

КОНСТРУИРОВАНИЕ ВЫСОКОСПЕЦИФИЧНЫХ СТРУКТУРНЫХ ФРАГМЕНТОВ ДЛЯ ФИЛЬТРАЦИИ СОЕДИНЕНИЙ С НЕЖЕЛАТЕЛЬНОЙ АКТИВНОСТЬЮ

П.И. Савосина*, Д.С. Дружиловский, Д.А. Филимонов, В.В. Поройков

Научно-исследовательский институт биомедицинской химии им. В.Н. Ореховича,
119121, Москва, Погодинская ул., 10; *e-mail: polina.savosina@ibmc.msk.ru

Одним из основных направлений, используемых в медицинской химии для разработки лекарственных средств, является конструирование химических соединений с заданными свойствами. Изменение структурной формулы может быть использовано и для уменьшения количества нежелательных видов биологической активности, которые проявляет разрабатываемое вещество. Для оценки проявляемых видов активности сконструированной структуры широко используются различные *in silico* подходы, которые позволяют уменьшать финансовые и временные затраты на исследование потенциально небезопасных веществ. Однако применение многих алгоритмов компьютерной оценки требует наличия сбалансированных обучающих выборок, содержащих достаточное количество отрицательных и положительных примеров, т.е. информации о структурных формулах веществ, не обладающих и обладающих требуемым видом активности. В случае невозможности использования стандартных прогностических подходов методом выбора может стать оценка наличия/отсутствия структурного фрагмента, ассоциированного с конкретным биологическим эффектом. Для конструирования таких фрагментов мы разработали метод, основанный на оценке вклада отдельных атомов с учетом их ближайшего окружения в проявление конкретного вида биологической активности. Применимость разработанного метода в решении практических задач, возникающих в ходе поиска новых лекарственных средств, была продемонстрирована с использованием информации о молекулярных мишенях, ассоциированных с развитием широкого спектра нежелательных лекарственных реакций. Дополнительное сравнение сформированных фрагментов с известными данными об участках структур молекул, ответственных за связывание с молекулярными мишенями, не только подтверждает надежность нашего метода, но и указывает на то, что он может быть использован для определения функционально важных структурных участков соединений, взаимодействующих с молекулярными мишенями без разрешенной трехмерной структуры.

Ключевые слова: *in silico* исследования; оценка вклада атомов; структурные фрагменты; молекулярные мишени; нежелательные лекарственные реакции

DOI: 10.18097/BMCRM00303

ВВЕДЕНИЕ

Конструирование химических соединений с заданными свойствами является одним из основных направлений медицинской химии [1]. С одной стороны, для улучшения фармакодинамических и фармакокинетических характеристик лекарственной субстанции в ее структурной формуле проводят добавления или изменения структурных фрагментов. Однако внесение этих изменений требует обязательной последующей оценки полученных новых химических соединений для подтверждения их безопасности и эффективности перед дальнейшими исследованиями в клинических испытаниях. Доказательство безопасности лекарственной субстанции на ранних стадиях исследований может быть основано на результатах анализа ее вторичной фармакодинамики – взаимодействия с другими мишенями, отличными от терапевтической и называемыми также “off-target”. С этой целью проводят серии экспериментов с использованием широкого перечня молекулярных мишеней, относящихся к различным классам и играющих роль в разных физиологических процессах [2]. Выявление у исследуемого соединения активности по отношению к мишеням, ассоциированным с развитием нежелательных

лекарственных реакций (НЛР), требует обязательной проверки полученных результатов *in vivo*; в случае их подтверждения может приниматься решение об остановке разработки этой субстанции и поиске более безопасной альтернативы [3].

В связи с этим целесообразно использовать предварительную оценку *in silico* для фильтрации наиболее опасных соединений и приоритизации дальнейших исследований веществ с благоприятным профилем безопасности [4]. В настоящее время инструментом для такой оценки может стать прогноз спектра видов биологической активности для химического соединения по его структурной формуле, который позволяет оценить вероятность наличия или отсутствия способности молекулы взаимодействовать с нежелательными мишенями. Однако его использование не всегда возможно, например по причине недостаточной точности прогностической модели, являющейся следствием высокого разнообразия положительных примеров в обучающей выборке.

В подобных случаях методом выбора может стать оценка структурного сходства [5] или наличия/отсутствия структурного фрагмента, строго характерного для соединений, обладающих нежелательным видом



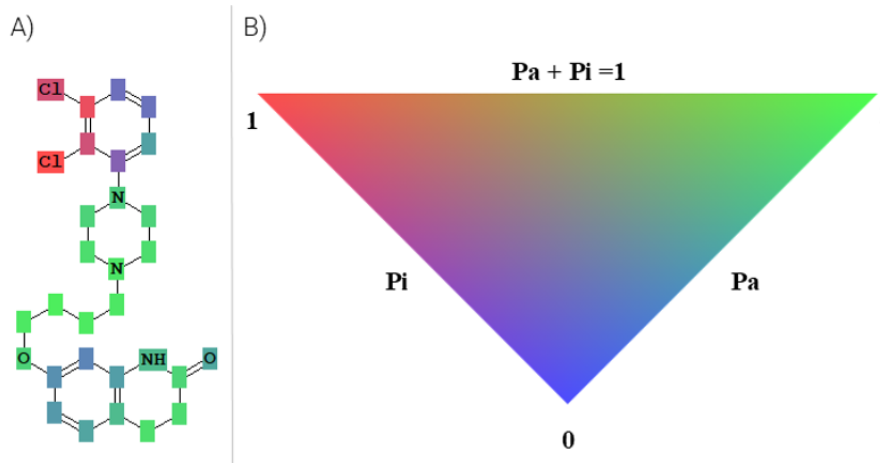


Рисунок 1. А) Пример цветовой окраски атомов структурной формулы соединения, полученной при прогнозе активности «Периферический вазодилататор» с использованием базы знаний PASS 2022. В) Цветовой градиент, используемый для визуализации оценки вкладов атомов в прогнозируемый вид активности.

биологической активности. Этот подход позволит еще на этапе конструирования химического соединения выполнять фильтрацию веществ, проявляющих нежелательные виды биологической активности, что значительно сократит как стоимость, так и время, необходимое для разработки нового лекарственного препарата.

Целью нашей работы была разработка метода конструирования высоко специфичных структурных фрагментов, которые могут быть использованы для идентификации соединений, не обладающих нежелательными видами активности, на примере взаимодействия с молекулярными мишенями, ассоциированными с развитием НЛР.

