

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тольяттинский государственный университет»

ИНСТИТУТ ХИМИИ И ИНЖЕНЕРНОЙ ЭКОЛОГИИ
(наименование института полностью)
Кафедра Химия, химические процессы и технологии
(наименование кафедры)
04.03.01 Химия
(код и наименование направления подготовки, специальности)
Органическая и биоорганическая химия
(направленность(профиль)/ специализация)

БАКАЛАВРСКАЯ РАБОТА

на тему: Индуцированный докинг производных 5-цианопиримидина в
активном сайте EGFR киназы

Студент	<u>Д.С. Азизмамадова</u> (И.О. Фамилия)	_____ (личная подпись)
Руководитель	<u>А.С. Бунев</u> (И.О. Фамилия)	_____ (личная подпись)
Консультанты	<u>О.А. Головач</u> (И.О. Фамилия)	_____ (личная подпись)

Допустить к защите

Заведующий кафедрой д.х.н., профессор Г.И. Остапенко
(ученая степень, звание, И.О. Фамилия) _____
(личная подпись)

« _____ » _____ 2019 г.

Тольятти 2019

АННОТАЦИЯ

Работа изложена на 50 страницах, включает в себе 3 таблицы. Объектами исследования в представленной работе являются производные 5-цианопиримидина и их биологическая активность в качестве ингибиторов EGFR-киназы.

В работе проведен обзор современных способов моделирования и докинга молекул относительно активных сайтов связывания белковых молекул. В частности рассмотрен индуцированный докинг, а также проведен расчет ряда производных 5-цианопиримидина в активном сайте EGFR. Оценена энергия взаимодействия и скоринг для докинга данных производных, выделены наиболее эффективные потенциальные ингибиторы, отмечена природа основных взаимодействий молекул с сайтом связывания.

Выпускная квалификационная работа может быть использована для синтеза соединений имеющих наибольший скоринг и дальнейшей оценки их биологической активности, в частности цитотоксичности против отдельных линий раковых клеток.

ABSTRACT

The title of the graduation project is « Induced docking of derivatives of 5-cyanopyrimidine in the active site of EGFR kinases».

The aim of the work is to give some information about study of features of interactions of 5-cyanopyrimidine derivatives in the EGFR active site within the induced docking and selection of potential leaders.

The object of the graduation project is 5-cyanopyrimidine.

The subject of graduation project is indicated by docking.

The goal of this graduation project is to investigate 5-cyanopyrimidine and their biological activity as inhibitors of EGFR kinases. The graduation project is divided into three chapters, each of which includes experimental data, theoretical review and results and their discussion.

We first discuss history of EGFR kinases inhibitors synthesis and improvement of this process in our days. We then analyze methods of docking, what the point of docking, and for which processes it get the best fit. We look at the various variant of docking, pick pros and cons of each ones, after this choose the best one - induced docking.

Next we study different molecules and their activity against proteins of cancer cells, especially EGFR. The results show that 5-cyanopyrimidine is the best fit as EGFR kinases inhibitors. Also we had analysed the role of bonding with MET793 aminoacid as most important bond between target and molecule.

Finally, we present work on the derivatives of 5- as inhibitors of EGFR kinases by modern methods of modeling and docking of molecules relative to the active binding sites of protein molecules is carried out. The interaction energy and scoring for docking of these derivatives have been estimated, the most effective potential inhibitors have been identified, the nature of the main interactions of molecules with the binding site has been noted.

The graduation project consists of an explanatory note on 50 pages, figures, 3 tables.

СОДЕРЖАНИЕ

Стр.

ВВЕДЕНИЕ	6
1. ЛИТЕРАТУРНЫЙ ОБЗОР	7
1.1.Эффективная генерация EPS с использованием моделей эластичных сетей (ENM)	11
1.2 Генерация EPS с использованием концепции лиганд-модель.....	11
1.3 Выбор «оптимального» EPS для докинга	12
1.4 Выбор EPS с использованием экспериментальных данных лиганда.....	13
1.5 Деформируемые сетки.....	13
1.6 Методы КМ и КМ / ММ для улучшения процесса подсчета очков в докинге. ..	14
1.7 Методы, основанные на КМ / ММ.	18
1.8 Методы линейного масштабирования и фрагментации.....	20
2. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ	23
3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ	44
3.1 Подготовка комбинаторной библиотеки	44
3.2 Подготовка молекулы белка	44
3.3 Молекулярный докинг.....	44
3.4 Индуцированный докинг.....	45
ЗАКЛЮЧЕНИЕ.....	47
СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ	48

ПРИНЯТЫЕ СОКРАЩЕНИЯ

В настоящей работе применяют следующие термины с соответствующими определениями:

EGFR–Рецептор эпидермального фактора роста

ErbB– Семейство тирозинкиназных рецепторов

Erk1/2–Регулируемая внеклеточными сигналами киназа 1/2

MAP–Митоген активируемая протеникиназа

PK–Фармококинетика

TIE2–Тирозинкиназа с иммуноглобулинподобными и EGFR- подобными доменами

VEGER–Фактор роста сосудистого эндотелия

PI3K–Фосфоинозитол- 3- киназа

s-Raf1–Протоонкогенсерин/ треонинпротеинкиназы

HВ4a.c5.2, HВ4a.c4.2, HВ4a– Клеточные линии

IC50–Концентрация 50% ингибирования

CDK–Циклин зависимаякиназа

NMP– Немелкоклеточный рак легкого

EPS–ансамбль белковых структур

QM– квантовой механики

MM– молекулярная механика моделей эластичных сетей (ENM)

ВВЕДЕНИЕ

Одной из важнейших задач современной медицинской химии является поиск низкомолекулярных органических соединений, несущих положительный фармакологический эффект. Ярким примером изыскания подобного рода веществ является поиск высокоэффективных ингибиторов протеинкиназ, эффективных агентов для лечения EGFR-мутантного немелкоклеточного рака лёгкого (НМРЛ). По прошествии определённого периода времени ингибиторы протеинкиназ теряют свою активность вследствие возникновения резистентности.

Поэтому возникает необходимость поиска новых ингибиторов EGFR-киназ путём осуществления молекулярного моделирования и последующего молекулярного докинга с киназой.

Производные 2-алкилсульфонил-4-амино-6-арил-5-цианопиримидина обладают высокой биологической активностью, поэтому они были выбраны для молекулярного докинга.

Цель работы:

Установить основные структурные закономерности, возникающие между базовым скаффолдом и структурой белка, с учетом конформационной подвижности для конструирования на основании полученных данных высокоаффинных лигандов.

Задачи:

1. Построение виртуальной комбинаторной библиотеки структур производных 2-алкилсульфонил-4-амино-6-арил-5-цианопиримидина;
2. Проведение индуцированного молекулярного докинга;
3. Анализ полученных результатов молекулярного докинга и оценка относительной аффинности лигандов методом MM/GBSA.

1. ЛИТЕРАТУРНЫЙ ОБЗОР

В последние несколько десятков лет докинг стал стандартной технологией для поиска лекарственных веществ путем виртуальной проверки библиотек существующих или гипотетических химических веществ с целью выявления новых активных хемотипов, которые связываются с целевыми белками, а также для предсказания способов связывания и сходства этих лигандов. Хотя методы докинга в целом достаточно эффективные инструменты для данных типов исследований, качество и надежность прогнозируемых взаимодействия в настоящее время ограничены рядом методологических недостатков, таких как трудности в моделирование гибкости белка при связывании лиганда [1–7] и количественное определение энергии связывания белок-лиганд с использованием упрощенных скоринговых функций [8,9]. В рамках литературного обзора рассмотрены методологические разработки, направленные на устранение этих недостатков за последние года, с особым акцентом на методы улучшения скоринга на основе квантовой механики (КМ) и смешанный метод квантовая механика/молекулярная механика (КМ/ММ).

Вычислительные концепции для моделирования гибкости белка во время докинга могут быть разделены на две категории (рис. 1) [10]. В первом классе методов, конформационный поиск лиганда и белка выполняется в тандеме, где обе конформации изменяются одновременно. Для ограничения количества комбинаторных конформаций, которые необходимо генерировать и энергетически оценивать, выбирается небольшое количество степеней свободы, которые описывают важный вклад в гибкость белковой структуры. Переменные, которые характеризуют величину конформационного изменения по выбранным степеням свободы белка оптимизируется параллельно со степенями свободы лиганда, такие как поступательная, вращательная и крутильные изменения. Второй класс методов генерирует ансамбль белковых структур (EPS) перед докингом. Каждый лиганд библиотеки впоследствии присоединяется к каждому члену EPS. Недавние исследования сделали значительный акцент на тестирование различных методов генерации EPS и выбор «оптимального» подмножества белковых структур из EPS, которые будут использоваться для прогнозирования положения связывания и виртуального скрининга (Таблица 1).

