ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ УНИТАРНОЕ ПРЕДПРИЯТИЕ
«ВСЕРОССИЙСКИЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ
АВТОМАТИКИ ИМ. Н.Л. ДУХОВА»

На правах рукописи

ЕВТЕЕВ СЕРГЕЙ АНТОНОВИЧ

РАЗРАБОТКА НОВЫХ ИНГИБИТОРОВ ПРОТЕИНКИНАЗ TTK, ATM И CAMKK2 С ПОМОЩЬЮ СОВРЕМЕННЫХ КОМПЬЮТЕРНЫХ ПОДХОДОВ

1.5.4. Биохимия
1.4.16. Медицинская химия

ДИССЕРТАЦИЯ
на соискание ученой степени кандидата биологических наук

Научный руководитель:
Иваненков Ян Андреевич,
кандидат биологических наук

Москва – 2026

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ…… 4
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ……..... 13
1.1. Роль киназ в биологических процессах и развитии заболеваний.… 13
1.2. Киназы ATM, CAMKK2 и TTK: биологическая роль и перспективы таргетирования…… 16
1.3. Современные подходы к поиску и разработке ингибиторов киназ.. 22
1.3.1. Ранее охарактеризованные ингибиторы ATM……. 24
1.3.1. Ранее охарактеризованные ингибиторы CAMKK2……. 29
1.3.1. Ранее охарактеризованные ингибиторы TTK…….. 31
1.4. Проблемы и ограничения современных вычислительных методов. 35
1.4.1. Ограничения метода AlphaFold 2……. 39
1.4.2. Ограничения современных генеративных методов……...…….. 45
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЙ……..... 54
2.1. Поиск сайтов связывания на белках…….. 54
2.2. Оптимизация структур белков, полученных с использованием AlphaFold 2……..…… 57
2.3. Генерация псевдолигандов, классификация сайтов связывания, оценка аффинности лигандов……. 59
2.3.1. Генерация псевдолигандов…….…….. 59
2.3.2. Классификация сайтов связывания……. 61
2.3.3. Оценка аффинности лигандов…… 63
2.4. Генерация новых структур на основе фрагментов…… 64
2.5. Схемы синтеза смоделированных соединений …….…… 68
2.6. Методики *in vitro* тестирования…….…….. 90
2.6.1. Анализ ингибирования киназы CAMKK2…… 90
2.6.2. Анализ ингибирования киназы ATM……. 91
2.6.3. Анализ ингибирования киназы DNA-PK…… 91
2.6.4. Анализ ингибирования киназы ATR…….…… 92

2.6.5. Анализ ингибирования киназы TTK…… 92
ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ…… 93
3.1. *In silico* валидация…… 93
3.1.1. Поиск сайтов связывания…… 93
3.1.2. Примеры применения метода поиска сайтов связывания…….. 96
3.1.3. Оптимизация геометрии сайта связывания…….. 100
3.1.4. Генерация псевдолигандов…… 104
3.1.5. Классификация сайтов связывания……..…… 106
3.1.6. Оценка аффинности лигандов…….. 109
3.1.7. Генерация молекулярных структур……. 111
3.2. Поиск ингибиторов TTK…….. 113
3.2.1. Отбор потенциальных ингибиторов TTK…….. 113
3.2.2. *In vitro* анализ ингибирования TTK…….. 114
3.3. Дизайн ингибиторов CAMKK2…….. 118
3.3.1. Виртуальная генерация структур QFASG-1, QFASG-2, QFASG-3. 118
3.3.2. Синтез соединений QFASG-1, QFASG-2, QFASG-3…….. 119
3.3.3. *In vitro* анализ ингибирования CAMKK2……..…….. 120
3.3.4. Виртуальная оценка селективности и ADMET…… 123
3.4. Дизайн ингибиторов ATM……. 125
3.4.1. Виртуальная генерация структур QFASG-4, QFASG-5, QFASG-6. 125
3.4.2. Синтез соединений QFASG-4, QFASG-5, QFASG-6…….. 126
3.4.3. *In vitro* анализ ингибирования ATM, ATR и DNA-PK…….. 128
3.4.4. Виртуальная оценка ADMET…… 131
ЗАКЛЮЧЕНИЕ……..…… 132
ВЫВОДЫ……..…….. 138
ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ……..……. 139
СПИСОК СОКРАЩЕНИЙ……..……. 140
СПИСОК ЛИТЕРАТУРЫ…….…… 141

ВВЕДЕНИЕ

Актуальность исследования

Разработка ингибиторов протеинкиназ является актуальной и важной составляющей современной медицинской химии. Протеинкиназы выступают в качестве ключевых регуляторов клеточных процессов, участвуя в контроле клеточного цикла, передаче сигналов, апоптозе, метаболизме и репарации ДНК [5]. Нарушение регуляторных механизмов, в которые вовлечены белки семейства протеинкиназ, лежит в основе многих патологий, таких как онкологические, нейродегенеративные и воспалительные заболевания. Несмотря на существенные успехи в выявлении биологических функций и структурной организации киназ, дизайн высокоселективных и эффективных ингибиторов остается нетривиальной задачей. В частности, это определяется вероятностью возникновения побочных эффектов ввиду высокого сходства в строении активных центров данного класса ферментов [9].

Экспериментальные подходы к разработке лекарств, включая органический синтез и биологическое тестирование, характеризуются высокой стоимостью и длительностью выполнения. Развитие и внедрение вычислительных методов позволяет ускорить и оптимизировать процессы поиска и дизайна новых биологически активных соединений. Традиционные вычислительные методы, включая классический молекулярный докинг и стандартные алгоритмы виртуального скрининга, имеют ряд ограничений по точности и прогнозирующей способности. Зачастую они основываются на системе правил и не учитывают ранее накопленные данные о лиганд-белковых взаимодействиях [49].

С появлением и развитием методов генерации трехмерных структур белков, в частности алгоритма AlphaFold 2, были открыты новые перспективы для работы с биомишенями без кристаллических данных, что позволило использовать сгенерированные структуры для определения сайтов связывания и последующего дизайна лигандов. Несмотря на значительные успехи в области структурной биологии, трехмерные модели AlphaFold 2 часто требуют дополнительной оптимизации для устранения ошибок боковых цепей и улучшения геометрии активных центров, что стало предметом активных исследований [33].

Одновременно с этим развиваются методы генерации структур новых лекарственных молекул с использованием машинного обучения. В настоящее время известно множество коммерческих и академических алгоритмов, направленных на решение данной проблемы. Однако большинство предлагаемых методов характеризуется рядом проблем в связи с ограниченными возможностями современных методов машинного обучения в понимании закономерностей медицинской химии. Ключевой проблемой существующих решений является генерация несвойственных для медицинской химии структур молекулярных соединений, которые содержат токсические функциональные группы, метаболически неустойчивы, синтетически недоступны или невалидны с точки зрения органической химии [34].

Учитывая существующие проблемы и ограничения, данное исследование описывает обнаружение структур новых ингибиторов протеинкиназ ATM, CAMKK2 и TTK с использованием разработанных вычислительных алгоритмов. В диссертации описаны разработка и применение алгоритмов для поиска сайта связывания (SiteRadar), оптимизации геометрии сайта связывания (AF Optimizer), автоматического конструирования молекулярных структур (QFASG) на основе фрагментов,

оценки лиганд-белковой аффинности (Skittles). Описаны способы химического синтеза смоделированных соединений и определены параметры ингибирования белков-мишеней *in vitro*. Разработка и применение указанных алгоритмов способствовало обнаружению новых ингибиторов, способных адресно взаимодействовать с протеинкиназами ATM, CAMKK2 и TTK — актуальными мишенями современной медицины, связанными со злокачественными новообразованиями и другими заболеваниями.

Обоснованность данного исследования подтверждается как актуальностью проблемы разработки высокоэффективных ингибиторов протеинкиназ, так и необходимостью интеграции современных вычислительных техник в многокомпонентный процесс разработки лекарств. В связи с тем, что успех фармацевтической разработки в значительной степени определяется качеством вычислительных моделей и эффективностью генеративных алгоритмов, выполненная работа по разработке и практической апробации новых компьютерных методов для обнаружения новых биологически активных соединений является особенно значимой и своевременной.

Таким образом, обнаружение новых ингибиторов протеинкиназ ATM, CAMKK2 и TTK с помощью новых вычислительных методов, совмещающих структурную биологию, машинное обучение и хемоинформатику, не только расширяет фундаментальные знания о компьютерной разработке лекарственных молекул, но и демонстрирует практическую значимость и потенциал применения данных методов для дальнейшего применения с целью повышения эффективности ранней разработки лекарств.

Степень разработанности темы исследования

Актуальность поиска новых ингибиторов киназ ATM, CAMKK2 и TTK обусловлена их ключевой ролью в регуляции процессов клеточного цикла, восстановления ДНК, апоптоза и передачи сигнала. Нарушения в работе этих киназ связаны с развитием онкологических, нейродегенеративных и воспалительных заболеваний. Выбор препаратов в данной области крайне ограничен, что свидетельствует о необходимости поиска новых, отличных с точки зрения химической структуры, ингибиторов [35, 37, 61].

Известны примеры применения вычислительных алгоритмов для дизайна ингибиторов киназ: глубокие генеративные модели применялись для дизайна структур новых ингибиторов киназ FLT3 [36], MerTK [31], DDR1 [88], TBK1 [78]. Применение указанных методов позволило обнаружить новые биологически активные соединения, однако их эффективность и вклад в обнаружение новых веществ неоднозначны. Анализ ранее опубликованных работ и полученных результатов показал, что большая доля сгенерированных молекулярных структур являются близкими аналогами ранее известных соединений. При этом значительную роль в выборе финальных структур для синтеза и бологического тестирования сыграли другие факторы, такие как экспертный анализ [34].

Существующие вычислительные методы дизайна лекарственных соединений, такие как молекулярный докинг, виртуальный скрининг и алгоритмы машинного обучения, не обладают достаточной точностью и интерпретируемостью. Это ограничивает их применимость для разработки новых биологически активных соединений. Дизайн новых лекарственных молекул является многокомпонентной задачей, требующей одновременного учета множества факторов: эффективность, селективность, безопасность и фармакокинетика. По этой причине развитие

и доработка компьютерных методов с повышенной точностью и подтвержденной эффективностью является важным и перспективным направлением [34].

В рамках настоящего исследования были виртуально сконструированы новые ингибиторы протеинкиназ ATM, CAMKK2 и TTK, которые были впоследствии синтезированы и показали активность в тестах ингибирования мишеней *in vitro*. Данные структуры удалось обнаружить благодаря использованию разработанных вычислительных алгоритмов для поиска и оптимизации геометрии лиганд-белковых сайтов связывания, автоматического конструирования молекулярных структур на основе фрагментов и расчета лиганд-белковой аффинности с использованием машинного обучения.

Цель исследования

Разработка и применение комплексных вычислительных методов для обнаружения ингибиторов протеинкиназ TTK, ATM и CAMKK2.

Задачи исследования

1. Разработать новые вычислительные методы для оптимизации трехмерных моделей белков, поиска сайтов связывания, генерации псевдолигандов, классификации сайтов связывания, прогнозирования лиганд-белковой аффинности, автоматического конструирования структур малых молекул на основе фрагментов.
2. Выполнить *in silico* валидацию предложенных методов на широкой выборке тестовых данных в сравнении с известными алгоритмами.
3. Применить разработанные вычислительные методы для обнаружения новых ингибиторов киназ ATM, CAMKK2 и TTK.

4. Сформировать библиотеку соединений и выполнить органический синтез шести целевых веществ на основании моделирования *in silico*.
5. Исследовать ингибирующую активность полученных веществ в отношении белков-мишеней *in vitro*.

Научная новизна исследования

Разработаны новые вычислительные методы для оптимизации трехмерных моделей белков, поиска сайтов связывания, генерации псевдолигандов, классификации сайтов связывания, прогнозирования лиганд-белковой аффинности, автоматического конструирования структур малых молекул на основе фрагментов.

Впервые проведена *in silico* валидация предложенных методов на широкой выборке тестовых данных, показавшая высокую точность работы в сравнении с известными алгоритмами.

Впервые с использованием описанных технологий разработаны, синтезированы и протестированы новые ингибиторы фармакологических мишеней — протеинкиназ ATM, CAMKK2 и TTK, подтвердившие биологическую активность *in vitro*.

Получен патент РФ № RU 2838984C1 от 24.04.2025 «Способ разметки лиганд-белковых сайтов связывания».

Теоретическая и практическая значимость работы

Теоретическая значимость работы заключается в разработке оригинальной вычислительной методологии, объединяющей оптимизацию трехмерных конформаций белков, применение графовых нейронных сетей для детального картирования сайтов связывания и автоматическое конструирование структур малых молекул на основе фрагментов с использованием квантово-химических расчетов. Совместное применение

предлагаемых методов позволяет атвоматизировать процесс дизайна новых биологически активных соединений.

Практическая значимость исследования определяется успешным применением предлагаемых вычислительных методов для обнаружения и синтеза новых ингибиторов протеинкиназ. Были обнаружены новые ингибиторы киназ ATM, CAMKK2 и TTK, которые играют ключевую роль в развитии онкологии и других заболеваний. Результаты биологических тестов *in vitro* продемонстрировали активность полученных соединений в наномолярном и низком микромолярном диапазоне концентраций. Это свидетельствует о перспективах дальнейшего развития и внедрения предложенных компьютерных алгоритмов в практику дизайна новых биологически активных соединений.

Разработанные вычислительные методы могут стать перспективными прикладными инструментами, способными быстро и ээфективно решать задачи медицинской химии, тем самым снижая затраты по сравнению с традиционными экспериментальными подходами.

Материалы и методы исследования

Диссертационная работа выполнена с применением компьютерного моделирования, включая молекулярный докинг, виртуальный скрининг, методы машинного обучения, фрагментный подход к поиску лекарств.

Основные положения, выносимые на защиту

Разработанные вычислительные методы позволяют оптимизировать трехмерные модели киназ, идентифицировать и классифицировать сайты связывания, генерировать псевдолиганды, прогнозировать аффинность, осуществлять дизайн структур малых молекул на основе фрагментов.

Проведенная *in silico* валидация предложенных методов на широкой выборке тестовых данных, показала высокую точность предлагаемых подходов в сравнении с известными алгоритмами. SiteRadar показал улучшение метрики топ-N+2 с 47% до 74%, AF Optimizer повысил успех обнаружения соединений-хитов с 28% до 40% в ретроспективном исследовании. Алгоритм Skittles и его применение для классификации сайтов связывания и оценки аффинности продемонстрировали уровень точности, сравнимый с современными решениями.

Предложенные вещества подтвердили способность ингибировать протеинкиназы ATM, CAMKK2 и TTK. Доля обнаруженных ингибиторов TTK с IC_{50} < 1 мкМ составила 25,6%. Доля синтезированных ингибиторов ATM и CAMKK2 с IC_{50} < 10 мкМ составила 50%.

Степень достоверности и апробация результатов

Результаты работы были представлены на российских и международных конференциях, включая HTSDD-2020, MCCMB 2023, Foresight Session 2023, AMTEXPO 2023, DataFusion 2024, Digital Innopolis Days 2024, Biomembranes 2024, Fall into ML 2024, ХимБиоSeasons 2025, Digital Innopolis Days 2025, РРЛС 2026, ХимБиоSeasons 2026, Сириус-Биотех 2026. Результаты подтверждены биологическими тестами и опубликованы в рецензируемых международных научных изданиях.

Внедрение результатов исследования

Результаты исследования внедрены в практику ФГУП «ВНИИА им. Н.Л. Духова».

Личный вклад автора

Автор разработал и протестировал новые вычислительные методы, выполнил планирование и контроль органического синтеза и

биологического тестирования, проанализировал и объяснил полученные результаты.

Соответствие диссертации паспорту научной специальности

Диссертация соответствует паспорту научной специальности 1.5.4. Биохимия (биологические науки), а именно пунктам: 2 — Биохимия белков. Протеомика. Белковая инженерия. Структурная биология; 3 — Энзимология. Механизмы катализа. Регуляция ферментативной активности. Моделирование ферментативных процессов. Диссертация соответствует паспорту научной специальности 1.4.16. Медицинская химия (биологические науки), а именно пунктам: 1 — Поиск, структурный дизайн и синтез соединений-лидеров — потенциальных физиологически активных (лекарственных) веществ, на основе: а) знания структурных параметров биомишени или особенностей патогенеза; б) анализа и модификации структур известных активных соединений; в) синтеза и биологического тестирования широкого разнообразия химических соединений; 2 — Использование фундаментальных методов математической химии (компьютерного молекулярного моделирования и QSAR) с целью прогнозирования возможности взаимодействия определенных химических соединений с предполагаемой биологической мишенью, а также для выявления взаимосвязи между химической структурой и физиологической активностью; 8 — Физико-химические исследования лиганд-рецепторных взаимодействий с целью фармакологической пригодности соединений. Использование методов докинга, рентгеноструктурного анализа, ЯМР спектроскопии, микрокалориметрии, поверхностного плазмонного резонанса для установления структурно-функциональных взаимоотношений потенциальных лекарственных средств.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Роль киназ в биологических процессах и развитии заболеваний

Протеинкиназы представляют собой один из крупнейших и наиболее разнообразных классов ферментов в организме. Они выполняют важную функцию катализа фосфорилирования белков. Этот посттрансляционный модификационный механизм регулирует активность, локализацию и взаимодействие целевых белков. Фосфорилирование является фундаментальным регуляторным механизмом в широком спектре клеточных процессов. Протеинкиназы участвуют в контроле клеточного цикла, передаче сигналов, апоптозе, метаболизме и устранении повреждений ДНК. Они обеспечивают поддержание гомеостаза и адаптивных реакций клетки на внешние и внутренние стимулы [68, 84]. До 30% белков в клетках животных подвергаются фосфорилированию под действием протеинкиназ. Это демонстрирует их ключевую роль и универсальность для регуляции биологических функций.

Геном человека содержит около 500 различных генов протеинкиназ. Это составляет примерно 2% его полного генетического материала и указывает на высокий уровень специализации и разнообразия этих ферментов. Химическая активность протеинкиназ основана на реакции переноса фосфатной группы с АТФ на гидроксильные группы аминокислотных остатков (серин, треонин и тирозин). Это приводит к изменению конформации белков и регуляции различных сигнальных путей. Эти процессы контролируются сложной системой регуляции, включая аутофосфорилирование, взаимодействие с белками-активаторами и ингибиторами, а также влияние малых молекул [5].

Протеинкиназы классифицируют по специфичности к субстратам. Выделяют серин-треониновые, тирозиновые киназы, киназы двойной специфичности и гистидин-специфичные киназы. Тирозиновые протеинкиназы выполняют ключевую роль в передаче сигналов, связанных с мембранными рецепторами. Они модулируют активацию белковых комплексов через фосфорилирование тирозиновых остатков и запускают каскады внутриклеточных сигнальных путей. Серин-треониновые киназы (например, протеинкиназа А) регулируют метаболические процессы через цАМФ-зависимые механизмы. Они участвуют в транскрипционной регуляции и обеспечивают адаптивный ответ клетки на изменения внутренней среды.

Протеинкиназы играют ключевую роль не только в функционировании клетки. Они также представляют собой важнейшие терапевтические мишени для лечения различных заболеваний. Изучение их функций, строения, механизмов регуляции и взаимодействия с биологически активными веществами важно для разработки эффективных и селективных ингибиторов.

Семейство PIKK включает такие важные представители, как ATM (ataxia telangiectasia mutated), ATR (ataxia telangiectasia and Rad3-related kinase) и DNA-PK (DNA-dependent protein kinase). Эти киназы играют критически важную роль в регуляции биологических функций и клеточных процессов. Указанные ферменты выступают в качестве ключевых сенсоров геномных повреждений, особенно повреждений ДНК. Они инициируют активацию сигнальных каскадов, направленных на репарацию поврежденного генетического материала. Их деятельность важна для поддержания стабильности и целостности генома. Это обеспечивает нормальное функционирование клеток и предотвращение патологических состояний [74].

ATM, ATR и DNA-PK активируются при возникновении различных структурных повреждений ДНК. Они взаимодействуют с уникальными сенсорными белковыми комплексами, такими как MRN для ATM, RPA/ATRIP для ATR и комплекс KU70/KU80 для DNA-PK. Это отражает их специализированные механизмы реагирования на повреждения. Эти киназы способны фосфорилировать широкий спектр аминокислотных остатков как в своем составе, так и на других белках. Это обеспечивает сложные взаимозависимые эффекты, способствующие эффективной репарации ДНК, контролю клеточного цикла и поддержанию гомеостаза. Нарушения функции этих белков, в том числе мутации или недостаточная экспрессия, связаны с развитием онкологических заболеваний, нейродегенеративных и иммунных расстройств. Это подчеркивает их значение в качестве терапевтических биомишеней для разработки лекартв [56].

Нарушение регуляции киназных сигналов лежит в основе множества патологических процессов. Среди них особое место занимают злокачественные опухоли. В клетках злокачественных новообразований нарушается сложный контроль клеточного цикла, ответственного за координацию последовательных этапов клеточного деления и репликации. Дисфункция киназ, регулирующих циклин-зависимые пути, приводит к утрате контроля над пролиферацией клеток, что приводит к неконтролируемому росту опухолевой ткани и формированию лекарственной устойчивости к терапии [71]. Такая устойчивость значительно затрудняет лечение и снижает эффективность существующих препаратов, тем самым увеличивая риск рецидивов и ухудшения прогноза течения заболевания.

Ингибиторы киназ считаются одними из наиболее перспективных фармакологических средств в онкологии и терапии других заболеваний.

Они направлены на восстановление контроля над клеточным циклом и сигнализацией, блокируя ключевые белки, ответственные за патологическую пролиферацию и выживание аномальных клеток [77]. Современные препараты, нацеленные на киназы, уже нашли широкое применение, однако по-прежнему ведутся интенсивные поиски новых ингибиторов с более высокой селективностью, эффективностью и сниженной токсичностью, что актуально для уменьшения побочных эффектов и повышения качества жизни пациентов [5, 13, 50, 80, 84, 89, 90].

1.2. Киназы ATM, CAMKK2 и TTK: биологическая роль и перспективы таргетирования

Киназы ATM, CAMKK2 и TTK играют ключевую роль в регуляции множества важных клеточных процессов.

ATM (ataxia telangiectasia mutated) — серин/треониновая протеинкиназа из семейства PIKK. Она состоит из 3056 аминокислот и имеет молекулярную массу приблизительно 350 кДа. Данный белок включает структурные домены HEAT, FAT, киназный домен, PIKK-регуляторный домен и С-концевой домен FATC. Каждый из этих доменов играет важную роль в регуляции активности и стабильности молекулы (рисунок 1).

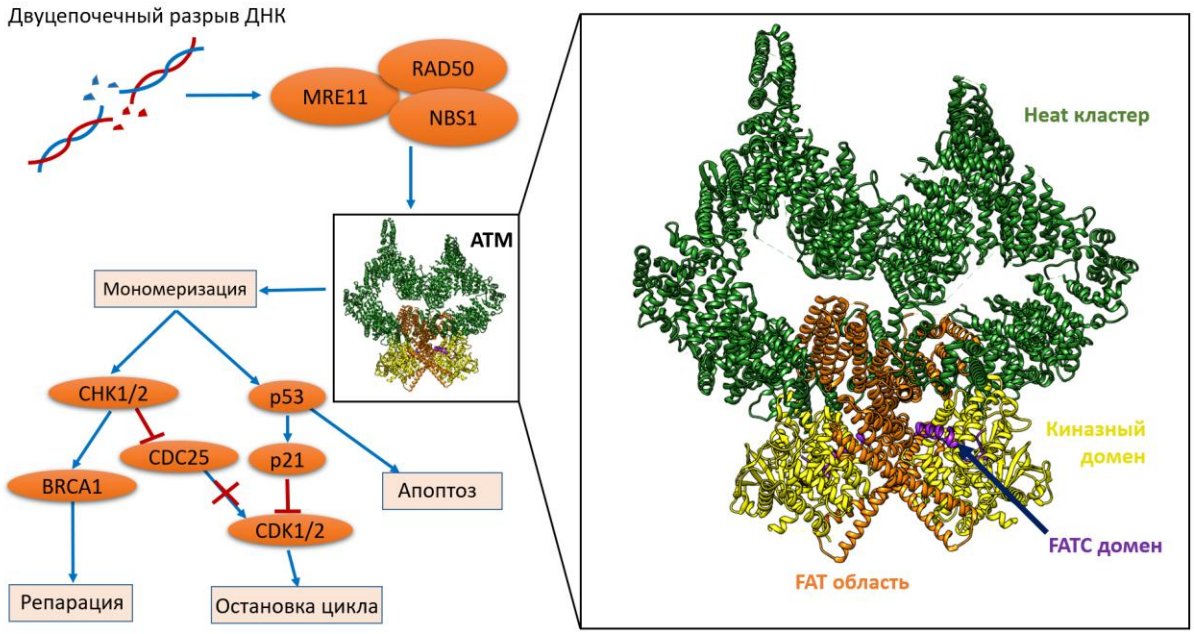


Рисунок 1. Роль ATM в репарации двухцепочечных разрывов ДНК. Структура киназы ATM.

Данный белок играет ключевую роль в клеточном ответе на двухцепочечные разрывы. Это обеспечивается благодаря способности связываться и активироваться непосредственно на концах поврежденной ДНК [46]. Активация ATM начинается с аутофосфорилирования по сериновому остатку Ser^{1981}. Это приводит к распаду неактивных димеров на активные мономеры. Для полноценной активации также требуется взаимодействие с комплексом MRN (Mre11-Rad50-NBS1) и дополнительное фосфорилирование остатков Ser^{367} и Ser^{1893}.

ATM осуществляет фосфорилирование ключевых эффекторов, таких как p53, CHK2 и BRCA1. Они контролируют остановку клеточного цикла в контрольных точках G1/S и G2/M, инициируют репарацию ДНК, и запускают апоптоз при необратимых повреждениях [10, 17, 73]. Такой механизм обеспечивает сохранение геномной стабильности и предотвращение мутагенеза. Это жизненно важно для нормального функционирования организма [74]. Кроме прямой роли в репарации ДНК,

ATM участвует в регуляции митохондриального гомеостаза и аутофагии. Это расширяет спектр его функций в клеточной биологии.

Мутации в гене ATM связаны с наследственным заболеванием атаксия-телеангиэктазия. Кроме того, они значительно повышают риск развития онкологических заболеваний, включая рак молочной железы, поджелудочной железы и простаты [35, 46, 54, 62]. Дефекты ATM приводят к нарушению механизмов контроля повреждений ДНК. Это провоцирует развитие геномной нестабильности, которая является предпосылкой к канцерогенезу [13, 95].

Кальций/кальмодулин-зависимая протеинкиназа киназа 2 (CAMKK2) — серин/треониновая киназа. Она принадлежит к семейству Ca^{2+}/CaM-зависимых протеинкиназ. Молекула CAMKK2 состоит из нескольких доменов: каталитический домен и регуляторные элементы. Они обеспечивают ее Ca^{2+}/кальмодулин-зависимую активацию. В структуре этой киназы наблюдается конформационная перестройка при связывании кальция и кальмодулина, что активирует каталитический центр и запускает фосфорилирование субстратов. Многочисленные исследования показали, что CAMKK2 образует димерные и мультимерные комплексы, что способствует кооперативной активации и дополнительной регуляции чувствительности к кальцию.

CAMKK2 регулирует активность нескольких ключевых киназ, включая CAMK1, CAMK4 и AMP-активируемую протеинкиназу (AMPK). Она вовлечена в различные физиологические и патологические процессы. Среди них следует отметить инсулиновый сигналинг, метаболический гомеостаз, воспаление, пролиферация и рост опухолевых клеток (рисунок 2) [61, 66].

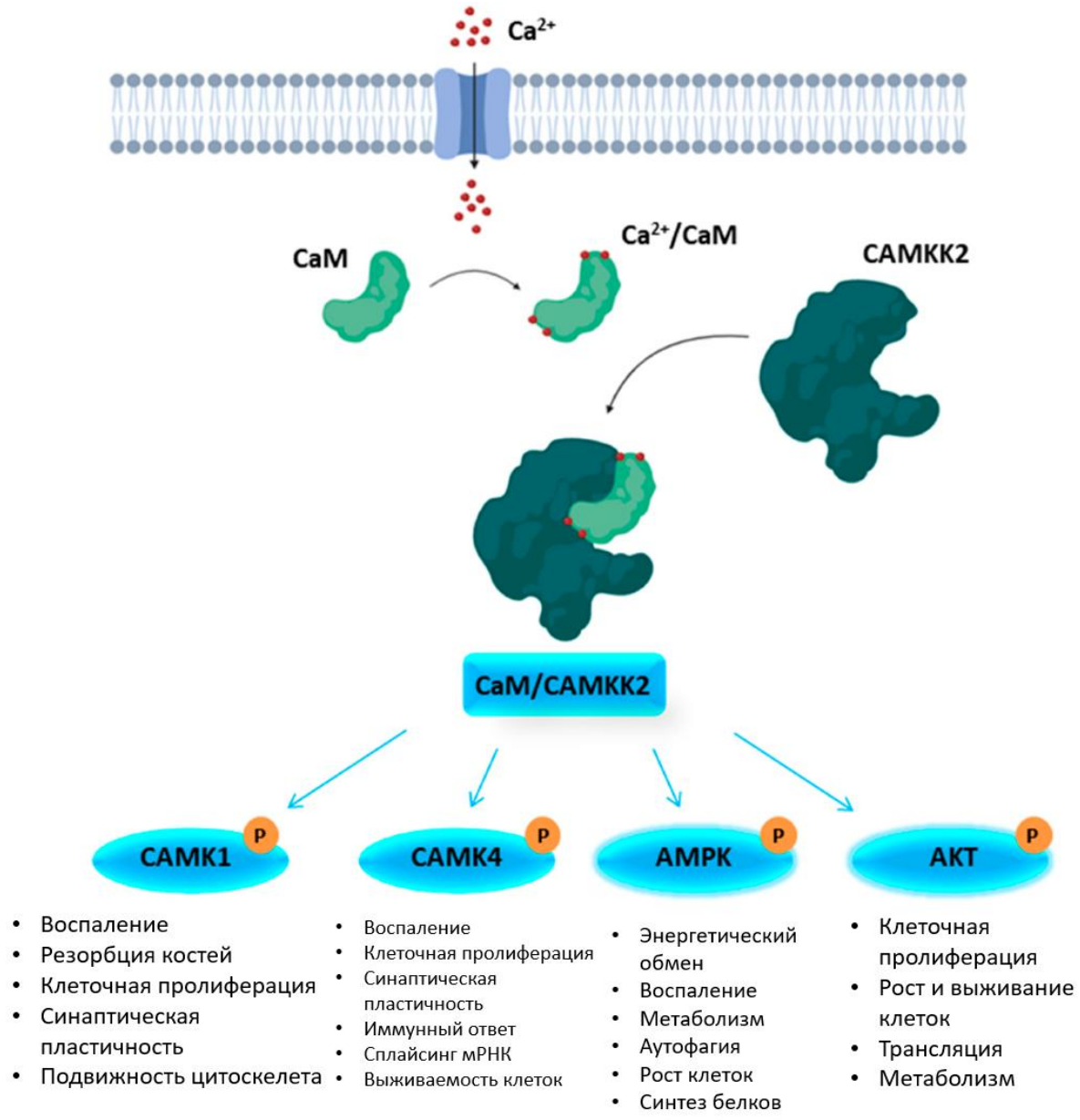


Рисунок 2. Сигнальный каскад и роль протеинкиназы CAMKK2 в физиологических и патологических процессах. Адаптировано из [61].

CAMKK2 гиперэкспрессируется во многих типах опухолей. Примерами являются рак предстательной железы, молочной железы, яичников, желудка и печени. Это связано с ее ролью в поддержке метаболической адаптации и выживании раковых клеток [61]. Такая гиперэкспрессия способствует активной пролиферации клеток, в связи с чем нокаут или ингибирование CAMKK2 в моделях на клетках и животных

приводят к существенному снижению опухолевой активности и их роста, что делает CAMKK2 привлекательной фармакологической мишенью для разработки новых противоопухолевых препаратов [79].

TTK (Thr/Tyr kinase), также известная как MPS1, является ключевой протеинкиназой, участвующей в регуляции контрольной точки сборки веретена (spindle assembly checkpoint, SAC) — критического механизма, обеспечивающего точное распределение хромосом при делении клеток [87].

TTK – это белок длиной 857 аминокислот с молекулярной массой около 97 кДа. Он состоит из из N-концевого каталитического домена и C-концевого регуляторного участка. TTK относится к киназам двойной специфичности, то есть способна фосфорилировать как тирозиновые, так и серин/треониновые остатки, что определяется особенностями строения активного центра. Переход между активной и неактивной конформацией осуществляется главным образом за счет изменения положения активационной петли и шарнирной области каталитического домена. В неактивном состоянии эти участки фиксируют каталитический домен в конформации, неблагоприятной для связывания АТФ и субстрата [30, 87].

TTK локализуется в кинетохорах и активируется в митотической фазе, что обеспечивает запуск сборки веретена деления. При этом TTK способствует формированию митотического контрольного комплекса, в который входят BUB3, BUBR1 и CDC20. Этот комплекс подавляет активность APC/C до того момента, когда все хромосомы будут корректно выровнены в плоскости метафазной пластинки (рисунок 3).

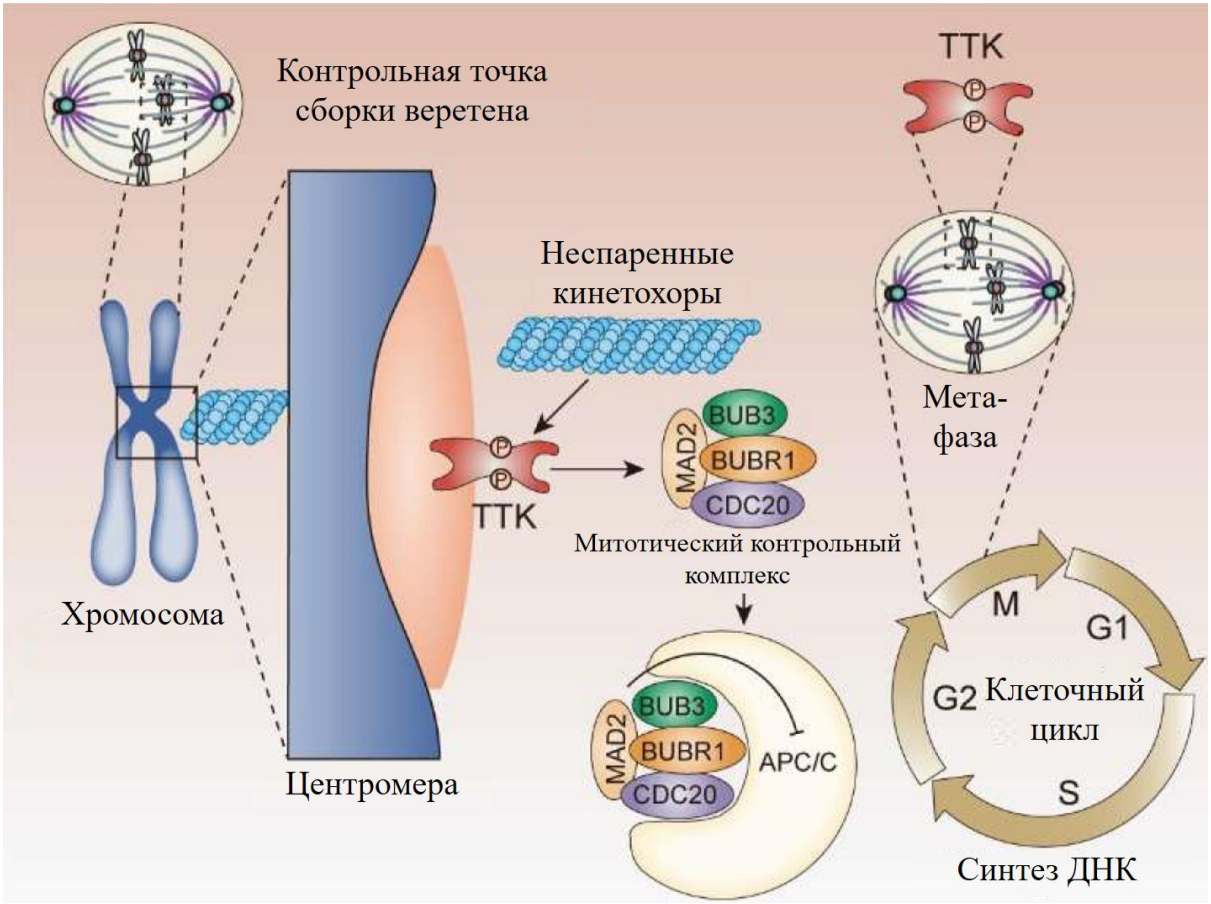


Рисунок 3. Сигнальный каскад протеинкиназы TTK. Адаптировано из [87].

TTK координирует процесс присоединения и ориентации хромосомных кинетохоров к митотическому веретену. Тем самым TTK препятствует преждевременному началу анафазы, пока все хромосомы не будут надежно закреплены. Это предотвращает возникновение анеуплоидии. Это состояние характеризуется неправильным числом хромосом. Оно часто наблюдается в злокачественных опухолях и приводит к геномной нестабильности [87].

Дисрегуляция активности TTK нередко выявляется в различных видах рака. Это коррелирует с агрессивным течением заболевания и неблагоприятным прогнозом. Повышенная экспрессия TTK связана с усиленным митотическим стрессом. Это способствует выживанию опухолевых клеток, несмотря на многочисленные хромосомные аномалии.

Именно поэтому селективные ингибиторы TTK считаются привлекательными терапевтическими агентами для противоопухолевой терапии. Они индуцируют митотический стресс, вызывают нарушения в клеточном цикле и приводят к апоптозу раковых клеток [30, 37].

TTK также играет важную роль в устранении ошибок прикрепления микротрубочек к кинетохору. TTK активирует сигнальные каскады, которые запускают механизмы исправления и обеспечивают сохранность генома [37]. Эта функция делает TTK незаменимым регулятором митоза. Данная киназа является важным фактором, поддерживающим геномную стабильность.

1.3. Современные подходы к поиску и разработке ингибиторов киназ

Современные методы поиска ингибиторов киназ базируются на комплексном использовании экспериментальных и вычислительных подходов. Это повышает точность и эффективность процесса разработки лекартсвенных молекул. Среди ключевых методик разработки лекарственных молекул следует отметить высокопроизводительный скрининг больших химических библиотек. Данный подход позволяет быстро определить перспективные соединения для биологического тестирования. Молекулярное моделирование и виртуальный скрининг позволяют прогнозировать возможные взаимодействия лигандов с белком на молекулярном уровне. Еще одним возможным инструментом является применение методов машинного обучения (ML). К ним, в частности, относятся алгоритмы на основе нейронных сетей. Они позволяют анализировать большие массивы данных, выявлять скрытые закономерности и предсказывать биологическую активность новых соединений на основе их химической структуры [5].

Одним из заметных достижений в области структурной биологии стало появление алгоритма AlphaFold. Данный алгоритм позволяет с высокой точностью моделировать трехмерные структуры белков [40]. Возможность предсказывать конформации белков с высокой точностью существенно повышает эффективность виртуального скрининга. Это важно при разработке новых молекул с высокой аффинностью и селективностью. Использование моделей для генерации трехмерной структуры белка позволяет применять структурно-обоснованный дизайн к фармакологическим мишеням с ранее неизученной трехмерной геометрией. Это позволяет улучшить качество оценки взаимодействия с лигандом и сократить количество ложноположительных результатов.

Фрагментный подход к поиску лекарств представляет собой одну из ключевых методологий в современной медицинской химии [23]. Данный подход направлен на конструирование новых ингибиторов посредством объединения фрагментов. Они представляют собой небольшие функциональные химические группы с высоким потенциалом связывания в пределах активного центра целевого белка. Принцип этого метода состоит в поэтапном подборе эффективных фрагментов, которые взаимно дополняют друг друга, обеспечивая высокую аффинность и селективность к целевым участкам белка. Такой подход позволяет не только повысить эффективность конструирования молекул-лигандов, но и существенно уменьшить экспериментальные затраты на стадии молекулярного дизайна за счет направленной генерации целевых соединений, оптимально подходящих для связывания с заданным участком белковой поверхности [22].