МЕТОДИКА

Конструирование структурных фрагментов

В качестве источника информации о зависимостях «структура-активность» мы использовали созданную нами ранее базу данных (БД) World Wide approved Drugs (WWAD), версия 2025.04 [6]. Она содержит информацию о структурных формулах, молекулярных мишенях, показаниях к применению и видах биологической активности для более чем 4500 низкомолекулярных активных фармацевтических субстанций (АФС), разрешенных к медицинскому применению. Информация о взаимодействии АФС с молекулярными мишенями собрана в WWAD из официальных документов, публикуемых 88 медицинскими ведомствами (общая характеристика лекарственного препарата, монография лекарства, инструкция по медицинскому применению, обзорные отчеты и пр.), и научных публикаций. Важно помнить, что при терапии одного заболевания нежелательным видом активности может являться такой эффект, который используется в качестве терапевтического для лечения другой нозологии. Это позволяет использовать базу данных WWAD для определения структурных фрагментов, специфичных как для терапевтических, так и для нежелательных эффектов.

Для конструирования структурных фрагментов, характерных для конкретного вида нежелательной активности, мы использовали компьютерную программу

PASS (Prediction of Activity Spectra for Substances), которая позволяет прогнозировать спектр биологической активности на основе структурной формулы исследуемого соединения [7]. В программе PASS биологическая активность химических соединений описывается качественно («активно» или «неактивно») и принимается, что соединение в обучающей выборке не обладает теми видами биологической активности, которые не указаны в списке его известных видов активности. Алгоритм построения зависимостей «структура-активность» для лекарственно-подобных соединений реализован в PASS на основе представления их структур специально разработанными дескрипторами MNA (многоуровневых окрестностей атомов) и модифицированного наивного Байесовского классификатора. Все соединения с определённым видом активности считаются положительными примерами, а все остальные – отрицательными. Это приближение не оказывает существенного влияния на результаты анализа зависимости «структура-активность» и выполняемого на этой основе прогноза благодаря тому, что алгоритм программы PASS обладает робастностью (статистической устойчивостью) по отношению к неполноте данных в обучающей выборке. Три положительных примера могут позволить получать статистически значимые оценки для такого вида активности. Результат прогноза PASS представляется в виде списка прогнозируемых видов биологической активности с соответствующими расчетными значениями вероятностей P_a («быть активным») и P_i («быть неактивным»).

Атомно центрированные фрагментные дескрипторы MNA позволяют оценивать вклад в биологическую активность соединения каждого атома его молекулы. Для визуальной интерпретации рассчитанных вероятностей в программу добавлена цветовая индикация, в которой интенсивность красной составляющей цвета пропорциональна $0.3+0.7P_i$, зелёной – $0.3+0.7P_a$, синей – $1-0.7*(P_a+P_i)$. Таким образом, красный цвет соответствует вероятностям $P_a=0$ и $P_i=1$ (атом характерен для неактивных соединений), зелёный цвет – $P_a=1$ и $P_i=0$ (атом характерен для веществ, обладающих активностью), синий цвет – $P_a=P_i=0$ (вклад в проявление активности данного атома несуществен) (рис. 1). При низком уровне уверенности в

Таблица 1. Характеристики основных этапов создания и валидации структурных фрагментов

Активность	IAP*	Кол-во АФС в WWAD использованных для построения фрагмента	Кол-во веществ в PubChem (найденных / обладающих анализируемой активностью)
Ингибитор АПФ	0.982	14	298 / 187
Ингибитор ГМГ-КоА редуктазы	0.937	8	458 / 257
Ингибитор EGFR	0.949	12	2988 / 1206
Антагонист Д2Р	0.976	13	1471 / 603
Антагонист МЗР	0.907	14	887 / 274

Примечание. * IAP – англ. Invariant Accuracy of Prediction, инвариантная точность прогноза.

прогнозе, например, $Pa=Pi=0.3$, цвет атома будет серым. Такая оценка позволяет выделять в структурной формуле соединения фрагменты, оказывающие положительное или отрицательное влияние на оценку его активности.

Разработанный нами метод конструирования структурных фрагментов можно представить в виде трёх этапов. На первом этапе с использованием специальной версии программы PASS Professional, которая позволяет создавать собственные SAR Base (база знаний о зависимостях «структура-активность»), на основе представленной в WWAD информации нами была построена собственная SAR Base для прогноза действия химических соединений на молекулярные мишени, ассоциированные с развитием НЛР. На втором этапе с использованием PASS выполнили прогнозирование НЛР только тех соединений, которые были использованы в качестве положительных примеров при создании SAR Base. На третьем этапе путем визуального сопоставления участков структур, окрашенных в зеленый цвет, создали структурный фрагмент (в виде т.н. структур Маркуша), общий для анализируемых соединений. Для сохранения созданных структурных фрагментов в машиночитаемом формате Marvin Document (MRV) [8] и в формате изображений использовали программу MarvinSketch от компании ChemAxon.

Валидация специфичности сконструированного фрагмента

Подтверждение специфичности структурных фрагментов для требуемого вида биологической активности выполняли на основе БД PubChem, которая содержит информацию о более чем 122 млн химических соединений [9], в том числе о результатах биологического тестирования 4884139 веществ [10]. С использованием программного комплекса Instant JChem (IJC) от ChemAxon была создана локальная копия, содержащая структурные формулы и идентификаторы той части соединений PubChem, которая была связана хотя бы с одним результатом биологического тестирования, представленного в этой БД. С помощью имплементированного в IJC подструктурного поиска в локальной копии были выявлены соединения, содержащие в своей структуре хотя бы один из созданных фрагментов. Далее были проанализированы наименования молекулярных мишеней, в отношении которых найденные соединения были протестированы. Информация о веществах, для которых в PubChem имелись сведения об отсутствии активности в отношении требуемой нежелательной мишени, была проверена вручную.

РЕЗУЛЬТАТЫ

Для иллюстрации разработанного нами метода конструирования структурных фрагментов, которые могут быть использованы для фильтрации соединений с нежелательными свойствами, нами были выбраны следующие пять молекулярных мишеней, взаимодействие с которыми не только используется для терапии различных заболеваний, но и ассоциировано с развитием широкого спектра НЛР:

1. Ангиотензин-превращающий фермент I (АПФ), ингибирование которого может приводить к снижению артериального давления, появлению кашля, развитию ангионевротического отека, панкреатита [3, 11, 12].

2. 3-Гидрокси-3-метилглутарил-кофермент А (ГМГ-КоА) редуктаза, при снижении активности которой могут развиваться такие серьезные НЛР как катаракта, миопатии, печеночная недостаточность и рабдомиолиз [13-15].

3. Рецептор эпидермального фактора роста (англ. epidermal growth factor receptor, EGFR) – тип тирозинкиназного рецептора, блокада которого сопровождается различными дерматологическими явлениями (зудом, сухостью кожи, акнеподобной сыпью, паронихиями, изменениями волос и ногтей) и гломерулопатиями [16, 17].