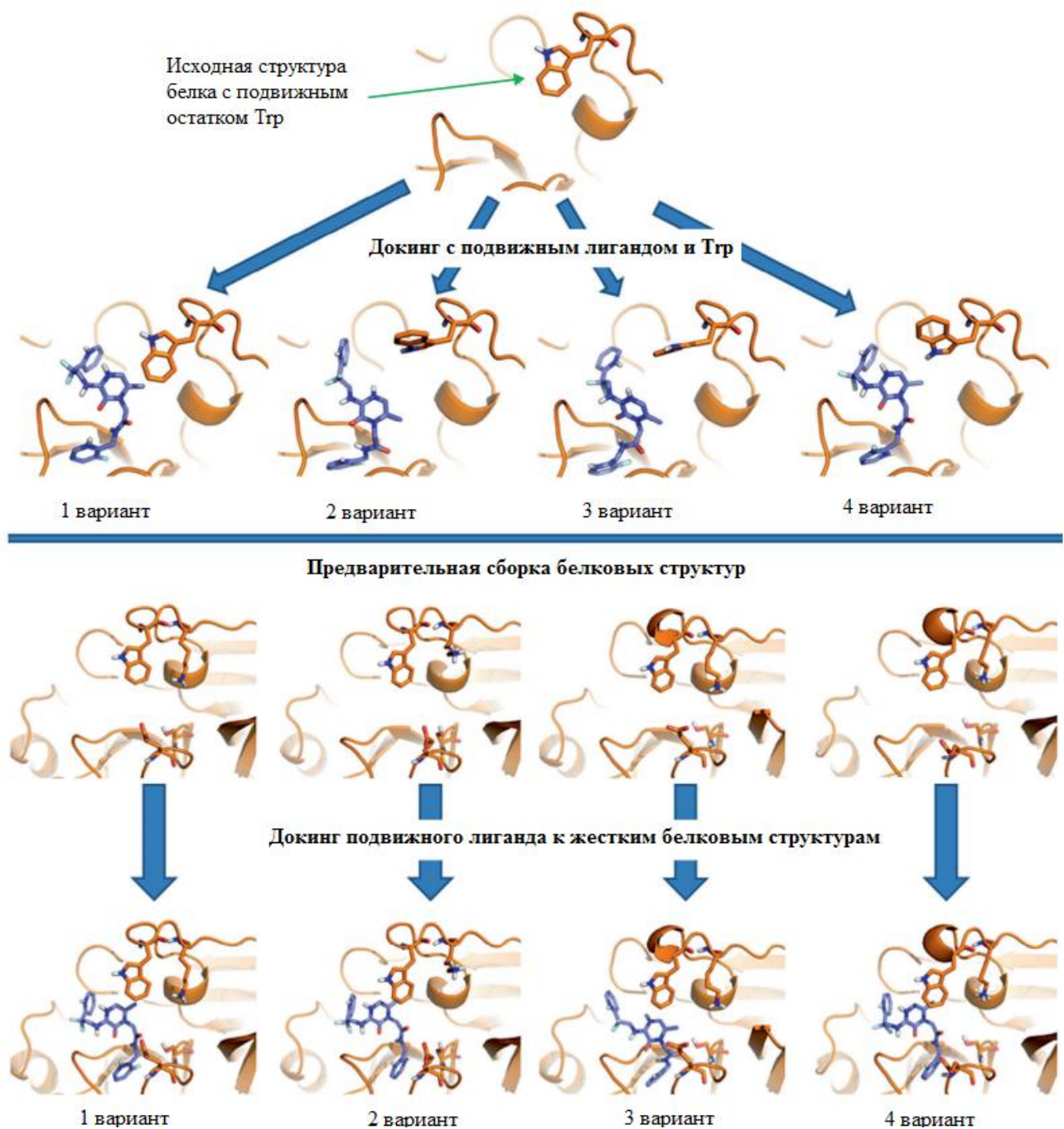


Рисунок 1. Обзор методик, которые включающие подвижность белка в симуляции докинга. Вверху: дополнительные выбранные степени свободы, описывающие подвижность белка (в данном примере: различные конформация боковой цепи Trp), изменяются в tandemе с лигандом, образуя различные варианты. Внизу: докинга выполняется с предварительно сформированным ансамблем белковых структур. Каждая структура белка обычно остается жесткой при докинге для повышения эффективности.

Таблица 1

Название конкретного типа технологии	Концепция лиганд-модель	ENM для генерации EPS	Выбор рентгеновских структур в качестве EPS	Отбор суб-EPS с использованием экспериментальных данных лиганда	Упругая потенциальная сетка	Интерполяция между сеточными представлениями различных белковых структур
1	2	3	4	5	6	7
Pros	Эффективная генерация белковых конформаций, которые индуцируются связыванием лиганда.	Эффективная генерация альтернативных белковых структур.	Улучшенные результаты докинга по сравнению с одноструктурным докингом. Минимизация рисков из-за низкой производительности виртуального скрининга.	Простая идентификация оптимального суб-EPS для виртуального скрининга. Позволяет эффективный виртуальный скрининг.	Эффективный алгоритм для включения гибкости белка в докинге на основе сетки. Требуется только одна начальная структура цели.	Непрерывное представление конформационных изменений. Позволяет изменять физико-химические свойства.

Продолжение таблицы 1

1	2	3	4	5	6	7
Cons	В настоящее время только небольшие изменения конформации.	Лишь несколько нормальных мод недостаточно для описания конформационных изменений, связанных со связыванием лиганда. Общие правила о том, сколько и какие режимы необходимо учитывать, неясны.	Лучший выбор рентгеновских структур, неизвестных до виртуального скрининга.	Требуются экспериментальные данные по нескольким активным лигандам.	Не позволяет моделировать изменения физико-химических свойств из-за конформационных изменений.	Требуется первоначальный отбор представителя EPS. В настоящее время предназначен только для белковых структур, разделенных одной конформационной переменной/модой.
Ссылка	[12]	[11]	[14,15]	[18]	[19]	[20,21]

1.1. Эффективная генерация EPS с использованием моделей эластичных сетей (ENM)

Самый простой подход к созданию EPS заключается в использовании набора рентгеновских структур одной и той же белковой системы, но разных методов молекулярного моделирования, таких как молекулярное динамическое (MD) моделирование, основанный на MD основной компонентный анализ (PCA) или анализ в нормальном режиме (NMA) также могут быть использованы для создания EPS. По сравнению с трудоемкими полностью атомистическими методами (например, методы на основе MD), использующие нормальные моды из моделей упругих сетей (ENM) являются эффективной альтернативой для создания EPS.

В недавнем исследовании Dietzen et al. [11] исследовали способность NMA в ENM прогнозировать конформационные изменения, связанные со связыванием малых молекул с различными белковыми молекулами и можно ли использовать этот метод для создания EPS для ансамбля докинга. Авторы использовали первые нормальные моды из апо-структуры для создания EPS который оптимально воспроизводил следовые количества Са от лиганд-связанных белковых структур. Используя это ограниченное количество нормальных режимов, авторы не смогли точно воспроизвести наблюдаемое конформационное изменение при связывании лиганда. Авторы также отметили, что в пределах одного и того же целевого белка амплитуды для каждого случая варьируются в зависимости от связываемого лиганда и не смогли определить общие правила о том, какие режимы были необходимы для успешного докинга. Однако они смогли показать, что результаты докинга могут быть улучшены, когда генерируемые EPS в нормальном режиме сочетаются с явной гибкостью боковой цепи на протяжении всего процесса. Следует отметить, что эта дополнительная гибкость белка значительно увеличивает необходимое время на компьютерное моделирование.

1.2 Генерация EPS с использованием концепции лиганд-модель

Альтернативная схема для генерации конформаций белка релевантной для докинга структурно разнообразных лигандов, известных как концепция Лимока модели-лиганда была разработана [12]. В методе Лимок MD симуляция выполняется с динамически изменяющимся набором ограниченного функционала группы в сайте связывания белка, по существу, представляющие большой гипотетический ансамбль различных химических видов, связывающиеся с одним и тем же целевым белком. Начиная с апо-структура, лиганд-модель подход был использован для получения EPS, используемый для докинга, и результаты превзошли докинг с EPS, сгенерированным из моделирования апо-MD. Кроме

того, метод был лишь немного менее успешным, чем докинг с экспериментально известной голографической формой каждого индивидуального белково-лигандного комплекса, демонстрирующий потенциал для создания голо-подобных белковых структур.

1.3 Выбор «оптимального» EPS для докинга

В исследованиях ансамбля докинга [13], было признано, что использования EPS только нескольких конформаций белка может увеличить вероятность успеха для правильного прогнозирования в виртуальном скрининге. Предполагалось, что с помощью очень большого EPS снижается производительность докинга из-за большого количества ложных срабатываний, генерируемых ненативных белковых структур, что, в свою очередь, снижает скорость скрининга.

Два недавних крупномасштабных исследования [14,15] описывают использование набора рентгеновских структур в качестве EPS, чтобы выяснить, в какой степени докинг EPS улучшает успех докинга по сравнению с докингом в единой структуре белка. В литературе описан [14] опыт выполнения ретроспективного виртуального скрининга на 36 фармацевтически релевантных белковых мишенях с использованием док-механизма ICM. Для подражания реальному применению виртуального скрининга докинга, где только соединения самого высокого ранга будут синтезированы и биологически протестированы, авторы сосредоточились на успешной идентификации активных веществ в самой высокой ранжированной фракции. Они проверили два разных протокола для ранжирования соединений в ансамбле докинга, ранжирования на основе лучших результатов среди всех белковых структур и на основе общего лучшего ранга среди разных моделирований докинга с одной структурой. Обе схемы ранжирования статистически превосходят среднюю структуру одного белка. Интересно, что докинг ансамбля позволил отметить не только обогащение активов в верхней фракции, но и химическое разнообразие выявленных активных скаффолдов. В соответствующих исследованиях [15] изучено влияние размера ансамбля на прогнозирование поз и качеству виртуального скрининга для восьми белковых мишеней с использованием GOLD с тремя различными скоринговыми функциями. Аналогично исследованию [14], в среднем докинг ансамбль работает лучше, чем одиночная структура. Увеличение производительности более важно для правильного прогнозирования нативных позиций лиганда по сравнению с виртуальным скринингом и сильно зависит от белковых мишеней и используемой балльной функцией. Также признано, что для данного размера ансамбля состав EPS сильно влияет на успех докинга. В этом контексте лучший одно структурный докинг регулярно превосходит лучшие результаты ансамбля. Однако самое важное, что докинг с

EPS выполняется почти всегда лучше, чем худшая структура одного белка; следовательно, докинг ансамбля сводит к минимуму риск плохой производительности виртуального скрининга, так как лучшие структуры белка для докинга, как правило, не известны до моделирования.