Преимуществом дизайна на основе фрагментов является его обоснованность с точки зрения межмолекулярных взаимодействий. В отличие от большинства известных методов, которые предлагают

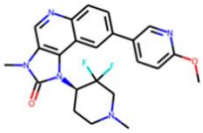
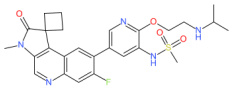
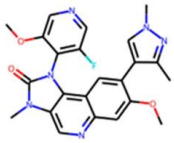
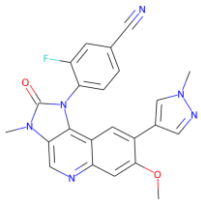
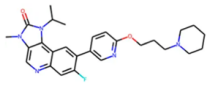
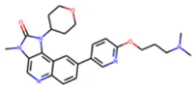
выполнить генерацию структур из гипотетически подходящего химического пространства, фрагментный подход позволяет выполнить конструирование молекулярных структур, наиболее оптимально подходящих для связывания в активном центре фермента. Это позволяет избежать анализа большого числа ложноположительных кандидатов и направленно сгенерировать молекулярные структуры под заданное макромолекулярное окружение.

В настоящее время активно ведутся исследования, направленные на разработку эффективных и безопасных ингибиторов для протеинкиназ, включая ATM, CAMKK2 и TTK.

1.3.1. Ранее охарактеризованные ингибиторы ATM

Несколько низкомолекулярных ингибиторов ATM разрабатываются и находятся на различных стадиях клинических испытаний, направленных на улучшение эффективности терапии, особенно при сочетании с лучевой и химиотерапией (таблица 1) [1, 55, 75]. Эти ингибиторы способны усилить эффект лечения за счет подавления репаративных механизмов опухолевых клеток, что делает исследования и разработку новых препаратов, нацеленных на ATM, актуальными. Несколько ранее обнаруженных ингибиторов ATM находятся на клинических исследованиях (таблица 1).

Таблица 1. Данные по ингибиторам ATM в клинических исследованиях.

КИ ID	ID соединения	Структура	Фаза КИ	Индикация	Компания	Годы
NCT05917145	WSD-0628		I	Рак мозга, мультиформная глиобластома	WayShine Biopharm	2024-2029
NCT05002140	XRD-0394		I	Солидные опухоли	XRad Therapeutics	2021-2023
NCT04882917	M-4076		I	Солидные опухоли	Merck KGaA	2021-2023
NCT03225105	M-3541		Прекращены	Солидные опухоли	Merck KGaA	2017-2020
NCT03423628	AZD-1390		I	Мультиформная глиобластома, саркома	AstraZeneca	2018-2026
NCT05678010						2023-2028
NCT02588105	AZD-0156		I	Солидные опухоли	AstraZeneca	2015-2022

WSD-0628 — это эффективный, селективный и перорально биодоступный ингибитор ATM, проникающий через гематоэнцефалический барьер (ГЭБ), что делает его особенно интересным для терапии опухолей ЦНС. В доклинических работах он показал выраженную радиосенсибилизацию при глиобластоме и метастазах в ЦНС,

а также увеличение выживаемости в комбинации с облучением. Авторы указывают, что была инициирована клиническая разработка в направлении терапии меланомы с метастазами в мозг. Его главное преимущество — сочетание проницаемости через ГЭБ, высокой селективности к ATM и хороших фармакокинетических свойств [67, 86].

XRD-0394 — двойной ингибитор киназ ATM и DNA-PK. Был разработан в качестве перорального радиосенсибилизатора и противоопухолевого препарата. В 2021 году была инициирована фаза Ia клинических испытаний у пациентов с метастатическими, локально-распространенными или рецидивирующими солидными опухолями в комбинации с лучевой терапией. Более поздняя публикация в 2024 году указала, что фаза Ia с лучевой терапией уже завершена. Препарат также показал синергизм с ингибиторами PARP и топоизомеразы I. Ключевое преимущество XRD-0394 — одновременное ингибирование киназ ATM и DNA-PK, который усиливает блокаду ответа на повреждение ДНК [27].

M3541 — субнаномолярный и высокоселективный ингибитор ATM с сильным радиосенсибилизирующим эффектом. Выполнено клиническое исследование фазы I с паллиативной лучевой терапией у пациентов с солидными опухолями. Была показана переносимость доз до 300 мг в день, однако оптимальное дозирование для последующих стадий установить не удалось, в связи с чем дальнейшие клинические исследования были прекращены. Клинические испытания были остановлены из-за неблагоприятного профиля фармакокинетики и фармакодинамики и отсутствия дозозависимого сигнала, в связи с чем программа разработки M3541 не получила дальнейшего развития [83, 95].

M4076 — высокоселективный, субнаномолярный ингибитор ATM из той же новой химической серии, что и M3541. В доклинических

исследованиях он эффективно подавлял репарацию двухцепочечных разрывов ДНК, усиливал противоопухолевый эффект облучения и демонстрировал синергию с ингибиторами PARP и топоизомеразы I. По опубликованным данным, M4076 находится в клинической разработке. Его преимуществами являются очень высокая селективность к ATM и хорошая пригодность для комбинирования с ДНК-повреждающими агентами и лучевой терапией [95].

AZD0156 — пероральный ATM-ингибитор с потенциалом повышения эффективности химио- и лучевой терапии. Проведена фаза I клинических испытаний у пациентов с солидными опухолями, где оценивались монотерапия и комбинации с олапарибом или комбинацией фторурацил + иринотекан. По доклиническим данным AZD0156 может избирательно усиливать радиочувствительность опухолевых клеток, включая меланому, при меньшем влиянии на здоровые фибробласты. Его преимущество — универсальность для комбинированных режимов, но он не обладает выраженной ЦНС-проницаемостью [63].

AZD1390 — ЦНС-проникающий, пероральный и очень селективный ингибитор ATM, изначально разработанный именно для внутричерепных опухолей. Является результатом модификации соединения AZD0156. В фазе I клинических исследований у пациентов с глиобластомой показаны приемлемая безопасность и предварительная эффективность. Его главное преимущество — сочетание высокой проницаемости через ГЭБ, высокой селективности и подтвержденной клинической переносимости в комбинации с лучевой терапией [63].

Таким образом, на данный момент на ранних стадиях клинических испытаний находится несколько ингибиторов ATM. Они схожи между собой по химической структуре, однако обладают различающимся

фармакокинетическим профилем, в частности разной проницаемостью через ГЭБ. Это определяет исследование рассмотренных соединений для разных нозологий, преимущественно в комбинации с существующими методами химио- и лучевой терапии.

Все указанные ингибиторы ATM имеют схожее химическое строение и характеризуются единообразным характером связывания с белком. Представленные соединения связываются в активном центре ATM и являются конкурентными ингибиторами ATM, мимикрирующими связывание АТФ. 1,3-дигидро-2H-имидазо[4,5-*c*]хинолин-2-оновый кор и его аналоги отвечают за связывание с шарнирным регионом, в том числе формируя водородную связь с Cys^{2770} [83, 95]. Подобное взаимодействие гетероароматических фрагментов с шарнирным регионом является классическим подходом к дизайну конкурентных ингибиторов киназ.

Одним из функциональных заместителей большинства представленных соединений является присоединенная к коровому фрагменту объемная алифатическая группа, такая как изопропил или морфолин. Данные функциональные группы направлены на выход из активного центра и не формируют ключевые взаимодействия с белком [67, 86].

Другим заместителем является пиридин с длинным подвижным линкером и терминальным ионизируемым атомом азота. Данная группа играет важную роль в связывание с аминокислотными остатками активного центра, включая Tyr^{2755}, Lys^{2717}, Asp^{2720} и Asp^{2725} [27, 63].

Имеющиеся данные свидетельствуют о том, что известные соединения являются близкородственными производными 1,3-дигидро-2H-имидазо[4,5-*c*]хинолин-2-она и его аналогов. Данный хемотип характеризуется высоким сродством к активному центру ATM, однако

поиск новых альтернативных хемотипов является перспективным направлением для расширения химического пространства ингибиторов ATM.

1.3.2. Ранее охарактеризованные ингибиторы CAMKK2

До настоящего времени было выделено и изучено несколько ингибиторов CAMKK2, среди которых наиболее изучены STO-609 и GSK650394, которые демонстрируют способность подавлять активность CAMKK2 и влиять на сигнальные пути, опосредованные этим белком [52, 65]. Эти молекулы обладают недостаточной селективностью и иногда вызывают побочные эффекты, что свидетельствует о необходимости поиска альтернативных ингибиторов. Соединения STO-609 и GSK650394 являются наиболее известными ингибиторами CAMKK2, в то же время они являются не клиническими кандидатами, а доклиническими инструментальными ингибиторами (рисунок 4).

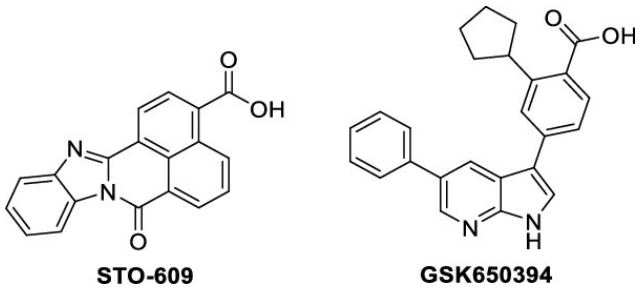


Рисунок 4. Структурные формулы соединений STO-609 и GSK650394.

STO-609 является классическим ингибитором кальций/кальмодулин-зависимой киназы киназы 2 (CaMKK2) и на протяжении длительного времени использовался в качестве одного из основных инструментов для изучения данного сигнального пути. Вместе с тем его применение ограничено недостаточной селективностью: соединение ингибирует не только CaMKK2, но и ряд других киназ, что повышает вероятность развития нежелательных побочных эффектов. В связи с этим STO-609

рассматривается прежде всего как экспериментальное химическое соединение, а не как перспективный кандидат для клинической разработки [52].

GSK650394 также относится к числу известных ингибиторов CaMKK2 и характеризуется выраженной ингибирующей активностью в отношении данного фермента. Однако его дальнейшее применение ограничивается неблагоприятным профилем селективности: в исследованиях перекрестного скрининга было показано, что соединение обладает активностью в отношении значительного числа других киназ, что существенно снижает его ценность как селективного ингибитора CaMKK2. В литературе GSK650394 рассматривается главным образом как исходная химическая структура для дальнейшей оптимизации, а не как препарат, достигший клинического этапа разработки [69].

Соединение GSK650394 является конкурентным ингибитором CAMKK2, связываясь в активном центре фермента. Азаиндольный кор соединения GSK650394 связывается с шарнирным регионом, формируя сеть водородных связей с аминокислотами Glu^{268} и Val^{270}. Бензольное кольцо направлено в сторону выхода из активного центра, формирую гидрофобные контакты с остатками Ile^{171} и Pro^{274}. Циклопентилбензойная кислота также формирует ряд ключевых контактов с остатками активного центра. Так, бензольное кольцо в его составе участвует в π-стэкинге с Phe^{267}, циклопентановый фрагмент формирует гидрофобные контакты с Ala^{329} и Gly^{174}, тогда как карбоксильная группа образует ионное взаимодействие с боковой цепью Lys^{194} [69]. Данный набор межмолекулярных взаимодействий обеспечивает высокую аффинность GSK650394 к CAMKK2 и определяет перспективы его использования в качестве прототипа для дальнейшего дизайна новых ингибиторов этой мишени.

1.3.3. Ранее охарактеризованные ингибиторы ТТК

Благодаря своей центральной роли в клеточном цикле и в контроле выполнения митоза ТТК стала объектом значительного интереса с точки зрения фармакологического воздействия. В настоящее время разрабатываются и тестируются несколько селективных ингибиторов этой киназы, которые уже продемонстрировали перспективность в *in vitro* тестах, доклинических моделях, а также в ранних клинических испытаниях (таблица 2, рисунок 5). Их применение направлено на преодоление лекарственной устойчивости и снижение агрессии опухолей за счет нарушения жизненно важных процессов клеточного деления [37, 45, 51].

Таблица 2. Данные по ингибиторам ТТК в клинических исследованиях.

КИ ID	ID соединения	Фаза КИ	Индикация	Компания	Годы
NCT02366949	BAY-1217389	I	Солидные опухоли	Bayer	2015-2019
NCT02138812	BAY-1161909	Прекращены	Солидные опухоли	Bayer	2014-2017
NCT03328494	BOS-172722	I	Негематологические злокачественные новообразования	Boston Pharmaceuticals	2017-2021
NCT02792465	CFI-402257	I	Солидные опухоли, рак молочной железы	University Health Network	2016-2027
NCT03568422		II	Рак молочной железы	Canadian Cancer Trials Group	2019-2026
NCT03411161	S81694	I	Метастазирующий рак молочной железы	Servier	2018-2020
		II			

NMS-P715 является первым селективным перорально биодоступным ингибитором TTK, разработанным в качестве химического зонда для изучения функции киназы. Соединение обладает IC_{50} = 182 нМ в отношении TTK и селективно подавляет рост опухолевых клеток, практически не влияя на нормальные клетки. В клеточных моделях NMS-P715 индуцирует апоптоз в клетках рака яичников, а также повышает радиочувствительность клеток глиобластомы. На ксенографтных моделях соединение демонстрирует 50% ингибирование роста опухолей яичников и меланомы. Клинических испытаний для NMS-P715 не проводилось: соединение находится на доклиническом уровне как инструментальный ингибитор TTK и послужил прототипом для дальнейшей оптимизации [81].

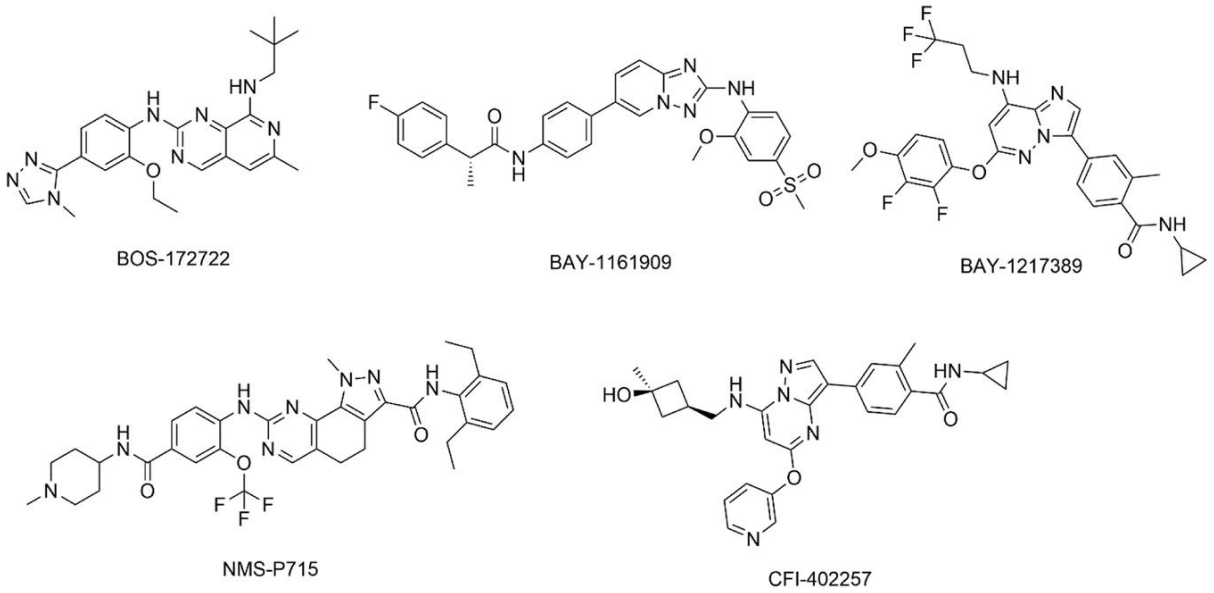


Рисунок 5. Структурные формулы ингибиторов TTK.

S81694 (**NMS-P153**) представляет собой эффективный и селективный ингибитор TTK, разработанный Nerviano Medical Sciences и Servier. Соединение является перорально биодоступным, вызывает хромосомный дисбаланс, запускает апоптоз и гибель раковых клеток [15]. Клиническая разработка проводилась в фазе I с 2015 по 2019 годы: зарегистрированные исследования оценивали безопасность и переносимость S81694 в

комбинации с паклитакселом при солидных опухолях. Servier принял стратегическое решение остановить развитие **S81694**, соединение не переходило в фазы II/III.

BOS-172722 (**CCT289346**) представляет собой эффективный, селективный и перорально биодоступный ингибитор TTK, разработанный на основе серии, включающей S81694. Соединение демонстрирует IC_{50} = 11 нМ в биохимических тестах и 63 нМ в клеточных моделях и обладает 100-кратной селективностью относительно CDK2. Основной целевой индикацией является тройной негативный рак молочной железы [2]. BOS-172722 в комбинации с паклитакселом показывает выраженную антипролиферативную активность в доклинических моделях, включая паклитаксел-резистентные опухоли. В рамках фазы I клинических испытаний были оценены безопасность и переносимость как монотерапии, так и комбинации с паклитакселом при раке молочной железы и раке яичников.

BAY-1161909 — пероральный ингибитор TTK на основе триазолопирдиновой серии с IC_{50} = 0,34 нМ. Был разработан компанией Bayer для терапии солидных опухолей. Молекула демонстрирует селективность по широкому спектру киназ, индуцирует полинуклеарность в клеточных моделях. В доклинических исследованиях в комбинации с паклитакселом BAY-1161909 превосходит монотерапию, включая паклитаксел-резистентные модели, и хорошо переносится [70]. Клиническое исследование было прекращено в пользу соединения BAY-1217389.

BAY-1217389 — клинический кандидат из имидазопиразиновой серии, селективный ингибитор TTK с IC_{50} < 10 нМ, разработанный Bayer как модификация BAY-1161909 в рамках оптимизации соединений с

изначально низкой метаболичной стабильностью. В доклинических исследованиях соединение демонстрирует умеренную эффективность в моно-терапии, но в комбинации с паклитакселом значительно усиливает противоопухолевую активность, включая паклитаксел-резистентные опухоли [70].

CFI-402257 — перорально биодоступный и селективный ингибитор TTK с IC_{50} = 1,7 нМ, обладающий высокой эффективностью в отношении рака молочной железы, яичников, простаты, поджелудочной железы, легких и толстой кишки. Соединение селективно относительно 262 киназ при 1 мкМ, увеличивает анеуплоидию в клетках HCT116 и снижает рост опухолей [12]. Клиническая разработка находится на фазе I/II: зарегистрировано исследование NCT03568422 по оценке безопасности и противоопухолевой активности CFI-402257 в комбинации с паклитакселом при раке молочной железы.

Указанные ингибиторы TTK являются конкурентными, связываясь в активном центре фермента. Разберем ключевые взаимодействия с мишенью на примере соединения BAY-1217389. Имидазо[1,2-*b*]пиридазин участвует в связывании с шарнирным регионом, формирую водородную связь с основной цепью Gly^{605}. Трифторпропиламино-группа направлена в сторону выхода из активного центра, находясь в окружении остатков Gln^{541}, Lys^{529} и Ile^{531}. Дифторметоксифенил расположен в гидрофобной полости, формируемой остатками Ala^{651}, Pro^{673} и Met^{671}. Метилбензамидный фрагмент с циклопропилом также формирует ряд выгодных межмолекулярных контактов. В частности, метилбензол участвует в образовании гидрофобных контактов с боковыми цепями Met^{602} и Ile^{663}, амидная группа образует водородную связь с основной цепью Ile^{663}, тогда как циклопропильный фрагмент формирует гидрофобные взаимодействия с Met^{600} и Leu^{575} [70].

Таким образом, активный центр протеинкиназы TTK является хорошо изученным, так же, как и межмолекулярные контакты, определяющие высокую аффинность известных ингибиторов данной фармакологической мишени. Это позволяет выполнять рациональный структурно-обоснованный дизайн и облегчает разработку новых ингибиторов TTK.

1.4. Проблемы и ограничения современных вычислительных методов

Современное моделирование лекарств основано на двух ключевых подходах: структурно-обоснованный дизайн лекарств (Structure-Based Drug Design, SBDD) и дизайн лекарств на основе лигандов (Ligand-Based Drug Design, LBDD). В случае дизайна ингибиторов малоизученной биомишени информация о потенциальных лигандах часто либо недостаточна, либо полностью отсутствует. При этом доступной для исследователей оказывается только трехмерная модель целевого белка [47]. В таких ситуациях SBDD является единственным возможным подходом для решения задачи дизайна новых молекулярных структур. С накоплением доступных кристаллографических данных, размещаемых в базе Protein Data Bank [6], важность методов SBDD приобретает все большее значение.

Начальным этапом в рамках SBDD является анализ трехмерной структуры биомишени и выявление участка, пригодного для связывания малой молекулы. Идентификация лиганд-белковых сайтов связывания расширяет число доступных белковых мишеней и создает новые перспективы для дизайна модуляторов. Биофизические методы обнаружения сайтов связывания включают направленный мутагенез и скрининг фрагментов. Однако данные подходы являются дорогостоящими и продолжаются длительное время.

На сегодняшний день существует множество вычислительных методов для обнаружения сайтов связывания лиганд-белок. Известны методы, основанные на сходстве строения (ProBiS), геометрии (FPocket), методах компьютерного зрения (BiteNet, DeepSite) и другие [39]. Несмотря на обилие доступных решений, их прогностическая способность несколько ограничена. Большинство ранее описанных методов способны обнаружить не более 40% сайтов связывания лиганд-белок [41]. Современные алгоримты с использованием машинного обучения не показали значительного прироста точности по сравнению с классическими компьютерными методами. Наиболее сложной задачей остается обнаружение ранее неохарактеризованных, поверхностных и аллостерических полостей, а также сайтов связывания ковалентных ингибиторов и полостей на интерфейсе белок–белок (PPI) [20].

Другим направлением, которое остается мало изученным с точки зрения применения методов машинного обучения, является генерация псевдолигандов.

Генерация псевдолигандов выполняется путем подбора наиболее подходящих типов атома для точек пространства внутри сайта связывания лиганд-белок. Данная методика обладает значительным спектром применения как в молекулярной биологии, так и в структурно-обоснованном дизайне лекарств. Примерами ее применения являются классификация сайтов связывания, предсказание лиганд-белковой аффинности и генерация новых структур биологически активных соединений. В частности, решение последней задачи может быть достигнуто путем использования псевдолигандов в качестве входных облаков точек для диффузионных нейронных сетей или через атом-обоснованный подход, когда специальная модель применяется для

предсказания наиболее подходящих типов атомов в данном макромолекулярном окружении.

Традиционные решения без применения ML-алгоритмов базируются на расчетах энергии взаимодействий [85]. Хотя некоторые ранее охарактеризованные генеративные подходы использовали ML для выбора подходящих типов атома для генерации структур малых молекул, прямое влияние этих методов на точность генерации псевдолигандов не было ранее изучено.

Другой ключевой проблемой в разработке лекарств является предсказание сродства разрабатываемого ингибитора к целевой биомишени. Резкое изменение активности при малых изменения структуры, многообразие межмолекулярных взаимодействий и несоответствие доступных данных из разных источников друг другу приводят к низкой точности предсказаний [44]. Хорошо известны методы расчета аффинности, реализованные в программном обеспечении докинга, таких как MOE, Glide Schrödinger [25], AutoDock и его модификации [18], LeadFinder [60] и других. Множество исследований сфокусированы в настоящее время на разработке ML-алгоритмов для ускорения и повышения точности оценки. К недавно появившимся алгоритмам относятся InteractionGraphNet (IGN) [38] и graphLambda [57].

Другим важным направлением является виртуальная генерация структур новых лекарственноподобных соединений. Генеративная химия является новой тенденцией в разработке лекарств, постепенно занимающей место наряду с виртуальным скринингом существующих соединений [32]. Большинство современных платформ для виртуального дизайна лекарственных молекул используют решения на основе глубокого обучения (DL). Несмотря на то, что алгоритмы DL являются ценным инструментом,

они часто характеризуются отсутствием интерпретируемости, что затрудняет понимание причины принимаемых решений. Недавно опубликованные алгоритмы также демонстрируют недостаток понимания в области медицинской химии, что приводит к генерации химически невалидных, нестабильных, токсичных, неактивных или ранее опубликованных структур [34].

В то же время на сегодняшний день отсутствуют полностью автоматизированные алгоритмы, прошедшие валидацию *in vitro*, которые могли бы генерировать структуры малых молекул, учитывая медицинскую химию и взаимодействия лиганда с белком. Виртуальный дизайн лекарственных соединений на основе фрагментов (virtual fragment based drug design, VFBDD) — подход, существующий уже несколько десятилетий [43]. Он обладает значительным потенциалом, но долгое время оставался недооцененным и недостаточно используемым в фармацевтической индустрии. Одним из ключевых преимуществ VFBDD является возможность эффективно исследовать большое химическое пространство с помощью относительно небольшого количества фрагментов. Фокусировка на небольших молекулярных фрагментах позволяет более эффективно выполнять скрининг и оптимизацию потенциальных кандидатов. Недавними примерами подобных алгоритмов являются MoleGear [14], AutoGrow и GenUI [76].

Существующие методы VFBDD имеют ряд ограничений, которые препятствуют их широкому применению. Многие известные методы разрабатывались для модификации скэффолдов, заранее расположенных сайте связывания, и не подходят для *de novo* дизайна. Другие проблемы включают низкое разнообразие генерируемых структур, создание химически нежелательных с точки зрения медицинской и органической химии соединений, высокие вычислительные требования. При работе с

большими библиотеками фрагментов это приводит к проблеме комбинаторного взрыва и неспособности исследовать требуемое химическое пространство. Еще одной проблемой является отбор хитов, требующий точной оценки аффинности связывания, синтетической доступности, лекарствоподобия и других критериев.

1.4.1. Ограничения метода AlphaFold 2

AlphaFold 2 — это алгоритм, разработанный компанией DeepMind, который применяет методы машинного обучения для предсказания трехмерной структуры белков по их аминокислотной последовательности. Благодаря использованию накопленных данных об известных белковых структурах и современных нейронных сетей с механизмом внимания (attention), AlphaFold 2 позволяет анализировать связи между аминокислотными остатками и прогнозировать их пространственное расположение. Это позволяет воспроизводить сложные процессы фолдинга белка. AlphaFold 2, продемонстрировал выдающиеся результаты в соревновании CASP-14. Данный алгоритм достиг точности предсказания, сравнимой с экспериментальными методами. Погрешность составила порядка 1–1,6 Å в RMSD для большинства белков, включая те, для которых отсутствуют близкие структурные гомологи.

Данный алгоритм рассматривает белок как граф, где вершинами являются аминокислотные остатки, а ребрами — их взаимосвязи. Это позволяет учитывать эволюционные и структурные взаимозависимости. Входные данные включают не только аминокислотную последовательность, но и информацию о множественном выравнивании похожих белков и вероятностных взаимодействиях между остатками. Обучение сети проводилось на базе свыше 170 тысяч экспериментально

определенных белковых структур, что обеспечивает высокую обобщающую способность модели.

AlphaFold 2 существенно ускорил и упростил процесс получения точных трехмерных моделей. Это критически важно для понимания биологических функций белков и анализа механизмов их действия. Также это может быть использовано для разработки новых лекарственных молекул. Классические биофизические методы занимают недели и месяцы. В то же время, AlphaFold 2 позволяет получить предсказание трехмерной структуры лиганд-белкового комплекса в течение нескольких минут.

Несмотря на революционный характер и высокую точность, AlphaFold 2 имеет определенные ограничения. Он моделирует статичную структуру белка без учета его динамики, конформационных изменений и влияния растворителей. Также не учитывается влияние кофакторов, взаимодействие с другими белками и малыми молекулами. В связи с этим, AlphaFold 2 имеет ограниченное применение в медицинской химии и разработке лекарств, несмотря на значительные успехи применения данного алгоритма в области структурной биологии.

В области медицинской химии точное знание трехмерной структуры белковой мишени является основой для успешного применения методов структурно-обоснованного дизайна. AlphaFold 2 позволяет получать трехмерные модели целевых белков, отсутствующие в базе данных Protein Data Bank (PDB). Это значительно расширяет домен применимости структурно-обоснованного виртуального скрининга. Такой подход повышает вероятность успешного рационального дизайна новых биологически активных соединений. Алгоритм использует глубокие нейронные сети и эволюционную информацию, полученную из множественных выравниваний последовательностей. Это позволяет

предсказывать двугранные углы и пространственные координаты атомов. При этом особое внимание уделяется взаимозависимости аминокислотных остатков, находящихся в непосредственной близости друг от друга. Это обеспечивает высокую точность прогноза геометрии пептидного остова, что важно для определения общей трехмерной структуры белка.

При использовании AlphaFold 2 в задачах медицинской химии наблюдаются определенные ограничения. Они могут выступать существенным ограничением при выполнении дизайна новых биологически активных соединений. Точность предсказания положения боковых цепей аминокислот, формирующих структуру сайта связывания, часто недостаточна для надежного проведения молекулярного докинга и виртуального скрининга. Пространственное размещение боковых групп играет критическую роль в формировании взаимодействий с лигандами, и малейшее отклонение может существенно исказить оценку сродства связывания. Кроме того, AlphaFold 2 прогнозирует статичную структуру, не учитывая конформационную подвижность белков. Конформационная подвижность имеет важное значение для отображения динамических изменений и альтернативных функциональных состояний белка. Такие изменения часто встречаются при связывании с лигандами. Как в случае ортостерических, так и аллостерических сайтов связывания. Отсутствие в моделях таких аспектов снижает применимость AlphaFold 2-моделей для реалистичного воспроизведения биологических процессов и усложняет задачи виртуального скрининга [33].

Также модели AlphaFold 2 не включают в себя кофакторы, молекулы растворителя и ионы металлов, которые могут существенно влиять на структуру и функцию активного центра, а также на связывание низкомолекулярных лигандов. Отсутствие этих составляющих может привести к ошибкам при оценке комплементарности лигандов, потере

активных соединений и возникновению ложноположительных результатов. Опираясь на доступные данные, AlphaFold 2 недостаточно хорошо подходит для полноценной подготовки трехмерных моделей белков с целью поиска новых биологически активных соединений. Однако, применение данного алгоритма может служить исходной точкой для дальнейшей оптимизации моделей, например, с помощью молекулярной динамики или методов гибкого докинга, что помогает компенсировать некоторые из перечисленных ограничений. Таким образом, несмотря на важный вклад AlphaFold 2 в структурную биологию и перспективы для медицинской химии, его применение в медицинской химии необходимо сочетать с использованием дополнительных вычислительных и экспериментальных методов и критической оценкой предлагаемых трехмерных структур.

Следует отметить, что за последнее время появились новые методы [3], включая те, которые прогнозируют трехмерную структуру комплексов лиганд-белок. К таким методам относятся, например, RosettaFold-All-Atom, AlphaFold 3, Boltz 2. Ограничения подобных методов еще мало изучены в связи с их новизной, однако с методологической точки зрения их применение должно решать часть проблем, описанных в данном разделе. В то же время, одним из достаточно частых источников трехмерных белковых моделей в настоящий момент служит база данных AlphaFold Protein Structure Database, полученная именно методом AlphaFold 2. В связи с этим важно понимать ограничение этого метода и полученных с его помощью белковых моделей, а также возможные пути их улучшения.

Таким образом, современные ограничения и вызовы применения AlphaFold 2 для задач медицинской химии включают ряд проблем, сформулированных ниже.

1. Низкая точность позиционирования боковых цепей в активном центре. AlphaFold 2 с высокой точностью предсказывает общую геометрию пептидного остова, но позиционирование боковых цепей аминокислот, что особенно важно в зоне активного центра, часто неоптимально для связывания лиганда. Ошибки в пространственной геометрии даже одного аминокислотного остатка могут привести к неправильной оценке сродства лиганда и потере эффективности молекулярного докинга.
2. Отсутствие учета конформационной подвижности аминокислотных остатков. Белки способны часто менять конформацию, в том числе в результате связывания малых молекул. AlphaFold 2-модели часто ближе к апо-формам, что затрудняет точное воспроизведение конформаций, необходимых для связывания лигандов, особенно при докинге. Моделирование аллостерических сайтов, формирующихся только в определенных конформациях, остается сложной задачей.
3. Отсутствие в моделях необходимых коферментов, молекул растворителя и ионов металлов. AlphaFold 2 не предсказывает положение молекул, не связанных с полипептидной цепью. Кофакторы, металлы и молекулы растворителя играют важную роль в стабилизации активных сайтов и обеспечении каталитической активности. Кроме того, они могут выполнять важную задачу по опосредованию лиганд-белковых взаимодействий. Это ограничивает применение моделей для выполнения докинга и дизайна биологически активных соединений.
4. Применение трехмерных белковых моделей из AlphaFold 2 для виртуального скрининга часто дает более слабые результаты с точки зрения доли обнаруженных активных соединений. Полученные в AlphaFold 2 белковые модели обычно уступают кристаллографическим структурам и гомологическим моделям по

качеству ранжирования и обнаружению активных соединений. Изучено применение дополнительной оптимизации белковых моделей методом гибкого докинга (induced fit docking, IFD-MD) или молекулярной динамики. Такие подходы существенно улучшают результаты, но требуют значительных вычислительных ресурсов. Это ограничивает их применение для виртуального скрининга обширных библиотек соединений.

5. Проблемы с моделированием сложных белковых комплексов и взаимодействий лиганд-белок. Несмотря на развитие новых версий AlphaFold (например, AlphaFold-3), обеспечивающих улучшенную точность предсказаний позиций как лигандов, так и взаимодействующих белков, сложные белковые системы остаются актуальной проблемой. Некоторые случаи разработки лекарств требуют более комплексных решений. Дизайн биологически активных соединений, нацеленных на связывание на интерфейсе взаимодействия нескольких белков, требует дальнейшего изучения.
6. Отсутствие биологической валидации выполненных виртуальных экспериментов. В большинстве исследований, использующих AlphaFold 2 для генерации белковых моделей, отсутствуют экспериментальные подтверждения прогнозов. Это затрудняет объективную оценку пригодности моделей для выполнения органического синтеза и биологического тестирования соединений.

1.4.2. Ограничения современных генеративных методов

Исторически поиск новых лекарственных молекул был исходно скорее процессом интуитивным и эмпирическим. С появлением высокопроизводительного скрининга удалось рационализировать этот

процесс и повысить эффективность лекарственной разработки. Затем в дополнение к нему появился виртуальный скрининг, без которого сейчас не обходится ни один проект по дизайну новых биологически активных соединений. На данный момент активно развиваются и внедряются в практику методы генеративной химии, позволяющие целенаправленно генерировать структуры малых молекул для воздействия на заданную мишень. Методы генеративной химии нацелены на создание новых молекул с желаемыми свойствами: высокой активностью, оптимальными показателями всасывания, распределения, метаболизма, выведения и фармакокинетики (ADME/PK), а также простотой синтеза с помощью автоматических или полуавтоматических методов. Такой подход помогает медицинским химикам и способствует обнаружению новых лекарственных соединений, многие из которых на данный момент находятся на различных этапах клинических исследований.

В научной литературе описаны различные алгоритмы на основе машинного обучения, включая сверточные нейронные сети (RNN) [11, 29], вариационные автоэнкодеры (VAE) [28], генеративные состязательные сети (GAN) [7], трансформеры и гибридные методы обучения с подкреплением (RL) [93]. Некоторые из них уже интегрированы в рабочие процессы ранней разработки лекарственных молекул и способствовали обнаружению новых биологически активных соединений.

Часто для оценки моделей используются метрики, такие как охват химического пространства, доля валидных сгенерированных химических представлений — SMILES-строк и 2D-структур, разнообразие и «новизна» (отличие от обучающей выборки). Несмотря на это, медицинские химики больше заинтересованы в реально синтезированных, современных, насыщенных атомами углерода в sp^3-гибридизации и лекарственноподобных соединениях без нежелательных структурных

фрагментов с низкими значениями концентрации полуингибирования (IC_{50}) и хорошими ADME/PK характеристиками. Поэтому для специалистов в области медицинской химии важна экспериментальная проверка, то есть синтез и биологическое исследование виртуально сгенерированных соединений.

Следует рассмотреть некоторые из недавно опубликованных работ с экспериментальной валидацией. Большинство генеративных моделей выполняют поставленную задачу и генерируют много виртуальных структур, но редко соответствуют всем требованиям современной разработки лекарств. Часто сгенерированные молекулярные структуры обладают высоким структурным сходством с обучающими данными, что ставит под сомнение патентоспособность сгенерированных структур.

Одним из примеров является разработка Джанг и соавторов малой молекулы PCW-A1001 [36], нацеленной на киназу FLT3 (рисунок 6). Исходное соединение PCW-1001 показало IC_{50} 13,6 и 1,83 мкМ против FLT3 и мутанта D835Y соответственно. Для оптимизации PCW-1001 использовали глубокую сверточную нейронную сеть с алгоритмом LSTM. Сгенерированные аналоги оценивались на токсичность, синтетическую доступность, новизну и лекарствоподобие, а затем проходили докинг и экспертную оценку. Соединения были синтезированы и протестированы *in vitro*. Лучшее соединение PCW-A1001 продемонстрировало IC_{50} 2,54 мкМ и 764 нМ против FLT3 WT и FLT3 D835Y соответственно, показав перспективу для дальнейшей разработки.

PCW-1001 тестировали *in vitro* против панели из 104 киназ при концентрации 10 мкМ. Он ингибировал активность TRKA (~55%), FLT (~45%), MELK (~35%), FGFR (~35%) и PKC (~35%), в то время как активности других киназ, включая MAPK2 и ALK, оставались без

изменений. Основной мишенью считался TRKA с IC_{50} 5,8 мкМ. Клеточные исследования подтвердили подавление TRKA-зависимых сигнальных путей.

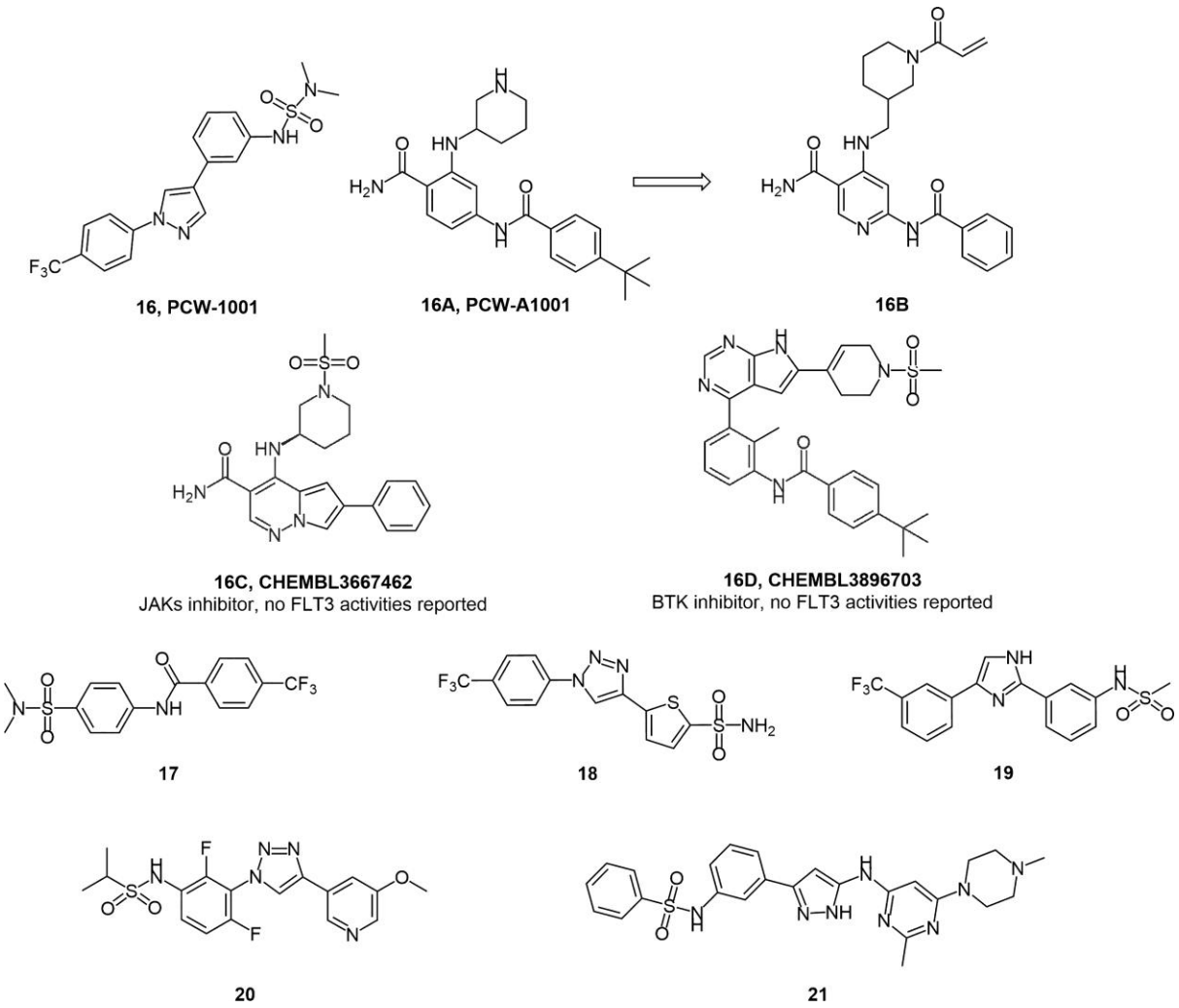


Рисунок 6. Сгенерированная структура PCW-A1001 и его сравнение с соединениями PCW-1001, 16B-D и 17-21.

