#### ООО «Димонта» Москва,

Научно-Исследовательский Вычислительный Центр МГУ имени М.В.Ломоносова

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

Квази-докинг с различными силовыми полями и методами квантовой химии







# Применение молекулярного моделирования улучшает эффективность рациональной разработки новых лекарств

Возможно ли использовать ДОКИНГ для точных вычислений энергии связывания белок-лиганд  $\Delta G_{bind}$ ?

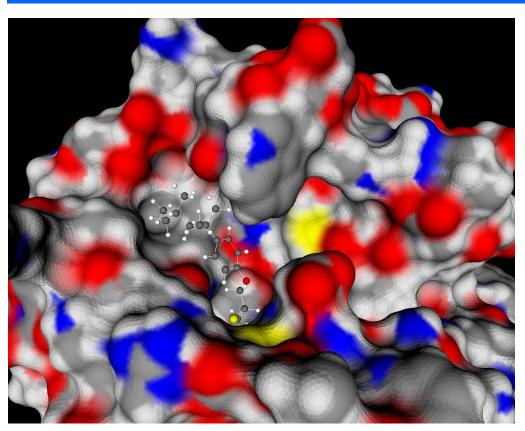
### **Д**окинг:

позиционирование лиганда в белке-мишени

- позиционь.

— расчет  $\Delta G_{bind}$ .

## ПАРАДИГМА ДОКИНГА: лиганд связывается в активном центре белка мишени вблизи глобального минимума энергии системы белок-лиганд



Докинг: нахождение спектра низкоэнергетических минимумов системы белок-лиганд

### Проблемы докинга

- Точность докинга при позиционировании лиганда неплохая, но есть много примеров, когда лучшая поза лиганда, найденная при докинге, сильно отличается от закристаллизованного положения лиганда в белке от нативной позы (RMSD > 2 Å)
- ightharpoonup Точность расчета  $\Delta G_{bind}$  плохая
- Для эффективного применения докинга точность расчетов  $\Delta G_{bind}$  должна быть лучше 1 ккал/моль
- Невозможно оптимизировать ингибиторы на основе докинга: различать слабые, средние и сильные ингибиторы по результатам докинга
- Во многих программах докинга используются подгоночные параметры – невозможно оценить точность а priory

### Тепловое движение лиганда в активном центре белка-мишени

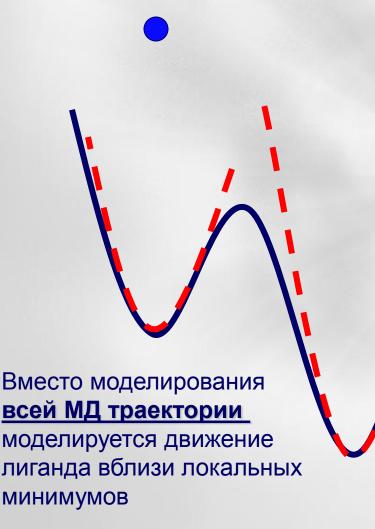


Свободная нергия Гиббса

$$G = -kT \ln Z$$

$$Z = \frac{1}{(2\pi\hbar)^{3n}} \int e^{-(U+W)/kT} dx_1 ... dx_{3n} dp_1 ... dp_{3n}$$

#### Конфигурационный интеграл



Энергетическая поверхность комплекса белок-лиганд представляется набором гармонических ям

### Высокая точность докинга определяется:

- Полнотой найденного спектра низкоэнергетических минимумов системы белок-лиганд (а также белка и лиганда по-отдельности)
- Высокой точностью позиционирования низкоэнергетические минимумы должны быть правильно локализованы в пространстве вблизи нативного положения лиганда в белке
- Высокой точностью вычисления относительных энергий низкоэнергетических минимумов
- Точностью вычисления конфигурационных интегралов в каждом минимуме
- Должна выполняться парадигма докинга

### Точность докинга зависит от:

- Моделирования внутри- и межмолекулярных взаимодействий – выбор силового поля:
- Какое силовое поле и какой метод квантовой химии лучше для докинга?
- Вакуум или Растворитель (вода) модель
- Модели белка-мишени и лигандов
- Алгоритм докинга
- Аппроксимации и метод вычисления свободной энергии

### Программа докинга FLM - Find Local Minima

I.V. Oferkin et al. Advances in Bioinformatics, vol. 2015, Article ID 126858

- FLM не использует предварительно рассчитанную сетку энергий пробных атомов лиганда в поле белка
- Жесткий протеин
- Локальная оптимизация энергии по всем атомам лиганда, начиная из случайных поз лиганда
- Силовое поле ММFF94 с неявной моделью растворителя РСМ
- Поиск спектра низкоэнергетических минимумов
- Получен реперный набор минимумов для тестирования: 16 комплексов белок-лиганд
- Параллельные многопроцессорные вычисления:
   8192 ядер несколько часов на суперкомпьютере Lomonosov FLM может считать сколь угодно пока множество найденных минимумов не перестанет обновляться