Важно подчеркнуть другую важную проблему, связанную с объединением докинга: какие структуры из EPS наиболее вероятно, принесет максимальный успех процесса? Разные стратегии, такие как кластеризация подобных структур [16] или выбор конформации белка, которые сохраняют относительные ориентации ядра аминокислот, важных для связывания [17], были разработаны, чтобы урезать EPS до наиболее соответствующих структур.

1.4 Выбор EPS с использованием экспериментальных данных лиганда

Сюй и соавт. [18] проверили две новые схемы выбора структуры белка для докинга ансамблей с использованием экспериментальных данных лигандов. В первой схеме были выбраны структуры белка, которые лучше всего воспроизводятся нативным режимом привязки из 1–3 известных белково-лигандных комплексных структур белка-мишени. Докинг ансамбля с этим суб-EPS немного превосходил EPS генерируемых путем кластеризации. Второй и лучший метод отбора заключался в выборе белковых структур, которые наилучшим образом отделяют активные вещества от обманок в предварительном скрининговом исследовании на небольшое количество лигандов. Предварительный скрининг EPS с использованием 10–20 активных лигандов и несколько неактивных лигандов-ловушек, что приводит к созданию вспомогательного EPS, содержащего лишь небольшое количество белковых структур (3–5) с оптимальными результатами докинга на полный набор данных виртуального скрининга.

1.5 Деформируемые сетки

Самым простым способом использования EPS является последовательный докинг всех членов ансамбля, как это было использовано во всех ранее описанных исследованиях. Эта процедура линейно увеличивает необходимое вычислительное время с размером EPS. Чтобы ускорить процесс докинга, в двух недавних исследованиях [19, 20] были изложены основанные на сетке подходы, в которых значения энергии взаимодействия белок-лиганд предварительно рассчитываются для каждой структуры белка на трехмерной сетке. Эти сетки затем либо деформируются, либо используются для быстрого вычисления альтернативных значений энергии взаимодействия из-за конформационных изменений белка, вызванных связанным лигандом. Изменения сетки выполняются параллельно с генерацией позы и не требуют последовательного докинга для

всех членов EPS. Значения энергии взаимодействия лиганда предварительно рассчитываются для каждой структуры белка на трехмерной сетке. Эти сетки затем либо деформируются, либо используются для быстрого вычисления альтернативных значений энергии взаимодействия из-за конформационных изменений белка, вызванных связанным лигандом. Изменения сетки выполняются параллельно с генерацией позы и не требуют последовательного докинга для всех членов EPS.

Leisetal. [19] сгенерировали семь альтернативных конформаций белка для протеинкиназы А вдоль мягкого нормального режима из ENM апо белка. Эти конформации были ранжированы в соответствии с их изменением вдоль нормальной моды и им была присвоена конформационная переменная. Потенциальные сетки, используемые в AutoDock, генерировались для каждой структуры белка, и на протяжении фазы позы конформации белка выбирались с использованием конформационной переменной. Нецелые значения этой переменной (например, 2.4) описывают промежуточные конформации белка и были представлены интерполяцией между потенциальными значениями в точках сетки двух дискретных структур белка, связанных со значениями связанного целого (например, 2 и 3). Такой подход позволяет объединить непрерывное конформационное пространство структуры белка с эффективным расчетом показателей взаимодействия белок-лиганд с использованием предварительно рассчитанных сеток.

Гольке и его коллеги разработали альтернативную концепцию включения гибкости белка в докинг на основе сетки [20,21]. В их подходе начальная потенциальная сетка вычисляется на основе апо-целевой структуры (РНК в их недавнем исследовании), а положения точек сетки изменяются в унисон с соседними атомами РНК. Принимая во внимание, что потенциальная сетка изменяет свою форму с конформационным изменением цели, подобным упругому телу, реагирующему на внешние силы, потенциальные значения остаются неизменными.

1.6 Методы КМ и КМ / ММ для улучшения процесса подсчета очков в докинге.

В дополнение к гибкости белка при связывании лиганда, точное количественное определение ассоциации белок-лиганд является серьезной проблемой для современных схем докинга. Методы квантовой механики (КМ) появились в структуре лекарственного средства на основе структуры, направленной на повышение точности прогнозирования взаимодействия белок-лиганд. Методы контроля качества имеют естественное преимущество в моделировании поляризации и переноса заряда, что является серьезной проблемой при использовании классических схем оценки на основе молекулярной

механики (ММ), которые предполагают фиксированные заряды во время связывания лиганда. Методы КМ, однако, требуют гораздо больше времени, чем стандартные схемы оценки на основе ММ. Чтобы ограничить вычислительную сложность, были разработаны различные схемы для уменьшения части системы белок-лиганд, которая детально рассматривается в КМ (рис. 2). Альтернативно, были разработаны схемы упрощения, такие как локальное приближение, которые позволяют линейное масштабирование компьютерного времени с размером молекулы по сравнению с поведением кубического или более высокого масштабирования стандартных численных методов КМ. Вторая схема позволяет обрабатывать полную белково-лигандную систему в деталях КМ. Последние разработки обеих схем освещены в следующих разделах (таблица 2).

Таблица 2. Сравнение методов, которые включают концепции КМ и КМ/ММ в процесс оценки докинга.

Название конкретного типа технологии	КМ на сайтах связывания лигандов	КМ/ММ	Линейное масштабирование КМ	Фрагментные методы: FMO и MFCC
Pros	Поляризация и перенос заряда могут быть смоделированы	Баланс между точностью и скоростью.	Избегайте использования любых классических параметров силового поля ММ.	Расчет можно легко распараллелить. Подходит для определения ключевых остатков из-за естественного разложения энергии взаимодействия.
Cons	Большая часть белка игнорируется. Дисперсионные эффекты обычно не включены точно.	Только лиганд и несколько остатков обрабатываются с помощью КМ. С границей КМ/ММ	Большинство исследований на полуэмпирическом уровне с низкой точностью. Включение динамических эффектов приведет к	Вычислительно интенсивно. Большой базовый набор не подходит для FMO.

через чрезвычайно
ковалентн высоким
ую связь вычислительным
трудно затратам.
иметь
дело

Ссылка [22] [23,24,28] [29,30,40] [33,34]