Структура PCW-1001 обладает структурным и топологическим сходством по крайней мере с пятью ранее описанными соединениями 17-21. В свою очередь, структура PCW-A1001 похожа на известный ингибитор JAK 16B с 2-амино-4-бензамидным фрагментом и их изостерными аналогами. По всей видимости, процесс генерации опирался на

трансформации этих структур путем комбинации с другими структурами, включая соединения 16C и 16D.

Данные по способности PCW-A1001 ингибировать другие киназы и примеры сгенерированных структур не предоставлены. Также отсутствует информация о новизне, лекарствоподобии и потенциальной токсичности сгенерированных соединений. Несмотря на умеренную *in vitro* активность, PCW-A1001 можно считать перспективным соединением среди низкомолекулярных ингибиторов FLT3.

Йошимори и соавторы применили глубокую генеративную модель с трехмерным фармакофорным моделированием для поиска новых ингибиторов DDR1 [88]. Было сгенерировано 570542 структуры, затем из них было отобрано 10694 по фармакофорному поиску и далее 4731 — по расчетной энергии связывания. Девять структур были отобраны, синтезированы и протестированы; при этом три из них были модифицированы вручную (рисунок 7). Три соединения показали субмикромолярную активность. Метод продемонстрировал себя как подходящий для поиска первичных хитов и замены скэффолда.

Менее 1% всех сгенерированных структур прошли фильтрацию по оценке расчетной аффинности и фармакофорной фильтрации, что косвенно указывает на низкую эффективность генеративного метода. Три из девяти финальных соединений были вручную модифицированы, поэтому реальный потенциал модели генерировать активные структуры сложно оценить. Например, структура 72A получена из 72 изменением положения ароматического азота в пиридиновом кольце, чтобы сформировать водородную связь с шарнирным регионом — ключевым взаимодействием для киназных ингибиторов. Однако соединение 72A показало лишь

микромолярную активность, поэтому маловероятно, что соединение 72 будет активно.

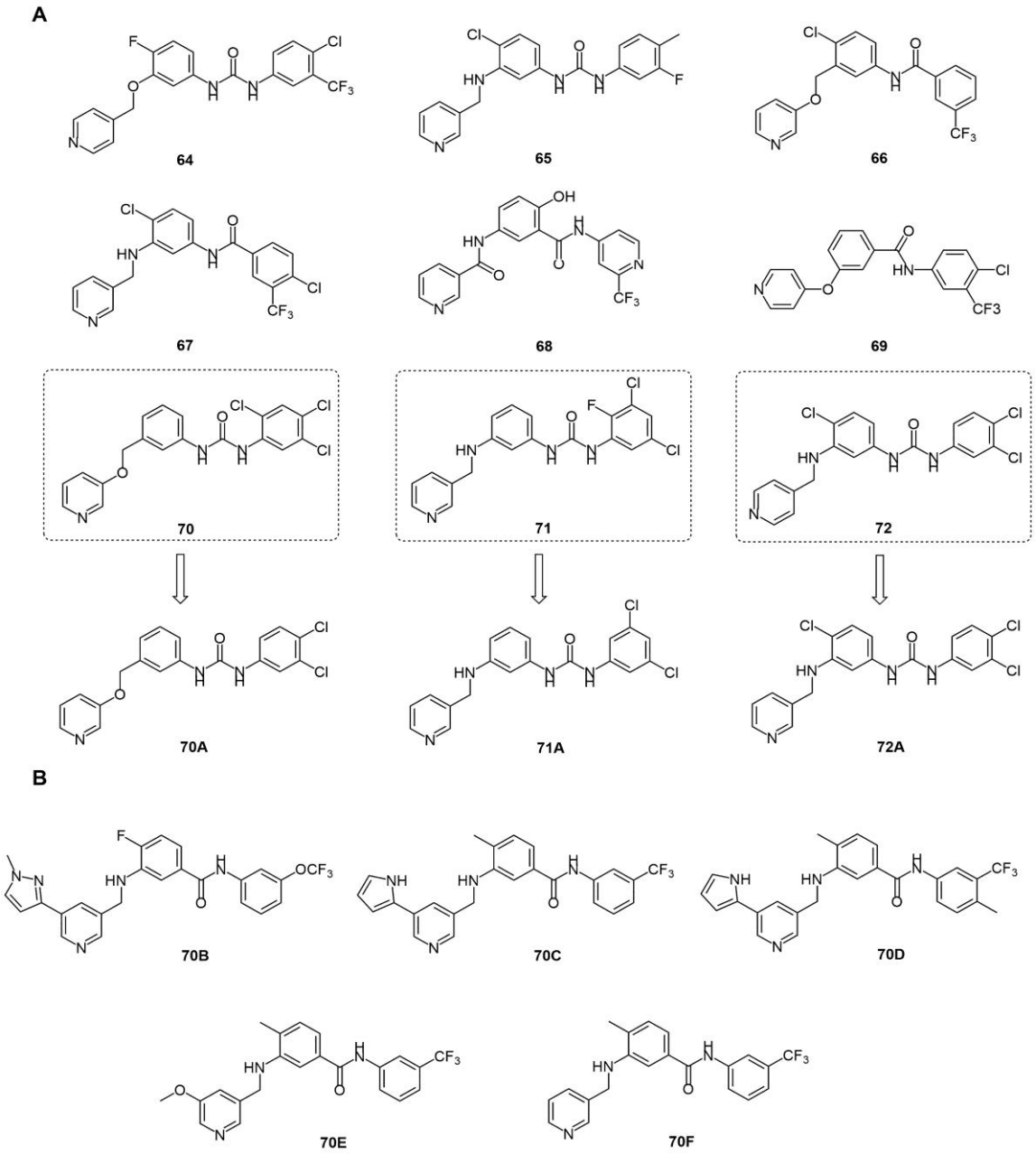


Рисунок 7. Структуры соединений, протестированные на предмет ингибирования DDR1и их ранее описанные аналоги.

Из протестированных соединений три (66, 67, 70A) показали умеренную субмикромолярную активность (IC_{50} 92,5, 186,7 и 171,3 нМ), однако один из этих ингибиторов (70A) получен с ручной доработкой.

Предположительно, удаление вручную хлора в *орто*-положении фенилмочевины является важным фактором, так как он ограничивает вращение и может препятствовать получению конформации, необходимой для связывания в случае исходного соединения 70.

Ранее были описаны низкомолекулярные ингибиторы DDR1 70B-70F с таким же скэффолдом. Скэффолд на основе 3-бензилоксипиридина, присутствующий в выбранных соединениях, известен более 10 лет как характерный для ингибиторов киназ. Таким образом, несмотря на полученные успехи, предложенный подход привел к получению серии структур с низким химическим разнообразием и высоким сходством с ранее описанными соединениями, при этом лишь малая часть сгенерированных молекул проявила активность против DDR1, что свидетельствует о необходимости дальнейшего развития данного метода.

Хуа и соавторы применили сверточную нейронную сеть для виртуальной генерации ингибиторов MerTK на основе комбинаторного синтеза [31]. Как уже было отмечено ранее, учет синтетического контекста в генеративной химии остается проблемой — синтетическая доступность новых структур часто не гарантируется.

Подход генерации новых реагентов с использованием сверточной нейронной сети, исходя из уже доступных коммерческих исходных материалов, кажется нецелесообразной, поскольку главная ценность известных реагентов — быстрая доступность, которую сложно обеспечить через генерацию реагентов. Из 20 000 сгенерированных структур была синтезирована только малая часть — 17 соединений успешным синтезом, но синтетическая доступность остальных не ясна. Малейшие структурные изменения коммерчески доступных реагентов могут сделать соединение недоступным.

Полученные в ходе докинга и анализа библиотеки пять структурных фрагментов (73-77, рисунок 8) не являются новыми: индазольные и пирролопиримидиновые структуры встречаются в сотнях патентов, включая киназные ингибиторы. В частности, соединение 81 почти полностью совпадает (98% сходства) с ранее опубликованным ингибитором JAK (82) с единственной заменой циклобутанового заместителя на аминоциклогексан. В связи с этим способность предложенного подхода генерировать новые структуры остается неподтвержденной.

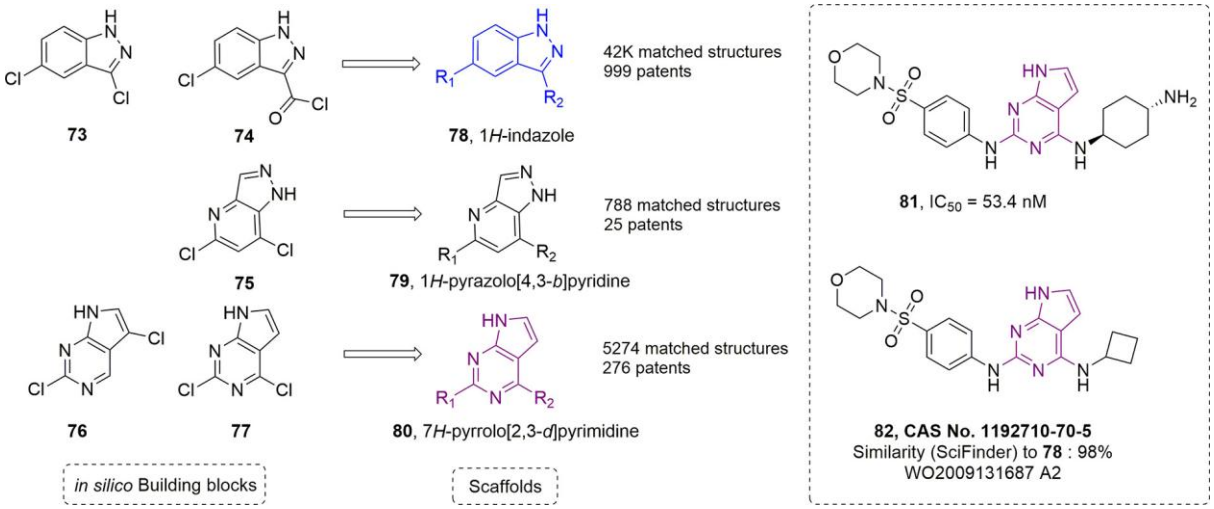


Рисунок 8. Сгенерированные и отобранные скэффолды и сравнение соединений 81 и 82.

Сонг и соавторы разработали ингибитор киназы TBK1 с помощью глубокой трансформерной нейронной сети SyntaLinker (рисунок 9) [78]. SyntaLinker основан на фрагментно-обоснованном дизайне и использует глубокое обучение на базе химического пространства из базы ChEMBL. Целью исследования являлась замена 2-аминопиримидинового скэффолда в известном ингибиторе MRT67307 (IC_{50} = 23,1 нМ).

Из молекулы MRT67307 удалили 2-аминопиримидиновое кольцо и загрузили полученные терминальные группы в SyntaLinker. Модель сгенерировала 1101 структуру, из которых 276 содержали обе

терминальные группы исходного соединения. Среди предложенных линкеров были 2-аминопиримидины с разными заместителями, которые не соответствовали цели замены скэффолда. Из уникальных скелетов были выделены два бициклических фрагмента: хинозолин-2-амин и пирроло[2,3-*d*]-2-аминопиримидин. Лучшие результаты докинга показало соединение 84 с пирроло[2,3-*d*]-2-аминопиримидиновым скелетом. Его синтезировали и оценили активность (IC_{50} = 66,7 нМ). Последующая оптимизация привела к более эффективному и селективному ингибитору 85 (IC_{50} 22,4 нМ), при этом SyntaLinker в дальнейшем не использовался.

Единственный вклад SyntaLinker — замена 2,4-аминопиримидина на 2-аминопирроло[2,3-*d*]пиримидин, что давно описано в литературе и патентах, с примерами с 2006 года. Таким образом, заявленная инновационность применения SyntaLinker вызывает сомнения, поскольку подобные структурные модификации можно выполнить вручную на базе известного соединения, без применения глубокого обучения. Генерация сотен структур и подбор одной с хорошо задокументированной трансформацией не выглядит оправданной.

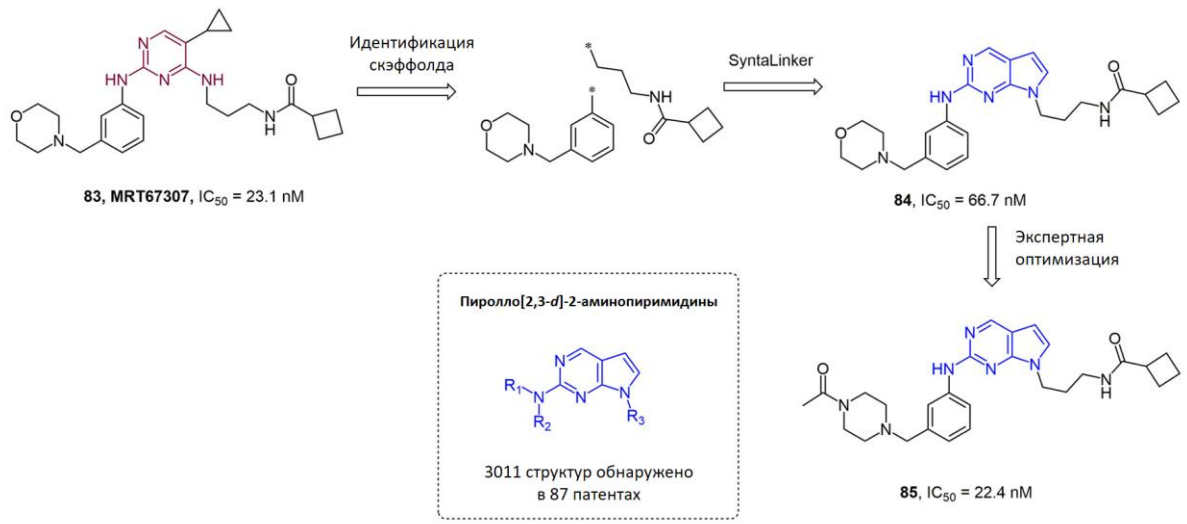


Рисунок 9. Процесс генерации соединения 84 с использованием алгоритма SyntaLinker и его последующая модификация.

Таким образом, применение новых вычислительных технологий является перспективным направлением развития методологической базы медицинской химии, однако требует рациональной оценки с точки зрения практических результатов. Новые компьютерные подходы, в том числе на основе машинного обучения, позволяют не только прогнозировать свойства целевых соединений, но также применяются для дизайна новых молекулярных структур. В то же время, современные вычислительные методы обладают рядом свойств, ограничивающих их практическое применение, в связи с чем требуется их дальнейшее развитие и валидация на практических задачах медицинской химии.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЙ

2.1. Поиск сайтов связывания на белках

Был разработан SiteRadar — метод, основанный на графовых нейронных сетях (graph neural network, GNN) для идентификации лиганд-белковых сайтов связывания. В данной модели структура белка представляется в виде графа, где каждый узел соответствует отдельному тяжелому атому. Ребра графа характеризуются межатомными расстояниями, а признаки узлов включают специфичную для аминокислот информацию (например, название атома, имя остатка, его тип). Модель обучалась на базе из 23 тысяч лиганд-белковых комплексов, отобранных из базы данных sc-PDB [17]. Данная модель была обучена распознавать макромолекулярное окружение заданной точки пространства для генерации вероятностных оценок для каждой точки грида около поверхности белка [20].

Подготовка обучающей выборки. Для обучения модели были использованы данные из базы sc-PDB, из которой были отобраны 23 тысячи лиганд-белковых комплексов, содержащих лекарственноподобные малые молекулы, связанные со своей биомишенью. Для подготовки обучающей выборки создавались трехмерные сетки с шагом 2 Å, покрывающие структуру белка, при этом точки сетки классифицировались по принадлежности к сайту связывания на основе референсного лиганда. Положительными определялись такие точки, которые находятся в радиусе не более 2 Å от любого атома лиганда, в то время как негативными примерами считались точки, характеризующиеся удаленностью более 2 Å [20] (рисунок 10).

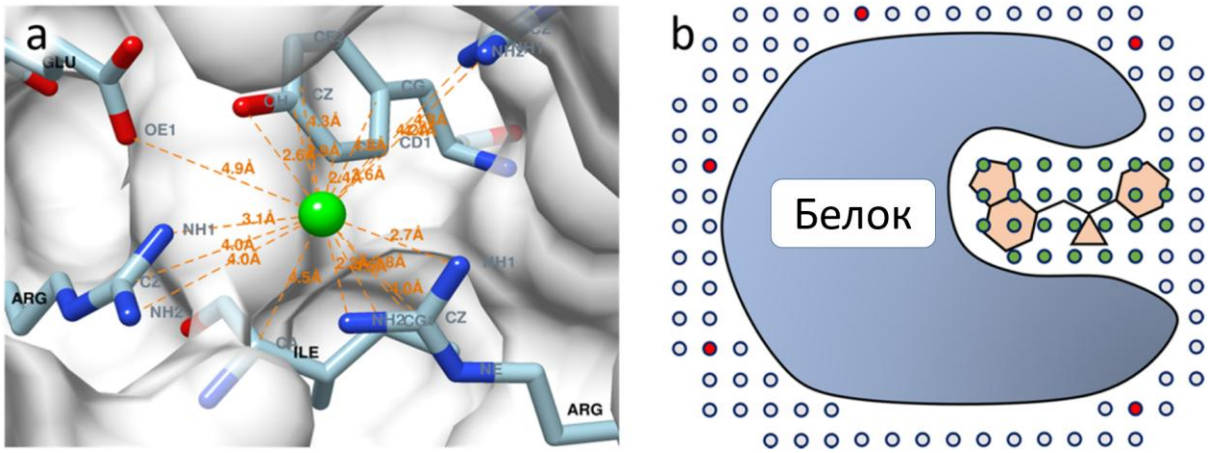


Рисунок 10. Способ сбора обучающих данных для поиска сайтов связывания.

Формирование графовых представлений. Каждая точка сетки представлялась как граф, в котором в качестве узлов выступали тяжелые атомы белка, находящиеся в пределах 7 Å от точки, а также сама точка сетки. Ребра графа соответствовали расстояниям между атомами белка, а также от атомов белка до точки сетки. Узлы графа описывались такими характеристиками как название атома, тип аминокислоты, химический класс остатка [20].

Виды полученных моделей. Было разработано две GNN модели. Геометрическая использовала только матрицу попарных расстояний. В свою очередь вторая, химическая, дополнительно учитывала химическую информацию об аминокислотном составе.

Обучение проводилось с использованием функции потерь кросс-энтропии с балансировкой весов классов. Оптимизация выполнялась методом Adam. Для обучения использовался графический процессор NVIDIA A100. Показатели качества модели оценивались по точности, полноте и F1-мере [20].

Обнаружение и кластеризация сайтов связывания лиганд-белок. После вычисления вероятностей принадлежности каждой точки сетки к сайту связывания выполнялась двухэтапная кластеризация.

- Агломеративная кластеризация с порогом расстояния 2,1 Å выделяла потенциальные области связывания малой молекулы.
- Для уточнения размеров кластера применялся алгоритм MeanShift с параметром bandwidth 5,8 Å, позволяющий отделять близлежащие сайты связывания и формировать компактные сущности, форма и размер которых соответствуют известным лиганд-белковым сайтам связывания.

Последовательность этапов идентификации сайтов связывания представлена на рисунке 11.

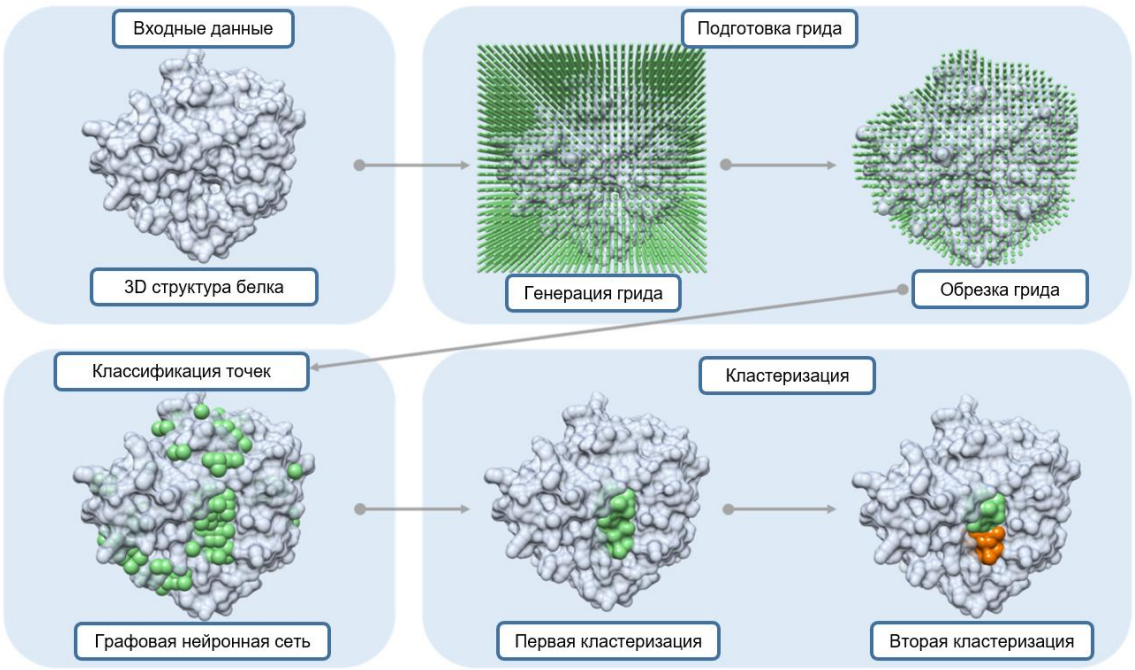


Рисунок 11. Алгоритм идентификации сайтов связывания.

Оценка и ранжирование выполнялись путем вычисления средней вероятности принадлежности точек кластера к сайту связывания [20].

2.2. Оптимизация структур белков, полученных с использованием AlphaFold 2

Был разработан компьютерный метод AF Optimizer, использующий глубокие нейронные сети и вычисления на основе силовых полей для автоматической оптимизации геометрии сайта связывания киназ (рисунок 12) [24]. Процесс начинается с идентификации остатков, вовлеченных в формирование сайта связывания, с помощью GNN модели, обученной на наборе из 1148 киназ в комплексе с малыми молекулами. Модель определяет вероятность принадлежности каждого атома к сайту связывания на основе макромолекулярного окружения и решетки искусственно сгенерированных точек.

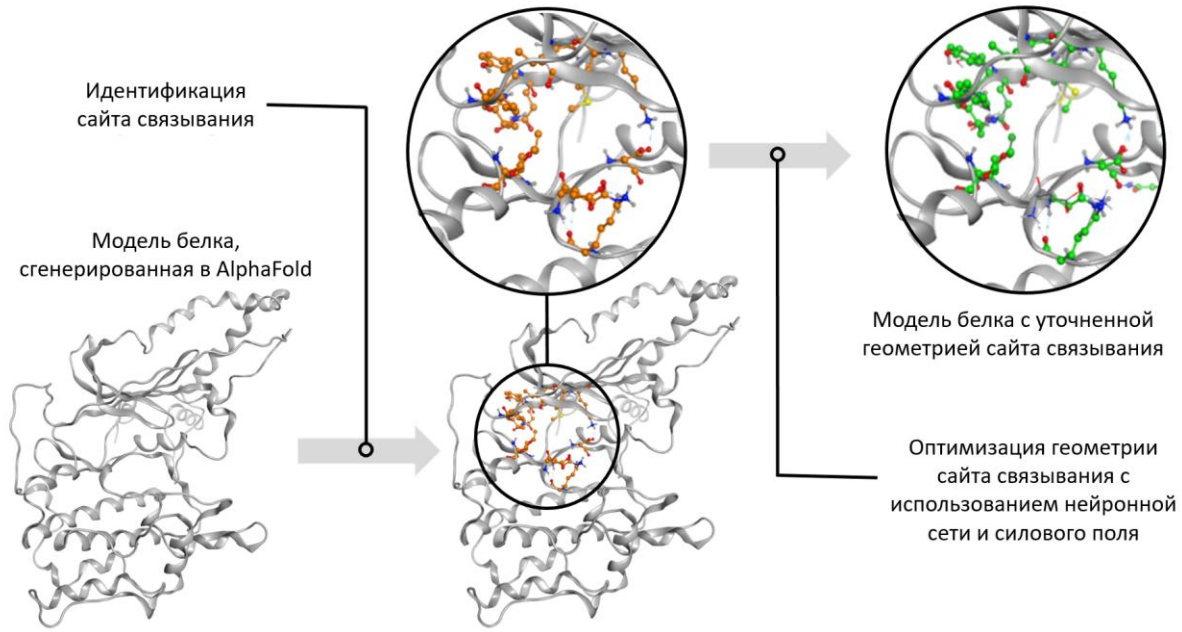


Рисунок 12. Алгоритм оптимизации геометрии сайта связывания.

Была обучена графовая нейронная сеть для оценки геометрии сайта связывания. Данная модель применяется для оценки вероятности нахождения атомов аминокислотных остатков в пространстве расположения потенциального лиганда. Для атомов с наиболее низкими оценками выбираются альтернативные положения, где вероятность

формирования неблагоприятных контактов с предположительной малой молекулой ниже. Далее выполняется оптимизация геометрии модифицированного аминокислотного остатка с использованием усовершенствованных силовых полей MMFF94 и алгоритма Adam. Процедура итеративно повторяется до достижения оптимального значения суммарной оценки аминокислот сайта связывания (до 300 итераций) (рисунок 13).

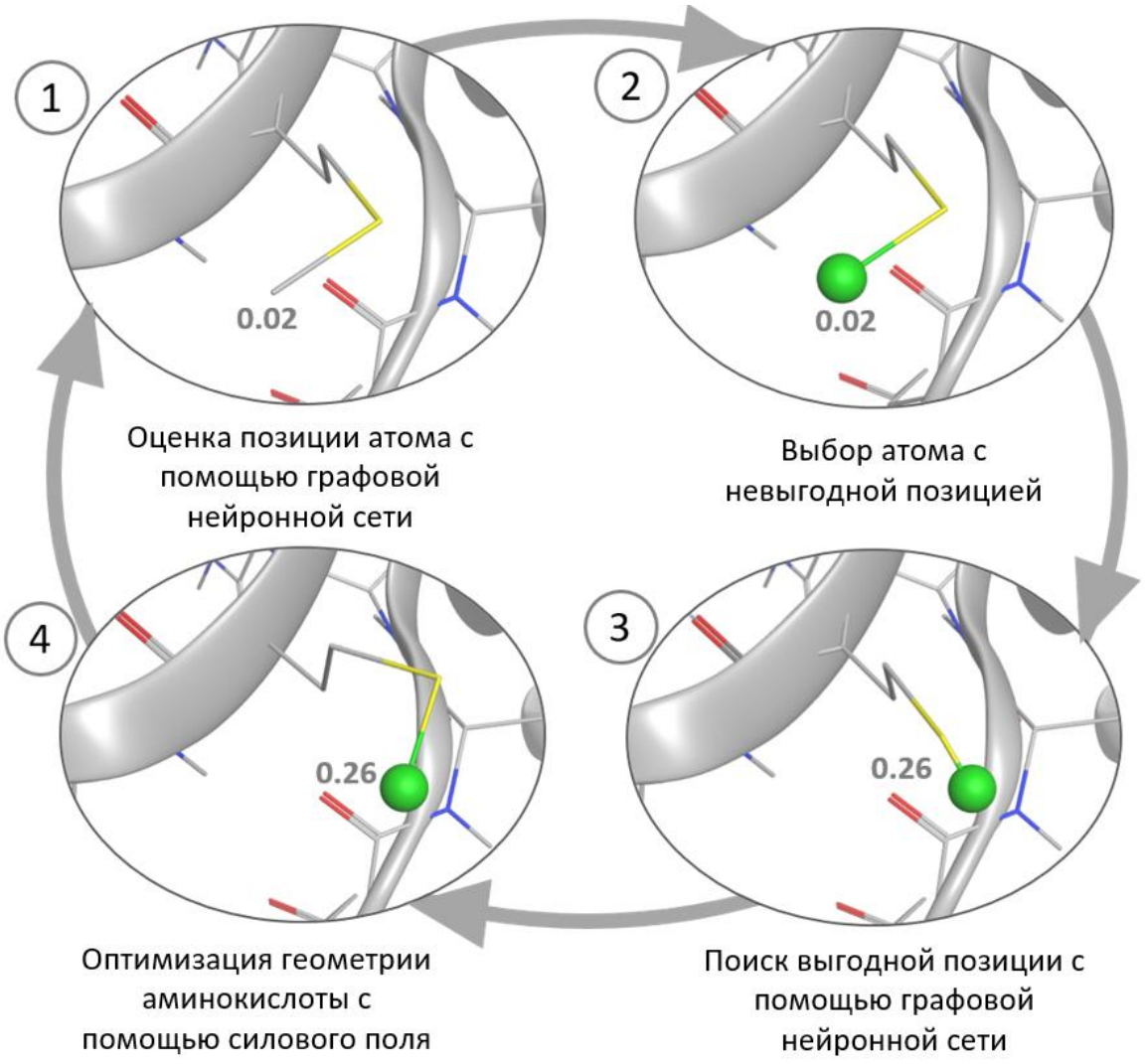


Рисунок 13. Стадии оптимизации геометрии сайта связывания.

2.3. Генерация псевдолигандов, классификация сайтов связывания, оценка аффинности лигандов

2.3.1. Генерация псевдолигандов

Был разработан алгоритм Skittles, состоящий из двух последовательных этапов: (1) генерация геометрии псевдолиганда и (2) предсказание типа атома для каждой точки псевдолиганда (рисунок 14). Оба этапа используют GNN.

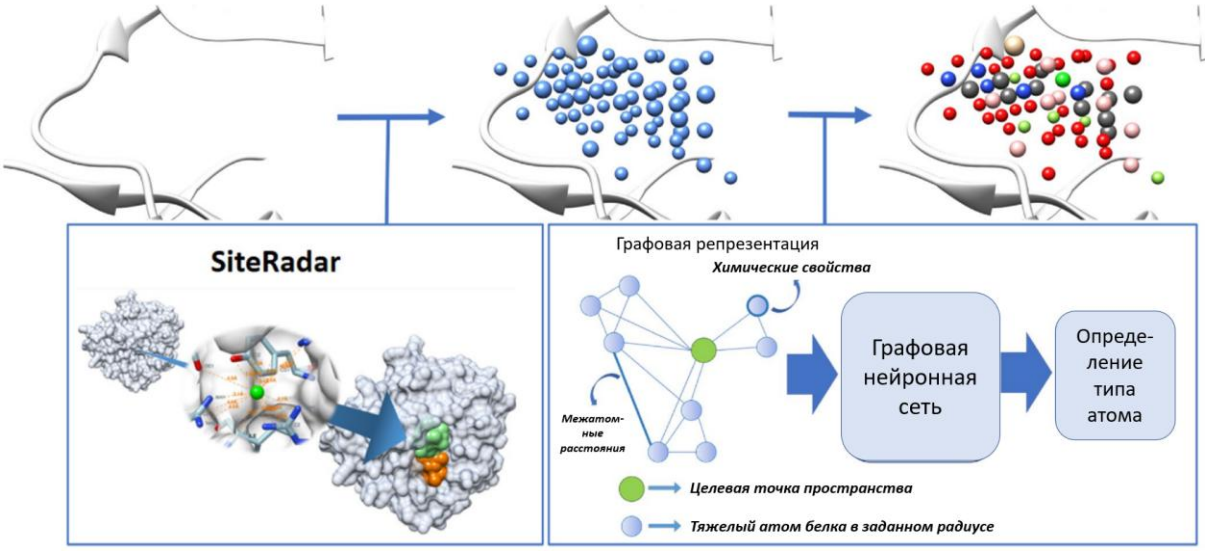


Рисунок 14. Схема алгоритма Skittles.

Первый этап принимает на вход трехмерную структуру белка и генерирует предполагаемые сайты связывания лигандов с белком. Для этого применяется описанный выше алгоритм SiteRadar [20]. Его результатом являются кластеры точек на поверхности белка (силуэт псевдолиганда), соответствующие предполагаемым сайтам связывания.

Второй этап направлен на предсказание подходящего типа атома для каждой точки псевдолиганда. Для этого было обучено 13 моделей, каждая из которых предсказывает вероятность размещения определенного типа атома в данной точке. Список типов атомов основан на свойствах, таких как

химический элемент, гибридизация, способность принимать участие в формировании водородных связей и заряд (положительный или отрицательный) (таблица 3).

Таблица 3. Список типов атомов с характеристиками.

Тип атома	Описание
$C_{аром.}$	Атом углерода в sp^2-гибридизации в составе ароматического цикла
$Csp^2_{цикл.}$	Атом углерода в sp^2-гибридизации в составе неароматического цикла
$Csp^2_{ацикл.}$	Атом углерода в sp^2-гибридизации, не входящий в состав цикла
$Csp^3_{цикл.}$	Атом углерода в sp^3-гибридизации в составе цикла
$Csp^3_{ацикл.}$	Атом углерода в sp^3-гибридизации, не входящий в состав цикла
Csp	Атом углерода в sp гибридизации
$N_{дон. ион.}$	Атом азота с присоединенным водородом, способный к ионизации и выступающий в роли донора водородной связи
$N_{дон. нейтр.}$	Нейтральный атом азота, связанный с атомом водорода, функционирующий как донор водородной связи
$N_{акц.}$	Нейтральный атом азота без непосредственно присоединённых атомов водорода, способный выступать акцептором водородной связи
$Osp^2_{акц.}$	Атом кислорода в sp^2-гибридизации; способен выполнять роль акцептора водородной связи
$Osp^3_{акц.}$	Атом кислорода в sp^3-гибридизации, не содержащий присоединенных к нему атомов водорода; способен выполнять роль акцептора водородной связи
$Osp^3_{дон.}$	Атом кислорода в sp^3-гибридизации, содержащий присоединенный к нему атом водорода и способный выполнять роль донора или акцептора водородной связи
$S_{двухвал.}$	Атом двухвалентной серы
$S_{шестивал.}$	Атом шестивалентной серы
Галоген	Любой галоген, кроме фтора

Эти 13 классов предложены вместо стандартных фармакофорных типов, поскольку функциональные группы с одинаковыми фармакофорными свойствами могут существенно различаться в предпочтениях по аминокислотам и пространственным параметрам межмолекулярных взаимодействий [8].

Для обучения моделей использовались 8150 кристаллических комплексов из RCSB PDB [6]. Лиганды были обработаны с помощью RDKit. Комплексы были случайно разделены на обучающую и валидационную выборки. Положительные и отрицательные примеры были определены для каждого типа атомов (например, ароматические атомы углерода против всех остальных классов). Макромолекулярное окружение в радиусе 5 Å было представлено графом с вершинами — атомами белка и их химическими свойствами. Модели GNN были обучены на этих графах, ансамбль моделей использовался для оценки вероятности размещения атома каждого типа в каждой точке псевдолиганда [21].

Для балансировки результатов бинарных моделей применялись экспоненциальные коэффициенты, оптимизированные с использованием библиотеки Sklearn для максимизации средней F1-меры.

В дополнение, были обучены 13 фармакофорных моделей с более строгими условиями отбора данных. Данные модели обучались только на тех атомах, которые образовывали межмолекулярные контакты (водородные связи, π-π стэкинг, ионные взаимодействия и другие).

2.3.2. Классификация сайтов связывания

Был разработан компьютерный метод классификации сайтов связывания лиганд-белок. Метод использует модели на основе GNN, а в качестве входных данных принимает псевдолиганды, полученные с

помощью метода Skittles. В данной работе были рассмотрены пять классов природных субстратов: нуклеотиды/АТФ, углеводы, гем, пептиды и нуклеиновые кислоты. Кроме того, прогнозировались два дополнительных свойства: возможность аллостерической регуляции и способность взаимодействовать с ковалентно связывающимися молекулами. Кроме того, была рассчитана способность сайта связывать лекарственноподобные молекулы [20].

Для подготовки моделей были собраны специализированные библиотеки комплексов белок–лиганд в качестве положительных примеров для каждой группы. На поверхности каждого белка были сгенерированы псевдолигандные структуры с использованием базовых моделей Skittles. Сохранялись только те псевдолиганды, которые совпадали с референсными лигандами по объемному пересечению. Данные внутри каждого класса были разделены на обучающую, тестовую и валидационную выборки на основе сходства аминокислотных последовательностей. Для каждого псевдолиганда строился граф, вершины которого отвечали вероятностям для каждого типа атомов, а ребра описывали расстояния между точками. Для каждой бинарной GNN-модели создавались положительные и отрицательные выборки. При обучении моделей прогнозирования химической природы субстратов отрицательные наборы формировались из примеров других химических классов. Для моделей, предсказывающих возможность ковалентного связывания, из отрицательных примеров исключались все ковалентные образцы из других классов. После обучения для балансировки прогнозов полученных моделей применялись экспоненциальные коэффициенты. Визуальное представление использования признаков, сгенерированных Skittles, для классификации сайтов связывания представлено на рисунке 15.

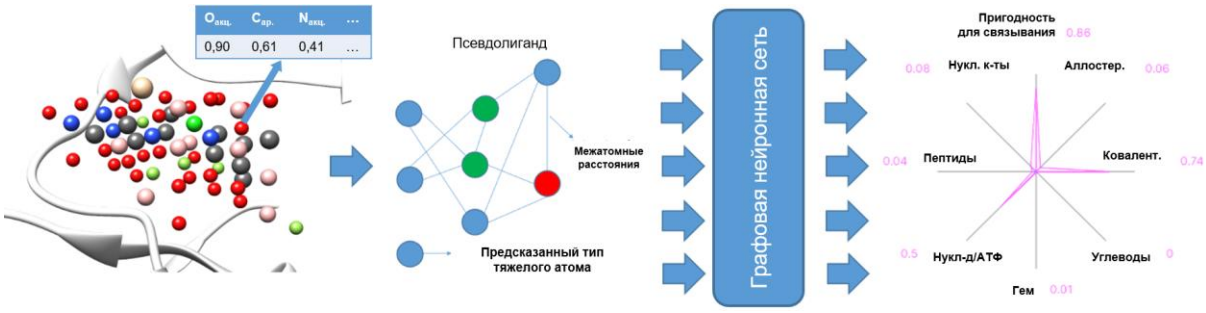


Рисунок 15. Алгоритм классификации сайтов связывания на основе признаков Skittles.

В результате была разработана система классификации сайтов связывания, рассчитывающая восемь свойств, включая способность связывать лекарственноподобные молекулы, способность к аллостерической регуляции, способность взаимодействовать с ковалентно связывающимися лигандами и вероятность связывания различных субстратов, таких как нуклеотиды/АТФ, углеводы, гем, пептиды и нуклеиновые кислоты.

2.3.3. Оценка аффинности лигандов

Был разработан метод для расчета лиганд-белковой аффинности с применением нейронных сетей Skittles. В основе алгоритма лежит предсказание вероятности размещения каждого атома малой молекулы в заданном макромолекулярном окружении. Затем с использованием полученных поатомных оценок с помощью отдельной модели выполняется предсказание аффинности связывания лиганда с белком.

Для обучения соответствующей модели были отобраны комплексы белок–лиганд из базы данных PDBBind [82]. Данные комплексы были обработаны с использованием RDKit и AutoDock Tools, после чего для каждого атома были вычислены вероятности с помощью нейросетей Skittles (как базовых, так и фармакофорных моделей). Далее были собраны графы молекул с оценками вероятностей на вершинах и метками наличия связей

на ребрах. Использовались дополнительные дескрипторы для характеристики межмолекулярных взаимодействий, такие как заполнение сайта связывания, количество межмолекулярных контактов (водородные связи, солевые мостики, π–π стэкинг, π–катионные взаимодействия). Схема алгоритма представлена на рисунке 16. Данные были разделены на обучающую, тестовую и валидационную выборки с учетом совпадения белковых последовательностей и сходства структур лигандов.