4. Дофаминовые рецепторы второго типа (Д2Р), антагонисты которых приводят к возникновению сонливости и судорог, развитию экстрапирамидных реакций, появлению гиперпролактинемии, обструкции дыхательных путей [3, 18].

5. Мускариновые рецепторы третьего типа (МЗР), блокировка которых способствует снижению секреции различных желез, запору, сухости во рту, помутнению поля зрения и другим расстройствам аккомодации, тахикардии [3, 19, 20].

Важно упомянуть, что перечисленные НЛР могут возникать вследствие взаимодействия не только с выбранными биомакромолекулами. Множество других рецепторов и ферментов были определены в качестве потенциальных звеньев патогенеза этих НЛР с разной степенью достоверности [3].

Для каждой из этих мишеней в WWAD представлены взаимодействующие с ними АФС, что позволило с использованием PASS Professional на основе БД WWAD создать SAR Base, обеспечивающую прогноз описанных выше пяти механизмов действия с инвариантной точностью прогноза 0.9 и выше (табл. 1). Используя эту базу знаний, нами был выполнен прогноз для 5 наборов соединений, в

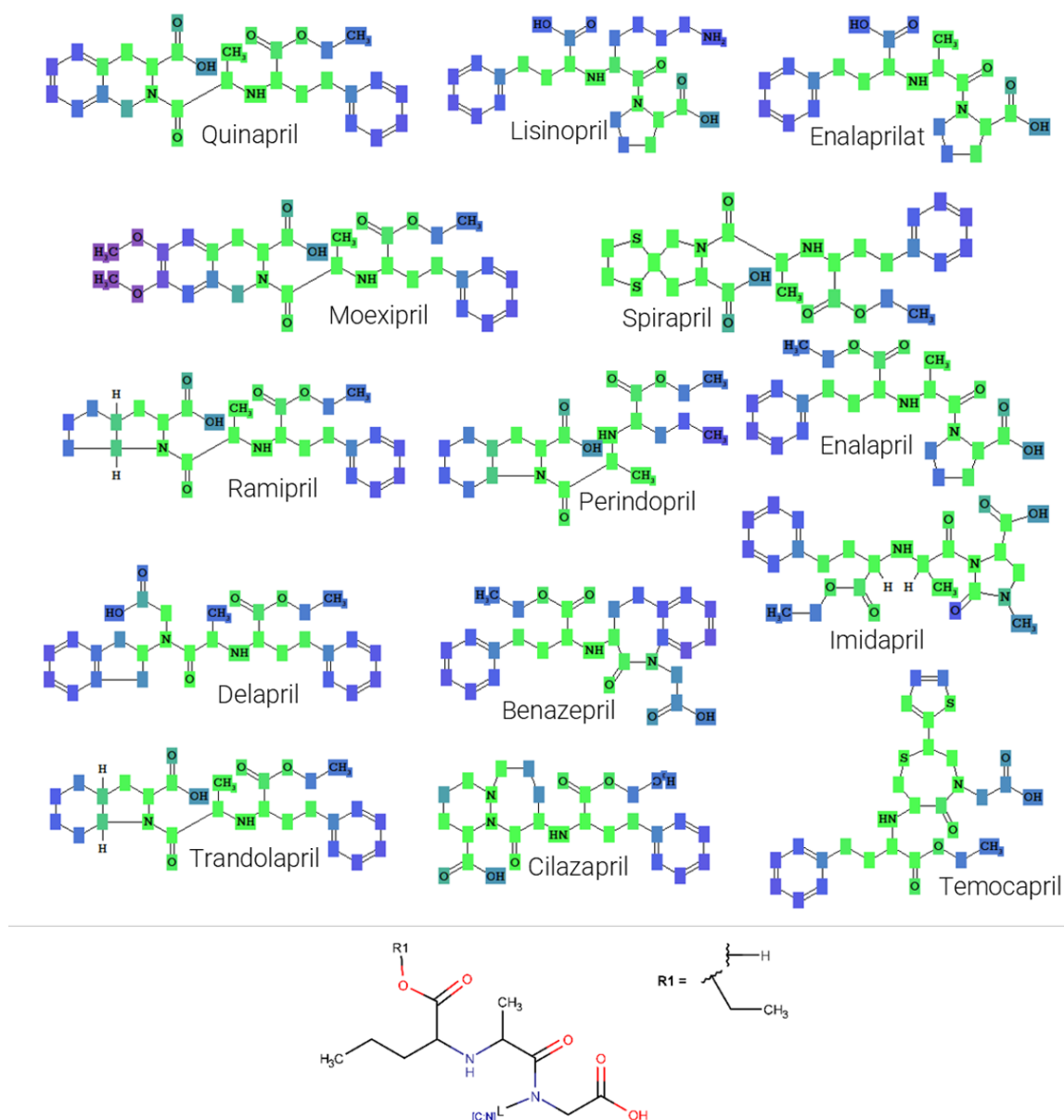


Рисунок 2. Сконструированный фрагмент, характерный для ингибиторов АПФ (снизу), и структурные формулы одобренных к медицинскому применению АФС, использованные для его создания.

каждом из которых присутствовали только АФС с требуемым видом активности. В результате мы получили изображения структурных формул веществ, на которых атомы были окрашены на основе их вклада в прогнозируемую активность с использованием цветовой схемы PASS (см. рис. 1Б). Пример таких изображений для ингибиторов АПФ представлен на рисунке 2А, для соединений из остальных групп – в Дополнительных материалах на рисунках S1-S4. Используя окрашенные в зеленый цвет участки структурных формул, мы построили структуры Маркуша, которые описывают важные для проявления активности участки структуры соединений.

Ингибиторы АПФ I

Первый структурный фрагмент был сконструирован на основе структурных формул ингибиторов АПФ (рис. 2). Используя его в качестве запроса с целью подструктурного поиска среди соединений PubChem, для которых существует информация о результатах биологических тестирований, мы

идентифицировали 298 молекул. Для 187 из них способность ингибировать АПФ была подтверждена хотя бы в одном *in vitro* исследовании. Ни для одного вещества не было показано отсутствия анализируемой активности. Остальные 111 соединений не были исследованы на наличие АПФ ингибирующей активности.

