находит спектр низкоэнергетических минимумов в силовом поле MMFF94
- ▶ Увеличение производительности суперкомпьютера в 1000 раз позволит проводить квантовый квазидокнинг тысяч лигандов
- ▶ Повысится точность докинга и увеличится эффективность разработки лекарств

Квантовый докинг

- ▶ **Квантовый докинг: энергия системы белок-лиганд рассчитывается с помощью квантовой механики (квантовой химии). Есть все составляющие для осуществления квантового докинга:**
- ▶ **Свободная программа MOPAC, в которой реализован квантово-химический полуэмпирический метод PM7 с учетом растворителя в континуальной модели COSMO. Высокая точность описания межмолекулярных взаимодействий: водородных связей, кулоновских и дисперсионных взаимодействий**
- ▶ **В модуле MOZYME программы MOPAC реализован метод локализованных молекулярных орбиталей, позволяющий проводить расчеты молекулярных систем, содержащих многие тысячи атомов, в том числе комплексов белок-лиганд**
- ▶ **Разработан и апробирован новый алгоритм докинга, основанный на методе глобальной оптимизации, использующем разложение тензоров в тензорные поезда – Tensor Trains**
- ▶ **Увеличение производительности суперкомпьютера в 1000 раз позволит проводить квантовый докинг тысяч лигандов**

Докинг в созвездии наук помогает разрабатывать новые лекарства

- ▶ Использование при Докинге последних достижений Математики, Физики, Химии, Молекулярной Биологии, Суперкомпьютеров
- ▶ Математика – алгоритмы глобальной оптимизации на сложной многомерной энергетической поверхности
- ▶ Физика, Квантовая механика – Аккуратное описание межмолекулярных взаимодействий, **переход к квантовому докингу**
- ▶ Химия – синтез новых ингибиторов, создание баз данных имеющихся в наличии и синтезируемых лекарственно подобных соединений
- ▶ Молекулярная биология – выбор терапевтических мишеней для лекарств, атомистические модели мишеней
- ▶ Суперкомпьютеры – скрининг с помощью докинга многих миллионов молекул – поиск ингибиторов; применение методов **молекулярной динамики** для скрининга

Спасибо за внимание



... Каждое лекарство есть инновация; а кто не хочет применять новые средства, должен ждать новых бед ...

**Francis Bacon (1561-1626)
OF INNOVATIONS**

**Работа поддержана грантом РФФ,
Соглашение №21-71-20031**