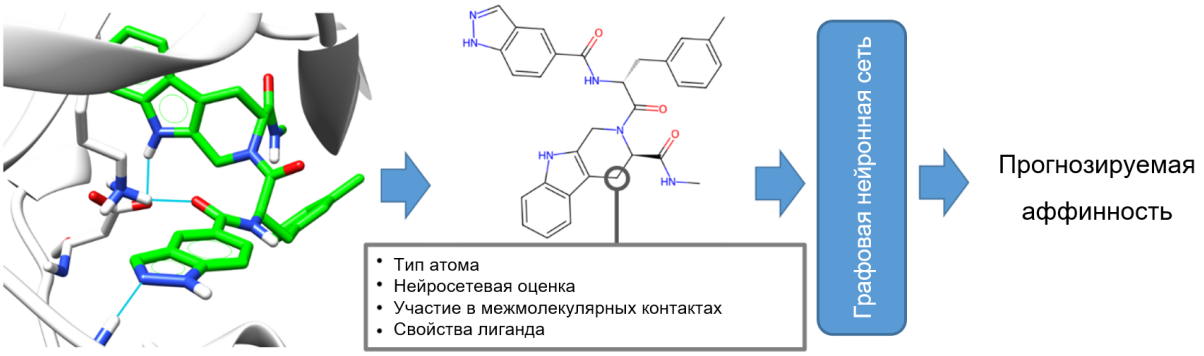


Рисунок 16. Схема алгоритма оценки аффинности на основе Skittles.

2.4. Генерация новых структур на основе фрагментов

Был разработан алгоритм QFASG, который состоит из нескольких взаимосвязанных модулей. Функционально алгоритм делится на две части: неитеративная часть алгоритма предназначена для размещения начальных фрагментов внутри целевого сайта связывания, а вторая (итеративная) реализует поэтапный рост структуры (рисунок 17).

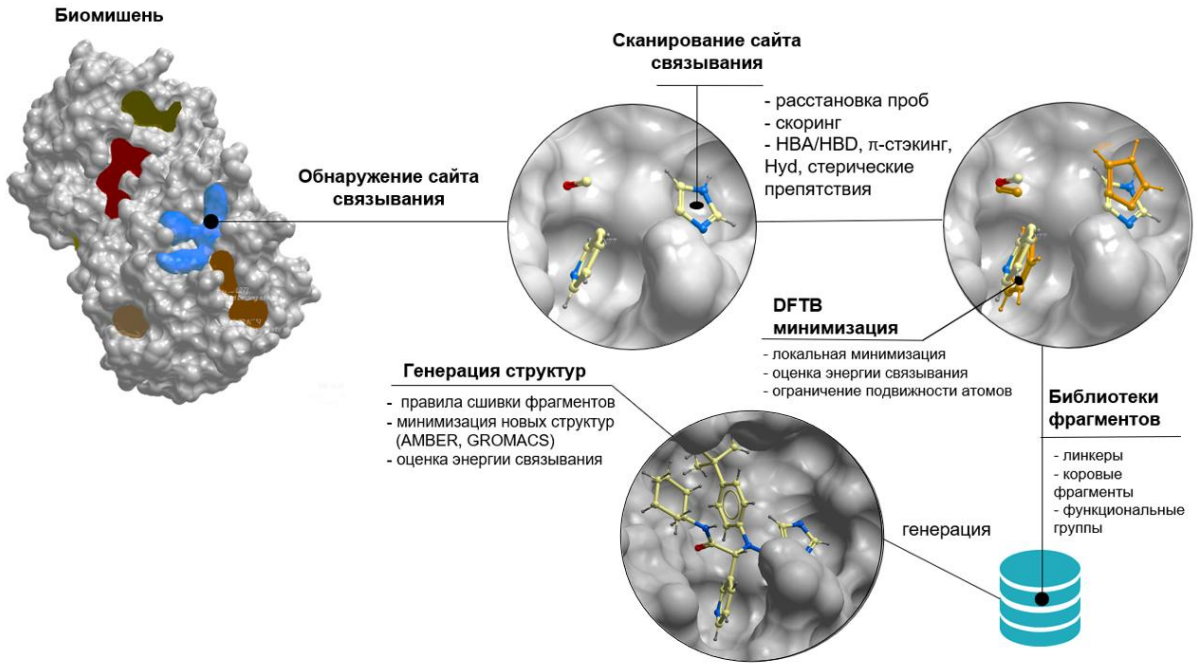


Рисунок 17. Схема работы алгоритма QFASG.

Позиционирование фрагментов. Позиционирование фрагментов проходит четыре этапа: выбор сайта связывания, позиционирование фрагментов на основе анализа фармакофоров, оптимизация позиций фрагментов с использованием метода GFN2-xTB [4] и финальный отбор позиций. Идентификация сайта связывания основана на положении референсного лиганда или выполняется *de novo*. Фрагменты позиционируются в сайтах связывания, фокусируясь на функциональных группах белка, таких как доноры и акцепторы водородных связей, ароматические кольца. Уточнение позиции выполняется с помощью квантово-химического метода GFN2-xTB. В системе используется реализация GFN2-xTB из программного решения xtb. После размещения фрагментов в сайте связывания проводится их отбор по фармакофорным оценкам и расчетным энергиям связывания.

Рост структуры. Итеративная часть включает этапы: присоединение фрагментов, генерация трехмерных структур, структурное выравнивание, локальный жесткий докинг и отбор. Новые структуры генерируются присоединением фрагментов к имеющимся в сайте связывания структурам, полученным на неитеративном этапе или из предыдущей итерации. Они проверяются с помощью фильтров медицинской химии. Далее для новых структур генерируются трехмерные конформации, которые затем выравниваются функцией «GetBestRMS» из библиотеки RDKit. Полученные позиции оптимизируются локальным жестким докингом с использованием улучшенной версии силового поля MMFF94.

Отбор сгенерированных соединений выполняется с использованием одной из двух оценочных функций: расчетная свободная энергия связывания или оценка взаимодействия белок–лиганд. Отбор может выполняться с кластеризацией по сходству позиций связывания или без нее. В первом случае структуры группируются по позициям связывания и взаимодействиям, после чего из каждого кластера выбирается лучшая позиция по оценочной функции, что обеспечивает разнообразие. Генерация продолжается до достижения ограничений по молекулярному весу или максимальному числу итераций.

Для генерации итогового файла структуры всех итераций проходят усиленную оценку с применением расширенного набора фильтров медицинской химии, учитывающих не только основной каркас, но и терминальные функциональные группы, а также другие пользовательские параметры (например, молекулярный вес). Кроме того, вычисляются различные дескрипторы, включая сходство формы лиганда с позицией референсной структуры, оценку межмолекулярных контактов и оценку синтетической доступности.

Правила соединения фрагментов. В QFASG используются несколько вариантов роста структуры. Во-первых, соединение двух фрагментов с помощью ковалентной связи, что применимо для пар уже расставленных фрагментов или в случае добавления фрагмента из библиотеки. Во-вторых, соединение двух фрагментов с помощью связующего линкера, предназначенного для объединения двух функциональных частей, находящихся на расстоянии друг от друга. Кроме того, допускается соединение циклов из разных фрагментов для образования конденсированных и аннелированных систем, а также формирование спироциклов (рисунок 18). Предлагаемый набор правил позволяет генерировать широкий спектр лекарственноподобных соединений, охватывая большое химическое пространство и в то же время позволяя получать новую разнообразную химию.

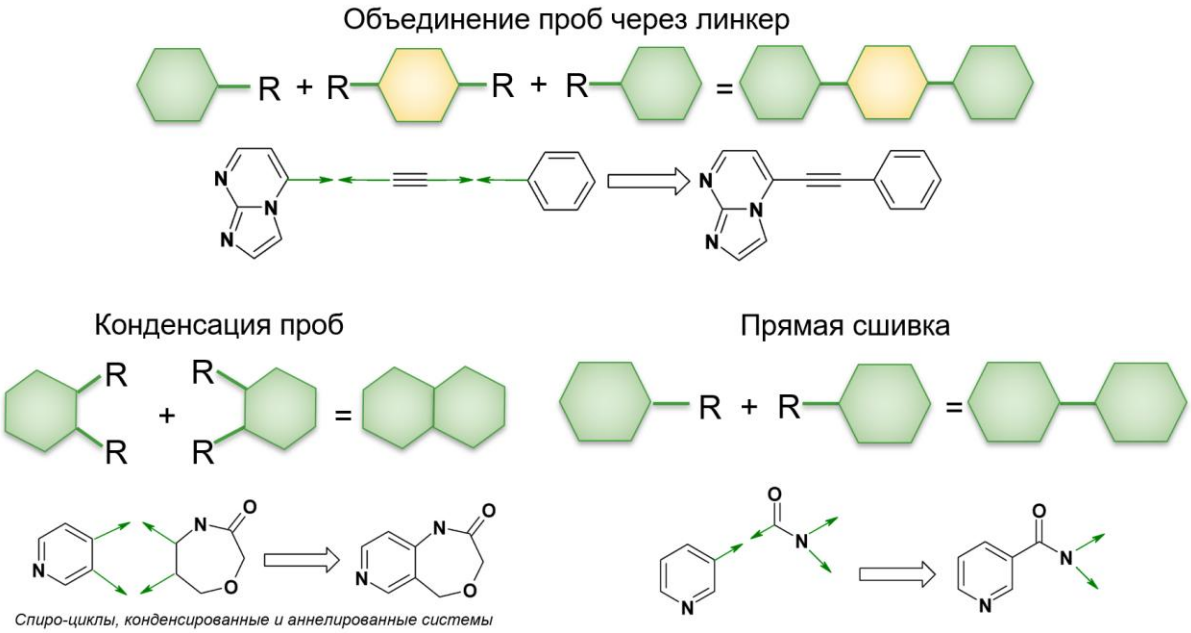


Рисунок 18. Правила соединения фрагментов в алгоритме QFASG.

2.5. Схемы синтеза смоделированных соединений

В рамках настоящей работы были синтезированы предполагаемые соединения-лидеры, отобранные на основании анализа виртуальных библиотек: QFASG-1, QFASG-2, QFASG-3, QFASG-4, QFASG-5, QFASG-6.

QFASG-1

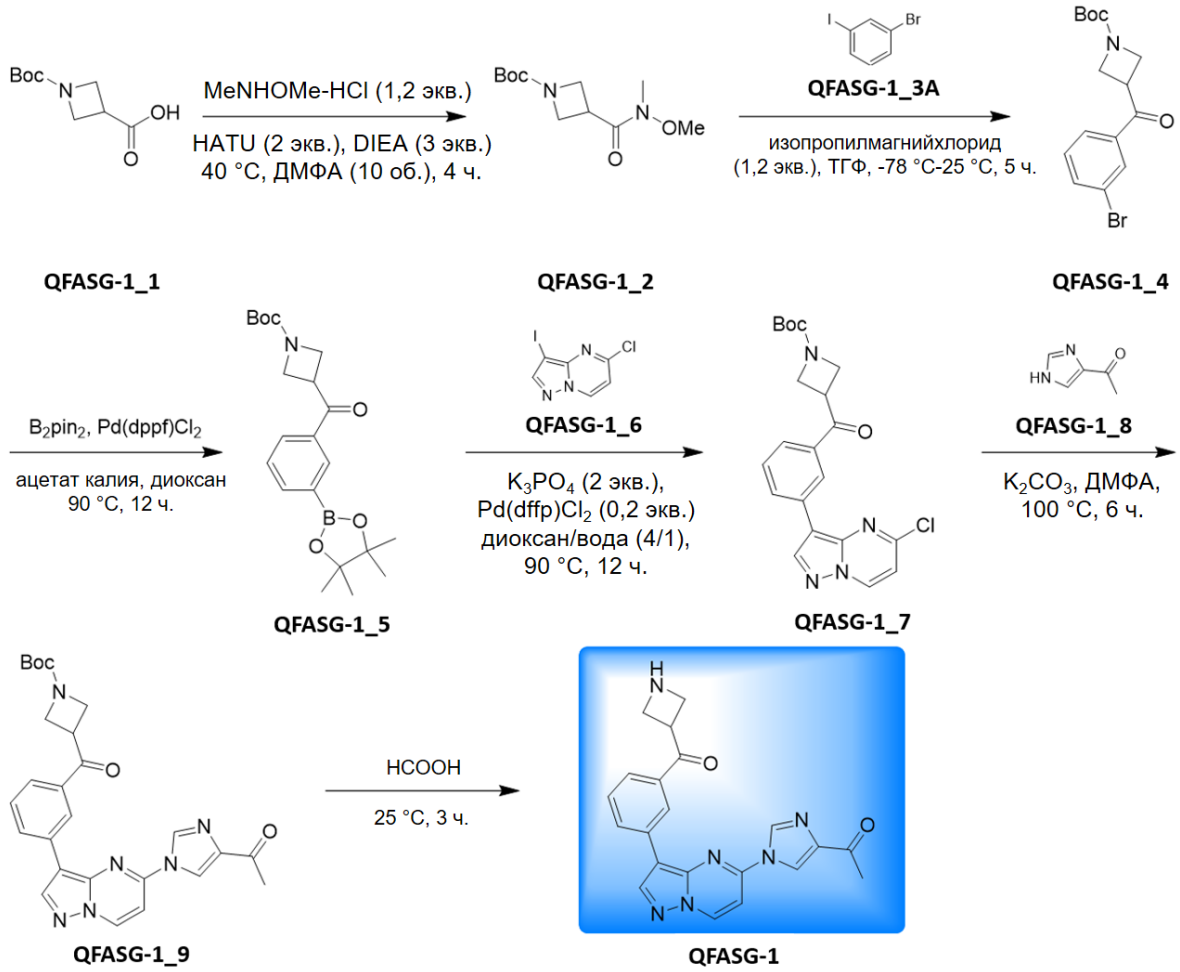


Рисунок 19. Схема синтеза соединения QFASG-1.

Стадия 1. Синтез QFASG-1_2.

К раствору QFASG-1_1 (10,0 г, 49,70 ммоль) в N,N-диметилформамиде (60 мл) при 25 °C добавляли 2-(3H-триазоло[4,5-b]пиридин-3-ил)-1,1,3,3-тетраметилизоуроний гексафторфосфат(V) (37,79 г, 99,39 ммоль), N-этил-N-изопропилпропан-2-амин (19,27 г, 149,09 ммоль, 25,97 мл) и N,O-

диметилгидроксиламин гидрохлорид (3,04 г, 49,70 ммоль). Смесь нагревали до 40 °C и перемешивали в течение 4 ч. По ТСХ (петролейный эфир/этилацетат = 4/1, R_f = 0,25) реагент был полностью израсходован и образовалось целевое вещество. Реакция была проведена дважды. Реакционные смеси охладили до комнатной температуры, затем при 0°C добавили воду (600 мл) и провели экстракцию этилацетатом (трижды по 200 мл). Органические фазы объединили, промыли насыщенным раствором NaCl (300 мл), высушили безводным Na_2SO_4 и отфильтровали. После удаления растворителя при пониженном давлении остаток подвергли колоночной хроматографии на силикагеле, используя градиент петролейный эфир/этилацетат (70/30 → 65/30, Rf = 0,25). В результате получили соединение QFASG-1_2 (12,0 г, выход 49%) в виде жёлтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, CD_3OD

δ = 4,11 - 4,01 (м, 4H), 3,85 - 3,74 (м, 1H), 3,70 (с, 3H), 3,20 (с, 3H), 1,44 (с, 9H)

Стадия 2. Синтез QFASG-1_4.

К раствору QFASG-1_3A (9,03 г, 31,92 ммоль, 4,07 мл) в тетрагидрофуране (70 мл) при −78 °C в атмосфере азота добавляли хлорид изопропилмагния (2 М, 14,73 мл, 1,2) и выдерживали 1 ч, после чего при −78 °C добавляли QFASG-1_2 (6,00 г, 24,55 ммоль). Реакционную смесь нагрели до 25 °C и выдерживали при перемешивании 5 ч. Контроль методом ТСХ (петролейный эфир/этилацетат = 3/1, Rf = 0,25) показал полное исчезновение исходного реагента и появление целевого продукта. При 0 °C к смеси прилили водный раствор хлорида аммония (100 мл) и трёхкратно экстрагировали этилацетатом (по 40 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (50 мл), высушили над безводным Na_2SO_4 и отфильтровали. После удаления растворителя при

пониженном давлении остаток подвергли колоночной хроматографии на силикагеле (градиент петролейный эфир/этилацетат 70/30 → 65/35, Rf = 0,25). Выход соединения QFASG-1_4 составил 7,5 г (90%), вещество получено в виде желтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 8,00 (т, *J* = 1,7 Гц, 1H), 7,87 (дд, *J* = 1,8, 7,9 Гц, 2H), 7,51 (т, *J* = 7,9 Гц, 1H), 4,46 - 4,35 (м, 1H), 4,13 (уш. с, 2H), 3,96 (уш. т, *J* = 6,7 Гц, 2H), 1,37 (с, 9H)

Стадия 3. Синтез QFASG-1_5.

К раствору QFASG-1_4 (3,7 г, 10,88 ммоль) в диоксане (80 мл) при 25 °C добавляли 4,4,4',4',5,5,5',5'-октаметил-2,2'-би(1,3,2-диоксаборолан) (4,14 г, 16,31 ммоль, 1,5), дихлорид [1,1’-бис(дифенилфосфино)-ферроцен]палладия (1,59 г, 2,18 ммоль) и ацетат калия (1,28 г, 13,05 ммоль) в атмосфере азота. Смесь нагревали до 90 °C и перемешивали в течение 4 ч. По ^{1}H ЯМР реагент был полностью израсходован и был обнаружен целевой продукт. По достижении комнатной температуры реакционную смесь обработали водным раствором хлорида аммония (100 мл) для гашения и трехкратно экстрагировали этилацетатом (по 40 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (50 мл), высушили над безводным сульфатом натрия и отфильтровали. После удаления растворителя при пониженном давлении остаток подвергли колоночной хроматографии на силикагеле, используя градиент петролейный эфир/этилацетат (70/30 → 65/35, Rf = 0,25). В результате выделено соединение QFASG-1_5 (4,2 г, выход 99,72%) в виде желтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 8,15 (с, 1H), 7,98 (тд, *J* = 1,5, 7,9 Гц, 1H), 7,92 (д, *J* = 7,4 Гц, 1H), 7,56 (т, *J* = 7,6 Гц, 1H), 4,48 - 4,35 (м, 1H), 4,12 (уш. с, 2H), 3,96 (уш. т, *J* = 6,3 Гц, 2H), 1,37 (с, 9H), 1,31 (с, 12H)

Стадия 4. Синтез QFASG-1_7.

К раствору соединения QFASG-1_5 (2,0 г, 5,16 ммоль, 1 экв.) в смеси диоксан/вода (30 мл/8 мл) при 25 °C добавили QFASG-1_6 (1,73 г, 6,20 ммоль), предварительно дегазированный трёхкратным продуванием азотом. Затем внесли дихлорид [1,1'-бис(дифенилфосфино)ферроцен]палладия (755,74 мг, 1,03 ммоль) и фосфат калия (2,19 г, 10,33 ммоль). Реакционную смесь нагрели до 90 °C и выдерживали 4 ч в атмосфере азота, после чего продолжали перемешивание при той же температуре еще 12 ч. Согласно данным ЖХ-МС, исходный реагент полностью прореагировал, зарегистрирован один доминирующий пик целевого продукта. По охлаждении до комнатной температуры смесь обработали водным раствором хлорида аммония (100 мл) и трёхкратно экстрагировали этилацетатом (по 40 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (50 мл), высушили над безводным Na_2SO_4 и отфильтровали. Растворитель удалили при пониженном давлении, остаток очистили колоночной хроматографией на силикагеле (градиент петролейный эфир/этилацетат 62/38 → 57/43, Rf = 0,25). Выделено соединение QFASG-1_7 (800 мг, 38%) в виде желтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,24 (д, *J* = 7,3 Гц, 1H), 8,93 (с, 1H), 8,54 (т, *J* = 1,4 Гц, 1H), 8,32 - 8,28 (м, 1H), 7,79 - 7,75 (м, 1H), 7,67 - 7,59 (м, 1H), 7.24 (д, *J* = 7,3 Гц, 1H), 4,51 - 4,40 (м, 1H), 4,22 (уш. д, *J* = 1,1 Гц, 2H), 4,09 - 4,01 (м, 2H), 1,38 (с, 9H)

Стадия 5. Синтез QFASG-1_9.

К раствору QFASG-1_7 (400 мг, 968,83 мкмоль) в N,N-диметилформамиде (6 мл) при 25 °C добавили карбонат калия (267,80 мг, 1,94 ммоль). Смесь нагревали до 100 °C и выдерживали в течение 6 ч. По данным ЖХ-МС реагент был полностью израсходован, выявлен один основной пиковый сигнал с целевым соединением. Реакция была проведена дважды. По охлаждении до комнатной температуры обе реакционные смеси при 0 °C обработали водой (40 мл) для гашения и трехкратно экстрагировали этилацетатом (по 10 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (15 мл), высушили над безводным сульфатом натрия и отфильтровали. После удаления растворителя при пониженном давлении остаток подвергли колоночной хроматографии на силикагеле, используя градиент петролейный эфир/этилацетат (3/97 → 1/99, Rf = 0,25). В результате выделено соединение QFASG-1_9 (300 мг, 32%) в виде желтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,46 (д, *J* = 7,5 Гц, 1H), 8,94 (с, 1H), 8,87 (д, *J* = 10,5 Гц, 2H), 8,67 (с, 1H), 8,42 (д, *J* = 7,8 Гц, 1H), 7,76 (уш. д, *J* = 7,5 Гц, 2H), 7,68 - 7,62 (м, 1H), 4,58 - 4,47 (м, 1H), 4,20 (уш. с, 2H), 4,03 (уш. д, *J* = 6,8 Гц, 2H), 2,53 (с, 3H), 1,38 (с, 9H)

Стадия 6. Синтез QFASG-1.

Раствор соединения QFASG-1_9 (300 мг, 616,62 мкмоль) в муравьиной кислоте (12 мл) выдерживали при перемешивании 3 ч при 25 °C. Согласно данным ЖХ-МС, исходный реагент полностью прореагировал, зарегистрирован один доминирующий пик целевого продукта. Реакционную массу упарили при пониженном давлении, остаток подвергли очистке методом препаративной ВЭЖХ. Выход целевого соединения

QFASG-1 составил 50 мг (21%), вещество получено в виде желтого твердого вещества.

Условия препаративной ВЭЖХ:

- Колонка: Phenomenex Luna C18, 100 × 30 мм, 3 мкм.
- Мобильная фаза: H_2O (0,2% муравьиная кислота) — ацетонитрил.
- Градиент: содержание фазы B увеличивали с 1 до 35% в течение 8,0 мин.

^{1}H ЯМР: 400 МГц, CD_3CN

δ = 9,04 (д, *J* = 7,6 Гц, 1H), 8,69 (с, 1H), 8,67 - 8,65 (м, 2H), 8,62 (д, *J* = 1,3 Гц, 1H), 8,36 (д, *J* = 8,0 Гц, 1H), 8,34 (с, 1H), 7,76 (д, *J* = 7,8 Гц, 1H), 7,68 - 7,63 (м, 1H), 7,39 (д, *J* = 7,5 Гц, 1H), 4,68 (квинтет, *J* = 8,5 Гц, 1H), 4,32 (с, 2H), 4,30 (с, 2H), 2,56 (с, 3H)

QFASG-2

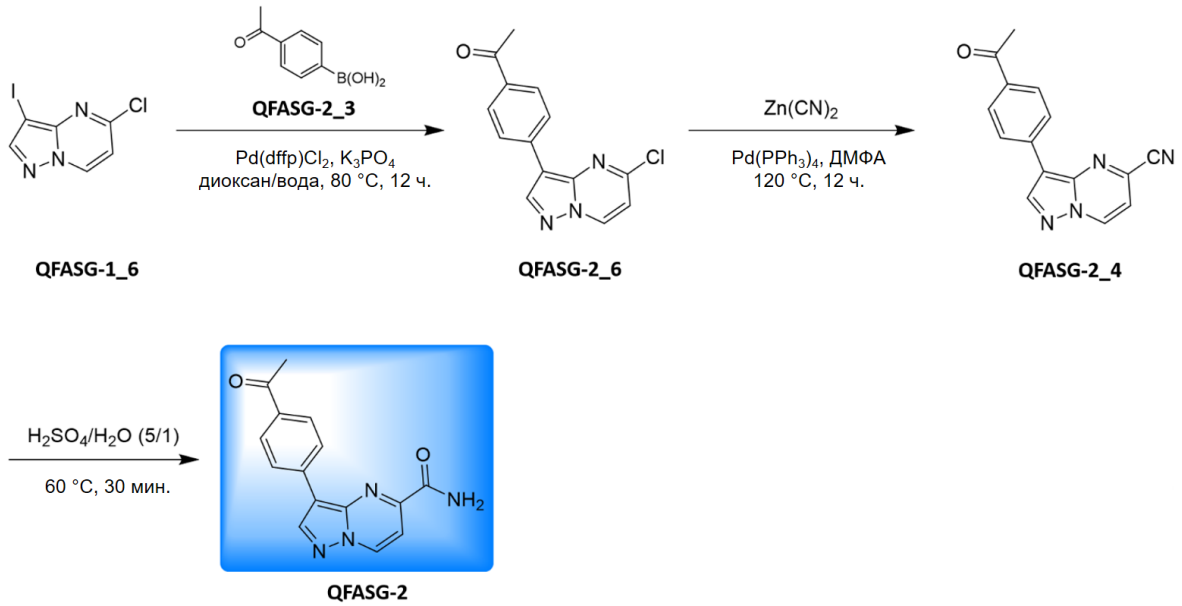


Рисунок 20. Схема синтеза соединения QFASG-2.

Стадия 1. Синтез QFASG-2_6.

К раствору QFASG-1_6 (2,5 г, 8,95 ммоль) и QFASG-2_3 (1,47 г, 8,95 ммоль) в смеси растворителей диоксан (50 мл) и вода (12,5 мл) при 25 °C в атмосфере азота добавляли дихлорид [1,1’-бис(дифенилфосфино)-ферроцен]палладия (1,31 г, 1,79 ммоль) и фосфат калия (3,80 г, 17,89 ммоль). Смесь нагрели до 80 °C и перемешивали при 80 °C в течение 12 ч. По ТЛФ (петролейный эфир/этилацетат = 2/1, R_f = 0,31) реагент был израсходован и образовался целевой продукт. Реакция была проведена дважды. По охлаждении до комнатной температуры обе реакционные смеси объединили, разбавили водой (дважды по 100 мл) и трехкратно экстрагировали этилацетатом (по 30 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (100 мл), высушили над сульфатом натрия и отфильтровали. Фильтрат упарили при пониженном давлении, остаток подвергли колоночной хроматографии на силикагеле с градиентом петролейный эфир/этилацетат (100/0 → 70/30). В результате выделено соединение QFASG-2_6 (2,0 г, выход 41%) в виде желтого твердого вещества.

^{1}H ЯМР: 400 МГц, $CDCl_3$

δ = 8,60 (д, *J* = 7,3 Гц, 1H), 8,51 (с, 1H), 8,15 -8,09 (м, 2H), 8,07 - 8,02 (м, 2H), 6,89 (д, *J* = 7,3 Гц, 1H), 2,64 (с, 3H)

Стадия 2. Синтез QFASG-2_4.

К раствору QFASG-2_6 (500 мг, 1,84 ммоль) в N,N-диметилформамиде (10 мл) при 25 °C в атмосфере азота добавляли цианид цинка (432,18 мг, 3,68 ммоль) и *тетракис*-(трифенилфосфин)палладий (425,31 мг, 368,05 мкмоль). Смесь нагревали до 120 °C и выдерживали в течение 12 ч. По ТЛФ (петролейный эфир/этилацетат = 1/2, R_f = 0,35) реагент был израсходован и образовался целевой продукт. По охлаждении до комнатной температуры

реакционную смесь разбавили водой (50 мл) и трехкратно экстрагировали этилацетатом (по 20 мл). Органические экстракты объединили, промыли насыщенным раствором NaCl (100 мл), высушили над сульфатом натрия и отфильтровали. После удаления растворителя при пониженном давлении полученный продукт подвергли колоночной хроматографии на силикагеле с градиентом петролейный эфир/этилацетат (100/0 → 50/50). Выход соединения QFASG-2_4 составил 220 мг (46%), вещество получено в виде желтого твердого вещества.

^{1}H ЯМР: 400 МГц, $CDCl_3$

δ = 8,87 (д, *J* = 7,3 Гц, 1H), 8,69 (с, 1H), 8,21 - 8,15 (м, 2H), 8,13 - 8,07 (м, 2H), 7,18 (д, *J* = 7,1 Гц, 1H), 2,66 (с, 3H)

Стадия 3. Синтез QFASG-2.

Раствор QFASG-2_4 (100 мг, 381,29 мкмоль) в смеси серной кислоты (0,5 мл) и воды (0,1 мл) нагрели до 60 °C и выдерживали при 60 °C в течение 30 мин. По данным ЖХ-МС реагент был израсходован и обнаружено целевое соединение. Реакция была проведена дважды. По охлаждении до комнатной температуры обе реакционные смеси объединили, добавили метанол (3 мл), после чего суспензию отфильтровали. Осадок высушили при пониженном давлении до получения продукта, который очистили препаративной ВЭЖХ и получили QFASG-2 (21 мг, выход 9,83%) в виде желтого твердого вещества.

Условия препаративной ВЭЖХ:

- Колонка: Phenomenex Luna C18, 100 × 40 мм, 3 мкм.
- Мобильная фаза: вода (0,2% муравьиная кислота) — ацетонитрил.

- Градиент: содержание фазы B увеличивали с 20 до 55% в течение 8,0 мин.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,32 (д, *J* = 7,1 Гц, 1H), 9,03 (с, 1H), 8,45 (д, *J* = 8,4 Гц, 2H), 8,40 (уш. с, 1H), 8,04 (д, *J* = 8,4 Гц, 3H), 7,63 (д, *J* = 7,3 Гц, 1H), 2,61 (с, 4H)

QFASG-3

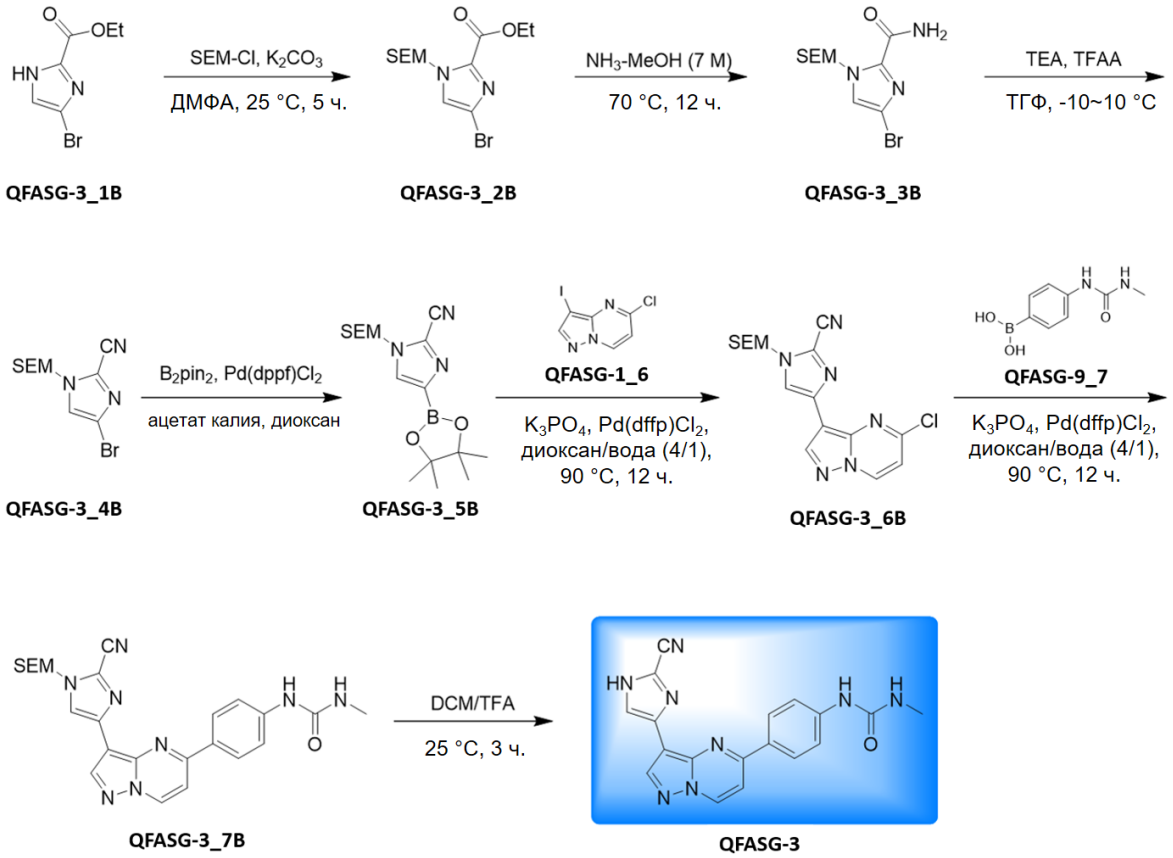


Рисунок 21. Схема синтеза соединения **QFASG-3.**

Стадия 1. Синтез QFASG-3_2B.

Раствор QFASG-3_1B (13,4 г, 61,18 ммоль) и карбоната калия (10,15 г, 73,41 ммоль) в ДМФА (270 мл) охлаждали до 0 °C, после чего добавляли 2-(триметилсилил)этоксиметилхлорид (12,24 г, 73,41 ммоль, 12,99 мл). Реакционную массу нагревали до 25 °C и перемешивали при этой температуре в течение 5 ч. По результатам ЖХ-МС-анализа исходный

реагент полностью исчезал, при этом содержание целевого соединения составляло около 60%. После окончания реакции к смеси добавляли воду (600 мл) и проводили трехкратную экстракцию этилацетатом (по 1 л). Органические экстракты объединяли, промывали насыщенным раствором хлорида натрия (1 л), осушали безводным сульфатом натрия и фильтровали. Полученный фильтрат концентрировали в вакууме. Неочищенный остаток разделяли колоночной хроматографией на силикагеле, используя петролейный эфир/этилацетат в соотношении 9:1 в качестве элюента (R_f =0,70). В результате выделяли QFASG-3_2B (9,2 г, выход 43%) в форме желтой маслянистой жидкости.

^{1}H ЯМР: 400 МГц, $CDCl_3$

δ = 7,25 (с, 1H), 5,76 (с, 2H), 4,41 (к, *J* = 7,1 Гц, 2H), 3,60 - 3,54 (м, 2H), 1,41 (т, *J* = 7,1 Гц, 3H), 0,96 - 0,90 (м, 2H), 0,00 - -0,04 (м, 9H)

Стадия 2. Синтез QFASG-3_3B.

Смесь QFASG-3_2B (2,3 г, 6,58 ммоль) в NH_3/MeOH (7 М, 20 мл) нагревали до 70 °C и перемешивали в течение 12 ч. По данным ЖХ-МС реагент был израсходован, и было обнаружено 90% целевого соединения. Было поставлено еще три аналогичных опыта. После доведения до комнатной температуры все четыре реакционные массы отфильтровали. Фильтраты объединили и удалили растворитель при пониженном давлении, получив QFASG-3_3B в виде желтого твердого вещества (9,0 г, выход 97,8%).

^{1}H ЯМР: 400 МГц, $CDCl_3$

δ = 7,23 (с, 1H), 5,86 (с, 2H), 3,64 - 3,50 (м, 2H), 1,00 - 0,91 (м, 2H), 0,06 - -0,06 (м, 9H)

Стадия 3. Синтез QFASG-3_4B.

Раствор QFASG-3_3B (4,0 г, 12,49 ммоль) в ТГФ (40 мл) охлаждали до −60 °C, затем последовательно вводили триэтиламин (3,16 г, 31,22 ммоль, 4,35 мл) и трифторуксусный ангидрид (3,15 г, 14,99 ммоль, 2,08 мл). После этого реакционную массу доводили до 5 °C и перемешивали при указанной температуре 2 ч. ЖХ-МС-контроль подтвердил полное расходование исходного вещества; содержание целевого продукта составляло примерно 90%. Смесь нагревали до комнатной температуры, разбавляли водой (120 мл) и экстрагировали этилацетатом в трёх порциях по 150 мл. Органические экстракты объединяли, промывали насыщенным раствором NaCl (300 мл), высушивали над сульфатом натрия и фильтровали. Растворитель из фильтрата удаляли при пониженном давлении. Соединение QFASG-3_4B получали в виде коричневой маслянистой жидкости (3,2 г, выход 80%).

^{1}H ЯМР: 400 МГц, $CDCl_3$

δ = 7.24 (s, 1H), 5.43 (s, 2H), 3.71 - 3.46 (m, 2H), 1.09 - 0.82 (m, 2H), 0.02 (s, 9H)

Стадия 4. Синтез QFASG-3_5B.

Смесь QFASG-3_4B (2 г, 6,62 ммоль) и 4,4,5,5-тетраметил-2-(4,4,5,5-тетраметил-1,3,2-диоксаборолан-2-ил)-1,3,2-диоксаборолана (2,52 г, 9,93 ммоль) в диоксане (20 мл) дегазировали и трижды продували азотом, затем добавляли ацетат калия (1,30 г, 13,23 ммоль) и дихлорид [1,1’-бис(дифенилфосфино)-ферроцен]палладия (968,37 мг, 1,32 ммоль). Реакционную смесь нагревали до 90 °C и перемешивали в течение 12 ч в атмосфере N_2. По данным ЖХ-МС реагент был израсходован, и было обнаружено 47% целевого соединения. После охлаждения до комнатной температуры реакционную смесь фильтровали, получая раствор QFASG-3_5B (2,1 г, выход 47%) в диоксане (20 мл).

ЖХ-МС (ESI+): R_t = 0,449 мин, m/z 268,2 (M+H)+.

Стадия 5. Синтез QFASG-3_6B.

Смесь QFASG-1_6 (1,52 г, 5,44 ммоль) и QFASG-3_5B (1,90 г, 5,44 ммоль) в диоксане (80 мл) и воде (20 мл) дегазировали трёхкратным продуванием азотом. Затем в реакционную систему вносили дихлорид [1,1′-бис(дифенилфосфино)ферроцен]палладия (796,00 мг, 1,09 ммоль) и фосфат калия (2,31 г, 10,88 ммоль). Смесь нагревали до 90 °C и выдерживали при перемешивании в атмосфере азота в течение 12 ч. ЖХ-МС-анализ показал полное расходование исходного реагента; доля целевого соединения составляла около 20%. После охлаждения до комнатной температуры в реакционную массу добавляли воду (200 мл) и проводили трехкратную экстракцию этилацетатом (по 300 мл). Объединённые органические фазы промывали насыщенным раствором NaCl (300 мл), осушали сульфатом натрия и фильтровали. Растворитель из фильтрата удаляли при пониженном давлении. Полученный остаток очищали колоночной хроматографией на силикагеле, применяя этилацетат в качестве элюента (R_f = 0,20). QFASG-3_6B выделяли как желтое твердое вещество в количестве 400 мг (выход 20%).

^{1}H ЯМР: 400 МГц, CD_3OD

δ = 8,90 (д, *J* = 7,3 Гц, 1H), 8,55 (с, 1H), 8,04 (с, 1H), 7,06 (д, *J* = 7,3 Гц, 1H), 5,61 (с, 2H), 3,72 - 3,67 (м, 2H), 0,98 (т, *J* = 7,9 Гц, 2H), 0,03 - 0,00 (м, 9H)

Стадия 6. Синтез QFASG-3_7B.