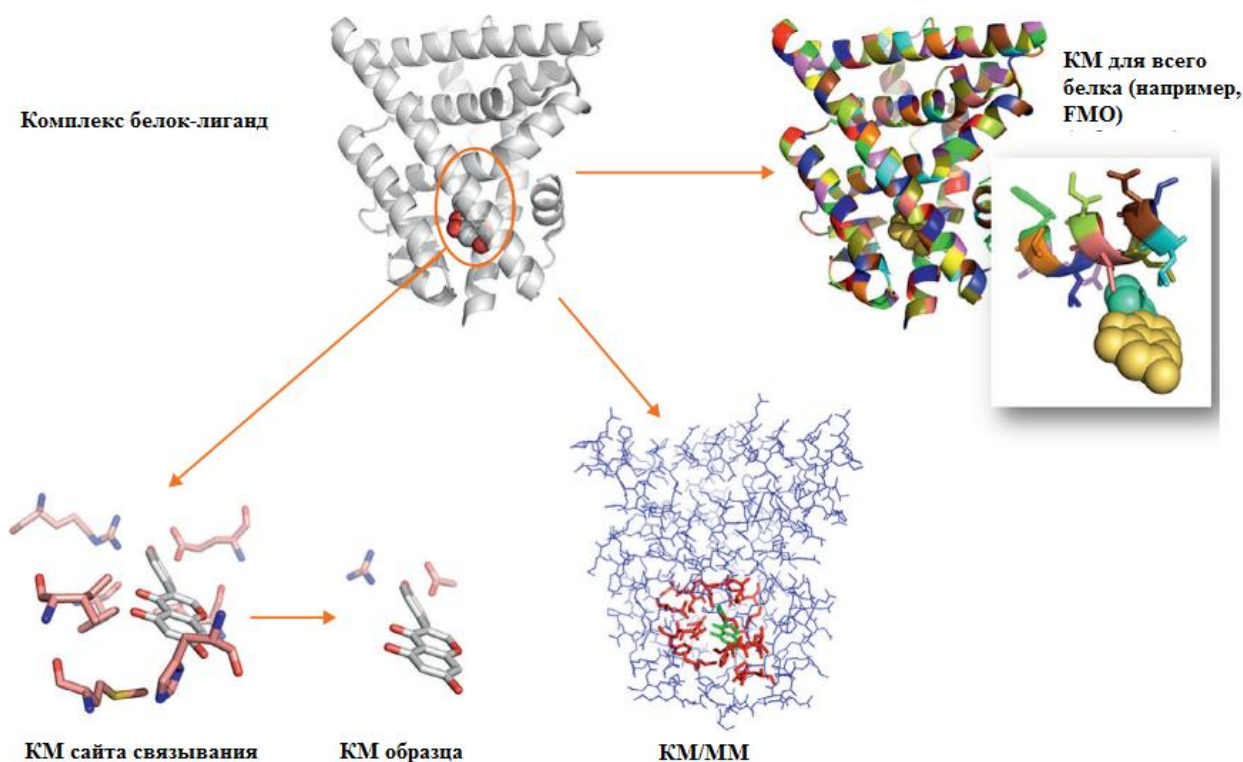


Рисунок 2. Схема методов обработки белково-лигандного комплекса в деталях КМ. Расчеты КМ могут быть выполнены для лиганда и нескольких остатков сайта связывания только (нижний левый рисунок). Аминокислоты могут даже быть уменьшены до небольших молекулярных образцов для повышения вычислительной эффективности. Методы КМ / ММ (ниже справа) обрабатывают сайт связывания в деталях КМ (белок: красный; лиганд: зеленый) и учитывают влияние оставшегося белка, используя представление ММ (синий). Методы линейного масштабирования или фрагментации могут обрабатывать весь комплекс белок-лиганд в деталях КМ. В методах фрагментации, например, FMO (верхняя справа) белково-лигандная система разделена на небольшие фрагменты, которые обрабатываются методами КМ, включая электростатические эффекты всего белка (подробный рисунок справа показывает спираль, взаимодействующую с лигандом; разные цвета представляют отдельные фрагменты).

1.7 Методы, основанные на КМ / ММ.

Некоторые подходы обрабатывают только сайт связывания лиганда в деталях КМ, а остальная часть системы в расчетах игнорируется.

Чжоу и Кафлиш [22] успешно использовали КМ-зонды в высокопроизводительном виртуальном экране для идентификации селективных ингибиторов. Авторы использовали зонды для представления полярных остатков в связывающем кармане, например, Сер / Тре / Тир, Асп / Глу, Лиз и Арг были представлены метанолом, ацетат-анионом, катионом метиламмония и катионом гуанидиния, соответственно. Полуэмпирическая энергия

взаимодействия КМ между лигандом и остаточными зондами была вычислена и соединена с гидрофобным согласующим членом и штаммом лиганда для вычисления общей энергии взаимодействия белок-лиганд. Они провели скрининг 2,7 миллиона соединений, где сходные позы были сгруппированы до расчета КМ, и идентифицировали ингибитор, селективный по отношению к киназе EphB4, из пула из 85 других протеинкиназ.

Альтернативными схемами для уменьшения вычислительной сложности без игнорирования больших фракций белка являются гибридные методы КМ / ММ, которые сочетают в себе точность КМ с эффективностью ММ. Эти методы обычно разделяют систему на два домена, КМ и область ММ. Слой КМ может содержать только лиганд [23] или лиганд в дополнение к остаткам, которые образуют критические контакты с лигандом в связанном состоянии [24–26]. В металлопротеины, содержащие переходные металлы, необходимо включить ион металла и его лигирующие группы в области КМ. Burgeretal. [23] сообщили, что моделирование короткой молекулярной динамики на позе закрепленного лиганда с последующей оптимизацией КМ / ММ значительно улучшает как прогнозирование нативных поз связывания, так и корреляцию между экспериментальной и расчетной связывающей свободной энергией. Ciancettaetal. [27] исследовали связывание серии из шести лигандов Ru (II) -РЕН (RAPTA) с catB, используя подход КМ / ММ с площадью поверхности Пуассона Больцмана (PBSA), чтобы неявно моделировать сольватацию системы белок-лиганд. Расчеты КМ были выполнены на уровне DFT. Наилучшая корреляция ($r^2 = 0,80$) между предсказанной и экспериментальной свободными энергиями связи была получена без учета энергий сольватации ММ и континуума, поскольку эти энергии не могли быть рассчитаны достаточно точно. В этом исследовании подчеркивается потенциальная трудность точного включения эффектов сольватации в оценку КМ / ММ.

Мерц и его коллеги [24] разработали полуэмпирическую оценочную функцию КМ/ММ AM1 и протестировали их подход на наборе из 23 комплексов металлофермента цинка. При оптимизации весовых коэффициентов было получено значение 2, равное 0,71, и стандартное отклонение 1,69 ккал / моль, это было достигнуто путем оптимизации весов отдельных сроков оценки. Они также обнаружили, что использование меры колебательной энтропии лиганда (то есть количества вращающихся связей) дало сопоставимые результаты для расчета частоты колебаний всей области КМ. Soderhjelmetal. [25] использовали модель КМ/ММ, в которой область КМ была вычислена на уровне MP2/cc-pVTZ для достаточного учета дисперсии и поляризации. Область КМ включает лиганд и все атомы в пределах 4\AA от лиганда. Чтобы получить сродство

связывания семи аналогов биотина к тетрамеру авидина, мультиполи вплоть до квадруполь и анизотропные дипольные поляризуемости, рассчитанные на уровне B3LYP/6-31G*, были объединены с сольватацией с поляризованной континуальной моделью (PCM) и энтропией из анализа в нормальном режиме. Энергии усреднялись по снимкам, отобранным с помощью моделирования 10 ns MD. Несмотря на использование энергий взаимодействия КМ высокого уровня, прогнозирование аффинности связывания было плохим со средним абсолютным отклонением от экспериментальных значений 26 ккал/моль. Это большое отклонение между прогнозируемыми и экспериментальными значениями потенциально является результатом неточностей в вычислении неполярных энергий сольватации для больших молекул с использованием текущей реализации PCM. Результаты также уступали как стандартным расчетам ММ / PBSA, так и результатам докинга с использованием AutoDock. Фонг и др. [26] оценили несколько функций оценки КМ/ММ и обнаружили, что функция оценки HF/6-31G*:: AMBER с поправкой на растворитель показала лучшую производительность, чем ChemScore и GoldScore, но была сопоставима только с функцией оценки ММ на растворителе с протеазой ВИЧ-1.

Анализ различных исследований КМ/ММ подчеркивает важную проблему, которая нуждается в дальнейшем анализе в будущих исследованиях: выбор базового набора оказывает сильное влияние на расчетную аффинность связывания. Soderhjelm et al. [28], например, показали, что энергия взаимодействия, рассчитанная между биотином и авидином, отличается на 38 ккал/моль между использованием 6-31G* и базового набора aug-cc-pVTZ.

1.8 Методы линейного масштабирования и фрагментации.

В некоторых подходах весь белок представлен в деталях КМ. Поскольку это представление является вычислительно дорогостоящим, обычно используются методы аппроксимации. Мерц и его коллеги [29] впервые применили метод линейного масштабирования КМ на полуэмпирическом уровне, используя подход «Разделяй и властвуй», чтобы разделить систему на перекрывающиеся подсистемы, которые можно обрабатывать индивидуально. Fanfrlik et al. [30] разработали схему восстановления, включающую оценку энтальпии с использованием полуэмпирического представления PM6-DH2 с поправкой на дисперсию и водородное связывание и включающие колебательную энтропию, энергии деформации и десольватации на основе ММ. Они смогли различить связывающие и не связывающие для 22 лигандов протеазы ВИЧ-1 и получили корреляцию ($r=0,84$) между теоретической и экспериментальной свободными энергиями связывания. Чжан и Соавт. [31] оценивали полуэмпирические функции оценки

КМ как на уровне РМ3, так и на уровне АМ1 при связывании 45 ингибиторов с протеинкиназой В. Обе модели РМ3 и АМ1 дали хорошие выходы с экспериментальными значениями (РМ3: $r^2 = 0,92$, набор для обучения, $r^2 = 0,71$ на тестовом наборе; АМ1: $r^2 = 0,92$ на тренировочном наборе, $r^2 = 0,70$ на тестовом наборе). Чжоу и Соавт. [32] модифицировали модель энергии линейного взаимодействия с помощью непрерывной электростатической сольватации (LIESE), заменив электростатический член на основе силового поля полуэмпирическим вычислением КМ с линейным масштабированием на уровне РМ1 (повторная параметризация АМ1). Они подчеркнули, что КМ необходим, если существуют значительные изменения в пространственных зарядах белка для разных ингибиторов из-за поляризации и переноса заряда.

Альтернативные подходы к обработке всего белка-лиганда. Система в деталях КМ - это методы фрагментации, которые разделяют систему на несколько фрагментов, которые индивидуально рассчитываются на уровне КМ, предварительно рассчитанные фрагменты затем объединяются с использованием явных поправок многих тел. Наиболее широко используемыми методами, основанными на фрагментах, являются методы молекулярной орбитали фрагментов (FMO) [33] и методы молекулярного фракционирования с использованием конъюгатных колпачков (MFCC) [34]. FMO рассчитывает распределение плотности мономеров и димеров фрагментов (и тримеров в FMO3) с учетом кулоновского потенциала всей системы, что позволяет фрагментацию без укупорки. В MFCC белок сегментируется на фрагменты, соединенные парами конъюгатных молекулярных колпачков. Энергия взаимодействия между лигандом и белком затем суммируется по всем вкладам отдельных фрагментов. FMO может быть связан с вычислениями Пуассона-Больцмана (PB) [35] и РСМ [36], чтобы включить эффекты сольватации. Ямагиши и Соавт. [37] изучали распознавание лигандом рецептора- γ , активируемого пролифератором эроксисомы (PPAR γ), с использованием FMO на уровне MP2/6-31G *. Основываясь на энергии взаимодействия, разлагаемой остатком, используя анализ энергии взаимодействия между фрагментами (IFIE) [38], они идентифицировали две совершенно разные картины связывания лиганда PPAR γ . Вследствие большого гидрофобного кармана PPAR γ автор предположил, что для разработки новых лигандов необходимы адекватные гидрофобные фармакофоры.

