

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

На правах рукописи

Загрибельный Богдан

**Платформа генеративной химии в моделировании структур
потенциальных лекарственных веществ**

1.4.16. Медицинская химия

1.4.3. Органическая химия

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата химических наук

Москва – 2025

Диссертация подготовлена на кафедре медицинской химии и тонкого органического синтеза химического факультета МГУ имени М.В.Ломоносова.

Научные руководители: **Палюлин Владимир Александрович**
кандидат химических наук, старший научный сотрудник

Иваненков Ян Андреевич
кандидат биологических наук

Официальные оппоненты: **Свитанько Игорь Валентинович**
доктор химических наук,
ФГБУН Институт органической химии им. Н.Д. Зелинского Российской академии наук, лаборатория молекулярного моделирования и направленного синтеза, заведующий лабораторией

Лагунин Алексей Александрович
доктор биологических наук, профессор РАН,
ФГАОУ ВО «Российский национальный исследовательский медицинский университет имени Н.И. Пирогова» Министерства здравоохранения Российской Федерации, Институт биомедицины (МБФ), заведующий кафедрой биоинформатики

Латышев Геннадий Владимирович
кандидат химических наук,
ФГБОУ ВО «Московский государственный университет имени М.В.Ломоносова», химический факультет, ведущий научный сотрудник кафедры органической химии

Защита диссертации состоится « 14 » ноября 2025 г. в 13 ч. 00 мин. на заседании диссертационного совета МГУ.014.7 Московского государственного университета имени М.В.Ломоносова по адресу: 119991, г. Москва, ГСП-1, Ленинские горы, д.1, стр.3, МГУ имени М.В.Ломоносова, химический факультет, аудитория 446.

E-mail: 7422990@mail.ru (e-mail совета)

С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки МГУ имени М.В.Ломоносова (Ломоносовский просп., д. 27) и на портале <https://dissovet.msu.ru/dissertation/3525>

Автореферат разослан « » _____ 2025 г.

Ученый секретарь
диссертационного совета МГУ.014.7,
кандидат химических наук

Н.А. Синикова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность и степень разработанности темы исследования.

Поиск новых малых лекарственных молекул остается краеугольным камнем разработки потенциальных лекарственных веществ. Малые лекарственные молекулы обладают уникальной способностью взаимодействовать с биологическими макромолекулами, в первую очередь, такими как белки, модулируя их функцию таким образом, что это может привести к терапевтическим эффектам. Эта универсальность делает малые молекулы незаменимыми в лечении широкого спектра заболеваний, от инфекционных болезней и рака до неврологических и аутоиммунных расстройств.

Развитие технологий искусственного интеллекта (ИИ) и машинного обучения значительно ускорило поиск новых малых лекарственных молекул, несмотря на то что пока созданные при помощи ИИ потенциальные лекарственные вещества не были зарегистрированы национальными профильными регуляторами. Тем не менее, эти технологии позволяют быстро анализировать огромные наборы данных, прогнозировать свойства и оптимизировать процесс разработки препаратов, снижая затраты и время, связанные с разработкой лекарств. В то же время, создание новых лекарственных веществ всё ещё является крайне трудоемкой и дорогой задачей, требующей многопараметрической оптимизации, которая помимо чисто химических и фармакологических требований к фармсубстанции, включает и факторы иной природы, такие как экономическую целесообразность (баланс между размером рынка и затратами на разработку и производство с учетом возможной конкуренции) и юридические аспекты в лице особенностей патентной конкуренции и регистрации препаратов. Интеграция учета всех факторов, прямо или косвенно влияющих на принятие решение в рамках разработки потенциальных лекарственных веществ, в рамках платформенных решений на основе искусственного интеллекта представляется для профессионального сообщества ключом к проблемам индустрии. Описанию одной из первых в мире таких платформ и созданию её ключевых узлов, связанных с моделированием структурных трендов медицинской химии и оценкой синтезируемости молекулярных структур посвящена настоящая диссертация.

Разработанная платформа генеративной химии Chemistry42 является первым в своем роде инструментом решения задач генеративной химии. По этой причине профессиональное сообщество не было консолидировано относительно того, какие сценарии моделирования структур потенциальных лекарственных веществ могут выполняться при помощи подобной платформы.

Предлагаемое в диссертации теоретическое определение понятия о синтетической доступности на языке теории вероятности ранее не было описано в литературе. Разработанный метод моделирования синтетической доступности молекулярных структур ReRSA (*Retrosynthesis-Related Synthetic Accessibility*, англ., синтетическая доступность, связанная с ретросинтезом) является первым описанным методом, учитывающим одновременно ретросинтетический, статистический и дескрипторный факторы, сочетает в

себе удовлетворительную точность и высокую скорость, делающую его первым подобным методом, в контексте задач генеративной химии.

Ранее в литературе не были описаны молекулярные дескрипторы, способные давать оценку соответствия молекулярных структур трендам, наблюдаемым в медицинской химии. Необходимость балансировать структурную новизну в терминах соответствия текущему состоянию развития медицинской химии и синтетическую доступность генерируемых молекулярных структур требовало создания вышеупомянутого молекулярного дескриптора.

Цель работы заключается в обеспечении разработанной платформы генеративной химии надежными алгоритмами оценки синтезируемости и соответствия структурным трендам медицинской химии, а также модельными экспериментами, иллюстрирующими функциональность платформы с позиции базовых сценариев моделирования структур потенциальных лекарственных веществ.

Для достижения указанной цели были поставлены следующие **задачи**:

1. Добиться того, чтобы модельные эксперименты покрывали большую часть базовых сценариев компьютеризированного дизайна малых лекарственных молекул и позволили пользователям эффективно овладеть функционалом платформы генеративной химии.

2. Учесть в рамках разработки нового метода моделирования синтетической доступности лучшие стороны существующих подходов и сделать для нового метода удобную визуализацию в рамках пользовательского интерфейса в целях повышения интерпретируемости результатов.

3. Создать набор данных из молекулярных структур, запатентованных крупнейшими фармацевтическими компаниями, с учётом хронологического порядка в целях более точного моделирования структурных трендов и соответствующую функцию, описывающие эти тренды.

Объектами исследования являлись молекулярные структуры зарегистрированных и потенциальных лекарственных веществ — малых лекарственных молекул.

Предметом исследования являлось моделирование синтетической доступности молекулярных структур зарегистрированных и потенциальных лекарственных веществ, моделирование структурных трендов, наблюдаемых в медицинской химии, а также моделирование молекулярных структур в соответствии с базовыми сценариями компьютеризированного дизайна малых лекарственных молекул.