Ингибиторы ГМГ-КоА редуктазы

Структурный фрагмент, характерный для ингибиторов ГМГ-КоА редуктазы (рис. 3), позволил найти в PubChem 458 соединений, для 257 из которых была подтверждена ингибиторная активность. Только для 4 структур наряду с положительными результатами тестирования присутствовали записи, в которых сообщалось об отсутствии активности в отношении ГМГ-КоА редуктазы. Среди них мы обнаружили разрешенное к медицинскому применению лекарство ловастатин, являющееся пролекарством, активный метаболит которого ингибирует анализируемую мишень, и два конъюгата, сконструированных из препарата

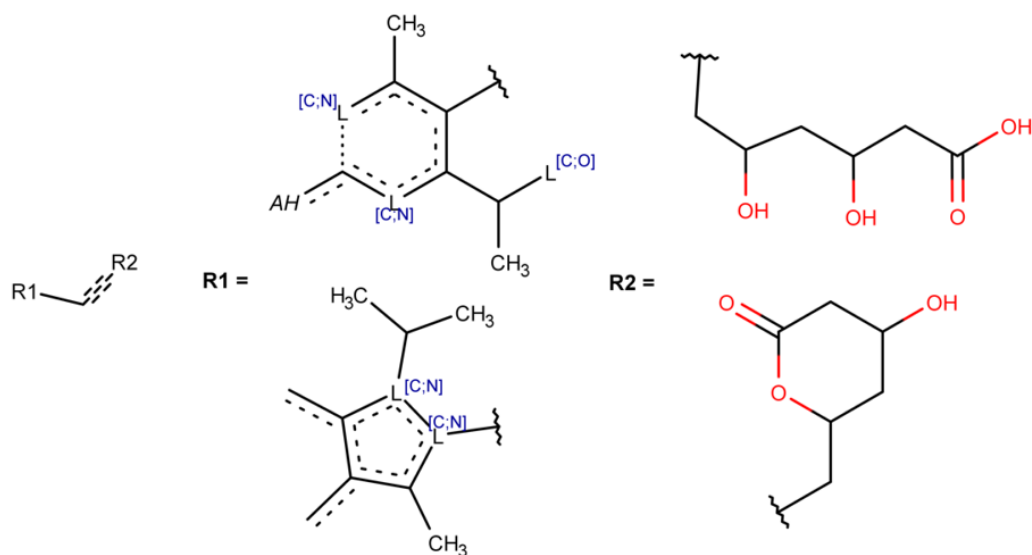


Рисунок 3. Структурный фрагмент, характерный для ингибиторов ГМГ-КоА редуктазы.

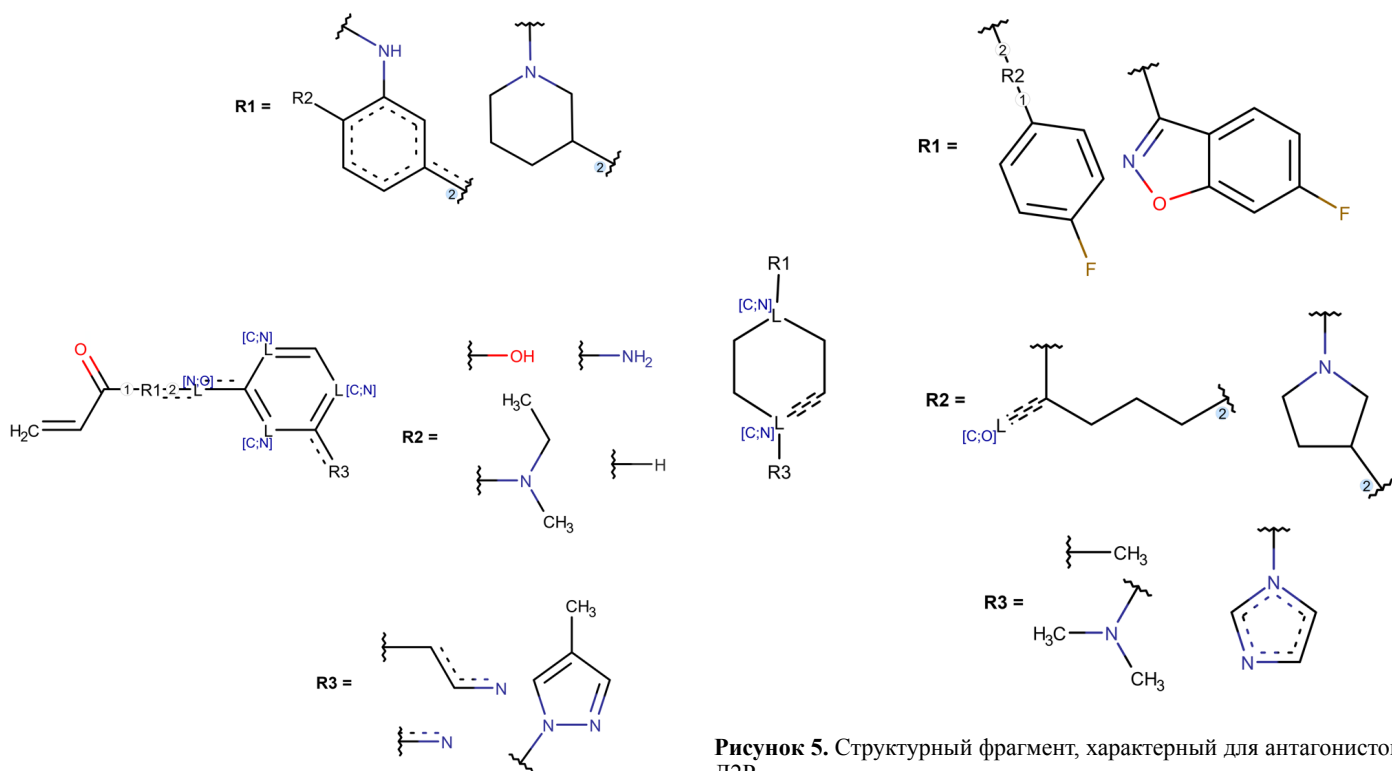


Рисунок 4. Структурный фрагмент, характерный для ингибиторов EGFR.

аторвастатина, для его адресной доставки. Остальные 198 веществ по данным PubChem не были протестированы в отношении ГМГ-КоА редуктазы.

Ингибиторы EGFR

Структурный фрагмент, специфичный для ингибиторов EGFR (рис. 4), присутствует в качестве подструктуры в 2988 соединениях PubChem. Наличие требуемой активности было подтверждено хотя бы в одном *in vitro* эксперименте для 1206 из них. Только для 20 найденных в PubChem веществ было упомянуто отсутствие ингибирования EGFR. При проверке первоисточников этой информации мы обнаружили, что в действительности эти 20 соединений могут блокировать действие требуемого тирозинкиназного рецептора, но их

Рисунок 5. Структурный фрагмент, характерный для антагонистов D2P.

активность невысока (полумаксимальная ингибирующая концентрация превышает 10 мкМ). Для остальных 1762 веществ в PubChem не содержится записей о тестировании в отношении EGFR.