В практических условиях разработки лекарств Mazanetzetal. [39] предсказали свободную энергию связывания для 28 ингибиторов циклинзависимой киназы (CDK2), используя метод FMO на уровне MP2 / 6-31G *. Четырнадцать лигандов с известными рентгеновскими структурами были выбраны в качестве обучающего набора, а оставшиеся

14 лигандов без существующей кристаллической структуры были использованы в качестве тестового набора. Производительность метода FMO на тренировочном наборе сравнивалась с пятью функциями подсчета MM в MOE. Корреляция между экспериментальной и прогнозируемой аффинностью связывания с использованием метода FMO (r^2 составляет 0,68) была лучше, чем корреляции трех функций MM, и немного хуже, чем у двух других (r^2 составляет 0,75 для наилучшей функции оценки).

Однако, добавив к энтальпийному вкладу в газовой фазе, полученному в результате расчета FMO, два поправочных условия для энергии сольватации и энтропии, была получена отличная корреляция как для тренировочного набора ($r^2=0,94$), так и для тестового набора ($r^2=0,84$), подчеркивая потенциал методов контроля качества для точного прогнозирования аффинности связывания при разработке лекарств.

Широко признано, что гибкость белка должна быть включена в исследования докинга для более реалистичной модели связывания белок-лиганд. Помимо оперативного включения гибкости боковых цепей, использование предварительно сгенерированных EPS в настоящее время является наиболее широко используемым подходом для моделирования белка при докинге. Недавние исследования продемонстрировали потенциал этой методологии, а также некоторые из важнейших вопросов, которые еще предстоит решить. Какой метод моделирования следует использовать для создания EPS? Сколько конформаций белка следует использовать для оптимальной эффективности и точности при виртуальном скрининге и каков наилучший метод выбора этих структур? Генерируется ли EPS без учета связанных лигандов для моделирования лигандиндуцированных конформационных изменений? Кроме того, новые подходы к эффективному включению гибкости белка в оценку на основе сетки может позволить более точное или более эффективное моделирование адаптации белка к связанным лигандам.

Еще одним важным направлением исследований в области докинга является улучшенная энергетическая количественная оценка белково-лигандного связывания. В последние годы многие исследования включают детали КМ в оценку белка-лиганда, подчеркивая потенциал этих методов для моделирования поляризации и переноса заряда при связывании лиганда. Несмотря на концептуальное преимущество методов КМ для более точной обработки электростатических взаимодействий между белком и лигандом по сравнению со стандартными функциями подсчета на основе MM и некоторыми индивидуальными историями успеха, пока не существует четких доказательств того, что методы КМ и КМ/MM всегда превосходят классические методы MM для всех систем. Значительные различия в количественной оценке белково-лигандных взаимодействий

между различными представлениями КМ и базовыми наборами были признаны, и необходимы дальнейшие исследования для более систематического изучения того, какое представление КМ наиболее точно вычисляет сродство к связыванию лигандов и насколько зависимыми от белка являются эти методы оценки. Следует также признать, что методы контроля качества намного более трудоемки, что потенциально ограничивает количество оцениваемых поз и лигандов при виртуальном скрининге. Несмотря на эти текущие ограничения, потенциал для более точной обработки электростатических взаимодействий между белком и лигандом по сравнению с методами ММ поддерживает значительные будущие усилия по оптимизации и адаптации концепций КМ для изучения систем белок-лиганд.

2. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В ходе ряда исследований по разработке новых эффективных цитостатических агентов в нашей лаборатории был установлен новый хемотип противоопухолевых соединений, а именно производные 4-амино-5-цианопиримидина. В представленном исследовании осуществлен высокопроизводительный индуцированный молекулярный докинг виртуальной комбинаторной библиотеки производных 2-алкилсульфонил-4-амино-6-арил-5-цианопиримидинов в активный сайт киназного внутриклеточного домена рецептора эпидермального фактора роста EGFR.

Исследуемая библиотека соединений состояла из 324 производных, построенных на основе варьирования структурных фрагментов 2-, 4- и 6-положений пиримидинового цикла, представленных в большинстве случаев остатками циклических аминов (пирролидин, морфолин, пиперазин, N-метилпиперазин), алкильных (метил, этил, пропил и изопропил) и арильных фрагментов.

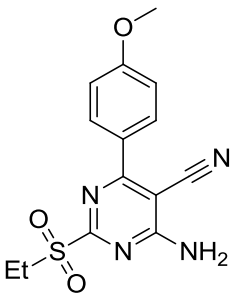
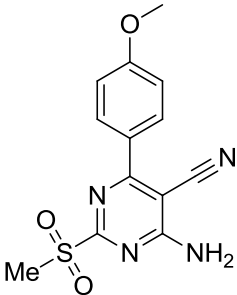
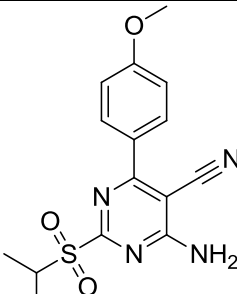
Кристаллическую структуру тирозинкиназы EGFR, связанную с низкомолекулярным лигандом, разрешение не более 2.4 Å, загружали из банка данных белковых структур Proteindatabank (<http://www.rcsb.org/pdb/>) (PDB id 4i23).

Стандартные протоколы молекулярного докинга не учитывают возможную конформационную подвижность исследуемого рецептора. Для решения данной проблемы и для получения более корректных данных по исследуемым соединениям нами проведен индуцированный докинг повышенной точности с учетом конформационной подвижности белковой мишени (EGFR-киназы). Данная возможность реализована в виде протокола *inducedfitmoleculardocking*, результаты реализации которого представлены в таблице 1.

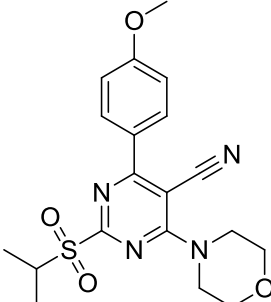
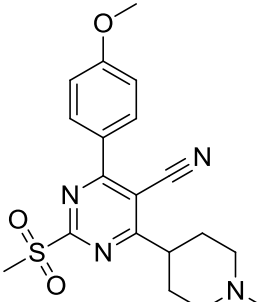
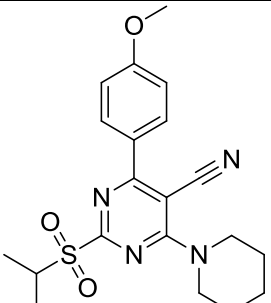
Как видно из представленных данных, учет гибкости белковой мишени приводит к значительному увеличению значений аффинности исследуемых соединений и к увеличению оценочной функции (скоринг-функция) эффективности стыковки.

Исследуемых соединений и к увлечению оценочной функции (скоринг функция) эффективности стыковки. Таблица 3.

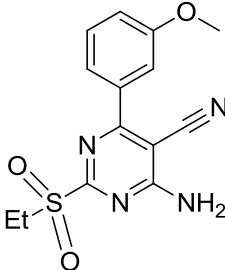
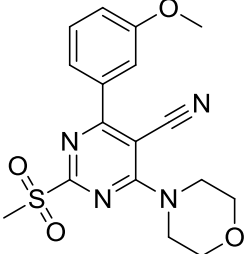
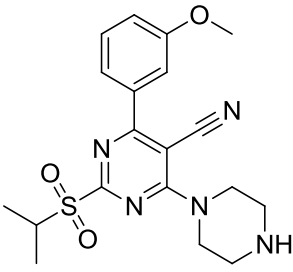
Результаты докинга

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD1	1		-8,850	-75,167	-8,850	MET793 (водородная) LEU788 (водородная) THR854(водородная)
AD2	2		-8,257	-68,287	-8,257	MET793(водородная) ASN842(водородная)
AD3	3		-8,221	-74,588	-8,221	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)

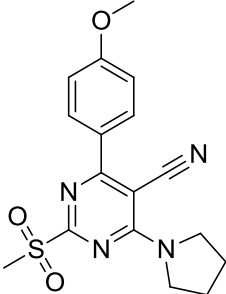
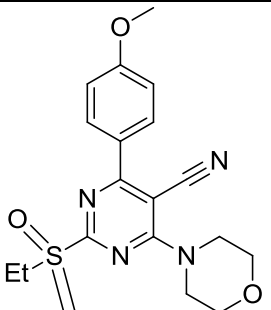
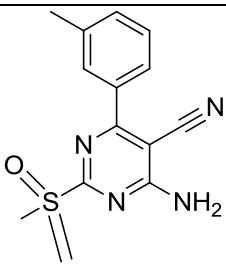
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD4	4		-8,638	-77,393	-8,638	LYS745(водородная) ASP855(водородная)
AD5	5		-8,792	-83,930	-8,616	MET793(водородная) APG841(водородная) LYS745(водородная, рiкатионая) THR790(водородная)
AD6	6		-8,734	-83,670	-8,734	MET793(водородная) LYS745(водородная)