QFASG-3_7 (206,98 мг, 1,07 ммоль) и QFASG-3_6B (400 мг, 1,07 ммоль) растворяли в смеси диоксана (3,2 мл) и воды (0,8 мл). Реакционную систему дегазировали с последующей трехкратной продувкой азотом, после чего добавляли дихлорид [1,1′-бис(дифенилфосфино)ферроцен]палладия (156,14 мг, 213,39 мкмоль) и фосфат калия (452,96 мг, 2,13 ммоль). Смесь

нагревали до 90 °C и перемешивали под азотом в течение 12 ч. По данным ЖХ-МС исходный реагент был полностью израсходован, при этом целевой продукт присутствовал в количестве около 50%. Реакционную массу охлаждали до комнатной температуры, разбавляли водой (5 мл) и экстрагировали этилацетатом (3 × 10 мл). Органические слои объединяли, промывали насыщенным раствором NaCl (20 мл), высушивали над сульфатом натрия и фильтровали. После удаления растворителя при пониженном давлении остаток очищали колоночной хроматографией на силикагеле с этилацетатом в качестве элюента (R_f = 0,20). В результате получали QFASG-3_7B (290 мг, 56%) в виде желтого твердого вещества.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,16 (д, *J* = 7,5 Гц, 1H), 8,93 (с, 1H), 8,55 (с, 1H), 8,33 - 8,27 (м, 3H), 7,66 (дд, *J* = 8,2, 14,3 Гц, 3H), 6,18 (к, J = 4,4 Гц, 1H), 5,72 (с, 2H), 3,68 (т, *J* = 8,1 Гц, 2H), 2,70 (д, *J* = 4,5 Гц, 3H), 0,95 (т, *J* = 8,0 Гц, 2H), 0,06 - -0,06 (м, 9H)

Стадия 7. Синтез QFASG-3.

Раствор QFASG-3_7B (135 мг, 276,29 мкмоль) в дихлорметане (15 мл) охлаждали до 0 °C и добавляли трифторуксусную кислоту (4,14 г, 36,35 ммоль, 2,70 мл). Далее реакционную смесь нагревали до 25 °C и перемешивали при этой температуре в течение 3 ч. Согласно результатам ЖХ-МС, содержание целевого соединения составляло около 30%. Эксперимент выполняли в двух параллельных реакциях. Обе реакционные массы концентрировали при пониженном давлении до получения сухих остатков. Полученные остатки очищали препаративной ВЭЖХ, выделяя QFASG-3 в виде желтого твердого вещества (16 мг, выход 8,08%).

Условия препаративной ВЭЖХ:

- Колонка: Phenomenex Luna C18, 100 × 30 мм, 3 мкм.

- Мобильная фаза: вода (0,2% муравьиная кислота) — ацетонитрил.

- Градиент: содержание фазы B увеличивали с 10 до 50% в течение 8,0 мин.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,00 (д, *J* = 7,5 Гц, 1H), 8,50 (с, 1H), 8,21 (д, *J* = 8,8 Гц, 2H), 8,03 - 7,87 (м, 1H), 7,62 (д, *J* = 7,5 Гц, 1H), 7,58 (д, *J* = 8,8 Гц, 2H), 2,65 (с, 3H)

QFASG-4

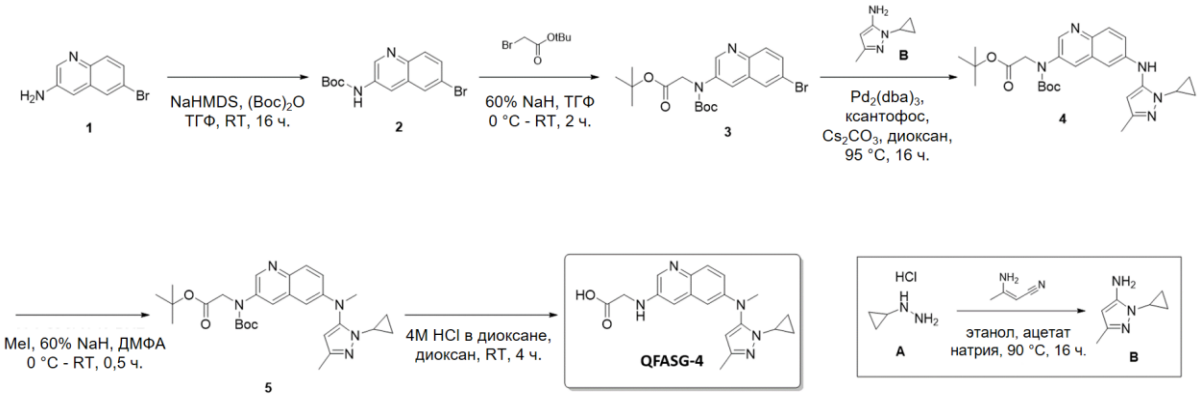


Рисунок 22. Схема синтеза соединения QFASG-4.

Стадия A. Синтез 1-циклопропил-3-метил-1H-пиразол-5-амина (B).

К перемешиваемому раствору гидрохлорида циклопропилгидразина (A) (500 мг, 4,629 ммоль) в этаноле (10 мл) добавляли (Z)-3-аминобут-2-еннитрил (379 мг, 4,629 ммоль) и ацетат натрия (949 мг, 11,57 ммоль) в герметично закрытой ампуле. Реакционную массу выдерживали при перемешивании при 90 °C в течение 16 ч; окончание превращения контролировали методом ТСХ. Затем растворитель удаляли на роторном испарителе при пониженном давлении. Остаток растворяли в EtOAc (70 мл) и последовательно промывали водой (50 мл) и насыщенным раствором NaCl (50 мл). После отделения органическую фазу осушали над Na_2SO_4 и

концентрировали в вакууме. В результате получали 1-циклопропил-3-метил-1H-пиразол-5-амин (B) в виде бесцветной жидкости (300 мг, чистота 76,35%).

ЖХ-МС (m/z): 76,35 % (M^{+1} = 138,06) при 2,80 R_t.

Стадия 1. Синтез *трет*-бутил(6-бромхинолин-3-ил)карбамата (2).

Раствор 6-бромхинолин-3-амина (1) (500 мг, 2,24 ммоль) в ТГФ (15 мл) охлаждали до 0 °C и при перемешивании добавляли 1 M раствор NaHMDS в ТГФ (5,6 мл, 5,60 ммоль). Затем вводили Boc_2O (0,77 мл, 3,36 ммоль). Смесь доводили до комнатной температуры и перемешивали в течение 16 ч. Завершение реакции устанавливали методом ТСХ. После этого реакционную массу обрабатывали насыщенным раствором NH_4Cl и экстрагировали этилацетатом (2 × 50 мл). Органическую фазу отделяли, осушали над Na_2SO_4 и удаляли растворитель при пониженном давлении. Остаток очищали колоночной хроматографией, используя 30% EtOAc в гексане в качестве элюента. В результате выделяли *трет*-бутил(6-бромхинолин-3-ил)карбамат (2) в виде коричневого твердого вещества (200 мг, 27,6%).

ЖХ-МС (m/z): 92,15 % (M^{+1} = 323,1) при 5,907 R_t.

Стадия 2. Синтез *трет*-бутил-N-(6-бромхинолин-3-ил)-N-(*трет*-бутоксикарбонил)глицината (3).

Раствор *трет*-бутил(6-бромхинолин-3-ил)карбамата (2) (250 мг, 0,773 ммоль) в ТГФ (5 мл) охлаждали до 0 °C. При перемешивании последовательно добавляли 60%-ный NaH (62 мг, 1,54 ммоль) и *трет*-бутил-2-бромацетат (301 мг, 1,54 ммоль). Далее смесь нагревали до комнатной температуры и перемешивали 2 ч; завершение реакции контролировали методом ТСХ. Реакционную массу гасили ледяной водой

(10 мл), после чего продукт экстрагировали этилацетатом (2 × 50 мл). Органическую фазу отделяли, высушивали над Na_2SO_4 и упаривали. Неочищенный остаток очищали колоночной хроматографией, используя 50% EtOAc в гексане в качестве элюента. Получали *трет*-бутил-N-(6-бромхинолин-3-ил)-N-(*трет*-бутоксикарбонил)глицинат (3) как коричневую жидкость (230 мг, выход 68%).

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 8,88 (д, *J* = 2,0 Гц, 1H), 8,29 (д, *J* = 2,4 Гц, 1H), 8,15 (д, *J* = 2,4 Гц, 1H), 7,95 (д, *J* = 8,4 Гц, 1H), 7,85 (дд, *J* = 2,0, 8,8 Гц, 1H), 4,37 (с, 2H), 1,43 (с, 18H).

Стадия 3. Синтез *трет*-бутил-N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)амино)хинолин-3-ил)глицината (4).

К перемешиваемому раствору *трет*-бутил-N-(6-бромхинолин-3-ил)-N-(*трет*-бутоксикарбонил)глицината (3) (250 мг, 0,572 ммоль) в 1,4-диоксане (7 мл) добавляли 1-циклопропил-3-метил-1H-пиразол-5-амин (117 мг, 0,858 ммоль), карбонат цезия (371 мг, 1,14 ммоль), трис((1E,4E)-1,5-дифенилпента-1,4-диен-3-он)дипалладий (10,4 мг, 0,011 ммоль) и [5-(дифенилфосфанил)-9,9-диметил-9H-ксан-4-ил]дифенилфосфан (16,5 мг, 0,028 ммоль) при комнатной температуре. Реакционную смесь дегазировали аргоном в течение 10 мин и нагревали до 90 °C в течение 16 ч. После подтверждения завершения реакции методом ТСХ смесь пропускали через слой целита, промывая фильтр этилацетатом (50 мл). Полученный фильтрат концентрировали при пониженном давлении. Продукт очищали колоночной хроматографией, элюируя 40% EtOAc в гексане, и получали *трет*-бутил-N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)амино)хинолин-3-ил)глицинат (4) (230 мг, выход 81%) в виде коричневой жидкости.

ЖХ-МС (m/z): 94,59 % (M^{+1} = 494,3) при 9,058 R_t.

Стадия 4. Синтез *трет*-бутил-N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)(метил)амино)хинолин-3-ил)глицината (5).

К перемешиваемому раствору *трет*-бутил-N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)амино)хинолин-3-ил)глицината (4) (230 мг, 0,466 ммоль) в DMF (8 мл) при 0 °C добавляли 60%-ный NaH (37,3 мг, 0,933 ммоль), затем добавляли CH_3I (0,43 мл, 0,699 ммоль). Реакционную массу выдерживали при перемешивании при комнатной температуре 30 мин; окончание реакции контролировали методом ТСХ. Затем её гасили ледяной водой (5 мл) и дважды экстрагировали этилацетатом (по 50 мл). Органическую фазу отделяли, осушали над Na_2SO_4 и удаляли растворитель при пониженном давлении. В результате выделяли *трет*-бутил-N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)(метил)амино)хинолин-3-ил)глицинат (5) в виде бесцветной жидкости (220 мг, выход 93%).

ЖХ-МС (m/z): 98,28 % (M^{+1} = 508,19) при 7,58 R_t.

Стадия 5. Синтез (6-((1-циклопропил-3-метил-1H-пиразол-5-ил)(метил)амино)хинолин-3-ил)глицина (QFASG-4).

К перемешиваемому раствору *трет*-бутил N-(*трет*-бутоксикарбонил)-N-(6-((1-циклопропил-3-метил-1H-пиразол-5-ил)(метил)амино)хинолин-3-ил)глицината (5) (230 мг, 0,453 ммоль) в 1,4-диоксане (3 мл) добавляли 4 М раствор соляной кислоты в 1,4-диоксане (3 мл) при 0 °C. Реакционную массу выдерживали при комнатной температуре при перемешивании в течение 4 ч. После подтверждения завершения реакции методом ТСХ растворитель удаляли при пониженном давлении, получая остаток. Продукт трижды растирали с диэтиловым эфиром (50 мл)

и н-пентаном (50 мл), в результате чего получали 130 мг (6-((1-циклопропил-3-метил-1H-пиразол-5-ил)(метил)амино)хинолин-3-ил)глицин (QFASG-4) в виде бледно-желтого твердого вещества (выход 81%).

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 13,001 (шир. с, 1H), 8,55 (д, *J* = 1.6 Гц, 1H), 7,88 (д, *J* = 8,2 Гц, 1H), 7,74 (с, 1H), 7,16 - 7,00 (м, 3H), 6,02 (с, 1H), 4,07 (с, 2H), 3,35 (с, 3H), 3,21 - 3,17 (м, 1H), 2,16 (с, 3H), 0,94 (д, *J* = 3,6 Гц, 2H), 0,75 (д, *J* = 5,2 Гц, 2H).

ЖХ-МС: 98,58 % (M^{+1} = 352,2) при 3,552 R_t.

ВЭЖХ: 96.80% чистоты при 6,673 R_t.

QFASG-5

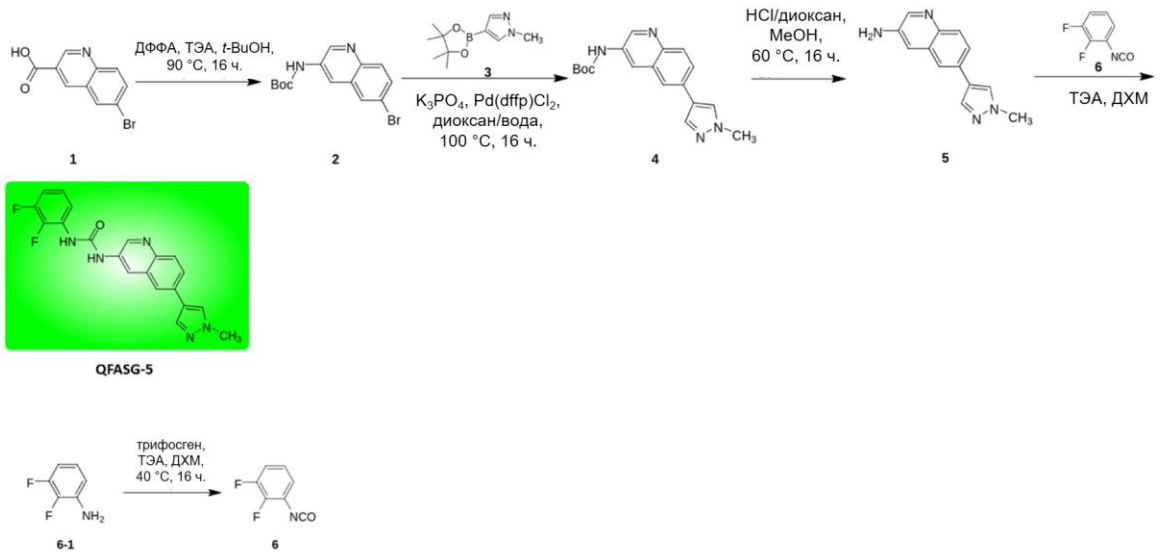


Рисунок 23. Схема синтеза соединения QFASG-5.

Стадия 1. Синтез *трет*-бутил-N-(6-бром-3-хинолил)карбамата (2).

Раствор 6-бромхинолин-3-карбоновой кислоты (4,5 г, 17,90 ммоль) и TEA (3,62 г, 35,80 ммоль, 5 мл) в *трет*-бутаноле (36 мл) дегазировали азотом в течение 5 мин, затем добавляли дифенилфосфорилазид (9,85 г, 35,80 ммоль, 8 мл). Реакционную массу выдерживали 4 ч при перемешивании при 90 °C. По данным ЖХ-МС реагент был полностью израсходован, и был обнаружен новый пик с целевым соединением. Раствор смешали с водой и этилацетатом с последующим разделением фракций. Водную фазу дважды экстрагировали этилацетатом (по 50 мл). Объединенные органические экстракты промывали насыщенным раствором $NaHCO_3$ (50 мл), затем насыщенным раствором NaCl (50 мл), осушали над Na_2SO_4 и фильтровали. После удаления растворителя при пониженном давлении полученный остаток дополнительно очищали хроматографией на силикагеле (80 г), используя 40–50% EtOAc в петролейном эфире в качестве элюента. Соединение *трет*-бутил-N-(6-бром-3-хинолил)карбамат (2,17 г, 6,71 ммоль, выход 38%) было получено в виде желтой маслянистой жидкости.

МС (ESI): расчетное 323,0 $[M\ (^{79}Br) + H]^+$, 325,0 $[M\ (^{81}Br) + H]^+$; измеренное 323,0 $[M\ (^{79}Br) + H]^+$, 325,0 $[M\ (^{81}Br) + H]^+$

Стадия 2. Синтез *Трет*-бутил-N-[6-(1-метилпиразол-4-ил)-3-хинолил]карбамата (4).

К раствору *трет*-бутил-N-(6-бром-3-хинолил)карбамата (2,0 г, 6,19 ммоль) и 1-метил-4-(4,4,5,5-тетраметил-1,3,2-диоксаборолан-2-ил)пиразола (1,55 г, 7,43 ммоль) в 1,4-диоксане (30 мл) и воде (2 мл) добавляли $Pd(dppf)Cl_2 \cdot CH_2Cl_2$ (505,37 мг, 618,84 мкмоль) и K_3PO_4 (3,28 г, 15,47 ммоль). Затем смесь дегазировали и продували азотом в течение 2 мин, почле чего смесь перемешивали в течение 16 ч при 100 °C. По данным ЖХ-МС

исходные вещества были полностью израсходованы, и был обнаружен новый пик с целевым соединением. К раствору добавляли воду (50 мл), после чего проводили двукратную экстракцию этилацетатом (по 50 мл). Органические экстракты объединяли, промывали насыщенным раствором хлорида натрия, осушали над Na_2SO_4 и фильтровали. Фильтрат концентрировали при пониженном давлении, получая остаток, который очищали на колонке с силикагелем (24 г), используя градиент 70–95% EtOAc в петролейном эфире. Соединение *трет*-бутил-N-[6-(1-метилпиразол-4-ил)-3-хинолил]карбамат (1,3 г, 4,00 ммоль, выход 64%) было получено в виде белого твердого вещества.

МС (ESI): расчетное 325,2 $[(M + H)^+]$; измеренное 325,2 $[(M + H)^+]$

Стадия 3. Синтез 6-(1-метилпиразол-4-ил)хинолин-3-амина (5).

К раствору *трет*-бутил-N-[6-(1-метилпиразол-4-ил)-3-хинолил]карбамата (0,50 г, 1,54 ммоль) в метаноле (10 мл) добавляли 4 М раствор HCl в 1,4-диоксане (4 мл). Реакционную массу выдерживали при 60 °C в течение 16 ч при перемешивании. После удаления растворителя в вакууме выделяли 6-(1-метилпиразол-4-ил)хинолин-3-амин в форме гидрохлорида (0,40 г, 1,53 ммоль, выход 99,53%) как белое твердое вещество. Полученный материал использовали в следующей стадии без дальнейшей очистки.

Стадия 4. Синтез 1-(2,3-дифторфенил)-3-[6-(1-метилпиразол-4-ил)-3-хинолил]мочевины (QFASG-5).

К раствору 2,3-дифторанилина (3,0 г, 23,24 ммоль, 2,36 мл) в толуоле (30 мл) добавляли TEA (5,88 г, 58,09 ммоль, 8 мл) и *бис*(трихлорметил)карбонат (2,34 г, 7,90 ммоль), после чего смесь перемешивали при 40 °C в течение 16 ч. Реакционную смесь фильтровали, а фильтрат собирали, получая 1,2-дифтор-3-изоцианатобензол (3,0 г) в виде

бесцветной жидкости, которую использовали на следующей стадии без очистки. Раствор 6-(1-метилпиразол-4-ил)хинолин-3-амина в форме гидрохлорида (0,40 г, 1,53 ммоль) в дихлорметане (20 мл) обрабатывали триэтиламином (466 мг, 4,60 ммоль), затем вводили 1,2-дифтор-3-изоцианатобензол (357 мг, 2,30 ммоль). Смесь перемешивали при 25 °C в течение 1 ч. ЖХ-МС-анализ показал полное исчезновение исходного вещества и появление нового пика, соответствующего целевому соединению. После завершения реакции к смеси добавляли воду (50 мл) и извлекали продукт дихлорметаном в трех порциях по 50 мл. Объединённые органические фазы промывали насыщенным раствором NaCl, осушали Na_2SO_4, фильтровали и концентрировали в вакууме. Остаток очищали препаративной ВЭЖХ на колонке Waters XBridge (150 × 25 мм, 5 мкм), используя систему вода с NH_4HCO_3/ацетонитрил и градиент 32–62% фазы B за 8 мин. После очистки материал лиофилизировали. Соединение 1-(2,3-дифторфенил)-3-[6-(1-метилпиразол-4-ил)-3-хинолил]мочевина (22,04 мг, 58,10 мкмоль, выход 3,79%) было получено в виде белого твердого вещества.

МС (ESI): расчетное 380,1 [$(M + H)^+$]; измеренное 380,1 [$(M + H)^+$]

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 9,54 (с, 1 H), 8,96 (с, 1 H), 8,72 (д, *J* = 2,4 Гц, 1H), 8,51 (д, *J* = 2,0 Гц, 1H), 8,30 (с, 1 H), 8,09 (д, *J* = 2,0 Гц, 1H), 8,04 (с, 1 H), 7,99 (с, 1 H), 7,93-7,90 (м, 1 H), 7,90-7,86 (м, 1 H), 7,20-7,18 (м, 1 H), 7,10-7,09 (м, 1 H) , 3,91 (с, 3 H).

QFASG-6

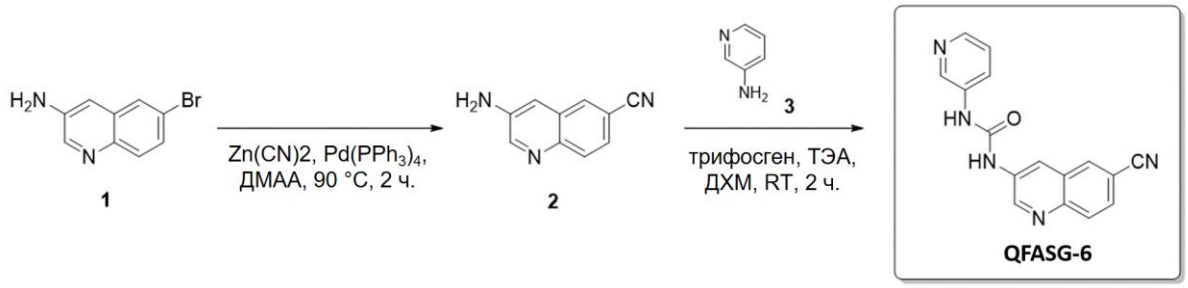


Рисунок 24. Схема синтеза соединения QFASG-6.

Стадия 1. Синтез 3-аминохинолин-6-карбонитрила (2).

К раствору 6-бромхинолин-3-амина (1) (100 мг, 0,448 ммоль) в диметилацетамиде (2 мл) при комнатной температуре при перемешивании добавляли $Zn(CN)_2$ (304 мг, 2,242 ммоль) и $Pd(PPh_3)_4$ (155 мг, 0,134 ммоль). Реакционную массу нагревали в микроволновом реакторе при 90°C в течение 2 ч. Завершение реакции контролировали методом ТСХ. Далее смесь разбавляли водой (10 мл) и этилацетатом (50 мл), фильтровали через слой целита и промывали фильтр этилацетатом (20 мл). Органическую фазу отделяли, осушали над Na_2SO_4 и концентрировали при пониженном давлении. Продукт очищали колоночной хроматографией, элюируя 40% EtOAc в гексане, и получали 3-аминохинолин-6-карбонитрил (2) в виде бледно-желтой жидкости (70 мг, чистота 84%).

ЖХ-МС: 84,51 % (M^{+1} = 170,10) при 4,44 R_t.

Стадия 2. Синтез 1-(6-цианхинолин-3-ил)-3-(пиридин-3-ил)мочевины (QFASG-6).

К перемешиваемому раствору 3-аминохинолин-6-карбонитрила (2) (120 мг, 0,71 ммоль) в дихлорметане (4 мл) при 0 °C добавляли трифосген (420 мг, 1,42 ммоль) и триэтиламин (0,18 мл, 1,42 ммоль), после чего в смесь

добавляли пиридин-3-амин (80 мг, 0,85 ммоль). Реакционную массу выдерживали при перемешивании при комнатной температуре в течение 2 ч; окончание реакции подтверждали методом ТСХ. Затем к смеси добавляли воду (10 мл) и проводили двукратную экстракцию дихлорметаном (по 50 мл). Органические экстракты объединяли, осушали над Na_2SO_4 и удаляли растворитель при пониженном давлении, получая продукт. Полученный продукт очищали методом препаративной ВЭЖХ с использованием 0,1% водного раствора муравьиной кислоты и ацетонитрила (градиент 0–30% ACN, колонка Luna C18, 250 × 50 мм, 10 мкм), в результате чего получали 10 мг 1-(6-цианхинолин-3-ил)-3-(пиридин-3-ил)мочевины (QFASG-6) в виде грязно-белого твердого вещества (выход 18,9%).

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 9,77 (с, 1H), 9,43 (с, 1H), 9,01 (с, 1H), 8,67 (д, *J* = 2 Гц, 2H), 8,62 (с, 1H), 8,24 (д, *J* = 4,4 Гц, 1H), 8,09 (д, *J* = 8,4 Гц, 1H), 7,99 (д, *J* = 8,0 Гц, 1H), 7,88 (д, *J* = 8,4 Гц, 1H), 7,37 – 7,34 (м, 1H).

ЖХ-МС: 98,88 % (M^{+1} = 352,2) при 3,524 R_t.

ВЭЖХ: 98.95% чистоты при 6,404 R_t.

2.6. Методики *in vitro* тестирования

В данном разделе приводятся методики тестирования ингибирующей активности обнаруженных соединений в отношении целевых протеинкиназ ATM, CAMKK2 и TTK, а также в отношении киназ ATR и DNA-PK, которые являются родственными ATM белками и позволяют оценить селективность ингибиторов ATM.

2.6.1. Анализ ингибирования киназы CAMKK2

CaMKK2 (человеческая) инкубировали в буфере для анализа, содержащем 20 мМ Tris/HCl, pH 8,5, 0,2 мМ ЭДТА, 0,5% БСА, 0,5 мМ

CaCl2, 0,016 мг/мл кальмодулина, 150 мкМ LSNLYHQGKFLQTFCGSPLYRRR, 10 мМ ацетата магния и [γ-33P]-АТФ (удельная активность и концентрация 10 мкМ). Реакцию запускали внесением раствора Mg/АТФ. Смесь инкубировали 40 мин при комнатной температуре, после чего процесс прекращали добавлением фосфорной кислоты до итоговой концентрации 0,5%. Далее часть реакционной смеси переносили на фильтр, который четырежды промывали 0,425%-ным раствором фосфорной кислоты в течение 4 мин на каждую промывку, затем однократно промывали метанолом. После высушивания фильтры подвергали сцинтилляционному счету.

2.6.2. Анализ ингибирования киназы ATM

Человеческую ATM инкубировали в аналитическом буфере в присутствии GST-cMyc-p53 (30 нМ) и Mg/АТФ (10 мкМ). Реакцию запускали внесением Mg/АТФ-смеси и проводили при комнатной температуре в течение 30 мин. Затем реакцию останавливали стоп-буфером, содержащим ЭДТА. Для последующей детекции вносили буфер с d2-меченным моноклональным антителом к GST и антителом к фосфо-Ser15 p53, меченным европием. После этого планшет регистрировали в режиме временно-разрешенной флуоресценции, а сигнал HTRF определяли по формуле: HTRF = 10000 × (Em665 нм / Em620 нм).

2.6.3. Анализ ингибирования киназы DNA-PK

Человеческую DNA-PK инкубировали в аналитическом буфере в присутствии GST-cMyc-p53 (30 нМ) и Mg/АТФ (10 мкМ). Реакцию запускали внесением Mg/АТФ-смеси и проводили при комнатной температуре в течение 30 мин. Затем реакцию останавливали стоп-буфером, содержащим ЭДТА. Для последующей детекции вносили буфер с d2-меченным моноклональным антителом к GST и антителом к фосфо-Ser15

p53, меченным европием. После этого планшет регистрировали в режиме временно-разрешенной флуоресценции, а сигнал HTRF определяли по формуле: HTRF = 10 000 × (Em665 нм / Em620 нм).

2.6.4. Анализ ингибирования киназы ATR

Человеческую ATR инкубировали в аналитическом буфере в присутствии GST-cMyc-p53 (30 нМ) и Mg/АТФ (10 мкМ). Реакцию запускали внесением Mg/АТФ-смеси и проводили при комнатной температуре в течение 30 мин. Затем реакцию останавливали стоп-буфером, содержащим ЭДТА. Для последующей детекции вносили буфер с d2-меченным моноклональным антителом к GST и антителом к фосфо-Ser15 p53, меченным европием. После этого планшет регистрировали в режиме временно-разрешенной флуоресценции, а сигнал HTRF определяли по формуле: HTRF = 10 000 × (Em665 нм / Em620 нм).

2.6.5. Анализ ингибирования киназы TTK

Растворы 2× АТФ и субстрата, а также 2× раствора киназы готовили в буфере для анализа, содержащем 50 мМ HEPES, 10 мМ MgCl2, 0,01% Brij35, 1 мМ EGTA, 2 мМ DTT и ddH2O. В лунки 384-луночного планшета с помощью Echo 655 переносили 40 нл соединения. Затем в лунку добавляли 2 мкл 2× раствора киназы и металла и инкубировали в 384-луночном планшете 10 мин при 25 °C. Затем в каждую лунку вносили 2 мкл двойного по концентрации раствора субстрата с АТФ и выдерживали планшет при 25 °C в течение 60 мин. После этого добавляли 4 мкл реагента ADP-Glo и инкубировали ещё 40 мин при 25 °C. На заключительном этапе вносили 8 мкл реагента Kinase Detection Reagent, после чего проводили инкубацию при 25 °C в течение 40 мин. Сигналы люминесценции регистрировали на микропланшетном ридере.

ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

3.1. *In silico* валидация

3.1.1 Поиск сайтов связывания

Для объективной оценки точности и обобщающей способности разработанного алгоритма SiteRadar была проведена сравнительная проверка точности обнаружения сайтов связывания с известными алгоритмами FPocket и PUResNet. Тестирование выполнялось на специально подготовленном наборе из 232 экспериментально охарактеризованных лиганд-белковых комплексов, обладающих низким сходством аминокислотных последовательностей с обучающей выборкой [60].

Результаты показали, что SiteRadar превосходит известные методы по способности идентифицировать лиганд-белковые сайты связывания. По показателю топ N+2, характеризующему долю успешно решенных комплексов, где удалось обнаружить истинный лиганд-белковый сайт связывания среди трех наиболее высокоранговых предсказаний модели, SiteRadar продемонстрировал точность 72–74% (рисунок 25) в зависимости от модели (учитывающая химические данные или геометрическая), тогда как FPocket и PUResNet демонстрировали 46% и 47% соответственно [60].

Вероятно, это достигается в том числе благодаря тому, что SiteRadar генерирует разумное количество лиганд-белковых сайтов связывания — в среднем 4,2 и 7,6 сайтов на белок для учитывающей химию и геометрической моделей соответственно (рисунок 25). При этом FPocket склонен генерировать чрезмерно большое число сайтов связывания (19,2), что не позволяет ему в дальнейшем определить, какой из сгенерированных сайтов является истинным. Напротив, PUResNet в среднем определяет один

сайт связывания для каждого белка, что ограничивает его возможность в обнаружении истинных полостей связывания малых молекул. Сбалансированность SiteRadar способствует высокой точности без избытка ложноположительных результатов. В то время как FPocket имеет тенденцию генерировать слишком крупные и многочисленные сайты, PUResNet, наоборот, ограничен небольшим количеством предсказанных полостей с менее чем 50% совпадением с истинными сайтами связывания малых молекул [60].

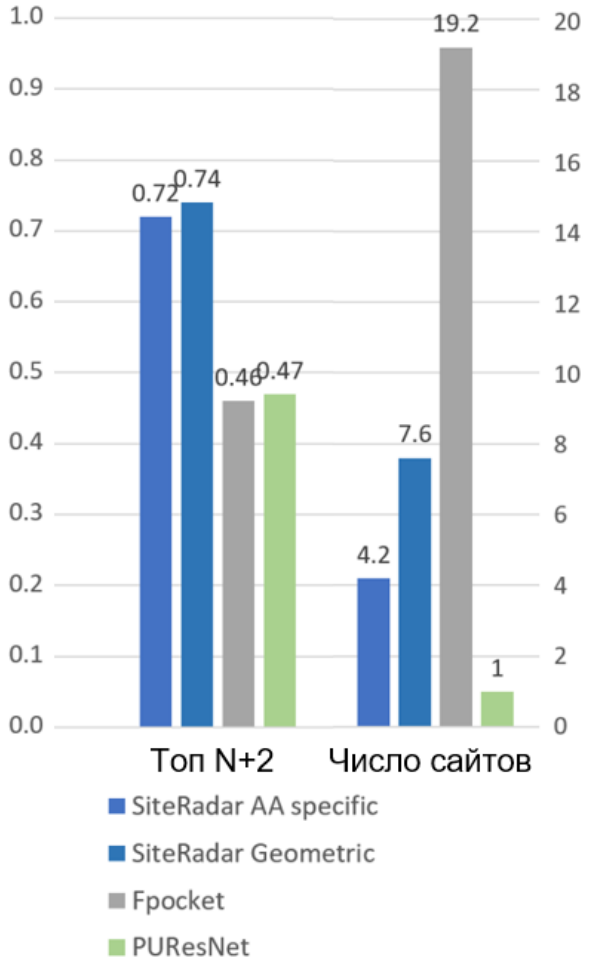


Рисунок 25. Сравнение точности по метрике топ N+2 и количеству предсказанных лиганд-белковых сайтов связывания для SiteRadar, FPocket и PUResNet.

Что касается точности воспроизведения объема и формы лиганд-белковых сайтов связывания, SiteRadar показал хорошие показатели по метрикам покрытия атомов лиганда (ligand coverage, LC) и покрытия сайта связывания (pocket coverage, PC). SiteRadar также отличается тем, что лучше разделяет большие предсказанные сайты связывания на логичные, отдельные полости, что улучшает качество предсказаний для крупных лигандов и мультифункциональных сайтов.

Также был выполнен анализ геометрических характеристик сайтов связывания, прогнозируемых алгоритмом SiteRadar. Оценка выполнялась по таким параметрам, как глубина и объем обнаруженной полости. Сравнение с известными сайтами связывания малых молекул показало, что большинство сайтов связывания, генерируемых алгоритмом, обладает характеристиками, свойственными для полостей, пригодных для связывания лекарственноподобных соединений (рисунок 26).

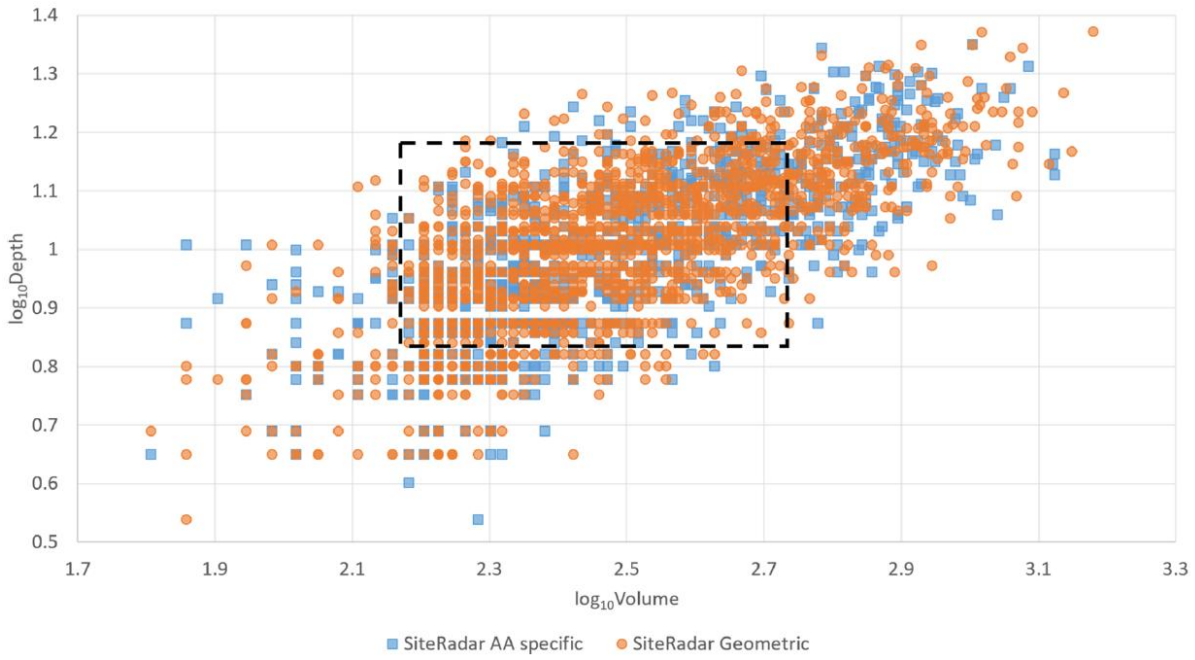


Рисунок 26. Визуализация объемных параметров для сайтов связывания, обнаруженных с помощью SiteRadar. Черный пунктир определяет пространство известных лиганд-белковых сайтов связывания.

Таким образом, SiteRadar демонстрирует более точное и избирательное выявление сайтов связывания по сравнению с FPocket и PUResNet, что делает его перспективным инструментом для интеграции в автоматизированные системы разработки лекарств.

3.1.2 Примеры применения метода поиска сайтов связывания

Разработанный алгоритм для поиска сайтов связывания был применен к трехмерным структурам различных белковых мишеней. Для анализа использовались белки, не имеющие аналогов в обучающей выборке с точки зрения сходства аминокислотных последовательностей (максимальная доля сходства не превышала 43%) [60].

Идентификация поверхностного, слабо выраженного сайта связывания в NADH-оксидазе. Сайт связывания, расположенный на поверхности NADH-оксидазы, известен своей трудностью обнаружения как геометрическими, так и химическими методами [26]. Модель SiteRadar, учитывающая аминокислотную информацию, эффективно картирует этот сайт связывания, исключая ложноположительные полости (рисунок 27a). Хотя геометрическая модель SiteRadar не способна полностью отобразить данный сайт связывания, она фиксирует его часть (рисунок 27b). Дополнительно, геометрическая модель обнаруживает вторую полость, прилегающую к основному сайту связывания. Этот пример подчеркивает преимущество использования химической информации при выявлении небольших или экспонированных в растворитель сайтов связывания [20].

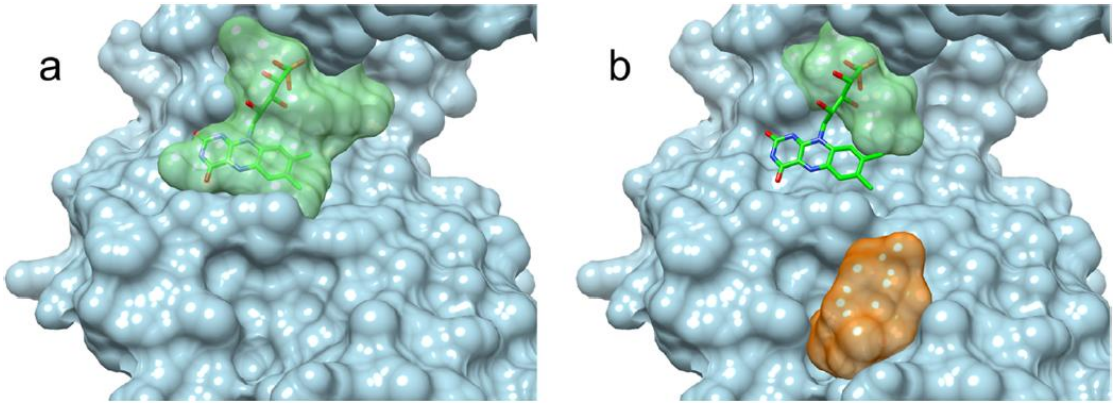


Рисунок 27. Визуализация сайта связывания, обнаруженного в структуре NADH-оксидазы (PDB ID 1NOX) с помощью химической (a) и геометрической (b) моделей SiteRadar.

Обнаружение аллостерического сайта в Polycomb белке EED. Обнаружение аллостерических сайтов является нетривиальной задачей для вычислительных методов [91]. Применение обеих моделей SiteRadar к структуре Polycomb белка EED (PDB ID 7QJG) показало успешное выявление аллостерического сайта и дополнительных полостей (одна для химической и пять для геометрической модели) (рисунок 28). Это указывает, что химическая модель более избирательна, в то время как геометрическая способна охватывать более обширные прилегающие области [20].