Антагонисты D2P

Четвертый фрагмент, сконструированный для антагонистов D2P (рис. 5), позволяет найти 1471 соединение из PubChem. Записи с отметкой «Активен» в экспериментах по оценке способности связываться или блокировать D2P существуют для 603 найденных веществ. Неактивны как при анализе способности связываться с D2P, так и в функциональном тестировании в отношении этого рецептора 39 найденных структур. После проверки информации в первоисточниках мы обнаружили, что 6 молекул обладают невысокой активностью (константа

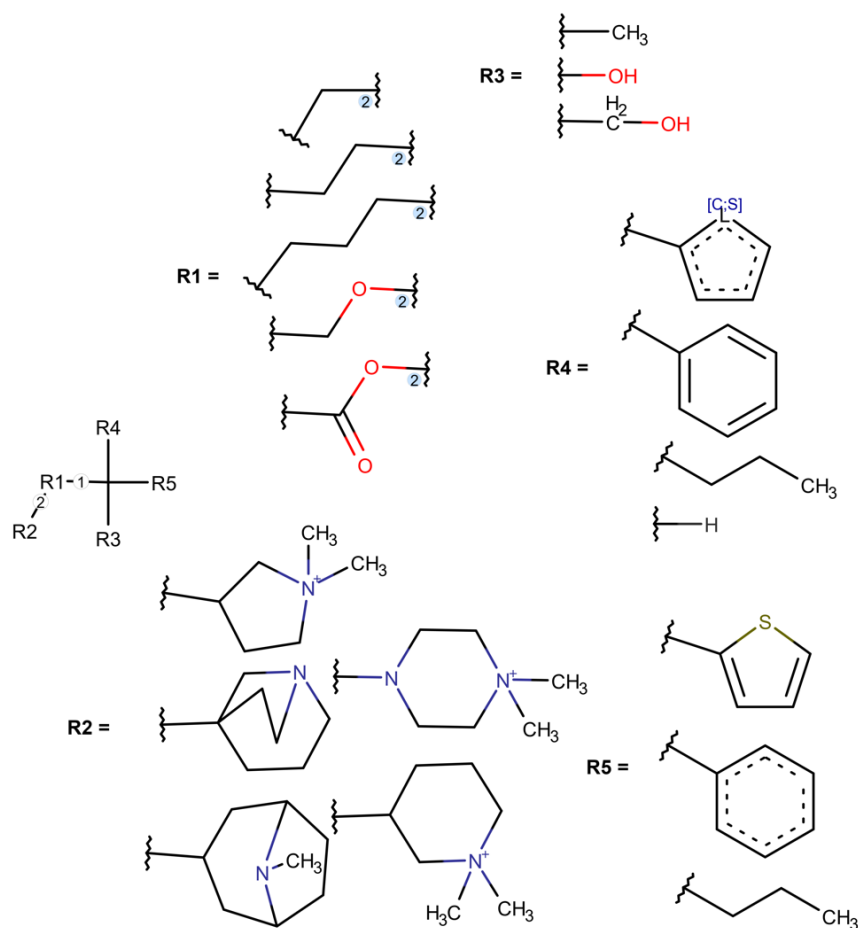


Рисунок 6. Структурный фрагмент, характерный для МЗР.

диссоциации превышает 10 мкМ). Еще 33 вещества были неактивны по результатам высокопроизводительных скрининговых компаний, проведенных NCATS (PubChem AID 485344 и 624463), для которых установлена высокая частота ложноотрицательных результатов. Например, в указанных отчетах о биологическом тестировании мы обнаружили отрицательные результаты для известных антагонистов Д2Р бенперидола, бромперидола, дроперидола, спиперона, галоперидола и др. Для остальных 829 из найденных соединений информация об экспериментальной оценке активности в отношении Д2Р в PubChem отсутствует.

Антагонисты МЗР

С использованием пятого фрагмента, сконструированного для антагонистов МЗР (рис. 6), в анализируемой части PubChem было найдено 887 веществ. Активными хотя бы в одном экспериментальном тестировании по оценке связывания или наличия активности антагониста МЗР были признаны 274 молекулы. Ни для одного из найденных соединений не сообщалось исключительно об отрицательных результатах. Для остальных 613 веществ информация об экспериментальной оценке активности в отношении МЗР в PubChem отсутствует.

ОБСУЖДЕНИЕ

С использованием разработанного нами метода на основе оценки вкладов отдельных атомов в проявление нежелательных видов активности нам удалось

сконструировать характерные для них структурные фрагменты. Предварительное создание SAR Base позволило выявить не просто общие для всех анализируемых соединений участки структуры, а наиболее часто встречающиеся атомы с учетом их ближайшего окружения, которые с высокой вероятностью важны для проявления требуемого вида активности. В этом состоит преимущество предложенного нами подхода конструирования высоко специфичных структурных фрагментов по сравнению, например, с поиском максимально общей подструктуры [21] для соединений с одним видом активности. Молекулярные мишени, выбранные для демонстрации разработанного нами подхода, не только используются для достижения терапевтического эффекта, но и включены в панели фармацевтических компаний для оценки вторичной фармакодинамики и безопасности разрабатываемой терапии [3, 22].

Составленные нами структурные фрагменты являются более специфичными по отношению к конкретным молекулярным мишеням и НЛР, чем предложенные ранее другими авторами (например, так называемые PAINS фильтры [23]). Эти фрагменты позволяют не просто определить наличие опасной функциональной группы в структурной формуле разрабатываемого соединения, но и точно указать, взаимодействие с какой мишенью определяет потенциальный риск.

Важно отметить, что сконструированные нами *in silico* высоко специфичные структурные фрагменты включают в себя участки структуры молекул, которые, по результатам исследования трехмерных комплексов