Методология и методы исследования.

Создание и первичная валидация программного кода MCE-18 (*Medicinal Chemistry Evolution*, англ., эволюция медицинской химии) и ReRSA выполнялось на основе хемоинформатической библиотеки RDKit на языке программирования Python. Вторичная валидация методов осуществлялась на платформе генеративной химии Chemistry42 в рамках модельных экспериментов.

Научная новизна.

В настоящей диссертации впервые предложен метод моделирования синтетической доступности, объединяющий элементы статистического анализа встречаемости фрагментов в референсном химическом пространстве, ретросинтетический анализ и дескрипторный подход к моделированию структурной сложности. Часть функционала метода, анализирующая встречаемость 5-членных ароматических гетероциклов, использует новаторские хемоинформатические идеи, в частности, автоматизированную генерацию больших библиотек SMARTS-подструктур, компрессию SMARTS-строк по атомным примитивам, иерархический поиск по SMARTS-подструктурам.

При разработке молекулярного дескриптора MCE-18 было впервые обозначено различие при анализе sp^3 -гибридизированных атомов углерода на предмет их включенности в кольцевые системы. Данное различие вошло в основу дескриптора NCSPTR, компонента дескриптора MCE-18, который выгодно отличается от классического дескриптора Fsp^3 , который, в свою очередь, ассоциируется с успехом в клинических испытаниях. В отличие от Fsp^3 , дескриптор MCE-18 эффективно описывает структурную эволюцию химического пространства потенциальных лекарственных веществ.

Теоретическая и практическая ценность работы.

Разработан теоретический аппарат на основе теории вероятности для сферы знаний о синтетической доступности. Метод моделирования синтетической доступности ReRSA запатентован и интегрирован в платформу генеративной химии Chemistry42, которой пользуются крупнейшие фармацевтические компании мира, включая Roche, Merck, Elly Lilly, BMS, Arvinas, UCB, Takeda и другие.

В рамках дизайна и валидации платформы генеративной химии Chemistry42 были созданы модельные эксперименты, иллюстрирующие базовые сценарии ранней разработки потенциальных лекарственных веществ. Данные модельные эксперименты могут быть взяты за основу для разработки и валидации любой другой платформы генеративной химии или быть использованы для создания бенчмаркинг-платформы, оценивающей эффективность генеративных моделей. Метод оценки соответствия структурным трендам медицинской химии MCE-18 интегрированный в платформу позволяет пользователям понять, насколько генерируемые молекулярные структуры соответствуют современным трендам в разработке потенциальных лекарственных веществ.

Положения, выносимые на защиту.

1. Модельные эксперименты, созданные в целях иллюстрации функционала разработанной платформы генеративной химии Chemistry42 в рамках базовых сценариев компьютеризированного дизайна малых лекарственных молекул, позволяют проводить обучение на платформе и её валидацию.

2. Разработанный метод моделирования синтетической доступности ReRSA позволяет эффективно генерировать на платформе молекулярные структуры с высокой ожидаемой синтетической осуществимостью.

3. Концепция полноподструктурного анализа молекулярных структур, которая была разработана в ходе создания метода моделирования синтетической доступности ReRSA и провалидирована на примере пятичленных ароматических гетероциклов, обладает большим потенциалом для хемоинформатической области и может быть распространена на любой другой класс подструктур (циклы, линкеры, периферические фрагменты).

4. Разработанный молекулярный дескриптор MCE-18 позволяет создавать на платформе генеративной химии Chemistry42 молекулярные структуры, соответствующие трендам современной медицинской химии, которые задают крупнейшие фармацевтические компании.

Достоверность полученных результатов обеспечивается использованием для расчетов стандартных, широко используемых, статистически обоснованных алгоритмов и программного обеспечения, публикациями в рецензируемых научных изданиях.

Личный вклад автора состоит в подборе, анализе и систематизации литературы, постановке промежуточных задач. Автор принимал непосредственное участие в создании важнейших элементов платформы генеративной химии, включая модули оценки синтезируемости молекулярных структур и оценки соответствия трендам современной медицинской химии, в обработке и интерпретации экспериментального материала, подготовке материалов к публикации в научных журналах, написании патентных заявок и в представлении полученных результатов на конференции. Во всех опубликованных в соавторстве работах по теме диссертационной работы вклад автора (Загрибельного Б.) является основополагающим, в том числе в работе [1], где автором проделана работа сбора обучающей выборки ингибиторов DDR1 киназы для генеративной модели GENTRL; в работе [2] и патенте [3], где автор проделал работу по моделированию структур-кандидатов потенциальных ингибиторов главной протеазы SARS-CoV-2 на платформе Chemistry42, отбору структур на синтез, подготовке патентной заявки, оптимизации соединений-хитов и, в

¹ Zhavoronkov A., Ivanenkov Y.A., Aliper A., Veselov M.S., Aladinskiy V.A., Aladinskaya A.V., Terentiev V.A., Polykovskiy D.A., Kuznetsov M.D., Asadulaev A., Volkov Y., Zholus A., Shayakhmetov R.R., Zhebrak A., Minaeva L.I., **Zagribelnyy B.**, Lee L.H., Soll R., Madge D., Xing L., Guo T., Aspuru-Guzik A. Deep learning enables rapid identification of potent DDR1 kinase inhibitors // *Nature Biotechnology* — 2019. — V. 37. — No. 9. — pp. 1038-1040, — DOI: 10.1038/s41587-019-0224-x. — EDN YKXOEF. Импакт-фактор 41.7 (JIF), 0.35 п.л., доля вклада 10%

² Sun J., Sun D., Yang Q., Wang D., Peng J., Guo H., Ding X., Chen Zh., Yuan B., Ivanenkov Y.A., Yuan J., **Zagribelnyy B.**, He Y., Su J., Wang L., Tang J., Li Zh., Li R., Li T., Hu X., Liang X., Zhu A., Wei P., Fan Y., Liu S., Zheng J., Guan X., Aliper A., Yang M., Bezrukov D.S., Xie Zh., Terentiev V.A., Peng G., Polykovskiy D.A., Malyshev A.S., Malkov M.N., Zhu Q., Aspuru-Guzik A., Ding X., Cai X., Zhang Man, Zhao J., Zhong N., Ren F., Chen X., Zhavoronkov A., Zhao J. A novel, covalent broad-spectrum inhibitor targeting human coronavirus Mpro // *Nature Communications* — 2025. — V. 16. — P. 4546, — DOI 10.1038/s41467-025-59870-4. — EDN DTDLMX. Импакт-фактор 15.7 (JIF), 1.15 п.л., доля вклада 25%

³ Патент № US20230174488A1. Sars-cov-2 inhibitors having covalent modifications for treating coronavirus infections: опубл. 08.06.2023 / Zhavoronkovs A., Ivanenkov Y.A., **Zagribelnyy B.**, 9.03 п.л., доля вклада 40%

целом, по руководству всем проектом по разработке упомянутых ингибиторов, начиная с идентификации соединений-хитов, вплоть до номинирования лидирующей серии соединений; в патенте [4], где автор проделал работу по химической концептуализации движка для автоматизированного ретросинтеза.