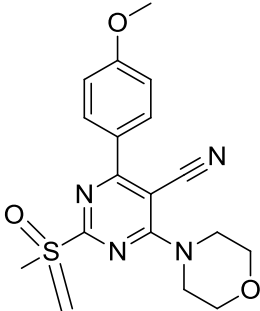
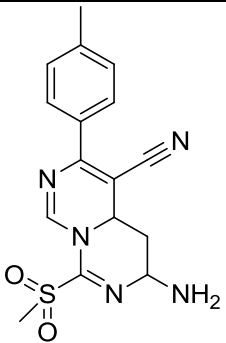
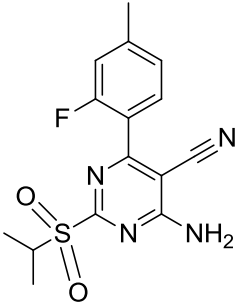
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействие
1	2	3	4	5	6	7
AD7	7		-8,358	-71,361	-8,358	THR790(водородная) THR854(водородная) ASP855(водородная) LYS745(водородная)
AD8	8		-9,279	-85,859	-9,279	LYS745(водородная)
AD9	9		-9,996	-92,496	-8,037	MET793(водородная) THF790(водородная) LYS745(водородная)

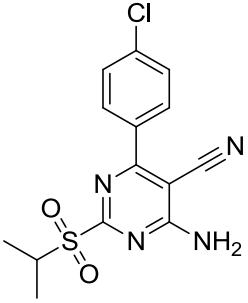
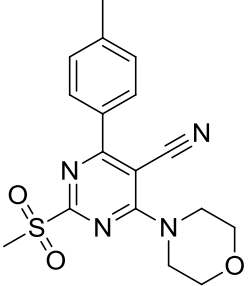
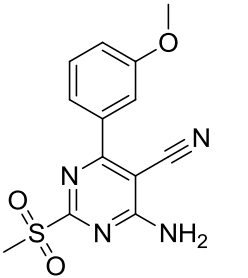
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD10	10		-7,891	-75,603	-7,891	МЕТ7993(водородная) LYS745(водородная)
AD11	11		-8,225	-81,885	-8,225	МЕТ793(водородная) LYS745(водородная) THR790(водородная) CYS797(водородная)
AD12	12		-8,622	-63,763	-8,622	МЕТ793(водородная) GLN791(водородная) THR750(водородная) THR854(водородная)

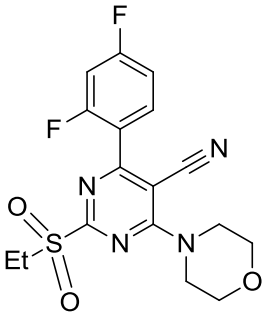
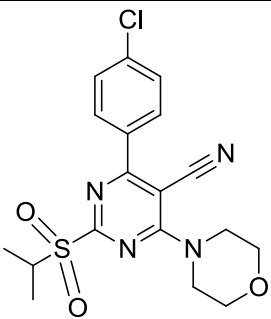
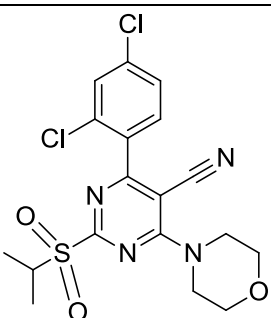
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD13	13		-7,367	-70,870	-7,376	МЕТ793(водородная)
AD14	14		-8,247	-64,594	8,247	МЕТ793(водородная) GLN791(водородная) THR790(водородная) THR854(водородная)
AD15	15		-8,453	-73,551	-8,452	МЕТ793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)

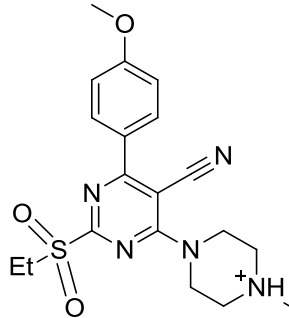
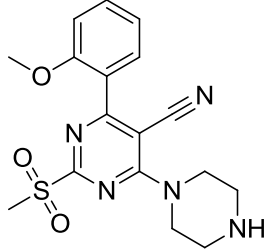
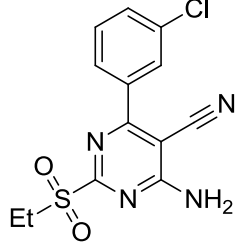
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD16	16		-8,116	-71,013	8,116	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)
AD17	17		-8,856	-83,205	-8,856	MET793(водородная) LYS745(водородная)
AD18	18		-7,537	-65,410	-7,537	MET793(водородная) LYS745(водородная) GYS797(водородная)

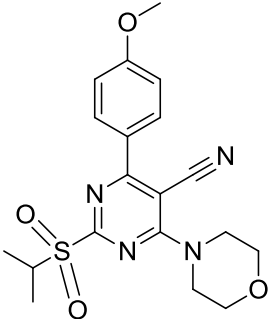
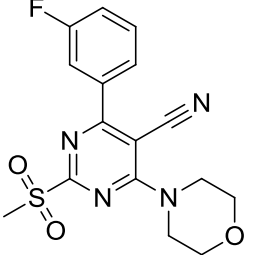
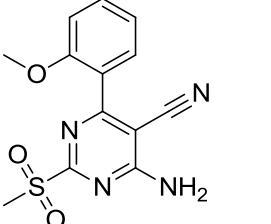
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD19	19		-9,351	-80,629	-9,351	MET793(водородная) THR790(водородная) ASP855(водородная) LYS745(водородная)
AD20	20		-7,605	-69,424	-7,605	MET793(водородная) LYS745(водородная)
AD21	21		-8,248	-69,003	-8,248	MET793(водородная) LYS745(водородная) THR854(водородная)

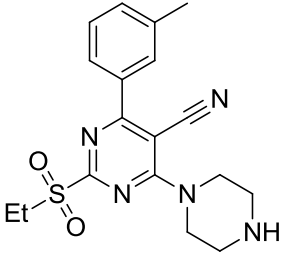
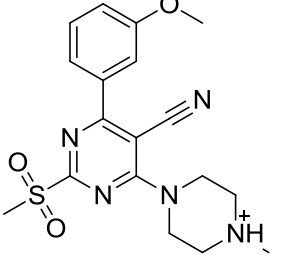
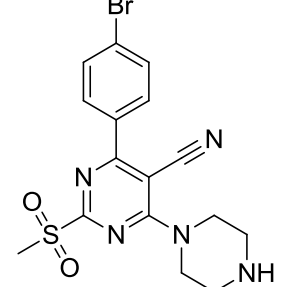
Продолженные таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD22	22		-7,554	-74,518	7,554	LYS745(водородная) LYS745(водородная)
AD23	23		-8,837	-82,938	-8,837	THR790(водородная)
AD24	24		-8,474	-75,180	-8,473	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)

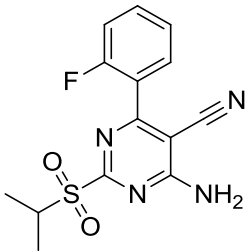
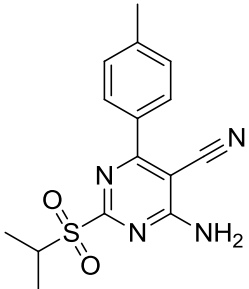
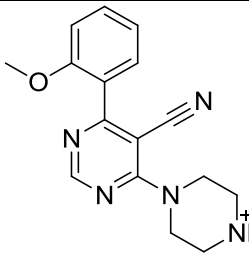
Продолженные таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD25	25		-9,309	-81,691	-9,133	МЕТ793(водородная) THR790(водородная) PHE856(водородная, п-катион) LYS745(водородная)
AD26	26		-8,602	-79,528	-8,580	МЕТ793(водородная) LYS745(водородная)
AD27	27		8,354	-73,347	-8,354	МЕТ793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)

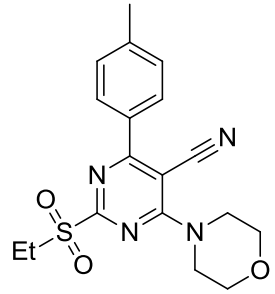
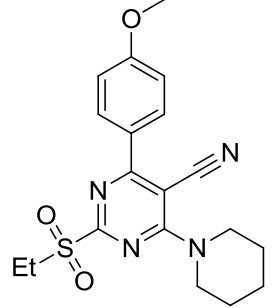
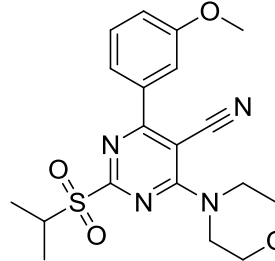
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD28	28		-7,981	80,046	-7,981	МЕТ793(водородная) LYS745(водородная)
AD29	29		-9,084	-82,640	-9,084	LYS745(водородная)
AD30	30		-8,506	-67,019	8,506	МЕТ793(водородная) GLN791(водородная) THR790(водородная) THR854(водородная) LYS745(водородная)

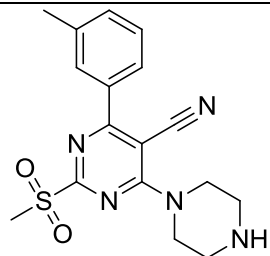
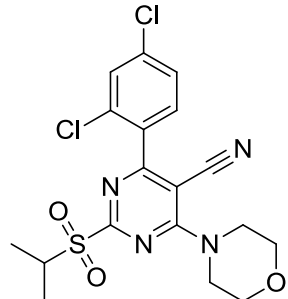
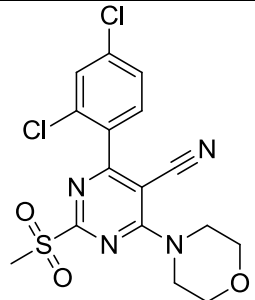
Продолженные таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD31	31		-8,356	-80,147	-8,334	LYS745(водородная) ARG841(водородная)
AD32	32		-7,706	-70,562	-7.530	MET793(водородная) ASP800(водородная)
AD33	33		-10,043	-89,609	-8,084	GLN791(водородная) THR790(водородная)