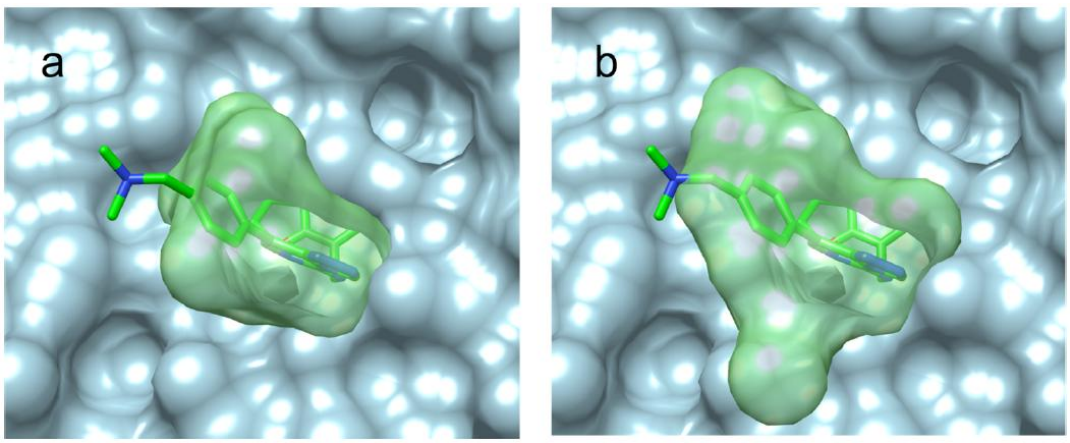


Рисунок 28. Визуализация обнаружения аллостерического сайта связывания в polycomb белке EED (PDB ID 7QJG) с помощью химической (a) и геометрической (b) моделей SiteRadar.

Идентификация сайта ковалентно-связывающегося лиганда CREB-белка. Ковалентные ингибиторы набирают все большую популярность [92]. Несмотря на то, что SiteRadar не был специально предназначен для идентификации сайтов ковалентно-связывающихся лигандов, он успешно идентифицировал такой сайт связывания в CREB-белке. Обе модели, геометрическая и химическая, выявили местоположение сайта ковалентного связывания, покрыв большинство тяжелых атомов с ко-кристаллизованным лигандом, за исключением ковалентно-связывающейся группы, выступающей в раствор (минимальное расстояние до ближайшего атома белка составляло 5,66 Å) (рисунок 29). В данном случае химическая модель точнее повторила форму лиганда, хотя различия с геометрической моделью были незначительны [20].

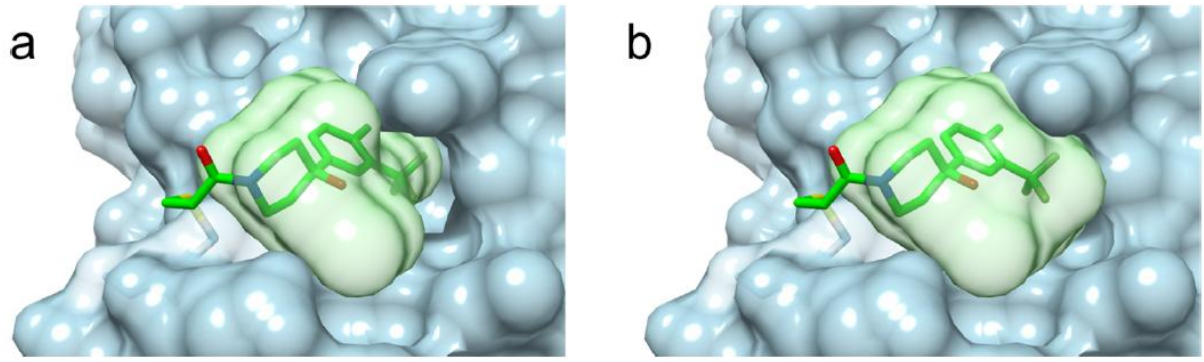


Рисунок 29. Сайт ковалентного связывания лиганда с белком CREB (PDB ID 4I9O), выявленный с помощью химической (a) и геометрической (b) моделей SiteRadar.

Идентификация сайта связывания на интерфейсе белок–белок (PPI) в репликационном белке A. Воздействие посредством связывания с PPI-сайтами приобретает все большую значимость [48]. Было проведено исследование по способности разработанного алгоритма обнаруживать такие сайты. Химическая и геометрическая модели были применены для анализа репликационного белка A. Химическая модель не полностью заполнила целевой PPI-сайт связывания. При этом были обнаружены два

раздельных кластера (рисунок 30a). Разрыв между кластерами сохранялся при любых изменениях параметров. Это может быть объяснено необычным аминокислотным окружением или присутствием ковалентно связанного остатка.

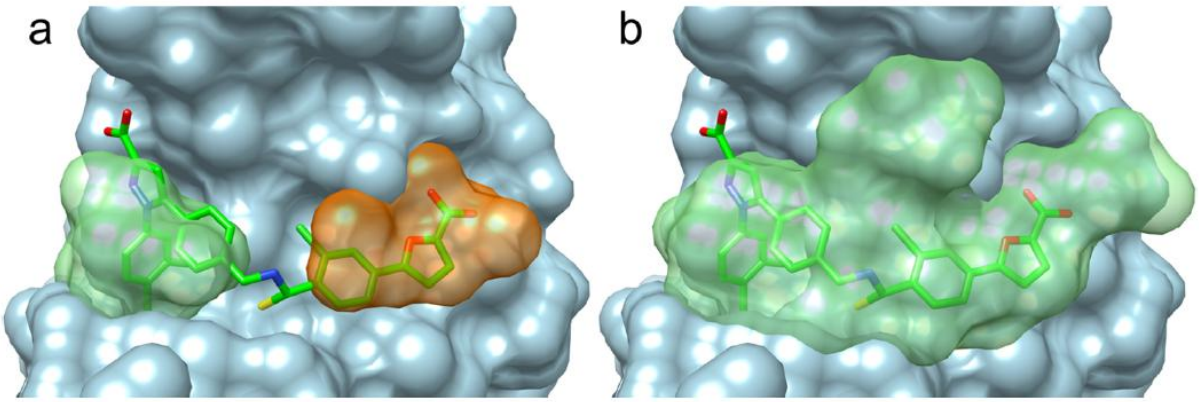


Рисунок 30. Визуализация сайта связывания PPI в репликационном белке A (PDB ID 4O0A) с помощью химической (a) и геометрической (b) моделей SiteRadar.

Геометрическая модель сформировала единую крупную область связывания (рисунок 30b). Это определяет преимущество для использования данной модели с целью дизайна химических соединений с оптимальным покрытием сайта связывания.

SiteRadar показал способность идентифицировать разнообразные типы сайтов связывания. Были проанализированы экспонированные в раствор, аллостерические, PPI и сайты связывания ковалентных лигандов. Химическая и геометрическая модели имеют свои преимущества и ограничения; химическая информация повышает точность идентификации сайтов нестандартной геометрии, тогда как геометрическая модель лучше справляется с нетипичными аминокислотными композициями [20].

Результаты показали более высокую точность SiteRadar по сравнению с известными методами, такими как FPocket и PUResNet, в идентификации истинных сайтов связывания для белков из выборки с низким сходством

аминокислотных последовательностей с данными обучающей выборки. Способность надежно идентифицировать лиганд-белковые сайты связывания при низком уровне ложноположительных результатов делает SiteRadar подходящим для использования на платформах автоматизированного дизайна лекарственных молекул. Кроме того, разработанная модель успешно выявляет экспонированные в растворитель, аллостерические, PPI сайты, а также участки связывания ковалентных лигандов, что обеспечивает ее универсальность.

3.1.3 Оптимизация геометрии сайта связывания

Было показано, что оптимизация белковых структур с использованием разработанного подхода значительно уменьшает количество стерических препятствий для лигандов, увеличивает точность докинга и улучшает результаты виртуального скрининга.

Выполненный анализ выявил, что структуры, обработанные с помощью AF Optimizer, показали снижение как абсолютного числа нежелательных межмолекулярных контактов, так и доли тяжелых атомов лиганда, формирующих стерические препятствия с белком (рисунок 31). Большинство как исходных, так и оптимизированных структур AF формировали до пяти межмолекулярных стерических препятствий, приводящих к невыгодным взаимодействиям у до 20% тяжелых атомов лиганда, но распределение в этом диапазоне значительно отличалось до и после оптимизации.

Количество оптимизированных структур без стерических затруднений было примерно вдвое выше, чем тех, у которых было от одного до пяти препятствий (до 20% тяжелых атомов лиганда), в то время как исходные белковые модели демонстрировали сопоставимое количество структур в этих группах. Напротив, применение только силового поля не улучшило

результат, так как оно направлено на оптимизацию межмолекулярных взаимодействий без учета возможности потенциального связывания лиганда. Этот эксперимент показал важность использования нейронных сетей, которые могут различать пространство потенциального связывания лиганда от той зоны, которую разрешено занимать остаткам сайта связывания.

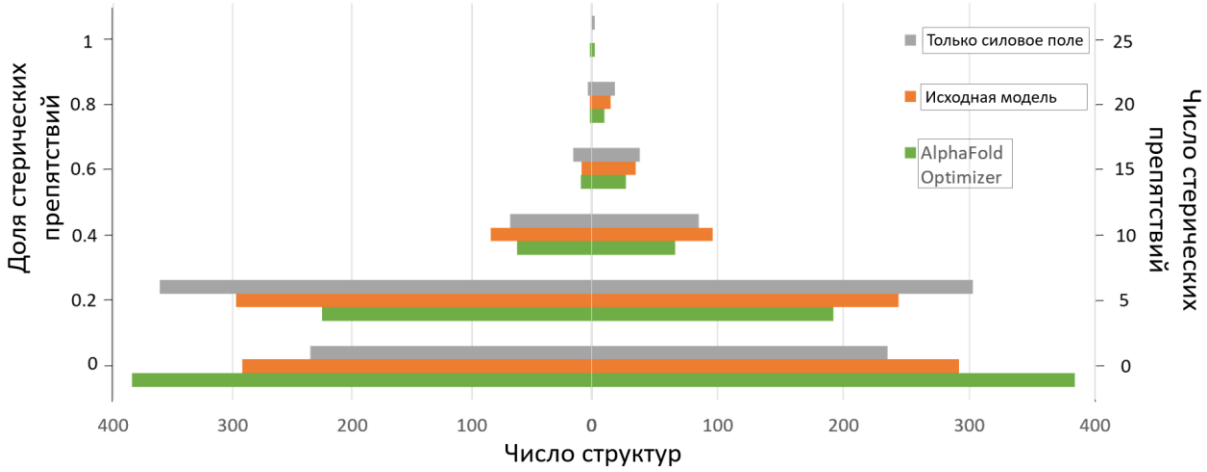


Рисунок 31. Результаты оптимизации стерических затруднений за счет применения силового поля или AF Optimizer.

Также было показано, что исходные модели AF не подходят для выполнения молекулярного докинга, так как лиганды неправильно располагаются в активном центре при сравнении с позициями референсных соединений. Напротив, оптимизированные структуры показали более оптимальные результаты по параметрам RMSD, расчетной энергии связывания и ранжированию позиций, наиболее похожих на позиции референсных лигандов (рисунок 32). Для citron Rho-interacting киназы основное различие в структуре белка до и после оптимизации сайта связывания заключалось в пространственном расположении боковых цепей Met^{174} и Tyr^{176}. В случае STK-33 смещения главным образом наблюдались для остатков Met^{245} и Met^{192}. В оригинальных белковых моделях AF некоторые из этих остатков находились в пространстве, занимаемом

атомами лиганда, и правильные позиции связывания не могли быть достигнуты при докинге. Напротив, уточненные структуры были более подходящими для докинга, и корректные положения были получены. В итоге, AF Optimizer позволил получить модели белков, более пригодные для молекулярного докинга и дизайна молекул на основе структуры.

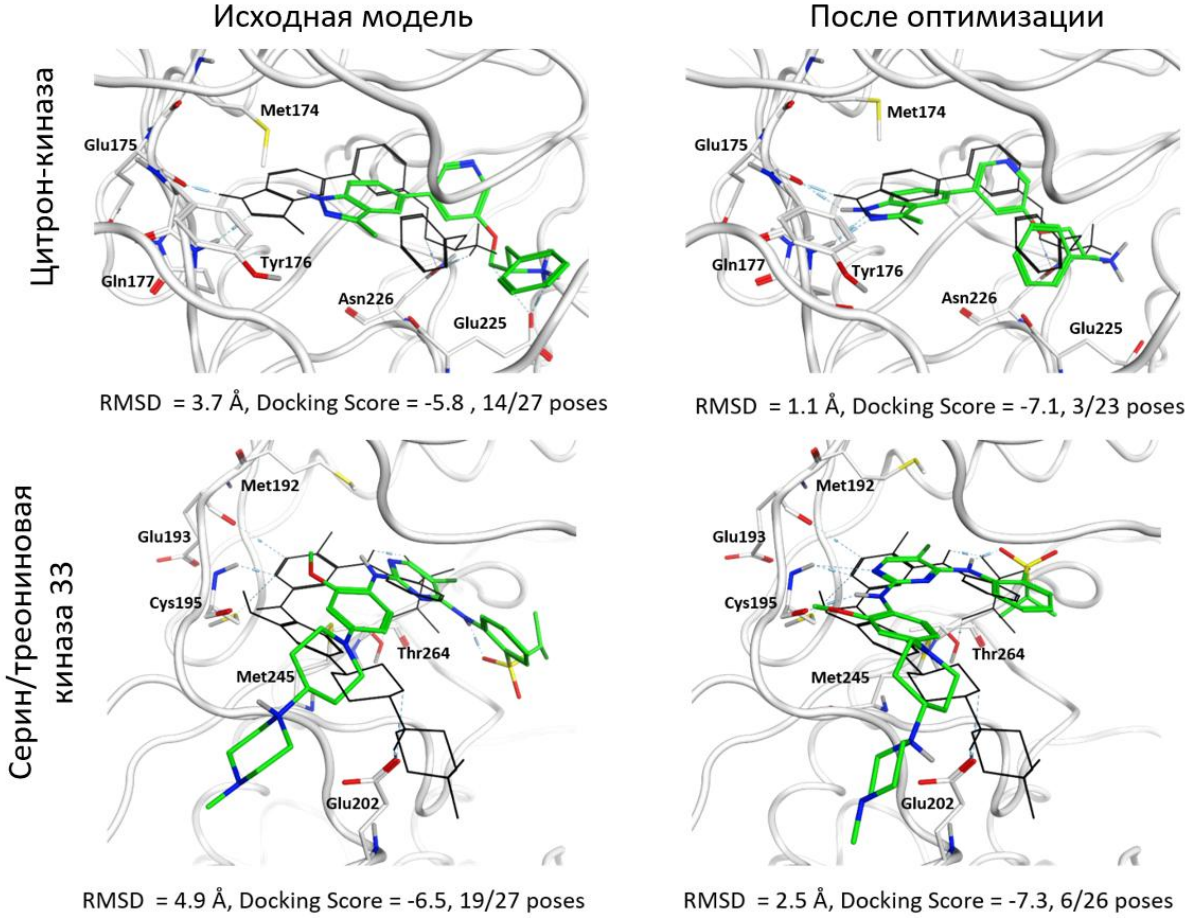


Рисунок 32. Результаты докинга малых молекул в сайты связывания до и после применения AF Optimizer.

Затем был проведен ретроспективный виртуальный скрининг, чтобы убедиться в практической применимости AF Optimizer для задач структурно-обоснованного дизайна. В качестве мишени была выбрана киназа CDK2, для которой описано большое число известных ингибиторов. Была выполнена оптимизация трехмерной модели киназы CDK2. Был выполнен молекулярный докинг 180 ее известных ингибиторов в

оригинальную и оптимизированную модели белка. Оптимизация позволила скорректировать трехмерное расположение важных остатков в области активного центра, таких как Lys^{89} и Tyr^{15}. Это прямо повлияло на повышение эффективности поиска ингибиторов и повышение числа обнаруживаемых соединений-хитов. Оптимизированная модель CDK2 позволила выявить 72 соединения-хита. Оригинальная модель позволила обнаружить только 51 ингибитор. Из их числа 41 структура успешно была докирована в обе модели (рисунок 33A). 10 – Только в оригинальную модель было успешно докировано 10 структур (рисунок 33B) и 31 – только в оптимизированную (рисунок 33C).

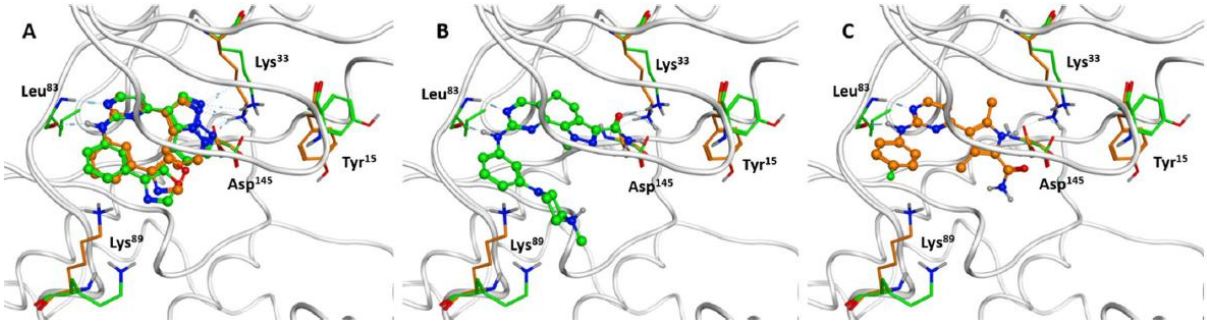


Рисунок 33. Примеры результатов ретроспективного виртуального скрининга.

AF Optimizer показал повышение точности при использовании сгенерированных с помощью AlphaFold 2 белковых моделей для разработки новых ингибиторов киназ. Такой подход снижает основные ограничения оригинальных трехмерных моделей и обеспечивает улучшенное качество пространственных структур в функционально важной зоне связывания лигандов. Следует заметить, что область применимости полученной модели ограничивается белками из семейства киназ. Это связано с тем, что обучающая выборка составлялась только из киназ в комплексе с их ингибиторами. Данное исследование было нацелено на проверку гипотезы о работоспособности и практического применения предложенного подхода.

Разработка аналогичного алгоритма, способного оптимизировать геометрию сайта связывания независимо от семейства рассматриваемого белка является потенциальным перспективным направлением продолжения данной работы.

3.1.4 Генерация псевдолигандов

Базовые модели тестировались как по отдельности, так и совместно (метрика топ-n, рисунок 34). Классификация отдельного класса может быть недостаточно неточной из-за наличия сходных фармакофорных ролей. Совместное использование бинарных моделей обеспечивает необходимую полноту и достоверность предсказаний.

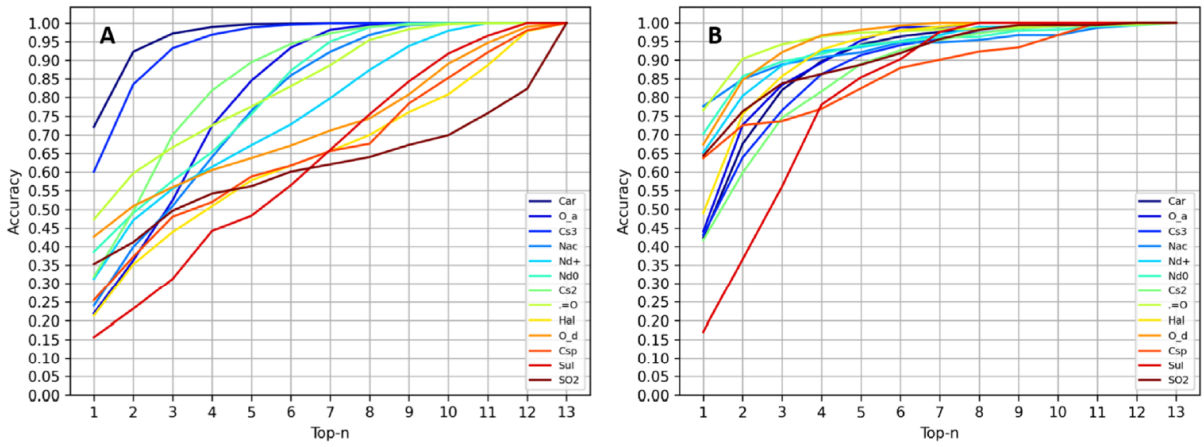


Рисунок 34. Метрики точности классификации для базовых (A) и фармакофорных (B) моделей.

Вероятностные карты для ароматического углерода и кислорода в составе карбонильной группы показаны на рисунке 35. Они демонстрируют соответствие кристаллическим данным.

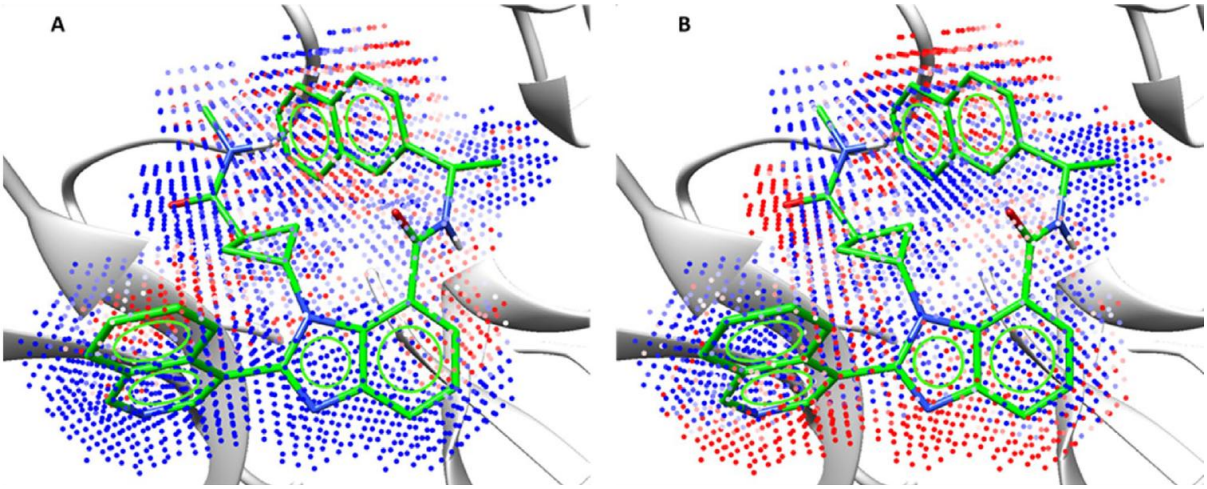


Рисунок 35. Вероятностные карты атомов ароматического углерода (A) и кислорода в составе карбонильной группы (B).

При сравнении с AutoSite и AutoLigand на наборе Astex Diverse Set метод Skittles показал сопоставимую или лучшую точность. Это наиболее заметно для акцепторов водородной связи и гидрофобных групп (таблица 4).

Таблица 4. Сравнительная точность с AutoSite и AutoLigand.

	Доля верно определенных типов атомов		
Модель	**HBD**	**HBA**	**HYD**
Разработанная модель	0.644	0.792	0.925
AutoSite	0.657	0.686	0.908
AutoLigand	0.201	0.447	0.717

3.1.5 Классификация сайтов связывания

Тестирование выполнялось на выборке, отличной по аминокислотной последовательности от обучающих данных. Предложенные модели показали относительно высокие показатели. ROC AUC оказался выше 0,9 для гемовых (0,921) и углеводных (0,915) сайтов. Значения от 0,7 до 0,9 наблюдались для нуклеотидных (0,799), пептидных (0,723) и аллостерических (0,897) сайтов. Значения в диапазоне от 0,6 до 0,7 были обнаружены для нуклеиновых кислот (0,693) и сайтов ковалентного связывания (0,606). Было выполнено сравнение с результатами других методов на их собственных тестовых наборах. Точность обнаружения сайтов связывания гема характеризовалась ROC AUC 0,74, 0,82 и 0,935 для алгоритмов BindSiteS-CNN, DeepDrug3D и BionoiNet соответственно. Точность обнаружения сайтов связывания нуклеотидов дала ROC AUC 0,89, 0,95 и 0,96 для тех же алгоритмов.

Несмотря на существование множества методов классификации сайтов связывания, включая ML-алгоритмы, их сравнение затруднено из-за ограниченного доступа, различий в классах и отсутствия единого тестового набора. Тем не менее, полученные модели были применены к тестовому набору BionoiNet для сравнения точности с BionoiNet [72]. Поскольку отсутствует доступ к предсказаниям и алгоритму BionoiNet для аннотирования данных, сравнивались только финальные метрики на оригинальной тестовой выборке.

На наборе BionoiNet точность обнаружения сайтов связывания гема разработанным методом характеризовалась значением ROC AUC 0,957. Для сайтов связывания нуклеотидов разработанный метод показал ROC AUC 0,966. Для метода BionoiNet были получены значения 0,935 и 0,960 для сайтов связывания гема и нуклеотидов соответственно. Для гемовых сайтов

ROC AUC была сравнима для обоих тестовых наборов и подтверждала сопоставимый уровень точности. Для нуклеотидных сайтов ROC AUC для разработанного метода на тесте BionoiNet оказалась сопоставима с методом BionoiNet. На непохожем на обучение тесте предлагаемый подход показал более низкую точность (0,799). Это может быть связано с большей сложностью предлагаемой выборки в сравнении с валидационной выборкой BionoiNet или большим отличием собранного внутреннего теста от обучающей выборки.

Для изучения совместного использования полученных моделей по предсказанию химической природы субстратов на тестовой выборке отбирался класс с максимальной вероятностью. Большинство классов было корректно предсказано, однако некоторые группы классифицировались неточно (рисунок 36).

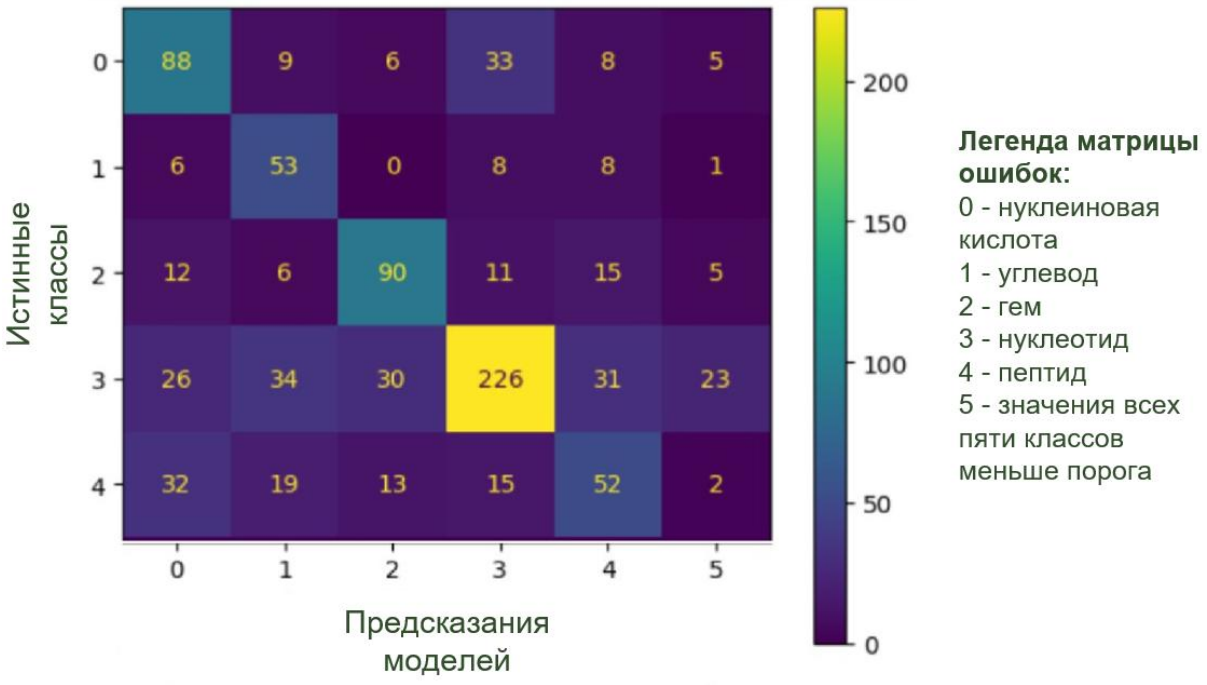


Рисунок 36. Точность классификации сайтов связывания лиганд-белок по природе субстрата.

В 22% случаев сайты связывания нуклеиновых кислот классифицировались как сайты связывания нуклеотидов. Сайты связывания пептидов в 24% случаев были определены как сайты связывания нуклеиновых кислот. Предлагаемый алгоритм генерирует псевдолиганд без учета формы и рамера природного субстрата. Выполняется только разметка предполагаемых типов атомов. Высокая структурная схожесть между нуклеотидами и нуклеиновыми кислотами, а также вариативность пептидов затрудняет точную классификацию.

Алгоритм был применен к 3D-структуре тирозин-киназы Брутона (PDB ID 8U2D). Сайт связывания АТФ был успешно определен с помощью SiteRadar. Была выполнена аннотация моделями Skittles и предсказание свойств сайта. Прогнозы показали высокую вероятность связывания с нуклеотидами/АТФ (0,5). Была показана высокая вероятность ковалентного связывания (0,74) и высокая способность связывания лекарственноподобных молекул (0,86). При этом остальные свойства получили низкие значения (рисунок 37). Эти результаты совпадают с известными фактами о связывании ковалентных ингибиторов в данном сайте.

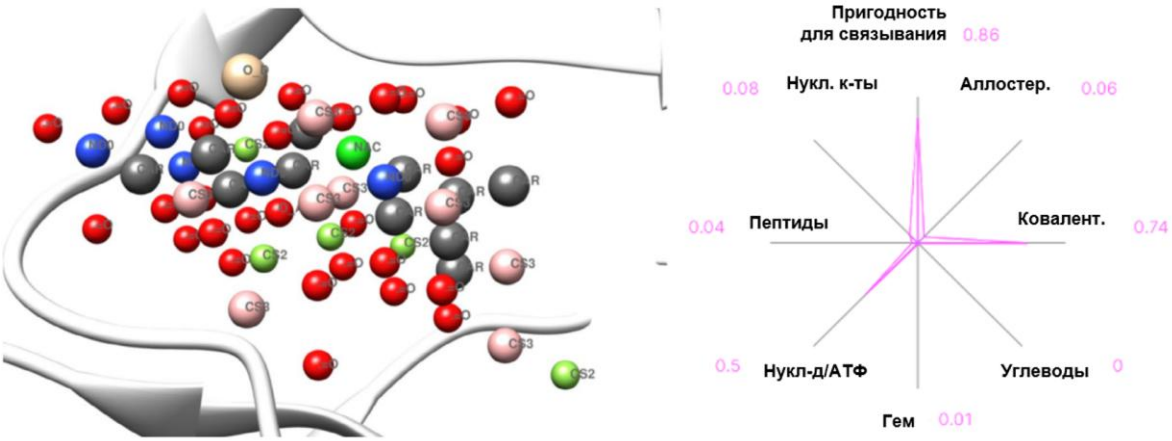


Рисунок 37. Аннотация сайта связывания АТФ тирозинкиназы Брутона.

3.1.6 Оценка аффинности лигандов

Цель применения данного метода — оценка аффинности белок–лиганд. В отличие от существующих методов на основе силовых полей и машинного обучения, в данной работе предложена новая методология, основанная на атомных вероятностях, предсказанных Skittles. Выполнено сравнение с Interaction Graph Net (IGN) — методом на основе GNN, алгоритмом GNINA 1.3 на основе сверточных нейронных сетей и традиционным подходом AutoDock Vina на основе потенциалов межатомных взаимодействий [18].

Полученные результаты показывают, что все протестированные методы машинного обучения статистически значимо превосходят AutoDock Vina по двум показателям качества прогноза — корреляции Пирсона и средней абсолютной ошибке (MAE): для каждого сравнения $p<10^{-5}$ (рисунок 38). По сравнению с IGN предлагаемое решение показало несколько более низкий коэффициент корреляции Пирсона — 0,7408 против 0,7486 ($p < 10^{-5}$). Вместе с тем значения MAE для двух подходов сопоставимы: 1,0 для IGN и 0,948 для предлагаемой модели ($p = 0,529$). При сравнении с GNINA разница по коэффициенту Пирсона (0,7408 против 0,7302, $p = 0,963$), и по MAE была незначительна (0,948 против 0,99, $p = 0,837$). Предложенный алгоритм находился в более строгих условиях из-за различия тестового и обучающего наборов, в то время как для IGN и GNINA такой отбор не проводился, в связи с чем вероятно, результаты для данного метода завышены. В целом, ML-подходы существенно улучшили точность по сравнению с AutoDock Vina, при этом показали сравнимые результаты между собой.

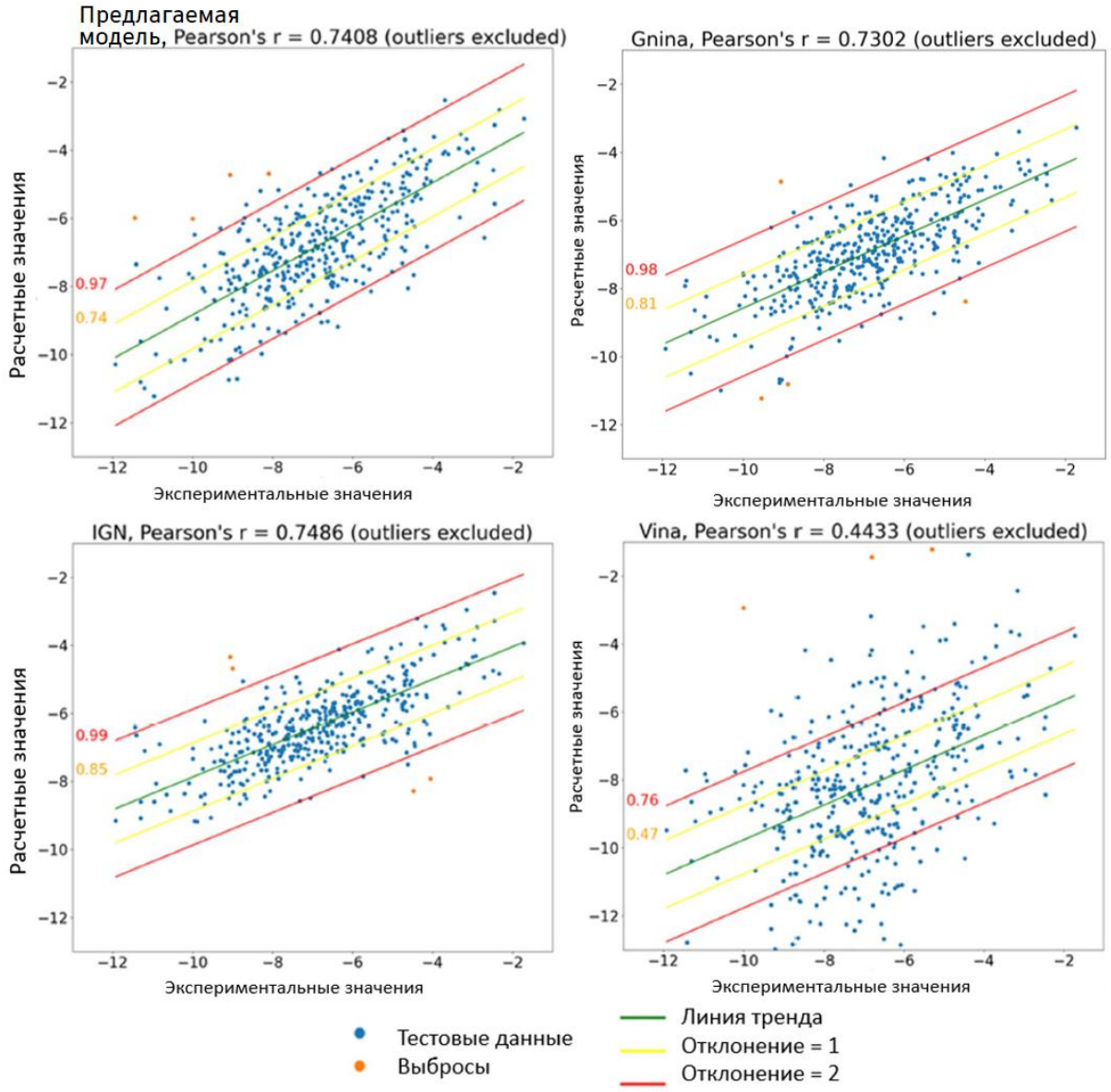


Рисунок 38. Сравнение точности предлагаемой модели, IGN, GNINA и AutoDock Vina.

Для дополнительной проверки значения признаков Skittles были обучены две модели: (1) без учета признаков Skittles, (2) исключительно на признаках Skittles. Исключение части дескрипторов в обоих случаях привело к снижению корреляции Пирсона по сравнению с моделью, использующей полный набор признаков. Различия по MAE оставались статистически незначимыми. Для модели без признаков Skittles коэффициент Пирсона составил 0,7188 ($p < 10^{-5}$), а MAE — 0,9571 ($p =$

0,994). Модель, обученная только на признаках Skittles, достигла коэффициента Пирсона 0,7325 ($p<10^{-5}$) при MAE 0,9575 ($p = 0,725$). При этом модель только на Skittles показала статистически значимое преимущество по Пирсону ($p<10^{-5}$) и отсутствие значительного отличия по MAE ($p = 0,7383$). Этот эксперимент свидетельствует о важности признаков Skittles в прогнозе аффинности и потенциал улучшения точности предсказаний при комплексном использовании с другими признаками.

3.1.7 Генерация молекулярных структур

Для оценки способности QFASG воспроизводить известные кристаллические структуры была сформирована выборка из 17 комплексов белок–лиганд. Данные были предварительно обработаны с использованием MOE v. 2022.02. Лиганды были разделены на фрагменты путем разрыва связей с циклами с помощью RDKit. Были выбраны стартовые фрагменты, формирующие водородные связи с белком. Остальные фрагменты использовались для генерации структуры. Результаты показали, что QFASG с успехом воспроизводит кристаллические позиции в 7 из 17 комплексов (RMSD менее 2 Å). Менее точно были воспроизведены положения для 8 комплексов (RMSD в интервале 2–4,5 Å). Не удалось определить корректную позицию для двух комплексов (RMSD выше 4,5 Å) (рисунок 39).

Была выявлена зависимость качества воспроизведения от количества тяжелых атомов в лиганде. Была достигнута высокая точность для молекул с числом тяжелых атомов до 25. При количестве около 30 тяжелых атомов точность снижается. Было установлено, что точность воспроизведения зависит от числа вращаемых связей. Поиск подходящего сайта связывания для подвижного стартового фрагмента является сложной задачей. Возможно, использование более современных алгоритмов, таких как

диффузионные модели, улучшит качество первичного позиционирования проб в дальнейшем.

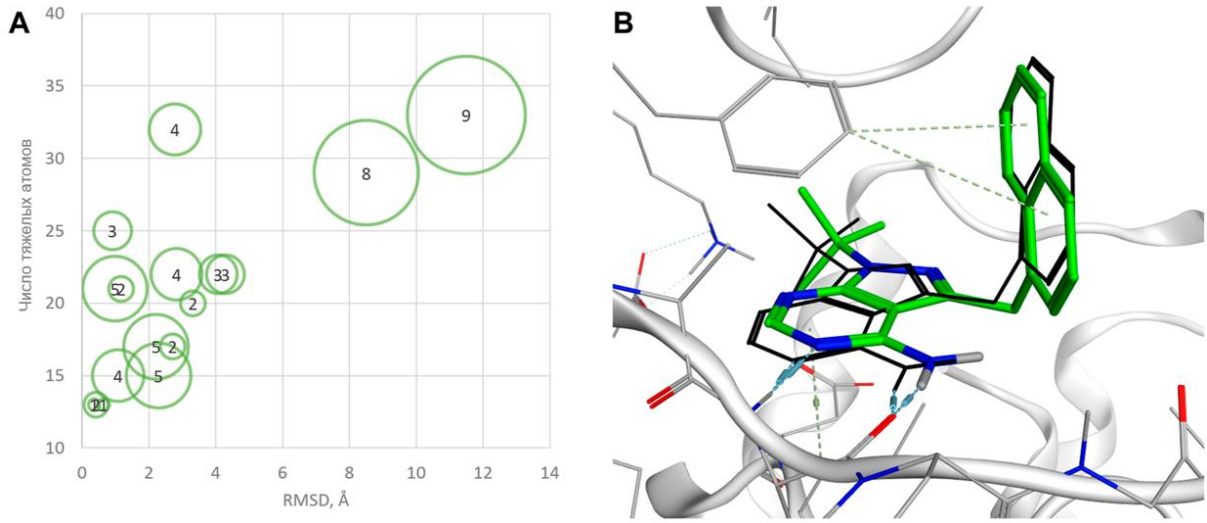


Рисунок 39. Результаты воспроизведения кристаллических позиций для известных комплексов белок–лиганд.