Апробация работы и публикации.

По результатам диссертации опубликовано 9 работ, в том числе 6 статей в рецензируемых научных журналах, индексируемых в базе ядра РИНЦ «eLibrary Science Index», международными базами данных (Web of Science, Scopus, RSCI) и рекомендованных для защиты в диссертационном совете МГУ для публикации результатов диссертационных работ по специальностям 1.4.16. Медицинская химия (химические науки) и 1.4.3. Органическая химия (химические науки) и 3 патента. Результаты, полученные в ходе проделанной работы, были представлены на XI Международной конференции молодых ученых по химии “Mendeleev–2019”, 9–13 сентября 2019. г. Петергоф. По результатам конференции автор настоящей диссертации был удостоен третьей премии “за лучший устный доклад”. Также по материалам диссертации был представлен доклад на IV Международном форуме об искусственном интеллекте, робототехнике, инновациях в образовании и подготовке кадров “Digital Innopolis Days 2024”, 2–4 октября 2024. Республика Татарстан, г. Иннополис.

Структура и объем работы.

Работа состоит из введения, трех глав, включающих обзор литературы, материалы и методы, результаты и их обсуждение, а также заключения, списка литературы и приложения. Общее количество страниц: 220, включая приложение. Основная часть работы содержит 77 иллюстраций (69 рисунков и 8 синтетических схем) и 41 таблицу; список литературы включает 196 наименований. Приложение, данное на 2 страницах, содержит 1 таблицу.

⁴ Патент № US20220172802A1. Retrosynthesis systems and methods: опубли. 02.06.2022 / Konstantinov A., Putin E.O., Zagribelnyy B., Ivanenkov Y.A., Zhavoronkovs A., 1.96 п.л., доля вклада 30%

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Chemistry42 Обзор генеративной платформы



1. Платформа генеративной химии Chemistry42

Разработанная платформа генеративной химии Chemistry42 представляет собой универсальное техническое решение для выполнения задач по автоматизированному моделированию структур потенциальных лекарственных веществ с заданным профилем свойств.

На архитектурном уровне платформа представляет собой конвейер из ансамбля генеративных моделей, производящих молекулярные структуры, и оценивающих и фильтрующих модулей, результат работы которых интегрируется в обобщенное значение функции награды (Рисунок 1). Данное значение передается генеративным моделям, обеспечивая непрерывный цикл их обучения на основании оценок, даваемых разнообразными 2D и 3D модулями, включая модули оценки синтезируемости, оценки подобия лекарствам, фармакофорный модуль, самоорганизующиеся карты Кохонена, модуль молекулярного докинга и многие другие.

Рисунок 1. Верхнеуровневое описание платформы Chemistry42

2. Базовые сценарии компьютеризированного дизайна малых лекарственных молекул и модельные эксперименты в рамках платформы Chemistry42

Разработка модельных экспериментов, которую проводили для платформы Chemistry42, носила исследовательскую, техническую и дидактическую ценности. Благодаря созданным модельным экспериментам удалось проиллюстрировать функционал платформы с точки зрения базовых стратегий компьютеризированного дизайна структур малых лекарственных молекул (см. табл.1). При помощи разработанных модельных экспериментов проводится валидация платформы на работоспособность в ходе каждого её обновления. Также на основании модельных экспериментов проводится обучение всех новых пользователей платформы.

Таблица 1. Модельные эксперименты платформы генеративной химии Chemistry42

#	Стратегия дизайна	Модельный эксперимент
1	Виртуальный скрининг	Виртуальный скрининг ингибиторов папаин-подобной протеазы коронавируса SARS-CoV-2
2	<i>De novo</i> дизайн	Генеративный <i>de novo</i> дизайн ингибиторов Jak3 киназы
3	<i>Hit-expansion</i>	Генеративный дизайн аналогов соединения-хита протеазы USP7
4	<i>Scaffold-hopping</i>	Генеративный <i>scaffold-hopping</i> дизайн ингибиторов CAMKK2 киназы
5	Дизайн R-групп	Генеративный дизайн заместителей ингибитора MPS1 киназы
6	FBDD	Генеративный дизайн ингибиторов главной протеазы коронавируса SARS-CoV-2 на основе знаний о связывании малого фрагмента

Для каждого из модельных экспериментов проводили серию запусков платформы с целью её стабилизации и получения эталонных результатов. Эталонные результаты систематизировали и представляли в руководстве пользователя платформы.

3. ReRSA — новый комбинированный метод моделирования синтетической синтетической доступности для решения задач генеративной химии

Ввиду недостатков существующих методов моделирования синтетической доступности (ММСД), в целях эффективной оценки синтезируемости молекулярных структур предложили новый ММСД, который комбинирует все преимущества существующих ММСД: как учет фактически синтезированного пространства в случае учёта подхода на основе анализа данных, так и учёт ретросинтетического фактора, как наиболее важного для наиболее адекватной оценки синтезируемости, а также фактор структурных дескрипторов, ассоциированных с понятием структурной сложности. Предлагаемый ММСД носит название **ReRSA** (*Retrosynthesis-Related Synthetic Accessibility*, связанная с ретросинтезом синтетическая доступность), ввиду того, что основной алгоритм метода в сущности представляет собой подобие ретросинтетического анализа.

Метод ReRSA представляет собой метод моделирования синтетической доступности молекулярных структур, включающий в себя следующие ключевые операции:

1. Фрагментация молекулярной структуры на синтоноподобные фрагменты и сравнение этих фрагментов с референсным набором фрагментов, извлекаемых из баз данных синтезированных веществ, например из ChEMBL. Разбиение на синтоноподобные фрагменты производится при помощи 52 квази-реакций, соответствующих наиболее распространенным реакциям в органическом синтезе потенциальных лекарственных веществ;

2. Конвертация синтоноподобных фрагментов, полученных на предыдущем этапе, в стартовые материалы и сравнение полученных синтетических эквивалентов с референсным набором коммерчески доступных исходных соединений (КДИС);

3. Расчёт молекулярных дескрипторов, описывающих структурную (молекулярную) сложность, для молекулярной структуры;

4. Агрегация результатов операций (1), (2) и (3) в единую оценку синтезируемости молекулы ReRSA (см. принципиальную схему на Рисунке 2).