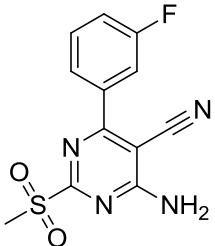
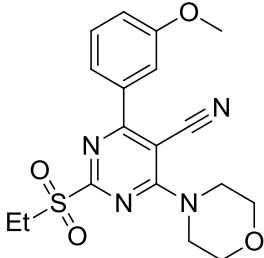
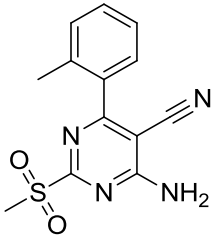
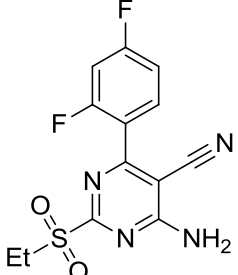
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD34	34		-8,788	-73,026	-8,788	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)
AD35	35		-8,602	-75,256	-8,602	MET793(водородная) GLN791(водородная) LYS745(водородная)
AD36	36		-8,718	-84,277	-8541	MET793(водородная) THR854(водородная) LYS754(водородная) ARG841(водородная, п-катионная)

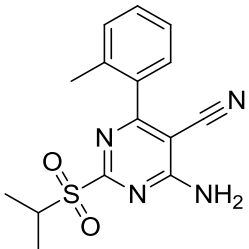
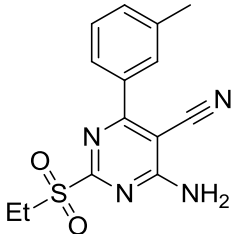
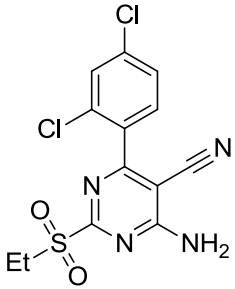
Продолженные таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD37	37		-7,138	-72,600	-7,138	MET793(водородная) LYS745(водородная) CYS797(водородная)
AD38	38		8,604	-84,928	-8,604	MET793(водородная) THR790(водородная) LYS745(водородная)
AD39	39		-9,071	-83,279	-9,071	MET793(водородная) THR790(водородная) ASP855(водородная) LYS745(водородная)

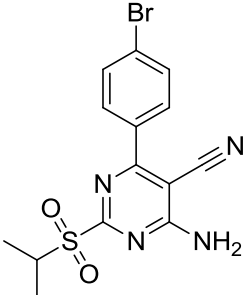
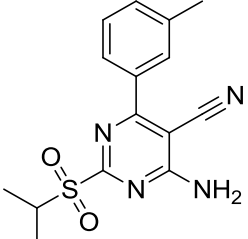
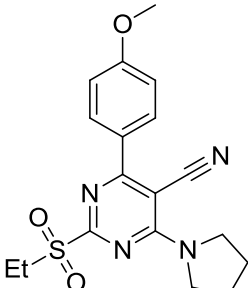
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD40	40		-7,918	-80,169	-7,896	MET793(водородная) ARG841(водородная)
AD42	42		7,917	-76,179	-7,917	MET793(водородная) CYS797(водородная)
AD43	43		-8,456	-71,661	-8,456	MET793(водородная) LYS745(водородная- катионная)
AD44	44		-8,506	-77,016	-8,506	CYS797(водородная) MET793(водородная)

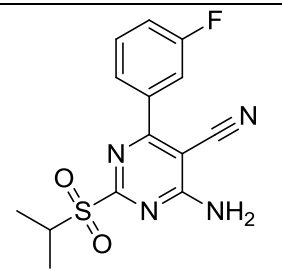
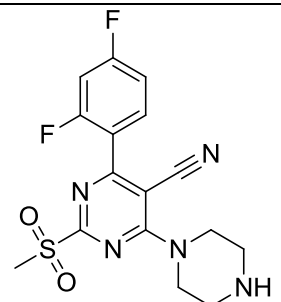
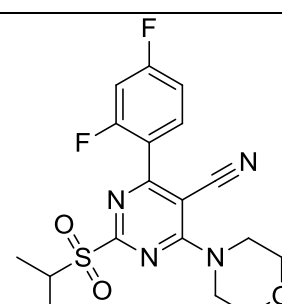
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD45	45		-8,976	-90,008	-8,976	LYS754(водородная)
AD46	46		-8,373	-64,456	-8,373	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)
AD47	47		-7,595	-64,952	-7,595	MET793(водородная) LYS745(водородная, п-катионная)

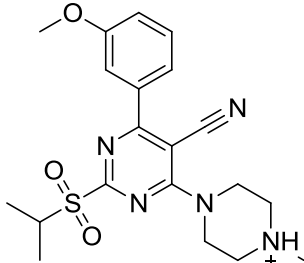
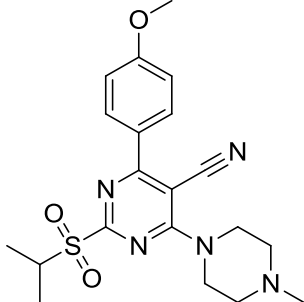
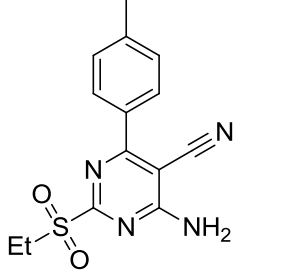
Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD48	48		-8,650	-67,534	8,650	MET793(водородная) GLN791(водородная) THR790(водородная) LYS745(водородная)
AD49	49		-8,194	-65,811	-8,194	MET793(водородная) GLN791(водородная) THR854(водородная) LYS745(водородная)
AD50	50		-8,402	-66,836	-8,402	MET793(водородная) GLN791(водородная) THR854(водородная)

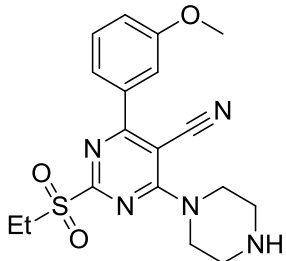
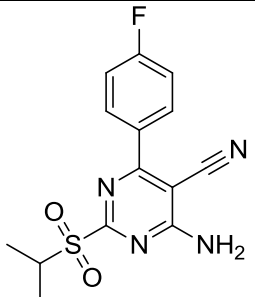
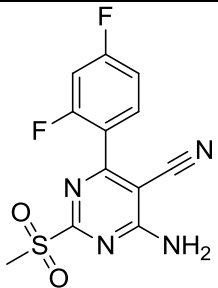
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD51	51		-7,631	-74,603	-7,631	MET793(водородная) GLN791(водородная) THR854(водородная)
AD52	52		-8,314	-73,470	-8,314	MET793(водородная) GLN791(водородная)
AD53	53		-8,161	-78,382	-8,161	MET793(водородная) GLN791(водородная) LYS745(водородная)

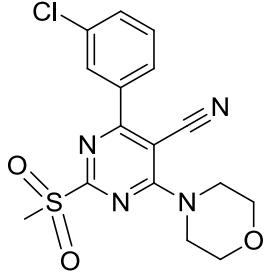
Продолженные таблица 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействие
1	2	3	4	5	6	7
AD54	54		-8,193	-66,638	-8,192	MET793(водородная) GLN791(водородная)
AD55	55		-7,981	-78,064	-7,959	MET793(водородная) LYS745(водородная, п-катионная) ARG841(водородная)
AD56	56		-9,708	-85,525	-9,708	LYS745(водородная)

Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD57	57		-8,049	-83,221	-7,873	МЕТ793(водородная) LYS745(водородная) ASP800 (солевой мостик)
AD58	58		-8,473	-82,333	-7,668	МЕТ793(водородная) LYS745(водородная)
AD59	59		-7,854	-67,036	-7,854	МЕТ793(водородная) GLN791(водородная) LYS745(водородная)

Продолжение таблицы 1

Код Соединение	№	Структура	Clidegscore	Clideemodel	Dockingscore	Ключевые взаимодействия
1	2	3	4	5	6	7
AD60	60		-9,216	-85,612	-7,257	MET793(водородная) LYS745(водородная)

3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Все расчеты в рамках выполненной работы проводили с применением программного пакета SchrodingerSuite 2017-1 (demo-лицензия).

3.1 Подготовка комбинаторной библиотеки

Исходные 3D-структуры исследуемых аллостерических лигандов были построены с использованием панели сборки Maestro в пакете Schrödinger. Модуль Ligprep был использован для генерации всех возможных состояний в физиологическом диапазоне pH 7 ± 2 с учетом наименьшей потенциальной энергии конформации лиганда, с использованием силового поля OPLS_2005. Все полученные структуры соединений были использованы для исследований молекулярной стыковки и свойств ADME с использованием модуля Qikprop.