# Квази-докинг: метод сравнения различных способов вычисления энергии белок-лиганд для докинга

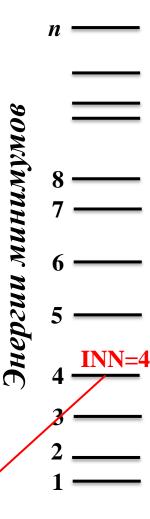
- С помощью программы докинга FLM для силового поля **MMFF94** с растворителем в модели Polarized Continuum Model получены наборы низкоэнергетических минимумов для тестового набора комплексов белок-лиганд: по 8192 минимума для каждого комплекса
- Энергии каждого минимума пересчитаны:
- с помощью силового поля **CHARMM**, силового поля **AMBER**, квантово-химических полуэмпирических методов **PM7** и **PM6**
- Проведён тщательный анализ парадигмы докинга: насколько близко к нативному закристаллизованному положению лиганда находится глобальный или близкие к нему по энергии минимумы
- Выявлены какие методы лучше для докинга

## Номенклатура локальных энергетических минимумов

### Индексы энергетических минимумов комплекса белок-лиганд,

#### индекс INN

- Каждый минимум имеет целочисленный индекс, соответствующий его положению в списке минимумов, отсортированных по возрастанию их энергий. Минимум с самой низкой энергией имеет индекс равный 1.
- INN (Index of Near Native) это индекс минимума с наинизшей энергией, в котором поза лиганда отличается от нативной позы меньше, чем на 2 Å



### Индекс INN

Complex PDB ID	MMFF+SGB	CHARMM +GBSW	AMBER+GB	PM7+COSMO	PM6-D3H4X+COSMO	
4FT0	6	76	1	1	1	
4FT9	1	7	1	1	1	
4FSW	3	1	1	1	1	
4FTA	97	8	7	1	1	
4FV5	5	1	1	1	5	
4FV6	24	1	1	1	3	
1DWC	8	1	8	2	8	
1TOM	1	1	2	1	1	
1C5Y	1	1	1	1	1	
1F5L	1	1	1	2	2	
1O3P	2	393	1	1	6	
1SQO	1	1	1	1	1	
1VJ9	14	1	74	14	61	
1VJA	1	1	40	2	6	
2P94	1	1	1	1	1	
3CEN	1	1	5	1	4	

### Индекс INN: вакуум-растворитель

Complex PDB ID	MMFF vacuum	MMFF PCM	MMFF SGB	CHARMM vacuum	CHARMM GBSW	AMBER vacuum	AMBER GB	PM7 vacuum	PM7 COSMO
4FT0	99	159	6	1	76	219	1	75	1
4FT9	125	1	1	221	7	283	1	22	1
4FSW	102	140	3	38	1	31	1	413	1
4FTA	Inf	187	97	2675	8	289	7	Inf	1
4FV5	134	3	5	14	1	1138	1	279	1
4FV6	289	68	24	1	1	1362	1	11	1
1DWC	114	35	8	102	1	141	8	106	2
1TOM	Inf	4	1	4010	1	245	2	877	1
1C5Y	1	1	1	1	1	1	1	1	1
1F5L	1	1	1	1	1	29	1	21	2
103P	62	1	2	20	393	298	1	2	1
1SQO	1	1	1	1	1	4	1	1	1
1VJ9	1	18	14	1	1	4838	74	17	14
1VJA	49	2	1	1	1	5798	40	6	2
2P94	1	1	1	1	1	204	1	7	1
3CEN	1	1	1	10	1	990	5	1	1

#### Выводы

- Парадигма докинга выполняется для некоторых тестовых комплексов при расчете энергии в вакууме с помощью силовых полей MMFF94, CHARMM, AMBER и с помощью квантово-химических полуэмпирических методов РМ7, РМ6-D3H4X
- Учет растворителя в соответствующих неявных моделях для всех этих методов расчета энергии существенно увеличивает процент комплексов, для которых парадигма докинга выполняется
- С растворителем силовое поле CHARMM лучше силовых полей MMFF94 и AMBER.
- С растворителем квантово-химический полуэмпирический метод РМ7 лучше всех трех рассмотренных силовых полей ММFF94, СНАRMM и AMBER
- Полуэмпирический метод РМ6-D3H4X заметно хуже метода РМ7
- Для каждого метода расчета энергии встречаются комплексы, для которых парадигма докинга не выполняется надо учитывать подвижность атомов белка