Техническую реализацию алгоритма ReRSA осуществляли при помощи хемоинформатической библиотеки RDKit, реализованную на языке программирования Python, путем существенной модификации и дополнения метода BRICSDecompose.

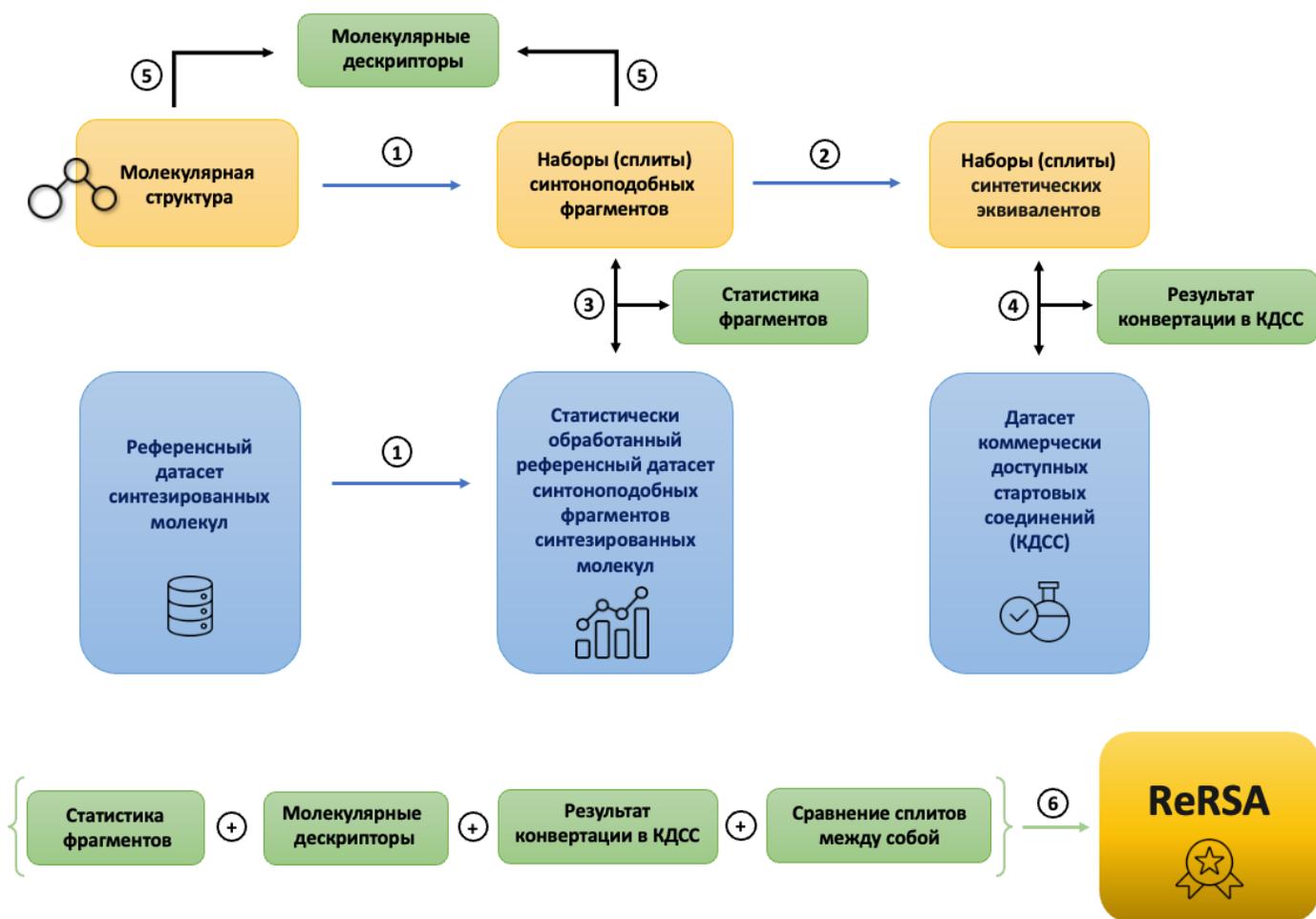


Рисунок 2. Принципиальная схема ММСД ReRSA.

Визуальный результат работы модуля оценки синтезируемости ReRSA представляет собой набор КДИС, индексированных при помощи идентификационных номеров CAS, найденных в ходе квази-ретросинтетической процедуры (см. Рисунок 3), а также агрегированное значение оценки синтезируемости, нормированное по шкале от 1 до 10, где 1 соответствует высокой синтетической доступности, а 10 — низкой синтетической доступности. На этапе ① (см. Рисунок 3) проходит квази-ретросинтетическая фрагментация молекулярной структуры барицитиниба, в ходе которой получают синтоноподобные фрагменты. На этапе ② проходит конвертация синтоноподобных фрагментов в синтетические эквиваленты, которые затем индексируются в базе данных коммерчески доступных исходных соединений (КДИС).