3.2 Подготовка молекулы белка

Кристаллическую структуру тирозинкиназы EGFR дикого типа, разрешение не выше 2.4 Å, загружали из банка данных белков RCSB (<http://www.rcsb.org/pdb/>) (PDB 3W32). Белок был приготовлен с использованием модуля мастера подготовки ProteinPreparationWizard. Все молекулы воды и гетероатомы, кроме нативного лиганда, были удалены и атомы водорода были добавлены к кристаллической структуре. Комплекс был оптимизирован, чтобы уменьшить стерические столкновения, с применением силового поля OPLS 2005. Модуль SiteMap использовался для оценки характеристики функций сайтов связывания. На его основе генерируются карты гидрофобных и гидрофильных взаимодействий, которые определяют донорные, акцепторные и гидрофобные области сайта связывания.

3.3 Молекулярный докинг

Активный сайт для стыковки лигандов определялся с использованием модуля Generate Grid Generation в Glide. В качестве модельного был выбран сокристаллизованный нативный лиганд, сетка генерировалась вокруг активного участка собственной области EGFR с использованием масштабирующего коэффициента ван-дер-ваальсовых взаимодействий равного 0.9. Для анализа режимов стыковки исследуемых ингибиторов использовались два разных протокола стыковки, протокол стандартной точности (SP) и протокол дополнительной точности (XP).

3.4 Индуцированный докинг

Для оценки конформационной подвижности молекулы белка использовался модуль InductedFitDocking модуль. Основные этапы протокола индуцированного докинга включали:

- Ограниченная минимизация рецептора (получение белка Glide) с отсечением RMSD 0.18 Å.
- Начальная Glide-стыковка каждого лиганда с использованием смягченного потенциала (масштабирование Ван-дер-Ваальса равно 1). В рамках используемого протокола сохраняется максимум 20 поз на лиганд, и сохраняемые позы должны иметь оценку Coulomb-vdW менее 100 и оценку Н-связи менее – 0.05.
- Один раунд прогнозирования конформации основной боковой цепи для каждого комплекса белок-лиганд на остатках в пределах заданного расстояния любой позы лиганда (равное 5 Å).
- Первичная минимизация одного и того же набора аминокислотных остатков и лиганда для каждой позы комплекса белок-лиганд. Структура рецептора в каждой позе теперь отражает индуцированное соответствие структуре и конформации лиганда.
- Конформационно «скользящая» конструкция каждого комплекса с характерной для него энергией низкой энергии (30 ккал/моль). Теперь лиганд строго пристыкован в структуру рецептора с индуцированной подгонкой.
- Оценка энергии связи (IFDScore) для каждой выходной позы.

ЗАКЛЮЧЕНИЕ

По результатам проделанной работы можно сделать следующие выводы:

- смоделированы и получены данные по индуцированному докингу ряда производных 5-цианопиримидина;
- выделены соединения, имеющие наибольшее значение скоринга;
- отмечено, что наибольший вклад в связывание с активным сайтом EGFR-киназы вносят водородные связи между описываемыми соединениями и аминокислотой MET 793, что сопоставляется с ранее описанной литературой по изучению ингибиторов EGFR-киназ.

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. Carlson, H.A. (2002) Protein flexibility and drug design: how to hit a moving target. *Curr.Opin. Chem. Biol.* 6, 447–452
2. Teodoro, M.L. and Kavraki, L.E. (2003) Conformational flexibility models for the receptor in structure based drug design. *Curr. Pharm. Design* 9, 1635–1648
3. Totrov, M. and Abagyan, R. (2008) Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr.Opin.Struc. Biol.* 18, 178–184
4. Beier, C. and Zacharias, M. (2010) Tackling the challenges posed by target flexibility in drug design. *Expert Opin. Drug Dis.* 5, 347–359
5. Rao, C. et al. (2009) Managing protein flexibility in docking and its applications. *Drug Discov. Today* 14, 394–400
6. Sotriffer, C.A. (2011) Accounting for induced-fit effects in docking: what is possible and what is not? *Curr.Top. Med. Chem.* 11, 179–191
7. Lin, J.H. (2011) Accommodating protein flexibility for structure-based drug design. *Curr.Top. Med. Chem.* 11, 171–178
8. Novikov, F.N. et al. (2011) CSAR scoring challenge reveals the need for new concepts in estimating protein–ligand binding affinity. *J. Chem. Inf. Model.* 51, 2090–2096
9. Smith, R.D. et al. (2011) CSAR Benchmark Exercise of 2010: combined evaluation across all submitted scoring functions. *J. Chem. Inf. Model.* 51, 2115–2131
10. Lill, M.A. (2011) Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry* 50, 6157–6169
11. Dietzen, M. et al. (2012) On the applicability of elastic network normal modes in small-molecule docking. *J. Chem. Inf. Model.* 52, 844–856
12. Xu, M. and Lill, M.A. (2011) Significant enhancement of docking sensitivity using implicit ligand sampling. *J. Chem. Inf. Model.* 51, 693–706
13. Barril, X. and Morley, S.D. (2005) Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. *J. Med. Chem.* 48, 4432–4443
14. Bottegoni, G. et al. (2011) Systematic exploitation of multiple receptor conformations for virtual ligand screening. *PLoS One* 6, e18845
15. Korb, O. et al. (2012) Potential and limitations of ensemble docking. *J. Chem. Inf. Model.* 52, 1262–1274
16. Amaro, R.E. et al. (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput. Aided Mol. Des.* 22, 693–705

17. Bolstad, E.S. and Anderson, A.C. (2009) In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking. *Proteins* 75, 62–74
18. Xu, M. and Lill, M.A. (2012) Utilizing experimental data for reducing ensemble size in flexible-protein docking. *J. Chem. Inf. Model.* 52, 187–198
19. Leis, S. and Zacharias, M. (2011) Efficient inclusion of receptor flexibility in grid-based protein–ligand docking. *J. Comput. Chem.* 32, 3433–3439
20. Kruger, D.M. et al. (2011) Target flexibility in RNA-ligand docking modeled by elastic potential grids. *ACS Med. Chem. Lett.* 2, 489–493
21. Kazemi, S. et al. (2009) Elastic potential grids: accurate and efficient representation of intermolecular interactions for fully flexible docking. *ChemMedChem* 4, 1264–1268
22. Zhou, T. and Caflisch, A. (2010) High-throughput virtual screening using quantum mechanical probes: discovery of selective kinase inhibitors. *ChemMedChem* 5, 1007–1014
23. Burger, S.K. et al. (2011) Quantum mechanics/molecular mechanics strategies for docking pose refinement: distinguishing between binders and decoys in cytochrome C peroxidase. *J. Chem. Inf. Model.* 51, 93–101
24. Hayik, S.A. et al. (2010) A mixed QM/MM scoring function to predict protein–ligand binding affinity. *J. Chem. Theory Comput.* 6, 3079–3091
25. Soderhjelm, P. et al. (2010) Ligand affinities estimated by quantum chemical calculations. *J. Chem. Theory Comput.* 6, 1726–1737
26. Fong, P. et al. (2009) Assessment of QM/MM scoring functions for molecular docking to HIV-1 protease. *J. Chem. Inf. Model.* 49, 913–924
27. Ciancetta, A. et al. (2011) A QM/MM study of the binding of RAPTA ligands to cathepsin B. *J. Comput. Aided Mol. Des.* 25, 729–742
28. Soderhjelm, P. et al. (2009) Calculation of protein–ligand interaction energies by a fragmentation approach combining high-level quantum chemistry with classical many-body effects. *J. Phys. Chem. B* 113, 11085–11094
29. Raha, K. and Merz, K.M., Jr (2004) A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding. *J. Am. Chem. Soc.* 126, 1020–1021
30. Fanfrlik, J. et al. (2010) A reliable docking/scoring scheme based on the semiempirical quantum mechanical PM6-DH2 method accurately covering dispersion and H-bonding: HIV-1 protease with 22 ligands. *J. Phys. Chem. B* 114, 12666–12678
31. Zhang, X. et al. (2010) Quantum mechanical pairwise decomposition analysis of protein kinase B inhibitors: validating a new tool for guiding drug design. *J. Chem. Inf. Model.* 50, 651–661

32. Zhou, T. et al. (2008) Is quantum mechanics necessary for predicting binding free energy? *J. Med. Chem.* 51, 4280–4288
33. Kitaura, K. et al. (1999) Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* 313, 701–706
34. Zhang, D.W. and Zhang, J.Z.H. (2003) Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein– molecule interaction energy. *J. Chem. Phys.* 119, 3599–3605
35. Watanabe, H. et al. (2010) Incorporation of solvation effects into the fragment molecular orbital calculations with the Poisson–Boltzmann equation. *Chem. Phys. Lett.* 500, 116–119
36. Sawada, et al. (2010) Role of the key mutation in the selective binding of avian and human influenza hemagglutinin to sialosides revealed by quantum-mechanical calculations. *J. Am. Chem. Soc.* 132, 16862–16872
37. Yamagishi, K. et al. (2010) Flexible ligand recognition of peroxisome proliferator-activated receptor-gamma (PPAR gamma). *Bioorg. Med. Chem. Lett.* 20, 3344–3347
38. Yamagishi, K. et al. (2010) Interactions between 1 alpha, 25(OH)(2)D-3 and residues in the ligand-binding pocket of the vitamin D receptor: a correlated fragment molecular orbital study. *J. Steroid Biochem.* 121, 63–67
39. Mazanetz, M.P. et al. (2011) Prediction of cyclin-dependent kinase 2 inhibitor potency using the fragment molecular orbital method. *J. Cheminf.* 3, 2
40. Soderhjelm, P. et al. (2008) Accuracy of typical approximations in classical models of intermolecular polarization. *J. Chem. Phys.* 128, 014102