### Учет подвижности атомов белка

- Разработан новый алгоритм, ТТ-докинг, позволяющий находить спектр низкоэнергетических минимумов в пространстве с большим числом измерений числом степеней свободы системы белоклиганд: показано для 157 степеней свободы (генетический алгоритм может это делать при 20-25 степеней свободы)
- ТТ-докинг основан на новых методах матричного анализа:
   разложения многомерных массивов в тензорные поезда и ТТ-Cross
   метода аппроксимации больших тензоров с помощью небольшого числа их элементов
- TT-докинг реализован в программе **SOL-P**: докинг гибких лигандов в белки с подвижными атомами до нескольких десятков подвижных атомов белка
- Учет подвижности атомов белка улучшает точность позиционирования лиганда: парадигма докинга выполняется для тех комплексов, для которых в модели жесткого белка она не выполняется

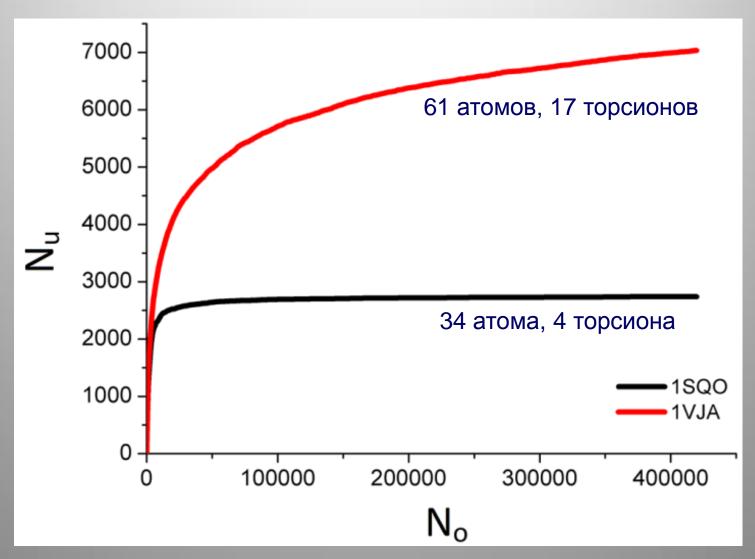
### Спасибо за внимание



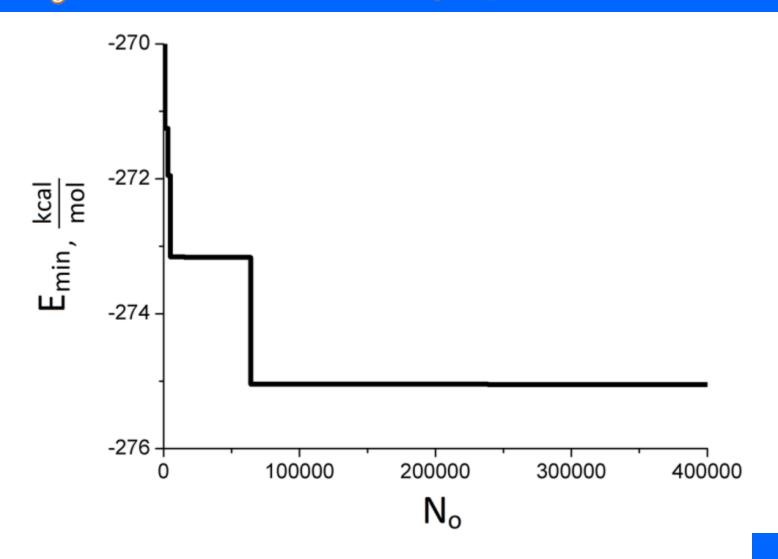
... Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils ...

Francis Bacon (1561-1626) OF INNOVATIONS

Работа поддержана грантом РНФ, Соглашение № 15-11-00025 Число обновлений локальных минимумов (*Nu*) в зависимости от числа сделанных оптимизаций (*No*) для комплексов 1SQO (black line) и 1VJA (red line)



### $E_{\min}$ — глобальный минимум энергии $N_o$ — число оптимизаций, комплекс 1VJA



Большинство программ докинга используют сетку потенциалов вместо белка-мишени — предварительно рассчитанные энергии пробных атомов лиганда в узлах 3х-мерной сетки в активном центре белка-мишени

### В рамках такой сеточной аппроксимации невозможно:

- Выполнить локальную оптимизацию по положениям атомов лиганда
- Учесть подвижность атомов белка в процессе докинга
- Использовать континуальные модели растворителя из-за нелокального характера взаимодействий поляризационных зарядов на поверхности растворителя с атомными зарядами