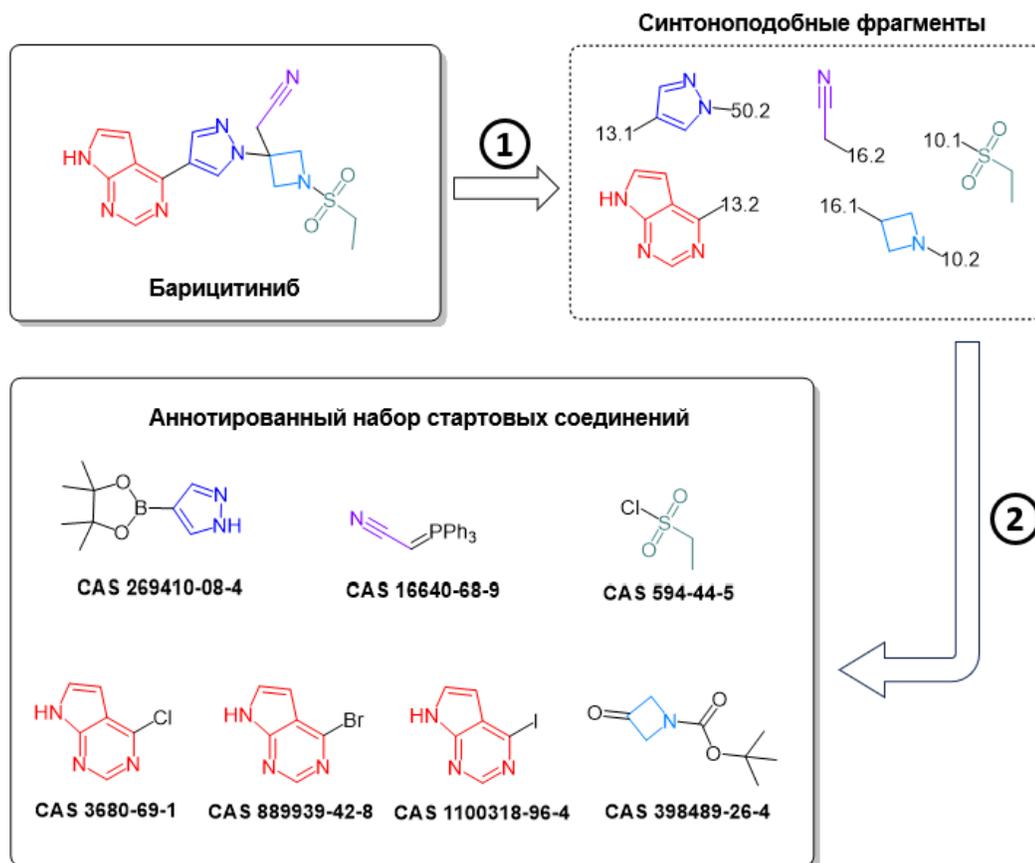


Рисунок 3. Квази-ретросинтетический анализ ингибитора Янус-киназы барцитиниба при помощи алгоритма ReRSA

В рамках разработки ММСД ReRSA предложена подход к полноподструктурному анализу молекулярных структур. Предлагаемый подход основан на фильтрации молекулярных структур, содержащих синтетически нерелевантные подструктуры, с применением иерархически организованной библиотеки SMARTS-подструктур, учитывающих локальное окружение фрагмента, что компенсирует недостатки существующих ММСД. Апробация разработанного подхода проведена на примере пятичленных ароматических гетероциклов. Применение алгоритма позволяет значительно снизить число ложноположительных предсказаний, часто возникающих при использовании существующих ММСД, и уже в текущем виде предложенный алгоритм может служить эффективным дополнительным фильтром. Впервые описан иерархический алгоритм фильтрации подструктур, а также предложен подход к автоматизированному объединению SMARTS-строк по атомным примитивам.

In silico валидацию метода ReRSA проводили на наборе данных из 80 молекулярных структур клинических кандидатов и зарегистрированных лекарственных веществ, отобранных экспертами в области медицинской и органической химии из компании WuXi AppTec. Основным критерием, который влиял на оценку для каждого из лекарственных веществ, была способность конвертировать молекулярную структуру

вещества в КДИС. Успешным выполнением процедуры считалась полная конвертация в КДИС. Помимо этого, внимание уделялось и тому, насколько предложенные методом ReRSA КДИС соответствуют тем исходным соединениям, которые были использованы для получения конечных веществ в рамках опубликованных ранее синтетических схем (см. Рисунок 4).

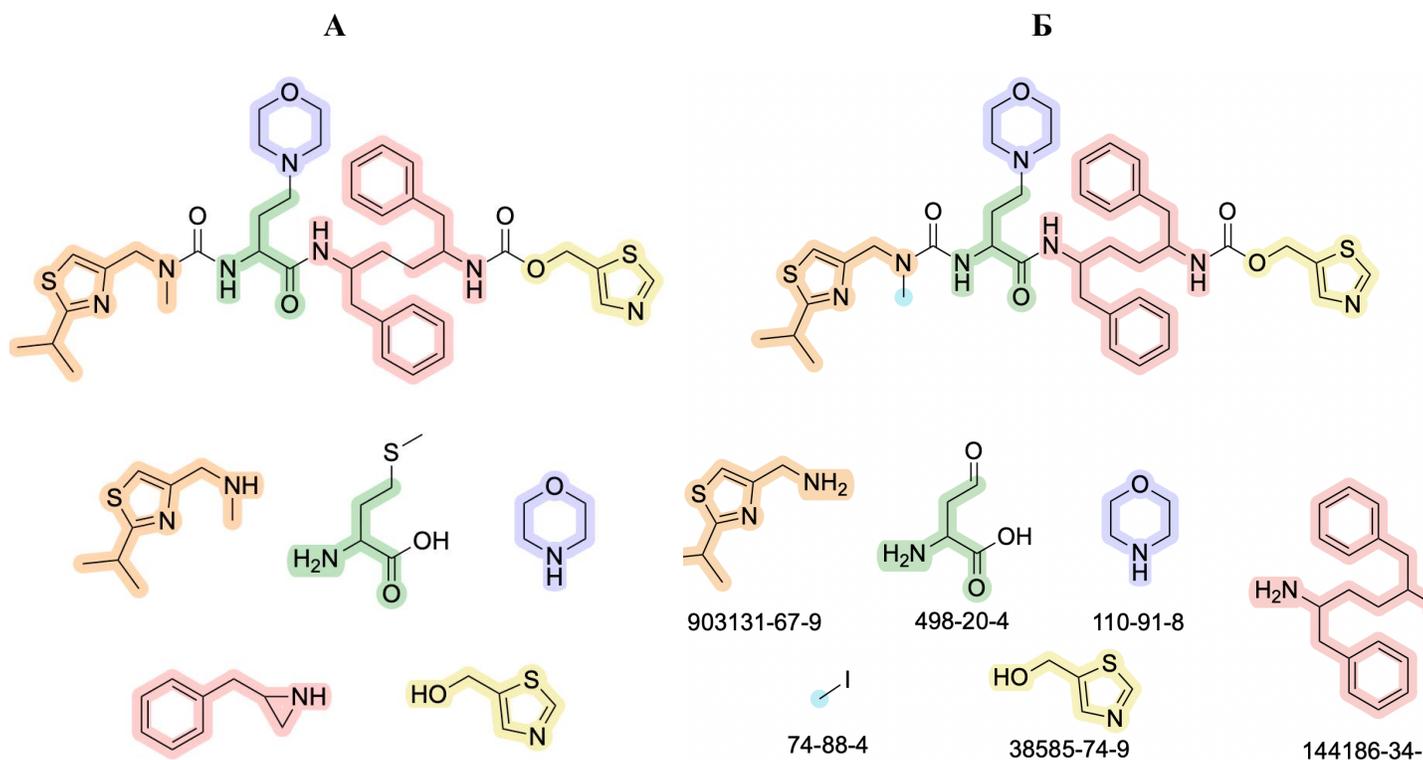


Рисунок 4. Исходные соединения, предлагаемые в оригинальной схеме синтеза кобицистата. (А) КДИС с идентификационными номерами CAS, предлагаемые MMCД ReRSA в ходе квазиретросинтетического анализа кобицистата (Б)

По итогам *in silico* эксперимента полная конверсия молекулярных структур в КДИС наблюдали для 45 соединений из 80. При этом для большей части соединений, которые не были полностью сконвертированы, степень конверсии оставалась высокой, то есть не были сконвертированы в КДИС единичные фрагменты, чаще всего циклической природы.