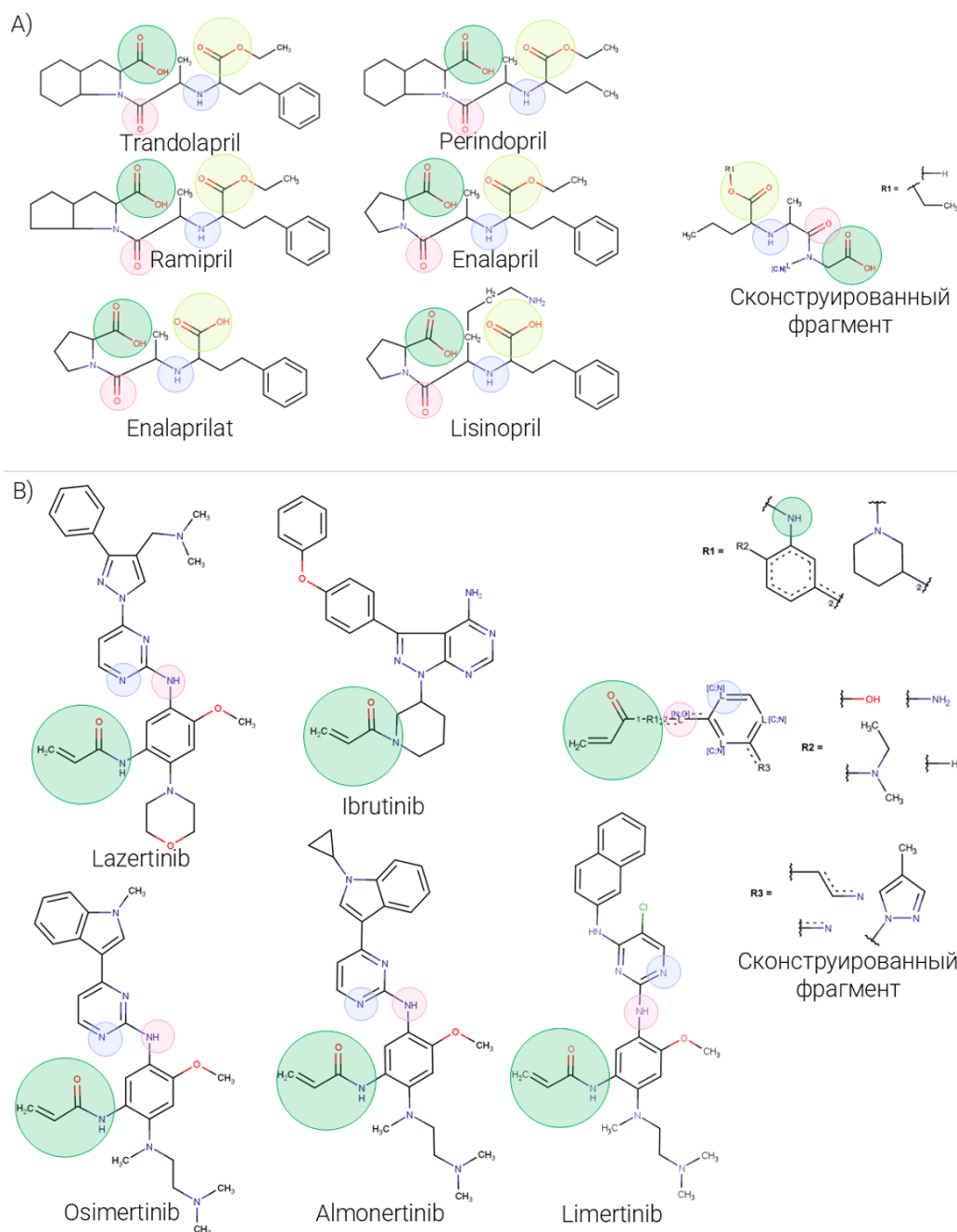


Рисунок 7. Структурные формулы ингибиторов А) АПФ и В) EGFR с построенными на их основе высоко специфичными фрагментами. Кругами отмечены структурные мотивы ингибиторов, необходимые для формирования связи с их молекулярными мишенями [24-27], и соответствующие им участки сконструированных фрагментов.

«мишень-лиганд», отвечают за их формирование. Например, все функциональные группы, обеспечивающие связь ингибиторов АПФ с этим ферментом [24], представлены в определенном нами фрагменте (рис. 7А). Аналогичная зависимость была обнаружена при поиске информации о способе связывания некоторых ингибиторов с EGFR (рис. 7В) [25-27].

Оценка присутствия сконструированных фрагментов в структурах фармацевтических субстанций, одобренных к медицинскому применению, показала их уникальность для соединений с определенным видом активности, несмотря на то, что для их создания были использованы структурные данные и результаты прогноза лишь части имеющихся данных. Подструктурный поиск в более широком химическом пространстве PubChem, содержащем 4884139 экспериментально протестированных веществ, подтвердил

высокую специфичность сконструированных фрагментов в отношении анализируемых молекулярных мишеней. Только 0.006-0.061% соединений содержали в своих структурных формулах построенные фрагменты. Однако менее половины веществ из найденных для каждого структурного фрагмента были экспериментально проверены на наличие требуемого вида активности. В то же время среди протестированных молекул либо все, либо подавляющее большинство взаимодействовали с мишенью, которая входила в список биомолекул, ассоциированных с анализируемой активностью.

Важно подчеркнуть, что в разработанном нами подходе для поиска фрагментов были использованы данные исключительно о фармацевтических субстанциях, разрешенных к медицинскому применению. С одной стороны, это является недостатком данной работы,

поскольку пространство одобренных АФС структурно менее разнообразно, чем все синтезированные на сегодняшний день вещества. В то же время, активность АФС была подтверждена на всех уровнях экспериментальных исследований (от биохимических тест-систем до организма человека), что исключает использование в предложенном подходе соединений, информация об активности которых является следствием ложноположительного результата тестирования.

Ранее группой исследователей из США также был разработан подход для определения структурных фрагментов соединений, связанных с возникновением различных нежелательных реакций [4]. Этот подход был основан на выявлении зависимостей «структура-активность» с использованием методов глубокого обучения. В качестве обучающего набора данных выступала информация о НЛР одобренных АФС из БД SIDER, в которой с использованием методов автоматизированного анализа текстов собрана информация из официальных документов (общих характеристик лекарственных препаратов, инструкций по медицинскому применению лекарственных препаратов) медицинских агентств США, Канады, Австралии, Великобритании и Европейского союза. Авторами были построены модели, позволяющие выполнять прогноз различных НЛР со средней точностью от 0.852 до 0.977, а также идентифицированы участки структурных формул соединений обучающей выборки, определенные во время обучения как вносящие наибольший вклад в прогнозируемые виды активности.

Позднее, исследователями из Китая также было опубликовано алгоритм прогноза некоторых видов токсичности для исследуемых веществ, который позволяет получать информацию о структурных участках, связанных с проявлением анализируемой активности [28]. Для этого авторами были созданы специальные обучающие наборы данных, содержащие информацию об АФС, одобренных к медицинскому применению в США и проявляющих один из трех видов неблагоприятных эффектов: тератогенность, удлинение интервала QT на электрокардиограмме или рабдомиолиз. Авторами были построены три модели с использованием сверточных нейронных сетей, выполняющие прогноз указанных видов токсичности. Последующий анализ позволил идентифицировать участки структурных формул соединений из обучающих выборок, которые в процессе обучения были отмечены как вносящие наибольший вклад в проявление прогнозируемого эффекта. Эти участки представляют собой отдельные атомы или функциональные группы размером до 3 неводородных атомов.