Таким образом, были разработаны вычислительные алгоритмы для решения различных задач на разных этапах разработки лекарств. Разработанные подходы нацелены на решение практических задач, таких как оптимизация геометрии белковой модели, поиск сайтов связывания, генерация псевдолигандов, классификация лиганд-белковых сайтов связывания, оценка аффинности связывания лиганд-белок, виртуальная генерация лекарственноподобных структур на основе фрагментов. Каждый из предложенных алгоритмов был провалидирован *in silico* в сравнении с другими известными методами и продемонстрировал возможность применения для решения практических задач для разработки новых лекарственных молекул. Следующий раздел посвящен применению данных методов для дизайна новых ингибиторов киназ TTK, ATM и CAMKK2 с последующим тестированием предложенных соединений *in vitro*.

3.2. Поиск ингибиторов TTK

3.2.1. Отбор потенциальных ингибиторов TTK

Прогностическая способность предложенного подхода AF Optimizer была оценена в ходе исследования, нацеленного на проверку способности разработанных методов выявлять ингибиторы TTK (MPS1). Сначала была собрана виртуальная библиотека, содержащая 4 793 519 соединений из WuXi, Angene и MolPort с последующей фильтрацией на основе физико-химических свойств и правил медицинской химии. Модель TTK, полученная с помощью AlphaFold 2, была оптимизирована с использованием разработанного алгоритма AF Optimizer. Консервативные молекулы воды были добавлены на основе опубликованных TTK ко-кристаллов и минимизированы с помощью MOE версии 2022.02 с фиксированными атомами кислорода, так как кластеры воды показали незначительные колебания после структурного выравнивания. Позиция референсного лиганда была получена путем структурного выравнивания комплекса TTK (PDB ID 7CHT) на оптимизированную модель. Затем был проведен молекулярный докинг в MOE версии 2022.02, при этом атом азота Gly^{90} был выбран как ключевая точка взаимодействия. Полученные позиции оценивали на основании значения оценочной функции (расчетная энергия связывания, ккал/моль), совпадения фармакофорных точек с гипотезой, сделанной на основе позиции референсного лиганда, и дополнительной экспертной оценки. На основании результатов докинга 39 соединений с высокой чистотой, доступные для заказа и соответствующие всем вышеуказанным критериям, были отобраны для дальнейшей оценки *in vitro*. Следует отметить, что новизна выбранных молекул не оценивалась при отборе соединений на тестирование, поскольку основная цель исследования заключалась в проверке способности предложенного подхода выявлять активные соединения.

3.2.2. *In vitro* анализ ингибирования TTK

Был выполнен экспериментальный анализ ингибирующей активности 39 соединений в отношении протеинкиназы TTK. По результатам выполненного биологического тестирования, 10 соединений показали высокую ингибирующую активность (> 50% ингибирования при 1 мкМ), то есть доля активных соединений среди всех отобранных составила 25,6%. Это значение определяет успешность выполненного скрининга и свидетельсвует о высокой способности к обнаружению ингибиторов TTK.

Среди обнаруженных ингибиторов TTK соединение 1 — ранее известный конкурентный ингибитор TTK с наномолярной активностью. Соединения 2, 3, 4, 5 (палбоциклиб), 7 и 8 (рибоциклиб) также ранее были описаны как конкурентные ингибиторы киназ, в основном CDK4 и CDK6. Однако ингибирование протеинкиназы TTK данными соединениями ранее не было изучено. Полученные данные свидетельствуют о наличии перекрестной активности проанализированных соединений в отношении указанных киназ, что говорит о наличии структурного сходства между их активными центрами.

Соединения 2-5 представляют группу близкородственных по структуре веществ на основе 8-циклопентил-2-[(5-пиперазин-1-ил-2-пиридинил)амино]пиридо[2,3-*d*]пиримидин-7-она. Показательно, что все соединения данной группы продемонстрировали высокую ингибирующую активность в отношении TTK более 80% ингибирования при концентрации 1 мкМ. Это свидетельствует о значительном потенциале дальнейшего изучения указанного хемотипа в отношении протеинкиназы TTK.

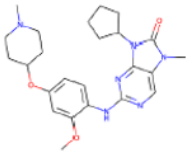
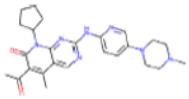
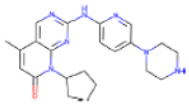
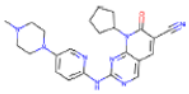
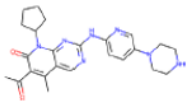
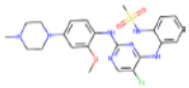
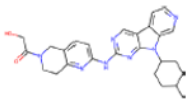
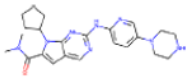
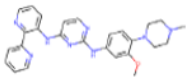
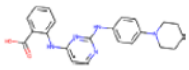
Информация об активности соединений 6, 9 и 10 отсутствует, но близкий структурный аналог соединения 10 показал наномолярную активность против киназ Aurora-A и c-Jun N-terminal. Все три указанные соединения содержат общую подструктуру 2-N-фенилпиримидин-2,4-

диамина. Соединение 6 является наиболее активным в данной серии, демонстрируя 84% ингибирования TTK при концентрации 1 мкМ. Это позволяет характеризовать данное вещество как высокоаффинный ингибитор протеинкиназы TTK. Вещества 9 и 10 продемонстрировали более слабую активность – 54% и 70% ингибирования TTK при концентрации 1 мкМ соответственно. Это характеризует соединения 9 и 10 как умеренные ингибиторы протеинкиназы TTK и свидетельствует о блее оптимальном подборе функциональных земестителей в случае соединения 6.

Большинство из десяти обнаруженных хитов продемонстрировали высокие значения ингибирования TTK в концентрации 1 мкМ. Так, соединения 1-7 характеризовались значениями в диапазоне от 84% до 100% ингибирования. Это наблюдение хорошо соотносится с тем фактом, что большинство данных соединений (1-5, 7) являются ранее охарактеризованными конкурентными ингибиторами различных киназ. Остальные три соединения (8, 9, 10) продемонстрировали более умеренные значения активности в диапазоне от 51% до 70% ингибирования при концентрации 1 мкМ.

Структуры выявленных хитов также были проверены с использованием оригинальной белковой модели, полученной с помощью AF 2. Только семь молекул показали успешные результаты докинга. Соединения 2, 4 и 8 также не было распознано оригинальной моделью белка из AF 2 как активный ингибитор TTK. Полученные результаты суммированы в таблице 5.

Таблица 5. Результаты *in vitro* тестирования соединений, отобранных в ходе виртуального скрининга.

Идентификатор	Структура	Ингибирование (%)	AlphaFold	AF Optimizer
1		94	+	+
2		84	–	+
3		93	+	+
4		92	–	+
5		86	+	+
6		84	+	+
7		100	+	+
8		51	–	+
9		54	+	+
10		70	+	+

Данный эксперимент продемонстрировал, что полученные с помощью AlphaFold 2 белковые модели могут не соответствовать оптимальной геометрии связывания с малыми молекулами. Применение разработанного алгоритма AF Optimizer позволило адаптировать пространственную геометрию сайта связывания, что привело к получению более оптимальных позиций связывания малых молекул и повысило эффективность виртуального скрининга. В результате были обнаружены биологически активные соединения, часть из которых была ранее описана в литературе.

В то же время были отобраны соединения, для которых активность в отношении TTK ранее не была изучена.

Основное отличие между исходной и оптимизированной структурой TTK было обнаружено для остатка Met^{602}, выступающего в роли привратника. Он расположен рядом с шарнирным регионом и известен способностью формировать гидрофобные контакты с ингибиторами TTK [30, 45, 51]. Дополнительно наблюдались небольшие изменения в пространственном расположении остатков Lys^{553}, Glu^{571}, Leu^{575}, Met^{600}. В оригинальной белковой модели остаток Met^{602} направлен внутрь сайта связывания и оказывает наибольшее влияние в создании помех для правильного связывания ингибиторов 2, 4 и 8. В ходе процедуры оптимизации атомы этого остатка сместились в позицию, более подходящую для оптимального связывания лигандов (рисунок 40).

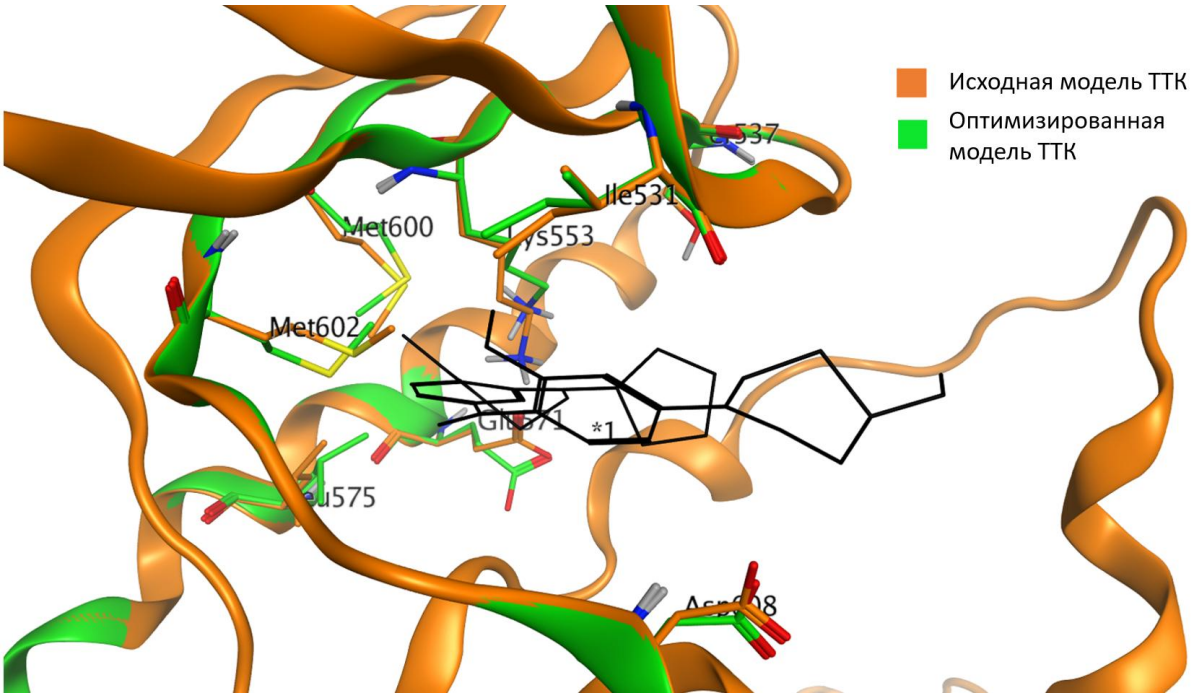


Рисунок 40. Отличия в структуре оригинальной (оранжевый) и оптимизированной (зеленый) модели TTK. Структура референсного лиганда окрашена в черный цвет.

3.3. Дизайн ингибиторов CAMKK2

3.3.1. Виртуальная генерация структур QFASG-1, QFASG-2, QFASG-3

Для оценки возможностей алгоритма QFASG была сгенерирована серия новых структур для ингибирования киназы CAMKK2. В качестве модели была использована кристаллическая структура комплекса CAMKK2 с ингибитором GSK650394 (PDB ID 6BKU). В качестве исходного фрагмента выбран пиразоло[1,5-*a*]пиримидин — фрагмент, широко известный способностью связывания с шарнирным регионом киназ [53]. Алгоритм позиционировал этот фрагмент в двух возможных вариантах, оба формировали водородные связи с ключевыми аминокислотными остатками.

В качестве библиотеки фрагментов были отобраны 43 молекулярных фрагмента, часто встречающиеся в современных ингибиторах киназ. Было проведено четыре итерации роста структуры. Выбор лучших структур осуществлялся на основе расчетной энергии связывания. Для всех сгенерированных структур вычислялись характеристики, такие как оценка межмолекулярных взаимодействий и сходство пространственной позиции с референсным лигандом. В итоге было сгенерировано 465 структур, из которых после оценки по формированию межмолекулярных контактов и пространственной позиции, а также по наличию ключевых водородных связей с аминокислотными остатками Val^{270} и Lys^{194} были отобраны десять структур для экспертного анализа. Из них три структуры (QFASG-1, QFASG-2 и QFASG-3) были выбраны на основе комплексной оценки расчетных характеристик, структурного разнообразия, медицинской химии и предполагаемой позиции связывания для синтеза и последующего *in vitro* тестирования. Был выполнен анализ новизны отобранных соединений. Структурные аналоги для QFASG-1 и QFASG-3 не были обнаружены, тогда

как QFASG-2 частично совпадает по подструктуре с ранее описанными в литературе ингибиторами киназы RETV804M [59].

3.3.2. Синтез соединений QFASG-1, QFASG-2, QFASG-3

Впервые была синтезирована группа предлагаемых соединений QFASG-1, QFASG-2 и QFASG-3 на основе пиразоло[1,5-*a*]пиримидинового каркаса в качестве потенциальных ингибиторов киназы CAMKK2. Все три целевые вещества были получены многостадийными синтезами: QFASG-1 — за 6 стадий, QFASG-2 — за 3 стадии, QFASG-3 — за 7 стадий, что отражает различную сложность встраивания соответствующих заместителей в скэффолд.

Низкие выходы на последних стадиях обусловлены препаративной ВЭЖХ-очисткой, необходимой для достижения высокой чистоты (>95%). Выходы отдельных стадий варьировались от 38% до 99,72%, с наиболее эффективными реакциями — боронирование (стадия 3, QFASG-1, 99,72%) и нуклеофильное ацилирование (стадия 2, QFASG-1, 90%).

Чистота всех синтезированных веществ подтверждена ^{1}H-ЯМР-спектроскопией. Структурное соответствие синтезированных веществ QFASG-1–3 было подтверждено наличием характеристических сигналов протонов центрального пиразоло[1,5-*a*]пиримидинового ядра в слабых полях спектров ^{1}H ЯМР, в частности, дублета протона в положении C(7) при δ = 9,00-9,32 (*J* = 7,1-7,6 Гц) и характеристического синглета при δ = 8,50-9,03. Молекула QFASG-1 надежно идентифицируется по сигналам алифатической области, где присутствуют два синглета метиленовых групп азетидинового цикла δ = 4,32 и 4,30), квинтет CH-протона (δ = 4,68), а также синглет ацетильной группы при δ = 2,56. В свою очередь, симметричное *пара*-замещенное бензольное кольцо в структурах QFASG-2 и QFASG-3 четко проявляется в виде двух дублетов с константами *J* = 8,4 Гц и *J* = 8,8 Гц соответственно. Наличие уширенного синглета амидных протонов (δ =

8,40 для QFASG-2) и метильного синглета мочевины (δ = 2,65 для QFASG-3) строго доказывает успешное формирование заявляемых функциональных групп.

QFASG-1.

^{1}H ЯМР: 400 МГц, CD_3CN

δ = 9,04 (д, *J* = 7,6 Гц, 1H), 8,69 (с, 1H), 8,67 - 8,65 (м, 2H), 8,62 (д, *J* = 1,3 Гц, 1H), 8,36 (д, *J* = 8,0 Гц, 1H), 8,34 (с, 1H), 7,76 (д, *J* = 7,8 Гц, 1H), 7,68 - 7,63 (м, 1H), 7,39 (д, *J* = 7,5 Гц, 1H), 4,68 (квинтет, *J* = 8,5 Гц, 1H), 4,32 (с, 2H), 4,30 (с, 2H), 2,56 (с, 3H)

QFASG-2.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,32 (д, *J* = 7,1 Гц, 1H), 9,03 (с, 1H), 8,45 (д, *J* = 8,4 Гц, 2H), 8,40 (уш. с, 1H), 8,04 (д, *J* = 8,4 Гц, 3H), 7,63 (д, *J* = 7,3 Гц, 1H), 2,61 (с, 4H)

QFASG-3.

^{1}H ЯМР: 400 МГц, ДМСО

δ = 9,00 (д, *J* = 7,5 Гц, 1H), 8,50 (с, 1H), 8,21 (д, *J* = 8,8 Гц, 2H), 8,03 - 7,87 (м, 1H), 7,62 (д, *J* = 7,5 Гц, 1H), 7,58 (д, *J* = 8,8 Гц, 2H), 2,65 (с, 3H)

3.3.3. *In vitro* анализ ингибирования САМКК2

Два из трех соединений проявили активность в низком микромолярном диапазоне: 3 мкМ с наклоном Хилла −0,85 и 6 мкМ с наклоном −1,05 для QFASG-2 и QFASG-3 соответственно, при этом концентрация полуингибирования QFASG-1 превысила 10 мкМ (рисунок 41). Следует отметить, что тестирование проводилось при концентрации АТФ равной 10 мкМ, что может снизить эффективность при сравнении с выполнением тестов при концентрации АТФ равной K_M.

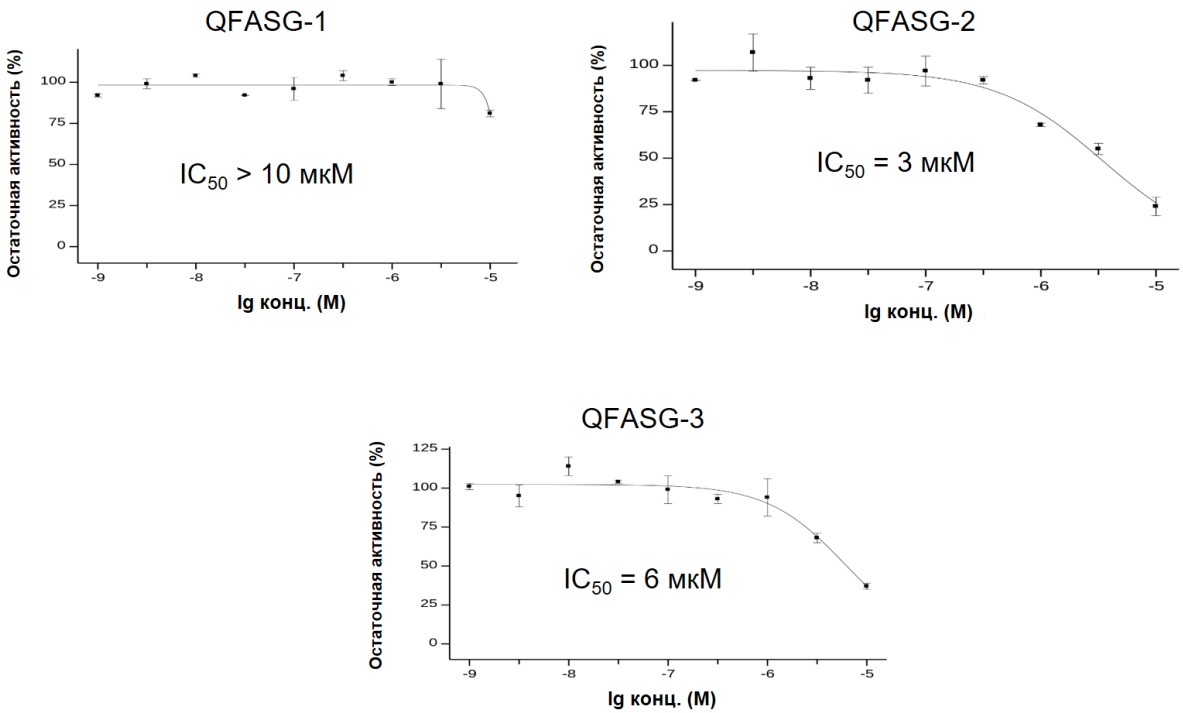


Рисунок 41. Результаты *in vitro* тестирования соединений на основе сгенерированных структур против CAMKK2.

Результаты биохимического тестирования показали способность соединений QFASG-2 и QFASG-3 ингибировать целевую киназу CAMKK2. Оба соединения показали концентрацию полуингибирования CAMKK2 в низком микромолярном диапазоне концентраций, что характеризует их как соединения-хиты для последующей оптимизации. Кривые титрования демонстрируют дозозависимый характер ингибирования. Согласно характеру кривых титрования, наблюдается классический механизм ингибирования с отсутствием выраженной кооперативности. Плавный сигмоидальный характер кривых титрования также говорит об отсутствии артефактов агрегации.

На рисунке 42 показаны предположительные позиции связывания данных соединений, полученные в ходе виртуальной генерации. На основе имеющихся данных предполагается, что соединения QFASG-2 и QFASG-3 являются конкурентными ингибиторами киназы CAMKK2, связываясь в ее

активном центре и формируя сеть водородных связей с аминокислотными остатками. Дополнительным доводом в пользу связывания в активном центре является наличие в структуре полученных соединений пиразоло[1,5-*a*]пиримидинового фрагмента, известного высоким сродством к шарнирному региону активного центра. Более детальное изучение механизма действия соединений данной серии является предметом изучения в рамках дальнейших исследований.

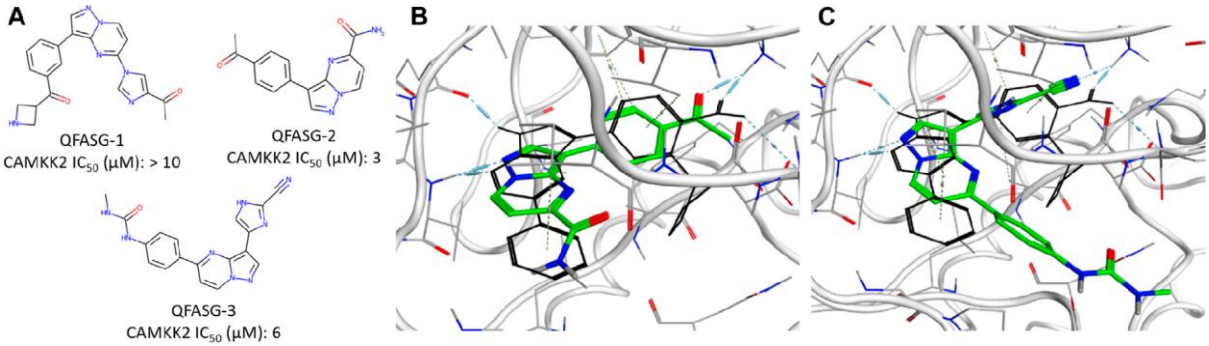


Рисунок 42. Структуры и позиции связывания протестированных соединений в активном центре CAMKK2.

Низкая активность соединения QFASG-1 может объясняться переоценкой алгоритмом расчетной энергии связывания в связи с ограниченной точностью оценочных функций. Возможно также, что в ходе транспортировки или хранения произошла биохимическая трансформация, например, окисление азота или гидролиз карбонильной группы с образованием геминального диола.

Было выполнено сравнение полученных структур с известными ингибиторами CAMKK2 STO-609 и GSK650394. Предложенные соединения обладают значительными структурными отличиями и представляют собой альтернативный, ранее неизученный в отношении CAMKK2 хемотип. Также следует отметить простоту синтеза соединения QFASG-2, которое может быть получено всего в три стадии органического синтеза.

Перспективы дальнейшего развития предложенных соединений заключаются в модификации функциональных групп для повышения сродства к белковой мишени. В частности, одним из возможных изменений предложенных соединений является введение в их структуру фрагмента 2-циклопентилбензойной кислоты, способствующего повышению аффинности малых молекул к активному центру CAMKK2.

3.3.4. Виртуальная оценка селективности и ADMET

Для предварительной оценки селективности действия соединений QFASG-2 и QFASG-3 был выполнен их виртуальный анализ на предмет активности в отношении 50 киназ [19]. Результаты выполненного анализа позволяют предполагать высокую селективность полученных соединений. Так, для структуры QFASG-2 не обнаружено других киназ, в отношении которых расчетное значение IC_{50} было бы менее 1 мкМ. В то же время для QFASG-3 была спрогнозирована потенциальная ингибирующая активность в отношении киназы KIT.

Для первичной оценки ADMET параметров был выполнен виртуальный анализ структур QFASG-2 и QFASG-3 с использованием веб-сервиса SwissADME [16]. Были оценены такие параметры как растворимость, всасывание из кишечника, оценка биодоступности, фильтры Вебера, PAINS и Бренка. Расчетные значения представлены в таблице 6.

Таблица 6. Расчетные параметры ADMET для соединений QFASG-2 и QFASG-3.

	QFASG-2	**QFASG-3**
Растворимость	1,01 мг/мл; $3,6*10^{-3}$ М	0,55 мг/мл; $1,5*10^{-3}$ М
Всасывание из кишечника	Высокое	Высокое
Оценка биодоступности	0,55	0,55
Фильтр Вебера	Удовлетворяет	Удовлетворяет
PAINS фильтры	Нежелательные фрагменты отсутствуют	Нежелательные фрагменты отсутствуют
Фильтры Бренка	Нежелательные фрагменты отсутствуют	Нежелательные фрагменты отсутствуют

Оба соединения QFASG-2 и QFASG-3 продемонстрировали схожий профиль расчетных ADMET параметров. Оба соединения относятся к классу растворимых, характеризуются высокими потенциалом всасывания из кишечника и оценкой биодоступности (0,55). Удовлетворение фильтру Вебера подтверждает соответствие критериям молекулярной массы ≤ 500 Да и площади полярной поверхности ≤140 Å², которые охватывают 85% пероральных препаратов FDA. Отсутствие нежелательных фрагментов в

PAINS-фильтрах (Pan-Assay INterference Compounds) и фильтрах Бренка (структурные фрагменты, ассоциированные с токсичностью) свидетельствует о высоком лекарствоподобии веществ и низком риске ложноположительных результатов в биохимических испытаниях. Эти наблюдения определяют благоприятный профиль расчетных ADMET параметров для соединений QFASG-2 и QFASG-3.

В данной диссертации были выполнены пилотные испытания алгоритма с использованием небольших библиотек фрагментов и базовых параметров. Были обнаружены соединения, ингибирующие целевые киназы в низком микромолярном диапазоне концентраций. При этом основное внимание уделялось отбору химически простых, легко синтезируемых структур. Предполагается, что использование более расширенных библиотек фрагментов, более продвинутых настроек и отбор более сложных структур позволит применять QFASG с еще более высокой эффективностью.

3.4. Дизайн ингибиторов ATM

3.4.1. Виртуальная генерация структур QFASG-4, QFASG-5, QFASG-6

Для дальнейшей проверки возможностей модели QFASG была проведена генерация новых структур для ингибирования киназы ATM. В качестве модели использовалась кристаллическая структура ATM в комплексе с ингибитором M4076 (PDB ID 7NI4). В качестве стартового фрагмента был выбран 3-аминохинолиновый каркас. Хинолиновый скэффолд широко известен в медицинской химии и конструировании ингибиторов киназ [42, 58, 90, 94]. В качестве библиотеки фрагментов было

отобрано 94 структурно разнообразных фрагмента. За четыре итерации была сгенерирована 191 молекулярная структура.

Из набора сгенерированных структур три были выбраны для синтеза и *in vitro* оценки, используя аналогичный подход, как и для CAMKK2. Был выполнен анализ новизны отобранных соединений. Структурные аналоги для QFASG-4 и QFASG-5 не были выявлены. При этом структура соединения QFASG-6 частично совпадала с ранее описанными ингибиторами киназы MALT1 [64].

3.4.2. Синтез соединений QFASG-4, QFASG-5, QFASG-6

В рамках данной серии впервые синтезированы новые вещества на основе 3-аминохинолинового каркаса — QFASG-4, QFASG-5 и QFASG-6. Синтетическая схема для QFASG-4 включала пять стадий, тогда как получение QFASG-6 осуществлялось в две стадии на основе 6-бромхинолин-3-амина. Для QFASG-4 ключевые этапы C–N-сочетания и N-метилирования протекали с хорошими выходами 81% и 93% соответственно. Конечный продукт QFASG-4 был получен в аналитически чистом виде: чистота по ЖХ-МС составила 98,58%, а по ВЭЖХ — 96,80%.

Синтез QFASG-5 включал четыре последовательные реакции: стадии карбаматной защиты, кросс-сочетания, C–N-сочетания и снятия защитных групп. Для QFASG-5 реакция получения целевого мочевинного производного завершалась с низким выходом на финальной стадии. Это отражает характерную сложность введения дифторфенильного изоцианатного фрагмента и необходимость препаративной ВЭЖХ-очистки. Вместе с тем целевое соединение было выделено в аналитически чистом виде.

QFASG-6 также характеризовался высокой чистотой по ЖХ-МС (98,88%) и ВЭЖХ (98,95%).

Однозначное соответствие экспериментальных спектров ^{1}H ЯМР целевым структурам QFASG-4–6 доказывается характерным набором сигналов протонов центрального хинолинового ядра. Для всех соединений фиксируются малоинтенсивные сигналы в позициях C(2) и C(4) пиридинового фрагмента хинолина (дублеты или синглеты при δ = 8,51-9,01 с константами $J \leq 2{,}4$ Гц). Структура QFASG-4 надежно верифицируется по синглету карбоксильного протона (δ = 13,00), синглетам метиленовой (δ = 4,07, N-CH_2-) и метильной (δ = 3,35, N-CH_3) групп, а также по характеристическим мультиплетам циклопропильного кольца (δ = 3,21-3,17, 0,94 и 0,75).

Образование несимметричных диарилмочевинных линкеров в соединениях QFASG-5 и QFASG-6 подтверждается парой выраженных синглетов протонов аминогрупп в диапазоне δ = 8,96-9,77. Метильный синглет при δ = 3,91 N-CH_3 подтверждает успешное включение пиразольного цикла в QFASG-5. Структура QFASG-6 валидируется за счет появления дополнительных протонов 3-пиридильного фрагмента при δ = 8,67, 8,24 и 7,35 на фоне исчезновения сигналов фторированного бензольного кольца.

QFASG-4

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 13,001 (шир. с, 1H), 8,55 (д, J = 1.6 Гц, 1H), 7,88 (д, J = 8,2 Гц, 1H), 7,74 (с, 1H), 7,16 - 7,00 (м, 3H), 6,02 (с, 1H), 4,07 (с, 2H), 3,35 (с, 3H), 3,21 - 3,17 (м, 1H), 2,16 (с, 3H), 0,94 (д, J = 3,6 Гц, 2H), 0,75 (д, J = 5,2 Гц, 2H).

ЖХ-МС: 98,58 % (M^{+1} = 352,2) при 3,552 R_t.

ВЭЖХ: 96.80% чистоты при 6,673 R_t.

QFASG-5

МС (ESI): расчетное 380,1 [$(M + H)^+$]; измеренное 380,1 [$(M + H)^+$]

^{1}H ЯМР: 400 МГц, ДМСО-d_6

δ = 9,54 (с, 1 Н), 8,96 (с, 1 Н), 8,72 (д, *J* = 2,4 Гц, 1Н), 8,51 (д, *J* = 2,0 Гц, 1Н), 8,30 (с, 1 Н), 8,09 (д, *J* = 2,0 Гц, 1Н), 8,04 (с, 1 Н), 7,99 (с, 1 Н), 7,93-7,90 (м, 1 Н), 7,90-7,86 (м, 1 Н), 7,20-7,18 (м, 1 Н), 7,10-7,09 (м, 1 Н) , 3,91 (с, 3 Н).

QFASG-6

^{1}Н ЯМР: 400 МГц, ДМСО-d_6

δ = 9,77 (с, 1Н), 9,43 (с, 1Н), 9,01 (с, 1Н), 8,67 (д, *J* = 2 Гц, 2Н), 8,62 (с, 1Н), 8,24 (д, *J* = 4,4 Гц, 1Н), 8,09 (д, *J* = 8,4 Гц, 1Н), 7,99 (д, *J* = 8,0 Гц, 1Н), 7,88 (д, *J* = 8,4 Гц, 1Н), 7,37 – 7,34 (м, 1Н).

ЖХ-МС: 98,88 % (M^{+1} = 352,2) при 3,524 R_t.

ВЭЖХ: 98.95% чистоты при 6,404 R_t.

3.4.3. *In vitro* анализ ингибирования ATM, ATR и DNA-PK

Из трех тестируемых соединений QFASG-6 показало IC_{50} 4 мкМ. При оценке селективности по отношению к киназам DNA-PK и ATR, QFASG-6 проявила умеренную активность против ATM и DNA-PK (IC_{50} 8 мкМ), при этом ингибирования ATR не наблюдалось (рисунок 43). Как и в случае с CAMKK2, тестирование проводилось при концентрации АТФ равной 10 мкМ.

QFASG-6 показало концентрацию полуингибирования ATM и DNA-PK в низком микромолярном диапазоне концентраций, что характеризует его как соединение-хит для последующей оптимизации. Кривые титрования демонстрируют дозозависимый характер ингибирования. Согласно характеру кривых титрования, наблюдается классический механизм ингибирования с отсутствием выраженной кооперативности. Плавный сигмоидальный характер кривых титрования свидетельствует об отсутствии артефактов агрегации.

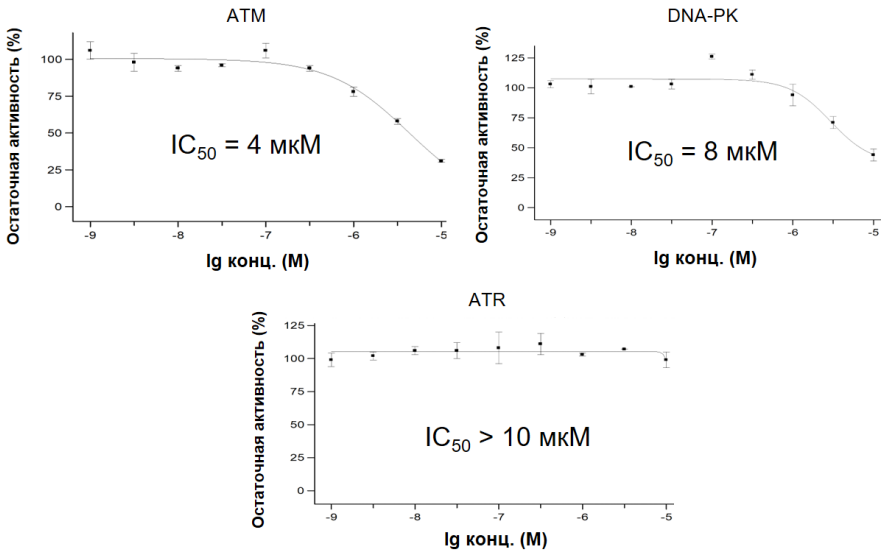


Рисунок 43. Результаты *in vitro* тестирования соединений на основе сгенерированных структур против киназы ATM.

Соединение QFASG-6 предположительно является конкурентным ингибитором киназ ATM и DNA-PK. В пользу этого факта свидетельствуют результаты молекулярного моделирования и наличие в структуре соединения 3-аминохинолинового фрагмента. Данный фрагмент распространен в дизайне конкурентных ингибиторов киназ. Более детальное изучение механизма действия соединений данной серии является предметом изучения в рамках дальнейших исследований.

Структуры QFASG-5 и QFASG-6 обладают значительным сходством, различаясь в основном одной функциональной группой: у QFASG-5 это N-метил имидазольная группа, тогда как у QFASG-6 — нитрильная группа (рисунок 44). Обе функциональные группы потенциально могут выступать акцепторами водородных связей и, предположительно, сконструированы для взаимодействия с положительно заряженной боковой цепью остатка Lys^{2717}. Однако размер N-метил имидазольной группы в QFASG-5, возможно, препятствует оптимальному контакту из-за наличия стерических препятствий.

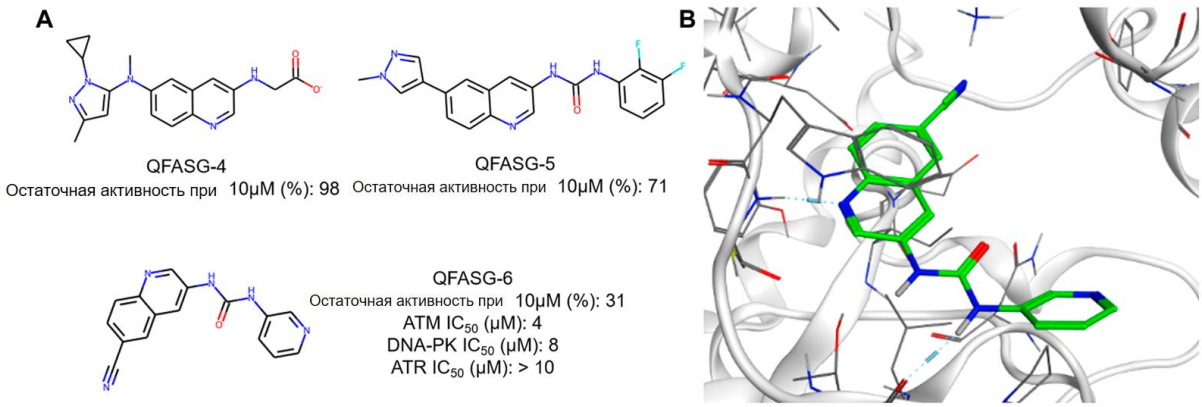


Рисунок 44. Структуры протестированных соединений и позиция связывания QFASG-6 в активном центре киназы ATM.

Был выполнен сравнительный анализ соединения **QFASG-6** с известными ингибиторами ATM WSD-0628, XRD-0394, M-4076, M-3541, AZD-1390, AZD-0156. Все указанные известные ингибиторы ATM представляют собой группу близкородственных соединений на основе 1,3-дигидро-2H-имидазо[4,5-*c*]хинолин-2-она, что свидетельствует о малом разнообразии изученного химического пространства и необходимости поиска новых альтернативных хемотипов. Соединение QFASG-6 принципиально отличается по структуре от известных ингибиторов ATM, что расширяет возможности дальнейшего молекулярного дизайна. Еще одним преимуществом соединения QFASG-6 является его высокая синтетическая доступность: соединение может быть получено в две стадии органического синтеза.

Другой важной отличительной чертой соединения QFASG-6 является отсутствие большого числа крупных заместителей. В совокупности с простотой синтеза это позволяет говорить о перспективах последующего использования данного соединения для разработки новых более высокоаффинных ингибиторов ATM на его основе. Введение дополнительных подвижных линкерных фрагментов с терминальным ионизируемым атомом азота предположительно должно повысить

аффинность рассматриваемой структуры к активному центру ATM за счет формирования дополнительных выгодных межмолекулярных контактов с более удаленными участками сайта связывания.

3.4.4. Виртуальная оценка ADMET

Был выполнен компьютерный анализ ADMET характеристик структуры QFASG-6 с использованием веб-сервиса SwissADME. Согласно полученным результатам, соединение QFASG-6 было отнесено к классу растворимых (0,47 мг/мл; $1,6*10^{-3}$ М), характеризуются высокими потенциалом всасывания из кишечника и оценкой биодоступности (0,55). QFASG-6 соответствует фильтрам Вебера, которые охватывают 85% пероральных препаратов FDA. Не выявлено алертов в PAINS-фильтрах и фильтрах Бренка, что говорит о высоком лекарствоподобии и низком риске ложноположительных результатов при *in vitro* тестировании. Выполненный анализ дополнительно характеризует соединение QFASG-6 как перспективный прототип для дальнейшей разработки.

Таким образом, в данной работе на примере поиска ингибиторов трех киназ — ATM, CAMKK2 и TTK — было продемонстрировано, как разработка и применение вычислительных методов позволяют повышать эффективность поиска новых биологически активных соединений. Результаты данной работы обладают значением как с теоретической точки зрения, так и для практического применения в разработке лекарств. В ходе проведенных исследований были получены новые алгоритмы, которые могут быть применены к различным биомишеням, а также были получены соединения, обладающие биологической активностью против терапевтически важных киназных белков.

ЗАКЛЮЧЕНИЕ

Выполненное исследование представляет собой комплексную научную работу в области компьютерного дизайна лекарственных молекул, направленную на разработку новых ингибиторов протеинкиназ с применением современных вычислительных методов и интеграцией последних достижений структурной биологии, хемоинформатики и машинного обучения. В работе предложены новые вычислительные решения, выполнена их *in silico* валидация и осуществлено их применение к решению практических задач. Применение разработанных методов способствовало обнаружению новых ингибиторов актуальных терапевтических мишеней — киназ ATM, CAMKK2 и TTK.