Другой сценарий *in silico* валидации алгоритма ReRSA включал в себя сравнении результатов моделирования синтетической доступности новых молекулярных структур алгоритмом ReRSA и движком автоматизированного ретросинтеза. Объектом эксперимента являлись новые молекулярные структуры, сгенерированные на платформе Chemistry42. Источниками послужили 5 генеративных модельных экспериментов (см., табл. 1), а также 9 SBDD генеративных экспериментов с новыми клинически релевантными мишенями лекарств. Доля успешно решенных соединений (*ретросинтетическая сходимость*, доля соединений, для которых ретросинтетический модуль нашел хотя бы 1 путь, сходящийся к КДИС в

рамках перечисленных выше ограничений поиска) для каждого из диапазонов ReRSA приведена на рис. 5. Диаграмма демонстрирует сильную корреляцию между оценкой ReRSA и ретросинтетической сходимостью. Ретросинтетическая сходимость снижается с ростом значений ReRSA.

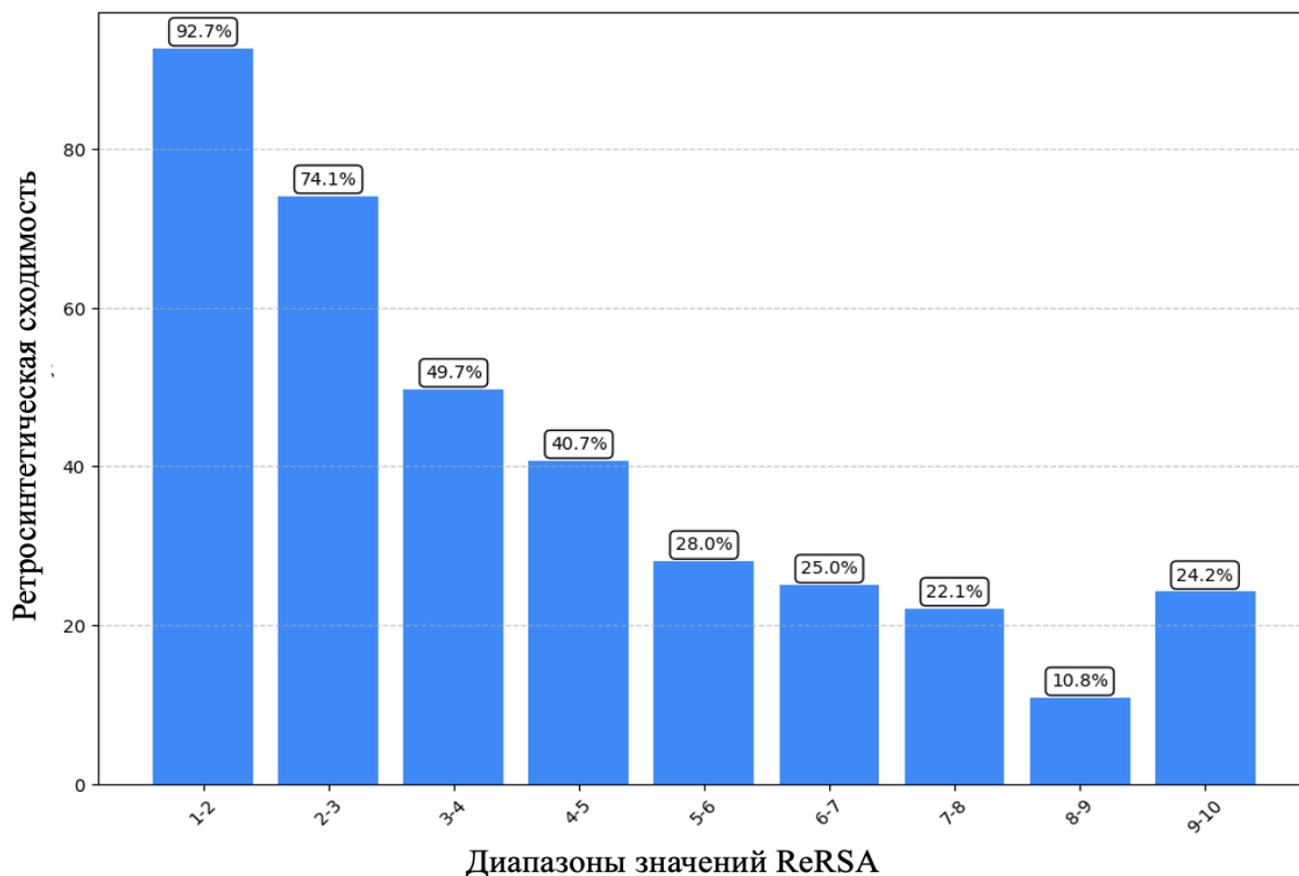


Рисунок 5. Ретросинтетическая сходимость молекулярных структур в зависимости от значений оценки синтезируемости в рамках модели ReRSA.

Настоящий эксперимент подтверждает гипотезу о том, что оценка ReRSA коррелирует с синтетической сложностью молекулярной структуры, даже несмотря на то, что “химический арсенал” модуля ReRSA, содержащий всего 52 вида квази-реакций органического синтеза (см. п. 2.2.4), заметно уступает таковому автоматизированного ретросинтетического движка, который содержит около 3000 ретро-реакций.

4. MCE-18 — новый молекулярный дескриптор для оценки соответствия структурным трендам медицинской химии

Для моделирования структурных трендов медицинской химии, наблюдаемых для молекулярных структур, запатентованных 23 крупнейшими фармацевтическими компаниями, собирали набор данных из соответствующих патентов с учетом хронологического порядка в отношении приоритетных дат подачи патентных заявок. Исходя из представлений о структурной сложности малых лекарственных молекул и

эволюционного восприятия картины, изменяющихся с годами структур потенциальных и зарегистрированных лекарственных веществ, предлагали формулу расчёта для нового молекулярного дескриптора MCE-18 (*Medicinal Chemistry Evolution by 2018*, эволюция медицинской химии по статусу на 2018 год):

$$\text{MCE-18} = \left(\text{AR} + \text{NAR} + \text{CHIRAL} + \text{SPIRO} + \frac{\overbrace{\text{sp}^3 + \text{Cyc} - \text{Acyc}}^{\text{NCSPTR}}}{1 + \text{sp}^3} \right) \times Q^1,$$

где AR обозначает наличие ароматического или гетероароматического кольца (0 или 1), NAR обозначает наличие алифатического или гетероалифатического кольца (0 или 1), CHIRAL обозначает наличие хирального центра (0 или 1), SPIRO обозначает наличие спиро-точки (0 или 1), sp^3 обозначает долю sp^3 -гибридизованных углеродных атомов (от 0 до 1), Cyc обозначает долю циклических углеродных атомов, которые являются sp^3 -гибридизованными (от 0 до 1), Acyc обозначает долю ациклических углеродных атомов, которые являются sp^3 -гибридизованными (от 0 до 1), и Q^1 обозначает нормализованный квадратичный индекс.