Предложенный нами метод имеет ряд преимуществ по сравнению с упомянутыми выше аналогичными работами. Во-первых, в качестве обучающего набора данных мы использовали БД WWAD, в которой представлено большее количество АФС, разработанных и используемых в намного большем числе стран. Это позволяет охватить более широкое фармакотерапевтическое химическое пространство при конструировании фрагментов. Во-вторых, в WWAD данные о взаимодействии АФС с молекулярными мишенями собраны посредством ручной проверки, что обеспечивает высокую достоверность используемых для обучения сведений в отличие от источника информации группы исследователей из США и Китая, для создания которого были применены

методы автоматизированного анализа текстов. Помимо этого, наивный Байесовский классификатор, который лежит в основе использованной в нашем методе программы PASS, обладает меньшей вычислительной сложностью и обеспечивает более высокую точность прогноза по сравнению с методами глубокого обучения, примененных в работах сравнения. Дополнительным преимуществом нашего подхода перед алгоритмом исследователей из Китая является получение более протяженных структурных фрагментов, ассоциированных с проявлением анализируемого вида биологической активности.

ЗАКЛЮЧЕНИЕ

Разработанный нами метод конструирования структурных фрагментов, высокоспецифичных для проявляющих определенный вид биологической активности веществ, может быть использован для фильтрации соединений с нежелательным типом активности, в том числе с целью разработки новых лекарственных средств. С помощью самого большого из находящихся в свободном доступе репозитория результатов биологических тестирований PubChem мы продемонстрировали, что построенные структурные фрагменты обнаруживаются в соединениях, для которых было экспериментально подтверждено взаимодействие с нежелательной мишенью. Приведенные в качестве примера молекулярные мишени, ассоциированные с различными НЛР, присутствуют в наборах для исследования вторичной фармакодинамики различных фармацевтических компаний. Это подтверждает актуальность и применимость предлагаемого нами подхода в решении практических задач, возникающих в ходе поиска новых лекарственных средств. Сравнение сформированных фрагментов с известными данными об участках структур молекул, ответственных за связывание с исследуемыми ферментами, указывает, что разработанный нами алгоритм оценки вклада отдельных атомов в проявление конкретного вида биологической активности может быть использован для определения функционально важных структурных участков соединений, которые взаимодействуют с молекулярными мишенями, трехмерная структура которых остается пока не охарактеризованной.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

В данной работе исследования с участием людей или использованием животных в качестве объектов не проводились.

БЛАГОДАРНОСТИ

Авторы выражают искреннюю признательность компании ChemAxon за предоставление лицензии на программный комплекс Instant JChem.

ФИНАНСИРОВАНИЕ

Работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021-2030 годы) (№ 122030100170-5).

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

К данной статье приложены дополнительные материалы, свободно доступные в электронной версии (<http://dx.doi.org/10.18097/BMCRM00303>) на сайте журнала.