Центральным элементом исследования стал дизайн новых биологически активных соединений путем разработки и применения вычислительной методологии, включающей несколько алгоритмов. Во-первых, оптимизация трехмерных структур белков, полученных с помощью AlphaFold 2, путем применения AF Optimizer, который уточняет пространственное расположение боковых цепей аминокислот и снижает число стерических затруднений. Во-вторых, обнаружение лиганд-белковых сайтов связывания с помощью алгоритма SiteRadar, который продемонстрировал значительно более высокую точность по сравнению с существующими методами. В-третьих, применение GNN для генерации псевдолигандов. Реализованные в алгоритме Skittles модели обеспечивают детальное картирование сайтов связывания с точной аннотацией различных типов атомов. Это было показано на задачах классификации сайтов связывания и оценки аффинности лиганд-белок. И, наконец, алгоритм генерации новых молекулярных структур на основе фрагментов. Алгоритм QFASG реализует итеративное конструирование молекулярных структур под заданный сайт связывания. При этом выполняется оптимизация

позиции и отбор полученных соединений на основе современных фильтров медицинской химии.

Эффективность применения предлагаемых решений продемонстрирована на протеинкиназах ATM, CAMKK2 и TTK. Эти киназы известны по своей ключевой роли в регуляции клеточных процессов и развитии различных патологий.

Одним из важных достижений в рамках выполненного исследования стала автоматическая оптимизация с помощью AF Optimizer трехмерных белковых моделей из AlphaFold 2. Данный подход позволил решить существенную проблему оригинальных белковых моделей. Были устранены неточности в пространственном расположении боковых цепей аминокислот, которые критичны для правильного формирования сайта связывания. Устранение этих ошибок снижает искусственные стерические затруднения и улучшает геометрию взаимодействия с лигандом. Это повышает качество результатов иолекулярного докинга и виртуального скрининга. Данный подход показал широкий диапазон применимости в тестах на различных киназах, что делает его универсальным для данного класса фармакологических мишеней. Было выполнено проспективное исследование на модели белка TTK. Полученные результаты подтверждают эффективность AF Optimizer и его потенциал для интеграции в системы структурно-обоснованного дизайна лекарственных соединений (SBDD).

Важные результаты были получены путем применения моделей GNN для идентификации сайтов связывания и генерации псевдолигандов. Алгоритм SiteRadar показал возможность обнаруживать лиганд-белковые сайты связывания на поверхности белка. Это необходимо для последующей разработки малых молекул, способных модулировать активность заданного белка посредством связывания с активным центром или аллостерическими

сайтами связывания. Был разработан алгоритм для аннотации обнаруженных сайтов связывания с точки зрения химических типов атомов. Для этого используется ансамбль бинарных моделей, одна для каждого типа атома. Это позволяет повысить управляемость и точность картирования. Такой подход открывает новые перспективы для характеризации сайтов связывания на основе химических и пространственных признаков и создает предпосылки для определения их свойств.

Разработанная на основе Skittles система классификации сайтов связывания обладает широким функциональным спектром. Она способна прогнозировать несколько важных характеристик сайта связывания. К ним относятся способность сайта связывания взаимодействовать с малыми молекулами, потенциал для аллостерической регуляции, возможность ковалентного связывания лигандов. Кроме того, прогнозируется химический профиль предполагаемого субстрата. Рассматриваемые классы включают нуклеотиды/АТФ, углеводы, гем, пептиды и нуклеиновые кислоты. Были выполнены исследования на различных тестовых выборках, включая белки, непохожие на обучающие данные. Высокие значения ROC AUC, достигнутые системой, говорят о высокой точности и обобщающей способности модели. Это делает метод особенно полезным в качестве инструмента для анализа биологических мишеней, характеризации сайтов связывания и эффективного подбора кандидатов при разработке новых лекарств.

Также в рамках данного исследования разработанный алгоритм Skittles был применен для предсказания аффинности связывания белок–лиганд. Особое внимание при этом уделяется локальным вероятностям атомных позиций в структуре лиганда. Предлагаемый метод предусматривает анализ вероятности нахождения каждого атома лиганда в конкретном макромолекулярном окружении. Такой подход демонстрирует результаты,

сравнимые с современными эталонными методами. Важной особенностью алгоритма является его способность анализировать межмолекулярные взаимодействия на основе статистики, собранной из реальных кристаллических данных. Это делает данную модель способной учитывать природные закономерности молекулярной биологии и позволяет избежать использования искусственных значений, таких как силовые поля и потенциалы межмолекулярного взаимодействия.

Был выполнен сравнительный анализ алгоритма Skittles с известными подходами. В качестве методов сравнения рассматривались программа классического докинга AutoDock Vina и современные методы с использованием машинного обучения IGN и GNINA. Результаты *in silico* валидации продемонстрировали высокую корреляцию предсказанных значений аффинности с экспериментальными данными. Это указывает на потенциал интеграции метода в высокопроизводительные системы виртуального скрининга на ранних стадиях поиска лекарственных соединений. Такой подход может позволить эффективно исключать неактивные соединения с меньшими затратами времени и ресурсов.

Другим важным элементом стал алгоритм QFASG. Данный метод предназначен для автоматического конструирования новых молекул на базе фрагментов с применением квантово-химических расчетов. QFASG реализует настраиваемый по химическим критериям процесс итеративного конструирования структур, интегрируя современные фильтры медицинской химии. Они обеспечивают генерацию разумных структур с точки зрения лекарствоподобия. Данный подход доказал свою эффективность в воспроизведении позиций известных молекулярных комплексов. Также была показана его эффективность в генерации структур новых соединений с последующим подтверждением их биологической активности.

In silico валидация продемонстрировала высокую воспроизводимость экспериментальных лиганд-белковых комплексов. Воспроизведение позиций наблюдалось для широкого спектра соединений, особенно для молекул с ограниченным количеством вращаемых связей и числом тяжелых атомов не более 25. Это определяет ограничения данного метода с точки зрения структуры и подвижности исследуемых молекул. Дальнейшее развитие предлагаемого метода является перспективным направлением для решения этих проблем.

Была выполнена генерация структур новых ингибиторов киназ ATM и CAMKK2. Практическая апробация на реальных биологических мишенях показала высокую эффективность QFASG. Для CAMKK2 было сгенерировано и отобрано три перспективных соединения. Для двух из них впоследствии была подтверждена *in vitro* активность в низком микромолярном диапазоне IC_{50}. В случае ATM из трех синтезированных молекул, одна показала умеренную активность на уровне 4 мкМ. При этом была оценена ее селективность при сравнении с другими родственными киназами DNA-PK и ATR, что важно для минимизации потенциальных побочных эффектов.

Результаты синтеза и биохимических тестов убедительно иллюстрируют работоспособность и применимость предложенной платформы, при этом выявлены ключевые направления для ее совершенствования. В результате исследования были найдены новые ингибиторы киназ TTK, ATM и CAMKK2, которые могут быть использованы в качестве прототипов для дальнейшей фармацевтической разработки.

В настоящей работе показано, как разработка и внедрение новых вычислительных подходов может способствовать ускорению и повышению

эффективности разработки лекарств. В ходе исследования обнаружены новые ингибиторы киназ ATM, CAMKK2 и TTK, что свидетельствует о практической значимости разработанных методов. Результаты биологического тестирования расширили понимание о влиянии различных химических модификаций на ингибирование изучаемых мишеней. Обнаруженные новые хемотипы могут быть далее использованы для последующей модификации и доработки с целью дальнейшей фармацевтической разработки.

Успешное применение разработанных алгоритмов для поиска новых ингибиторов киназ ATM, CAMKK2 и TTK способствовало обнаружению новых активных соединений. Это свидетельствует о потенциале дальнейшего применения данного алгоритма для дизайна новых модуляторов других фармакологических мишеней. Таким образом, данная работа раскрывает примеры дизайна новых биологически активных соединений с использованием компьютерных алгоритмов и демонстрирует перспективы применения вычислительных технологий для дизайна новых лекарственных молекул.

ВЫВОДЫ

1. В ходе исследования разработаны новые вычислительные алгоритмы для решения задач ранней стадии разработки лекарств, включая оптимизацию пространственной структуры белковых моделей, поиск сайтов связывания, генерацию псевдолигандов, классификацию сайтов связывания, прогнозирование аффинности лиганд-белок и автоматическое конструирование структур новых биологически активных соединений на основе фрагментов.
2. Выполнена *in silico* валидация разработанных алгоритмов, в том числе в сравнении с существующими решениями, при этом SiteRadar показал улучшение метрики топ-N+2 с 47% до 74%, AF Optimizer повысил успех обнаружения соединений-хитов с 28% до 40% в ретроспективном исследовании. Алгоритм Skittles и его применение для классификации сайтов связывания и оценки аффинности продемонстрировали уровень точности, сравнимый с современными решениями.
3. С использованием разработанных алгоритмов выполнен поиск ингибиторов TTK и дизайн ингибиторов ATM и CAMKK2
4. Предложен и осуществлен органический синтез серии производных пиразоло[1,5-*a*]пиримидина; получена серия соединений на основе хинолинового каркаса.
5. Выполнено *in vitro* исследование ингибирующей активности полученных веществ в отношении белков-мишеней. Доля обнаруженных ингибиторов TTK с IC_{50} < 1 мкМ составила 25,6%. Десять соединений продемонстрировали от 51% до 100% ингибирования TTK при концентрации 1 мкМ. Доля обнаруженных ингибиторов ATM и CAMKK2 с IC_{50} < 10 мкМ составила 50%. Соединения QFASG-2 и QFASG-3 ингибировали CAMKK2 с IC_{50}

равным 3 мкМ и 6 мкМ соответственно. Соединение QFASG-6 ингибировало ATM и DNA-PK с IC_{50} равным 4 мкМ и 8 мкМ соответственно.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

В результате исследования обнаружены биологически активные соединения, демонстрирующие способность ингибировать киназы ATM, CAMKK2, TTK. В частности, путем виртуального скрининга выявлены ингибиторы TTK с концентрацией полуингибирования менее 1 мкМ. В ходе рационального дизайна на основе пиразоло[1,5-*a*]пиримидинового каркаса разработаны соединения QFASG-2 и QFASG-3, ингибирующие киназу CAMKK2 в низком микромолярном диапазоне концентрации. Также, с использованием компьютерного моделирования на основе хинолинового каркаса разработано соединение QFASG-6, ингибирующее ATM и DNA-PK в низком микромолярном диапазоне концентрации.

Полученные в ходе выполненного исследования данные расширяют возможности для изучения химического пространства ингибиторов протеинкиназ ATM, CAMKK2 и TTK и открывают перспективу для дизайна высокоаффинных ингибиторов указанных мишеней на основе обнаруженных хемотипов.

СПИСОК СОКРАЩЕНИЙ

- АТФ — аденозинтрифосфат
- AF — AlphaFold (алгоритм предсказания структур белков)
- ATM — Ataxia Telangiectasia Mutated (протеинкиназа ATM)
- ATR — Ataxia telangiectasia and Rad3-related kinase (киназа ATR)
- CAMKK2 — Calcium/Calmodulin-Dependent Protein Kinase Kinase 2 (кальций/кальмодулин-зависимая протеинкиназа)
- DNA-PK — DNA-dependent protein kinase (ДНК-зависимая протеинкиназа)
- GNN — Graph Neural Network (графовая нейронная сеть)
- IC_{50} — Half Maximal Inhibitory Concentration (концентрация ингибитора, снижающая активность фермента на 50%)
- IGN — InteractionGraphNet (алгоритм на базе графовых нейросетей)
- MAE — Mean Absolute Error (средняя абсолютная ошибка)
- ML — Machine Learning (машинное обучение)
- PDB — Protein Data Bank (база структур белков)
- PIKK — Phosphatidylinositol 3-Kinase-Related Kinase (семейство протеинкиназ)
- QFASG — Quantum-assisted Fragment-based Automated Structure Generator (квантово-химический автоматический генератор структур на основе фрагментов)
- RMSD — Root Mean Square Deviation (среднеквадратичное отклонение)
- SBDD — Structure-Based Drug Design (структурно-обоснованный дизайн лекарств)
- TTK — Thr/Tyr Kinase, также известная как MPS1 (тирозинкиназа)

СПИСОК ЛИТЕРАТУРЫ

1. Ahmed, Z. The brain-penetrant ATM inhibitor, AZD1390, promotes axon regeneration and functional recovery in preclinical models of spinal cord injury / Z. Ahmed, R.I. Tuxworth // Clin Transl Med. – 2022. – V. 12, № 7. – e962. DOI: 10.1002/ctm2.962.
2. Anderhub, S.J. High Proliferation Rate and a Compromised Spindle Assembly Checkpoint Confers Sensitivity to the MPS1 Inhibitor BOS172722 in Triple-Negative Breast Cancers / S.J. Anderhub, G.W. Mak, M.D. Gurden, et al. // Mol. Cancer Ther. – 2019 – V. 18, №10 – P. 1696–1707. DOI: 10.1158/1535-7163.MCT-18-1203.
3. Baek, M. Accurate prediction of protein structures and interactions using a three-track neural network / M. Baek, F. DiMaio, I. Anishchenko, et al. // Science. – 2021. – V. 373, № 6557. – P. 871–876. DOI: 10.1126/science.abj8754.
4. Bannwarth, C. GFN2-xTB–An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions / C. Bannwarth, S. Ehlert, S. Grimme // J Chem Theory Comput. – 2019. – V. 15, № 3. – P. 1652–1671. DOI: 10.1021/acs.jctc.8b01176.
5. Baldi, S. Advancements in protein kinase inhibitors: from discovery to clinical applications / S. Baldi, N. Long, S. Ma, et al. // Research (Wash D C). – 2025. – V. 8. – 0747. DOI: 10.34133/research.0747.
6. Berman, H.M. The Protein Data Bank / H.M. Berman, J. Westbrook, Z. Feng, et al. // Nucleic Acids Res. – 2000. – V. 28, № 1. – P. 235–242. DOI: 10.1093/nar/28.1.235.
7. Bian, Y. Deep Convolutional Generative Adversarial Network (dcGAN) Models for Screening and Design of Small Molecules Targeting Cannabinoid Receptors / Y. Bian, J. Wang, J.J. Jun, X.-Q. Xie // Mol.

Pharm. – 2019 – V. 16, №11 – P. 4451–4460. DOI: 10.1021/acs.molpharmaceut.9b00500.

8. Bissantz, C. A medicinal chemist's guide to molecular interactions / C. Bissantz, B. Kuhn, M. Stahl // J Med Chem. – 2010. – V. 53, № 14. – P. 5061–5084. DOI: 10.1021/jm100112j.
9. Biswas, B. The prospect of substrate-based kinase inhibitors to improve target selectivity and overcome drug resistance / B. Biswas, Y.H. Huang, D.J. Craik, C.K. Wang // Chem Sci. – 2024. – V. 15, № 33. – P. 13130–13147. DOI: 10.1039/d4sc01088d.
10. Blackford, A.N. ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response / A.N. Blackford, S.P. Jackson // Mol Cell. – 2017. – V. 66, № 6. – P. 801–817. DOI: 10.1016/j.molcel.2017.05.015.
11. Chen, N. Recurrent neural network (RNN) model accelerates the development of antibacterial metronidazole derivatives / N. Chen, L. Yang, N. Ding, et al. // RSC Adv. – 2022. – V. 12, № 35. – P. 22893–22901. DOI: 10.1039/d2ra01807a.
12. Chan, C.Y. CFI-402257, a MPS1 Inhibitor, Effectively Suppresses Hepatocellular Carcinoma / C.Y. Chan, D.K. Chiu, V.W. Yuen, et al. // Proc. Natl. Acad. Sci. U.S.A. – 2022 – V. 119, №32 – P. e2119514119. DOI: 10.1073/pnas.2119514119.
13. Cheng, B. Recent advances in DDR (DNA damage response) inhibitors for cancer therapy / B. Cheng, W. Pan, Y. Xing, Y. Xiao, J. Chen, Z. Xu // Eur J Med Chem. – 2022. – V. 230. – 114109. DOI: 10.1016/j.ejmech.2022.114109.
14. Chu, Y. MoleGear: A Java-Based Platform for Evolutionary De Novo Molecular Design / Y. Chu, X. He // Molecules. – 2019. – V. 24, № 7. – 1444. DOI: 10.3390/molecules24071444.
15. Colombo, R. Preclinical Characterization of the Novel TTK Kinase Inhibitor S81694 for the Treatment of Triple Negative Breast Cancer / R.

Colombo, M. Burbridge, M. Rodriguez, F. Cantero, M. Caldarelli, M.L. Giorgini, F. Sola, D. Ballinari, M. Ciomei, R. Bosotti, A. Montagnoli, A. Isacchi, D. Donati, A. Galvani // Cancer Res. – 2015 – V. 75, №15 – P. 1638–1638.

16. Daina, A. SwissADME: a Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules / A. Daina, O. Michielin, V. Zoete // Sci. Rep. – 2017 – V. 7 – P. 42717. DOI: 10.1038/srep42717.
17. Desaphy, J. sc-PDB: a 3D-database of ligandable binding sites--10 years on / J. Desaphy, G. Bret, D. Rognan, E. Kellenberger // Nucleic Acids Res. – 2015. – V. 43 (Database issue). – D399–D404. DOI: 10.1093/nar/gku928.
18. Eberhardt, J. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings / J. Eberhardt, D. Santos-Martins, A.F. Tillack, S. Forli // J Chem Inf Model. – 2021. – V. 61, № 8. – P. 3891–3898. DOI: 10.1021/acs.jcim.1c00203.
19. Ereshchenko, A. Enhancing Accuracy of Virtual Kinase Profiling via Application of Graph Neural Network to 3D Pharmacophore Ensembles / A. Ereshchenko, S. Evteev, A. Malyshev, et al. // J. Comput. Aided Mol. Des. – 2025 – V. 39, №1 – P. 86. DOI: 10.1007/s10822-025-00666-5.
20. Evteev, S.A. SiteRadar: Utilizing Graph Machine Learning for Precise Mapping of Protein-Ligand-Binding Sites / S.A. Evteev, A.V. Ereshchenko, Y.A. Ivanenkov // J Chem Inf Model. – 2023. – V. 63, № 4. – P. 1124–1132. DOI: 10.1021/acs.jcim.2c01413.
21. Evteev, S. Skittles: GNN-Assisted Pseudo-Ligands Generation and Its Application for Binding Sites Classification and Affinity Prediction / S. Evteev, A. Ereshchenko, D. Adjugim, et al. // Proteins. – 2025. – V. 93, № 7. – P. 1269–1280. DOI: 10.1002/prot.26816.

22. Evteev, S. Quantum-assisted fragment-based automated structure generator (QFASG) for small molecule design: an in vitro study / S. Evteev, Y. Ivanenkov, I. Semenov, et al. // Front Chem. – 2024. – V. 12. – 1382512. DOI: 10.3389/fchem.2024.1382512.
23. Evteev, S. Bifunctional Inhibitors of Influenza Virus Neuraminidase: Molecular Design of a Sulfonamide Linker / S. Evteev, D. Nilov, A. Polenova, V. Švedas // Int J Mol Sci. – 2021. – V. 22, № 23. – 13112. DOI: 10.3390/ijms222313112.
24. Evteev, S. AlphaFold Kinase Optimizer: Enhancing Virtual Screening Performance Through Automated Refinement of AlphaFold-Based Kinase Structures / S. Evteev, Y. Ivanenkov, A. Aiginin, et al. // Proteins. – 2025. DOI: 10.1002/prot.70056.
25. Friesner, R.A. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy / R.A. Friesner, J.L. Banks, R.B. Murphy, et al. // J Med Chem. – 2004. – V. 47, № 7. – P. 1739–1749. DOI: 10.1021/jm0306430.
26. Gagliardi, L. SHREC 2022: protein–ligand binding site recognition / L. Gagliardi, A. Raffo, U. Fugacci, et al. // Comput Graph. – 2022. – V. 107. – P. 20–31.
27. Gilmer, T.M. A Novel Dual ATM/DNA-PK Inhibitor, XRD-0394, Potently Radiosensitizes and Potentiates PARP and Topoisomerase I Inhibitors / T.M. Gilmer, C.H. Lai, K. Guo, et al. // Mol. Cancer Ther. – 2024 – V. 23, №6 – P. 751-765. DOI: 10.1158/1535-7163.MCT-23-0890.
28. Gómez-Bombarelli, R. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules / R. Gómez-Bombarelli, J.N. Wei, D. Duvenaud, et al. // ACS Cent. Sci. – 2018 – V. 4, №2 – P. 268–276. DOI: 10.1021/acscentsci.7b00572.
29. Grisoni, F. Bidirectional Molecule Generation with Recurrent Neural Networks / F. Grisoni, M. Moret, R. Lingwood, G. Schneider // J. Chem.

Inf. Model. – 2020 – V. 60, №3 – P. 1175–1183. DOI: 10.1021/acs.jcim.9b00943.

30. Gurden, M.D. Naturally occurring mutations in the MPS1 gene predispose cells to kinase inhibitor drug resistance / M.D. Gurden, I.M. Westwood, A. Faisal, et al. // Cancer Res. – 2015. – V. 75, № 16. – P. 3340–3354. DOI: 10.1158/0008-5472.CAN-14-3272.
31. Hua, Y. Effective reaction-based de novo strategy for kinase targets: a case study on MERTK inhibitors / Y. Hua, X. Fang, G. Xing, et al. // J Chem Inf Model. – 2022. – V. 62, № 7. – P. 1654–1668. DOI: 10.1021/acs.jcim.2c00068.
32. Ivanenkov, Y.A. Chemistry42: an AI-driven platform for molecular design and optimization / Y.A. Ivanenkov, D. Polykovskiy, D. Bezrukov, et al. // J Chem Inf Model. – 2023. – V. 63, № 3. – P. 695–701. DOI: 10.1021/acs.jcim.2c01191.
33. Ivanenkov, Y. AlphaFold for a medicinal chemist: tool or toy? / Y. Ivanenkov, S. Evteev, A. Malyshev, et al. // Russ. Chem. Rev. – 2024. – V. 93, № 3. – RCR5107.
34. Ivanenkov, Y. The hitchhiker’s guide to deep learning driven generative chemistry / Y. Ivanenkov, B. Zagribelnyy, A. Malyshev, S. Evteev, et al. // ACS Med Chem Lett. – 2023. – V. 14, № 7. – P. 901–915.
35. Ivanenkov, Y.A. Small molecule ATM inhibitors as potential cancer therapy: a patent review (2003–present) / Y.A. Ivanenkov, A.S. Malyshev, V.A. Terentiev, et al. // Expert Opin Ther Pat. – 2025. – V. 35, № 2. – P. 111–136. DOI: 10.1080/13543776.2024.2446228.
36. Jang, S.H. PCW-A1001, AI-assisted de novo design approach to design a selective inhibitor for FLT-3(D835Y) in acute myeloid leukemia / S.H. Jang, D. Sivakumar, S.K. Mudedla, et al. // Front Mol Biosci. – 2022. – V. 9. – 1072028. DOI: 10.3389/fmolb.2022.1072028.

37. Jelluma, N. Chromosomal instability by inefficient Mps1 auto-activation due to a weakened mitotic checkpoint and lagging chromosomes / N. Jelluma, A.B. Brenkman, I. McLeod, et al. // PLoS One. – 2008. – V. 3, № 6. – e2415. DOI: 10.1371/journal.pone.0002415.
38. Jiang, D. InteractionGraphNet: a novel and efficient deep graph representation learning framework for accurate protein–ligand interaction predictions / D. Jiang, C.Y. Hsieh, Z. Wu, et al. // J Med Chem. – 2021. – V. 64, № 24. – P. 18209–18232. DOI: 10.1021/acs.jmedchem.1c01830.
39. Jiménez, J. DeepSite: protein-binding site predictor using 3D-convolutional neural networks / J. Jiménez, S. Doerr, G. Martínez-Rosell, A.S. Rose, G. De Fabritiis // Bioinformatics. – 2017. – V. 33, № 19. – P. 3036–3042. DOI: 10.1093/bioinformatics/btx350.
40. Jumper, J. Highly accurate protein structure prediction with AlphaFold / J. Jumper, R. Evans, A. Pritzel, et al. // Nature. – 2021. – V. 596, № 7873. – P. 583–589. DOI: 10.1038/s41586-021-03819-2.
41. Kandel, J. PUResNet: prediction of protein-ligand binding sites using deep residual neural network / J. Kandel, H. Tayara, K.T. Chong // J Cheminform. – 2021. – V. 13, № 1. – 65. DOI: 10.1186/s13321-021-00547-7.
42. Kim, H.J. Discovery of new quinolone-based diarylamides as potent B-RAFV600E/C-RAF kinase inhibitors endowed with promising in vitro anticancer activity / H.J. Kim, J.W. Park, S. Seo, et al. // Int J Mol Sci. – 2023. – V. 24, № 4. – 3216. DOI: 10.3390/ijms24043216.
43. Kumar, A. Fragment based drug design: from experimental to computational approaches / A. Kumar, A. Voet, K.Y. Zhang // Curr Med Chem. – 2012. – V. 19, № 30. – P. 5128–5147. DOI: 10.2174/092986712803530467.

44. Landrum, G.A. Combining IC50 or Ki values from different sources is a source of significant noise / G.A. Landrum, S. Riniker // J Chem Inf Model. – 2024. – V. 64, № 5. – P. 1560–1567. DOI: 10.1021/acs.jcim.4c00049.
45. Langdon, S.R. Scaffold-focused virtual screening: prospective application to the discovery of TTK inhibitors / S.R. Langdon, I.M. Westwood, R.L. van Montfort, N. Brown, J. Blagg // J Chem Inf Model. – 2013. – V. 53, № 5. – P. 1100–1112. DOI: 10.1021/ci400100c.
46. Lavin, M.F. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks / M.F. Lavin // Oncogene. – 2007. – V. 26, № 56. – P. 7749–7758. DOI: 10.1038/sj.onc.1210880.
47. Leelananda, S.P. Computational methods in drug discovery / S.P. Leelananda, S. Lindert // Beilstein J Org Chem. – 2016. – V. 12. – P. 2694–2718. DOI: 10.3762/bjoc.12.267.
48. Li, H. Computational method to identify druggable binding sites that target protein-protein interactions / H. Li, V. Kasam, C.S. Tautermann, D. Seeliger, N. Vaidehi // J Chem Inf Model. – 2014. – V. 54, № 5. – P. 1391–1400. DOI: 10.1021/ci400750x.
49. Li, Y. Decoding the limits of deep learning in molecular docking for drug discovery / Y. Li, J. Yi, H. Li, et al. // Chem Sci. – 2025. – V. 16, № 37. – P. 17374–17390. DOI: 10.1039/d5sc05395a.
50. Li, Y. Generative deep learning enables the discovery of a potent and selective RIPK1 inhibitor / Y. Li, L. Zhang, Y. Wang, et al. // Nat Commun. – 2022. – V. 13, № 1. – 6891. DOI: 10.1038/s41467-022-34692-w.
51. Longo, L.V.G. TTK/MPS1 inhibitor OSU-13 targets the mitotic checkpoint and is a potential therapeutic strategy for myeloma / L.V.G. Longo, T. Hughes, B. McNeil-Laidley, et al. // Haematologica. – 2024. – V. 109, № 2. – P. 578–590. DOI: 10.3324/haematol.2023.282838.

52. Ma, Z. Growth inhibition of human gastric adenocarcinoma cells in vitro by STO-609 is independent of calcium/calmodulin-dependent protein kinase kinase-beta and adenosine monophosphate-activated protein kinase / Z. Ma, D. Wen, X. Wang, et al. // Am J Transl Res. – 2016. – V. 8, № 2. – P. 1164–1171.
53. Mathison, C.J.N. Efficacy and tolerability of pyrazolo[1,5-a]pyrimidine RET kinase inhibitors for the treatment of lung adenocarcinoma / C.J.N. Mathison, D. Chianelli, P.V. Rucker, et al. // ACS Med Chem Lett. – 2020. – V. 11, № 4. – P. 558–565. DOI: 10.1021/acsmedchemlett.0c00015.
54. McCabe, N. Mechanistic rationale to target PTEN-deficient tumor cells with inhibitors of the DNA damage response kinase ATM / N. McCabe, C. Hanna, S.M. Walker, et al. // Cancer Res. – 2015. – V. 75, № 11. – P. 2159–2165. DOI: 10.1158/0008-5472.CAN-14-3502.
55. Mei, L. Ataxia telangiectasia and Rad3-related inhibitors and cancer therapy: where we stand / L. Mei, J. Zhang, K. He, J. Zhang // J Hematol Oncol. – 2019. – V. 12, № 1. – 43. DOI: 10.1186/s13045-019-0733-6.
56. Mohan, M. Unravelling the role of protein kinase R (PKR) in neurodegenerative disease: a review / M. Mohan, A. Mannan, T.G. Singh // Mol Biol Rep. – 2025. – V. 52, № 1. – 377. DOI: 10.1007/s11033-025-10484-5.
57. Mqawass, G. graphLambda: Fusion graph neural networks for binding affinity prediction / G. Mqawass, P. Popov // J Chem Inf Model. – 2024. – V. 64, № 7. – P. 2323–2330. DOI: 10.1021/acs.jcim.3c00771.
58. Musiol, R. An overview of quinoline as a privileged scaffold in cancer drug discovery / R. Musiol // Expert Opin Drug Discov. – 2017. – V. 12, № 6. – P. 583–597. DOI: 10.1080/17460441.2017.1319357.
59. Newton, R. Discovery and optimization of wt-RET/KDR-selective inhibitors of RETV804M kinase / R. Newton, B. Waszkowycz, C.

Seewooruthun, et al. // ACS Med Chem Lett. – 2020. – V. 11, № 4. – P. 497–505. DOI: 10.1021/acsmedchemlett.9b00615.

60. Novikov, F.N. Lead Finder docking and virtual screening evaluation with Astex and DUD test sets / F.N. Novikov, V.S. Stroylov, A.A. Zeifman, O.V. Stroganov, V. Kulkov, G.G. Chilov // J Comput Aided Mol Des. – 2012. – V. 26, № 6. – P. 725–735. DOI: 10.1007/s10822-012-9549-y.

61. O'Byrne, S.N. In depth analysis of kinase cross screening data to identify CAMKK2 inhibitory scaffolds / S.N. O'Byrne, J.W. Scott, J.R. Pilotte, et al. // Molecules. – 2020. – V. 25, № 2. – 325. DOI: 10.3390/molecules25020325.

62. Paulsen, R.D. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability / R.D. Paulsen, D.V. Soni, R. Wollman, et al. // Mol Cell. – 2009. – V. 35, № 2. – P. 228–239. DOI: 10.1016/j.molcel.2009.06.021.

63. Pike, K.G. Identification of Novel, Selective Ataxia-Telangiectasia Mutated Kinase Inhibitors with the Ability to Penetrate the Blood-Brain Barrier: The Discovery of AZD1390 / K.G. Pike, T.A. Hunt, B. Barlaam, et al. // J. Med. Chem. – 2024 – V. 67, №4 – P. 3090–3111. DOI: 10.1021/acs.jmedchem.3c02277.

64. Pissot Soldermann, C. Discovery of potent, highly selective, and in vivo efficacious, allosteric MALT1 inhibitors by iterative scaffold morphing / C. Pissot Soldermann, O. Simic, M. Renatus, et al. // J Med Chem. – 2020. – V. 63, № 23. – P. 14576–14593. DOI: 10.1021/acs.jmedchem.0c01245.

65. Price, D.J. An orally available, brain-penetrant CAMKK2 inhibitor reduces food intake in rodent model / D.J. Price, D.H. Drewry, L.T. Schaller, et al. // Bioorg Med Chem Lett. – 2018. – V. 28, № 10. – P. 1958–1963. DOI: 10.1016/j.bmcl.2018.03.034.

66. Racioppi, L. Calcium/calmodulin-dependent protein kinase kinase 2: roles in signaling and pathophysiology / L. Racioppi, A.R. Means // J Biol Chem.

– 2012. – V. 287, № 38. – P. 31658–31665. DOI: 10.1074/jbc.R112.356485.

67. Rathi, S. Preclinical systemic pharmacokinetics, dose proportionality, and central nervous system distribution of the ATM inhibitor WSD0628, a novel radiosensitizer for the treatment of brain tumors / S. Rathi, J.H. Oh, W. Zhang, et al. // J Pharmacol Exp Ther. – 2024. – V. 390, № 2. – P. 260–275. DOI: 10.1124/jpet.123.001971.

68. Roskoski, R. Jr. Properties of FDA-approved small molecule protein kinase inhibitors: a 2025 update / R. Roskoski Jr. // Pharmacol Res. – 2025. – V. 216. – 107723. DOI: 10.1016/j.phrs.2025.107723.

69. Santiago, A.D.S. Structural Analysis of Inhibitor Binding to CAMKK1 Identifies Features Necessary for Design of Specific Inhibitors / A.D.S. Santiago, R.M. Couñago, P.Z. Ramos, et al. // Sci. Rep. – 2018 – V. 8, №1 – P. 14800. DOI: 10.1038/s41598-018-33043-4.

70. Schulze, V.K. Treating Cancer by Spindle Assembly Checkpoint Abrogation: Discovery of Two Clinical Candidates, BAY 1161909 and BAY 1217389, Targeting MPS1 Kinase / V.K. Schulze, U. Klar, D. Kosemund, et al. // J. Med. Chem. – 2020 – V. 63, №15 – P. 8025–8042. DOI: 10.1021/acs.jmedchem.9b02035.

71. Schwartz, G.K. Targeting the cell cycle: a new approach to cancer therapy / G.K. Schwartz, M.A. Shah // J Clin Oncol. – 2005. – V. 23, № 36. – P. 9408–9421. DOI: 10.1200/JCO.2005.01.5594.

72. Shi, W. BionoiNet: ligand-binding site classification with off-the-shelf deep neural network / W. Shi, J.M. Lemoine, A.A. Shawky, et al. // Bioinformatics. – 2020. – V. 36, № 10. – P. 3077–3083. DOI: 10.1093/bioinformatics/btaa094.

73. Shibata, A. ATM's role in the repair of DNA double-strand breaks / A. Shibata, P.A. Jeggo // Genes (Basel). – 2021. – V. 12, № 9. – 1370. DOI: 10.3390/genes12091370.

74. Shiloh, Y. The ATM protein kinase: regulating the cellular response to genotoxic stress, and more / Y. Shiloh, Y. Ziv // Nat Rev Mol Cell Biol. – 2013. – V. 14, № 4. – P. 197–210. DOI: 10.1038/nrm3546.
75. Shu, J. ATM inhibitor KU60019 synergistically sensitizes lung cancer cells to topoisomerase II poisons by multiple mechanisms / J. Shu, X. Wang, X. Yang, G. Zhao // Sci Rep. – 2023. – V. 13, № 1. – 882. DOI: 10.1038/s41598-023-28185-z.
76. Sicho, M. GenUI: interactive and extensible open source software platform for de novo molecular generation and cheminformatics / M. Sicho, X. Liu, D. Svozil, G.J.P. van Westen // J Cheminform. – 2021. – V. 13, № 1. – 73. DOI: 10.1186/s13321-021-00550-y.
77. Sledge, G.W. Jr. The role of angiogenesis inhibition in the treatment of breast cancer / G.W. Sledge Jr., H.S. Rugo, H.J. Burstein // Clin Adv Hematol Oncol. – 2006. – V. 4, Suppl 21. – P. 1–12.
78. Song, S. Application of deep generative model for design of Pyrrolo[2,3-d] pyrimidine derivatives as new selective TANK binding kinase 1 (TBK1) inhibitors / S. Song, H. Tang, T. Ran, et al. // Eur J Med Chem. – 2023. – V. 247. – 115034. DOI: 10.1016/j.ejmech.2022.115034.
79. Subbannayya, Y. Calcium calmodulin dependent kinase kinase 2 - a novel therapeutic target for gastric adenocarcinoma / Y. Subbannayya, N. Syed, M.A. Barbhuiya, et al. // Cancer Biol Ther. – 2015. – V. 16, № 2. – P. 336–345. DOI: 10.4161/15384047.2014.972264.
80. Tan, X. Discovery of pyrazolo[3,4-d]pyridazinone derivatives as selective DDR1 inhibitors via deep learning based design, synthesis, and biological evaluation / X. Tan, C. Li, R. Yang, et al. // J Med Chem. – 2022. – V. 65, № 1. – P. 103–119. DOI: 10.1021/acs.jmedchem.1c01205.
81. Waenphimai, O. Multiple Actions of NMS-P715, the Monopolar Spindle 1 (MPS1) Mitotic Checkpoint Inhibitor in Liver Fluke-Associated Cholangiocarcinoma Cells / O. Waenphimai, P. Mahalapbutr, K.

Vaeteewoottacharn, et al. // Eur. J. Pharmacol. – 2022 – V. 922 – P. 174899. DOI: 10.1016/j.ejphar.2022.174899.

82. Wang, R. The PDBbind database: methodologies and updates / R. Wang, X. Fang, Y. Lu, C.Y. Yang, S. Wang // J Med Chem. – 2005. – V. 48, № 12. – P. 4111–4119. DOI: 10.1021/jm048957q.
83. Waqar, S.N. Phase I trial of ATM inhibitor M3541 in combination with palliative radiotherapy in patients with solid tumors / S.N. Waqar, C. Robinson, A.J. Olszanski, et al. // Invest New Drugs. – 2022. – V. 40, № 3. – P. 596–605. DOI: 10.1007/s10637-022-01216-8.
84. Wu, X. Protein kinases in neurodegenerative diseases: current understandings and implications for drug discovery / X. Wu, Z. Yang, J. Zou, et al. // Signal Transduct Target Ther. – 2025. – V. 10, № 1. – 146. DOI: 10.1038/s41392-025-02179-x.
85. Xu, H. The slow but steady rise of binding free energy calculations in drug discovery / H. Xu // J Comput Aided Mol Des. – 2023. – V. 37, № 2. – P. 67–74. DOI: 10.1007/s10822-022-00494-x.
86. Xue, Z. The novel brain penetrant ataxia-telangiectasia mutated inhibitor WSD0628 provides robust radiosensitization of brain tumor patient-derived xenografts / Z. Xue, A.C. Mladek, S. Rathi, et al. // Neuro Oncol. – 2025 – V. 27, №9. – P. 2313–2325. DOI: 10.1093/neuonc/noaf102.
87. Yao, W. TTK: A promising target in malignant tumors / W. Yao, M. Jiang, M. Zhang, et al. // J. Cell. Signal. – 2021 – V. 2, №3. – P. 212–220.
88. Yoshimori, A. Design and synthesis of DDR1 inhibitors with a desired pharmacophore using deep generative models / A. Yoshimori, Y. Asawa, E. Kawasaki, et al. // ChemMedChem. – 2021. – V. 16, № 6. – P. 955–958. DOI: 10.1002/cmdc.202000786.
89. Yu, Y. A novel scalarized scaffold hopping algorithm with graph-based variational autoencoder for discovery of JAK1 inhibitors / Y. Yu, T. Xu, J.

Li, et al. // ACS Omega. – 2021. – V. 6, № 35. – P. 22945–22954. DOI: 10.1021/acsomega.1c03613.

90. Zaraei, S.O. Design and synthesis of new quinoline derivatives as selective C-RAF kinase inhibitors with potent anticancer activity / S.O. Zaraei, N.N. Al-Ach, H.S. Anbar, et al. // Eur. J. Med. Chem. – 2022. – V. 238. – 114434. DOI: 10.1016/j.ejmech.2022.114434.

91. Zhang, W. Identification of the binding site of an allosteric ligand using STD-NMR, docking, and CORCEMA-ST calculations / W. Zhang, R. Li, R. Shin, et al. // ChemMedChem. – 2013. – V. 8, № 10. – P. 1629–1633. DOI: 10.1002/cmdc.201300267.

92. Zhang, Y. Identification of covalent binding sites targeting cysteines based on computational approaches / Y. Zhang, D. Zhang, H. Tian, et al. // Mol. Pharm. – 2016. – V. 13, № 9. – P. 3106–3118. DOI: 10.1021/acs.molpharmaceut.6b00302.

93. Zhavoronkov, A. Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors / A. Zhavoronkov, Y.A. Ivanenkov, A. Aliper, et al. // Nat. Biotechnol. – 2019 – V. 37, №9 – P. 1038–1040. DOI: 10.1038/s41587-019-0224-x.

94. Zhou, Y. Identification of novel quinoline analogues bearing thiazolidinones as potent kinase inhibitors for the treatment of colorectal cancer / Y. Zhou, X. Xu, F. Wang, et al. // Eur. J. Med. Chem. – 2020. – V. 204. – 112643. DOI: 10.1016/j.ejmech.2020.112643.

95. Zimmermann, A. A new class of selective ATM inhibitors as combination partners of DNA double-strand break inducing cancer therapies / A. Zimmermann, F.T. Zenke, L.Y. Chiu, et al. // Mol. Cancer Ther. – 2022. – V. 21, № 6. – P. 859–870. DOI: 10.1158/1535-7163.MCT-21-0934.