Предложенную функцию валидировали на обучающем наборе данных из патентных записей фармацевтических компаний. Было показано, что дескриптор F_{sp^3} (средневзвешенное за год значение для запатентованных в указанном году молекулярных структур) не чувствителен к очевидным структурным изменениям, в то время как аналогично найденное средневзвешенное за год значение дескриптора MCE-18 моделирует восходящий тренд, который соответствует структурному усложнению представителей химического пространства, патентованным крупными фармацевтическими компаниями (см. Рисунок 6). Значения дескрипторов даны на правой вертикальной шкале. Соответствующее нормализованное число патентов (столбчатая диаграмма, левая вертикальная шкала) иллюстрирует научно-практическую активность 23 крупнейших фармацевтических компаний. Так, установка диапазона по умолчанию 75–150 для структурного дескриптора MCE-18 позволяет направлять генеративные алгоритмы платформы Chemistry42 создавать молекулярные структуры, соответствующие трендам, которые задают крупнейшие фармацевтические компании (см. Рисунок 7).

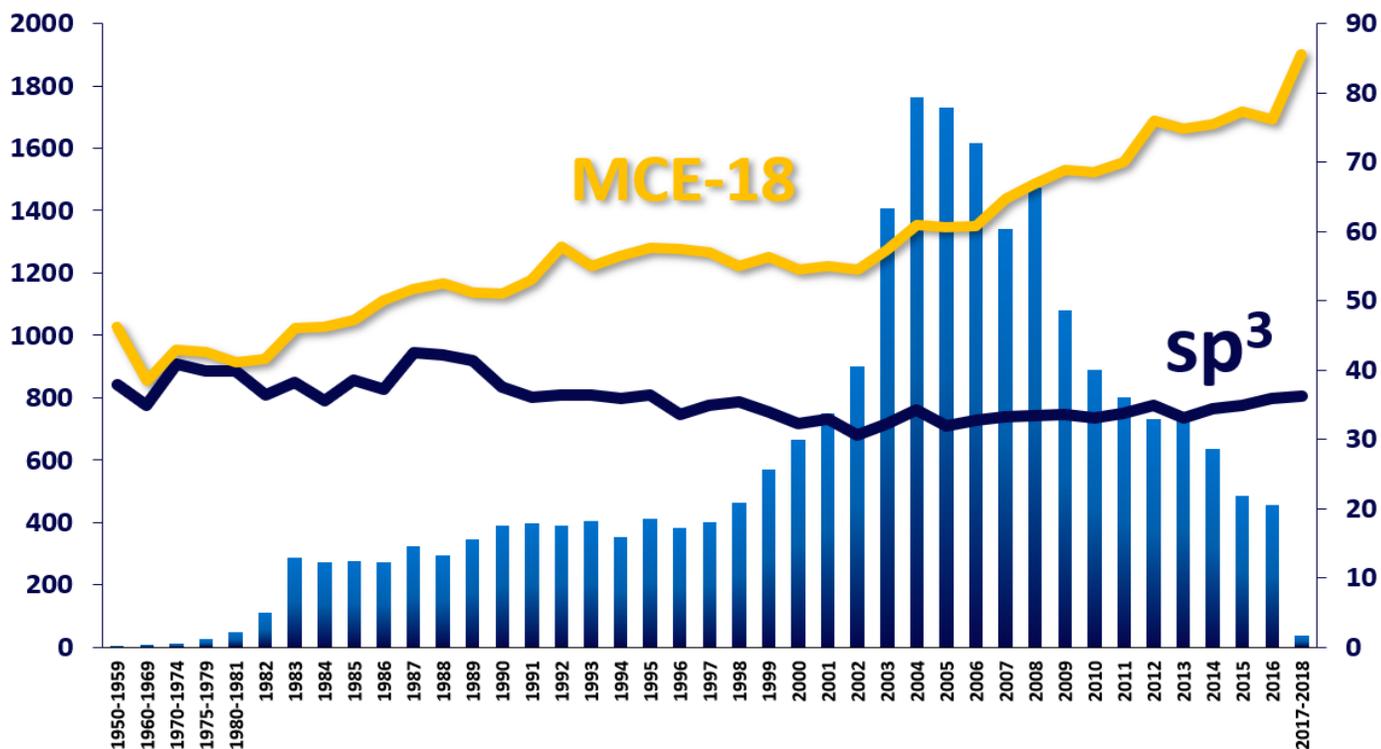


Рисунок 6. Сравнение структурных дескрипторов MCE-18 (желтая ломаная) и Fsp³ (темно-синяя ломаная) на временной шкале (горизонтальная ось).