ЛИТЕРАТУРА

1. Pennington, L.D., Hesse, M.J., Koester, D.C., McAtee, R.C., Qunies, A.M., Hu, D.X. (2024) Property-Based Drug Design Merits a Nobel Prize. *Journal of Medicinal Chemistry*, **67**(14), 11452–11458. DOI: 10.1021/acs.jmedchem.4c01345
2. Jenkinson, S., Schmidt, F., Rosenbrier Ribeiro, L., Delaunois, A., Valentin, J.P. (2020) A practical guide to secondary pharmacology in drug discovery. *Journal of Pharmacological and Toxicological Methods*, **105**, 106869. DOI: 10.1016/j.vascn.2020.106869
3. Lynch, J.J. 3rd, Van Vleet, T.R., Mittelstadt, S.W., Blomme, E.A.G. (2017) Potential functional and pathological side effects related to off-target pharmacological activity. *Journal of Pharmacological and Toxicological Methods*, **87**, 108–126. DOI: 10.1016/j.vascn.2017.02.020
4. Dey, S., Luo, H., Fokoue, A., Hu, J., Zhang, P. (2018) Predicting adverse drug reactions through interpretable deep learning framework. *BMC Bioinformatics*, **19**(Suppl 21), 476. DOI: 10.1186/s12859-018-2544-0
5. Druzhilovskiy, D.S., Stolbov, L.A., Savosina, P.I., Pogodin, P.V., Filimonov, D.A., Veselovsky, A.V., Stefanisko, K., Tarasova, N.I., Nicklaus, M.C., Poroikov, V.V. (2020) Computational Approaches To Identify A Hidden Pharmacological Potential In Large Chemical Libraries. *Supercomputing Frontiers and Innovations*, **7**(3), 57–76. DOI: 10.14529/jsfi200306
6. Savosina, P., Druzhilovskiy, D., Filimonov, D., Poroikov, V. (2024) WWAD: the most comprehensive small molecule World Wide Approved Drug database of therapeutics. *Frontiers in Pharmacology*, **15**, 1473279. DOI: 10.3389/fphar.2024.1473279
7. Filimonov, D., Druzhilovskiy, D., Lagunin, A., Gloriovova, T., Rudik, A., Dmitriev, A., Pogodin, P., Poroikov, V. (2018) Computer-aided Prediction of Biological Activity Spectra for Chemical Compounds: Opportunities and Limitations. *Biomedical Chemistry: Research and Methods*, **1**(1), e00004. DOI: 10.18097/BMCRM00004
8. ChemAxon. Cross-product documentation. Document formats. Marvin Documents – MRV. Retrieved October 16, 2025, from: https://docs.chemaxon.com/display/docs/formats_marvin-documents-mrv.md
9. PubChem database. Retrieved October 1, 2025, from: <https://pubchem.ncbi.nlm.nih.gov/>
10. PubChem database. Table of Contents. BioAssay Results. Retrieved October 1, 2025, from: <https://pubchem.ncbi.nlm.nih.gov/classification/#hid=72&search=BioAssay+Results>
11. Smolinska, S., Antolin-Amérigo, D., Popescu, F.D. (2023) Bradykinin Metabolism and Drug-Induced Angioedema. *International Journal of Molecular Sciences*, **24**(14), 11649. DOI: 10.3390/ijms241411649
12. Weissman, S., Aziz, M., Perumpail, R.B., Mehta, T.I., Patel, R., Tabibian, J.H. (2020) Ever-increasing diversity of drug-induced pancreatitis. *World Journal of Gastroenterology*, **26**(22), 2902–2915. DOI: 10.3748/wjg.v26.i22.2902
13. Macías Saint-Gerons, D., Bosco Cortez, F., Jiménez López, G., Castro, J.L., Tabarés-Seisdedos, R. (2019) Cataracts and statins. A disproportionality analysis using data from VigiBase. *Regulatory Toxicology and Pharmacology* : RTP, **109**, 104509. DOI: 10.1016/j.yrtph.2019.104509
14. Safitri, N., Alaina, M.F., Pitaloka, D.A.E., Abdulah, R. (2021) A Narrative Review of Statin-Induced Rhabdomyolysis: Molecular Mechanism, Risk Factors, and Management. *Drug, Healthcare and Patient Safety*, **13**, 211–219. DOI: 10.2147/DHPS.S333738
15. Zeng, W., Deng, H., Luo, Y., Zhong, S., Huang, M., Tomlinson, B. (2025) Advances in statin adverse reactions and the potential mechanisms: A systematic review. *Journal of Advanced Research*, **76**, 781–797. DOI: 10.1016/j.jare.2024.12.020
16. Guggina, L.M., Choi, A.W., Choi, J.N. (2017) EGFR Inhibitors and Cutaneous Complications: A Practical Approach to Management. *Oncology and Therapy*, **5**, 135–148. DOI: 10.1007/s40487-017-0050-6
17. Izzedine, H., Perazella, M.A. (2017) Adverse kidney effects of epidermal growth factor receptor inhibitors. *Nephrology, Dialysis, Transplantation* : Official Publication of the European Dialysis and Transplant Association - European Renal Association, **32**(7), 1089–1097. DOI: 10.1093/ndt/gfw467
18. Kapur, S., Zipursky, R., Jones, C., Remington, G., Houle, S. (2000) Relationship between dopamine D(2) occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *The American Journal of Psychiatry*, **157**(4), 514–520. DOI: 10.1176/appi.ajp.157.4.514
19. Ehler, F.J., Pak, K.J., Griffin, M.T. (2012) Muscarinic agonists and antagonists: effects on gastrointestinal function. *Handbook of Experimental Pharmacology*, **208**, 343–374. DOI: 10.1007/978-3-642-23274-9_15
20. Lin, S., Kajimura, M., Takeuchi, K., Kodaira, M., Hanai, H., Kaneko, E. (1997) Expression of muscarinic receptor subtypes in rat gastric smooth muscle: effect of M3 selective antagonist on gastric motility and emptying. *Digestive Diseases and Sciences*, **42**(5), 907–914. DOI: 10.1023/a:1018808329603
21. Cao, Y., Jiang, T., Girke T. (2008) A maximum common substructure-based algorithm for searching and predicting drug-like compounds. *Bioinformatics (Oxford, England)*, **24**(13), i366–i374. DOI: 10.1093/bioinformatics/btn186
22. Scott, C., Dodson, A., Saulnier, M., Snyder, K., Racz, R. (2022) Analysis of secondary pharmacology assays received by the US Food and Drug Administration. *Journal of Pharmacological and Toxicological Methods*, **117**, 107205. DOI: 10.1016/j.vascn.2022.107205
23. Huggins, D.J., Venkitaraman, A.R., Spring, D.R. (2011) Rational methods for the selection of diverse screening compounds. *ACS Chemical Biology*, **6**(3), 208–217. DOI: 10.1021/cb100420r
24. Zheng, W., Tian, E., Liu, Z., Zhou, C., Yang, P., Tian, K., Liao, W., Li, J., Ren, C. (2022) Small molecule angiotensin converting enzyme inhibitors: A medicinal chemistry perspective. *Frontiers in Pharmacology*, **13**, 968104. DOI: 10.3389/fphar.2022.968104
25. Patel, K.B., Heppner, D.E. (2025) Lazertinib: breaking the mold of third-generation EGFR inhibitors. *RSC Medicinal Chemistry*, **16**(3), 1049–1066. DOI: 10.1039/d4md00800f
26. Sutanto, F., Konstantinidou, M., Dömling, A. (2020) Covalent inhibitors: a rational approach to drug discovery. *RSC Medicinal Chemistry*, **11**(8), 876–884. DOI: 10.1039/d0md00154f
27. Damghani, T., Chitnis, S.P., Abidkun, O.A., Patel, K.B., Lin, K.S., Ouellette, E.A., Lantry, A.M., Heppner, D.E. (2025) Profiling and Optimizing Targeted Covalent Inhibitors through EGFR-Guided Studies. *Journal of Medicinal Chemistry*, **68**(16), 17917–17932. DOI: 10.1021/acs.jmedchem.5c01661
28. Lin, J., He, Y., Ru, C., Long, W., Li, M., Wen, Z. (2024) Advancing Adverse Drug Reaction Prediction with Deep Chemical Language Model for Drug Safety Evaluation. *International Journal of Molecular Sciences*, **25**(8), 4516. DOI: 10.3390/ijms25084516

Поступила: 05.11.2025
После доработки: 16.11.2025
Принята к публикации: 17.11.2025

DESIGN OF HIGHLY SPECIFIC STRUCTURAL FRAGMENTS FOR FILTERING COMPOUNDS WITH UNDESIRABLE ACTIVITIES

P.I. Savosina, D.S. Druzhilovskiy, D.A. Filimonov, V.V. Poroikov*

Institute of Biomedical Chemistry, 10 Pogodinskaya str., Moscow, 119121 Russia; *e-mail: polina.savosina@ibmc.msk.ru

The design of chemical compounds with desired properties is a key approach in medicinal chemistry for drug development. Modifying structural formulas it is possible to reduce undesirable biological activities exhibited by the substance under study. Various *in silico* methods are widely used to predict the types of activity in designed molecules, thereby reducing both the financial and time costs associated with experimental screening of potentially unsafe substances. However, the performance of many computer-based assessment algorithms depends on well-balanced training datasets containing a sufficient number of both positive and negative examples (information on the structural formulas of compounds that do and do not exhibit the desired type of activity). When standard predictive approaches are not applicable, an alternative strategy involves assessing the presence or absence of structural fragments associated with specific biological effects. To support this approach, we have developed a method that estimates the contribution of individual atoms to a given biological activity, taking into account their local structural environment. The applicability of this method to practical drug discovery tasks was demonstrated by analyzing molecular targets linked to a broad range of adverse drug reactions. Moreover, comparison of the constructed fragments with known structural motifs responsible for molecular target binding confirms the robustness of our method and highlights its potential for identifying functionally relevant structural regions in compounds that interact with molecular targets lacking resolved three-dimensional structures.

Key words: *in silico* studies; estimating atomic contributions; structural fragments; molecular targets; adverse drug reactions

FUNDING

The study was performed in the framework of the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) (No. 122030100170-5).

Received: 05.11.2025, revised: 16.11.2025, accepted: 17.11.2025