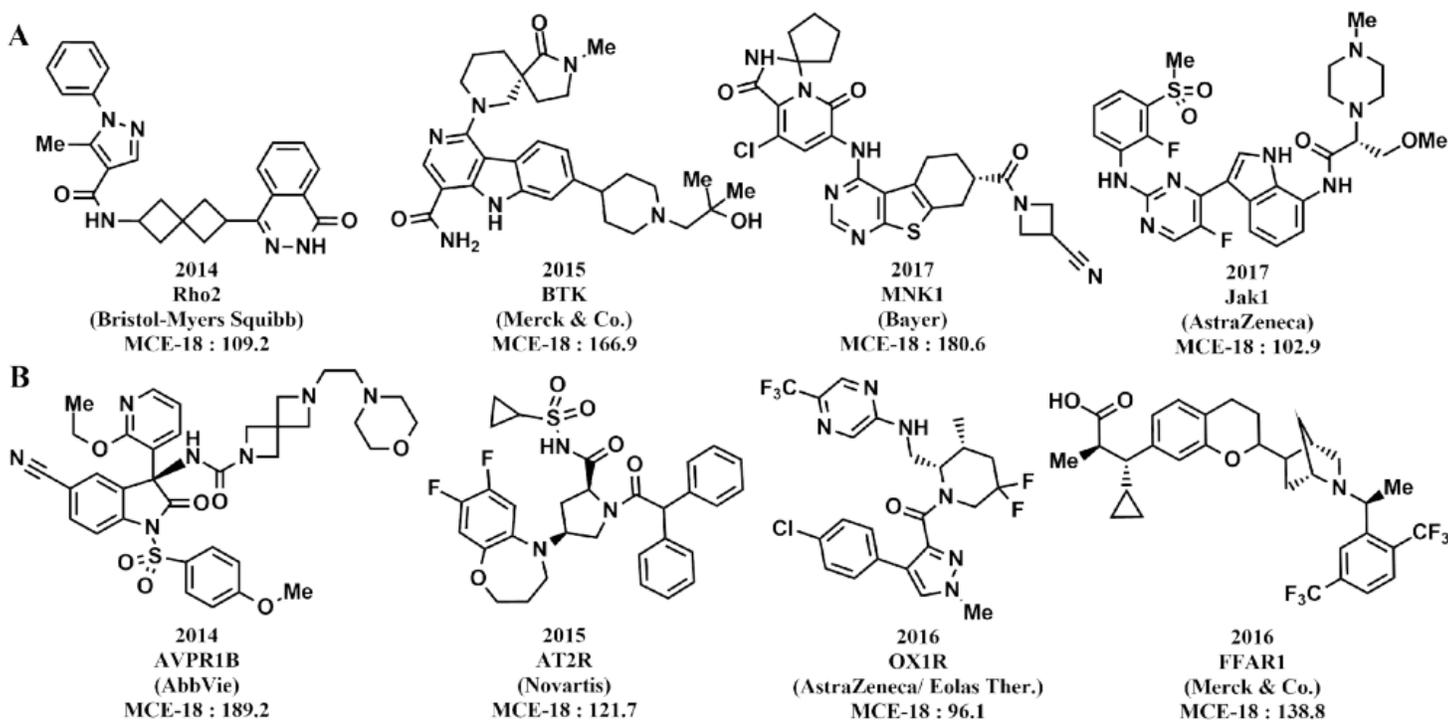


Рисунок 7. Типичные примеры структур с высокими значениями MCE-18, заявленные в патентах крупными фармацевтическими компаниями за последние 10 лет: **A**) ингибиторы протеиновых киназ, **B**) лиганды GPCR.

5. Применение платформы генеративной химии для решения практических задач медицинской химии

За последние 4 года был опубликован целый ряд статей, иллюстрирующих то, как при помощи функционала Chemistry42 могут быть проведены кампании по разработке потенциальных лекарственных веществ. В рамках этих проектов, например, разрабатывались ингибиторы циклин-зависимой киназы 8 (CDK8) (Рисунок 8а) [5], соль-индуцируемой киназы 2 (SIK2) (Рисунок 8б) [6], ДНК-полимеразы низкой точности тета (Polθ) (Рисунок 8в):

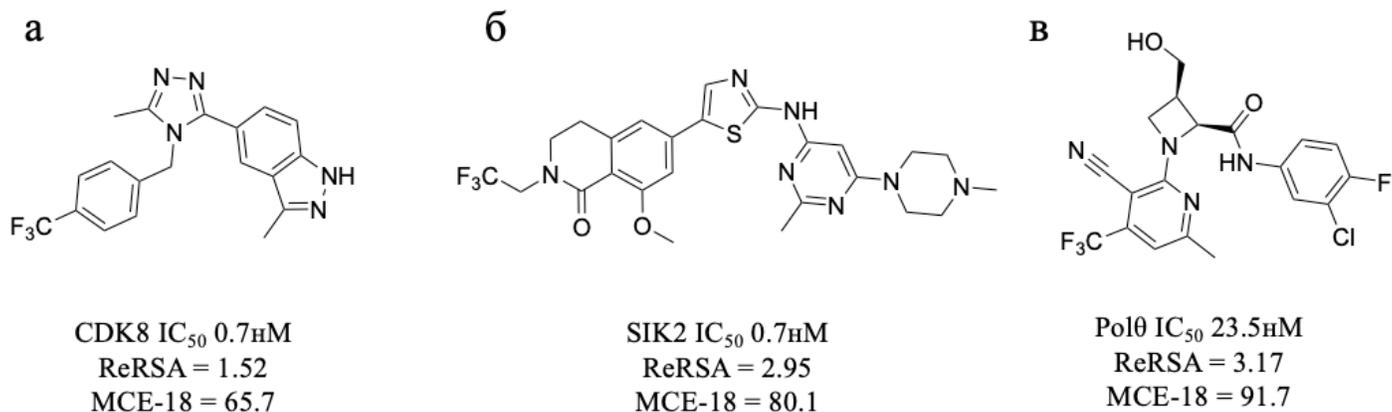


Рисунок 8. Малые лекарственные молекулы, разработанные при помощи платформы Chemistry42.

Впервые при помощи платформы было продемонстрировано, что алгоритмы генеративной химии сочетаемы с успехами в области прогнозирования фолдинга белков, у которых нет информации о трехмерных структурах, кроме как апо-структура, спрогнозированная при помощи алгоритма AlphaFold2 [7].

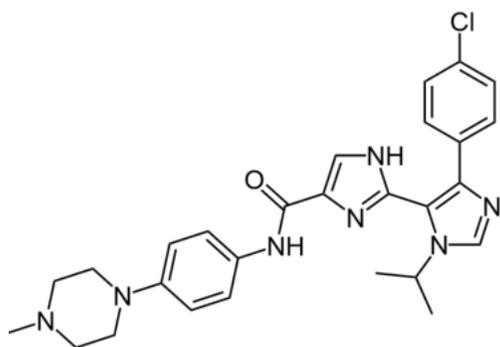
Наконец, именно используя функционал сначала прототипов Chemistry42, а затем и полноценную пользовательскую версию, удалось разработать ведущие клинические кандидаты, находящиеся в портфеле группы компаний Insilico Medicine. В их числе ингибитор TNIK киназы **INS018_055**, исследуемый в настоящий момент во II фазе клинических испытаний в качестве терапевтического агента в лечении

⁵ Li Y., Liu Y., Wu J., Liu X., Wang L., Wang J., Yu J., Qi H., Qin L., Ding X., Ren F., Zhavoronkov A. Discovery of Potent, Selective, and Orally Bioavailable Small-Molecule Inhibitors of CDK8 for the Treatment of Cancer // *Journal of Medicinal Chemistry*. — 2023. — Т. 66. — № 8. — С. 5439–5452.

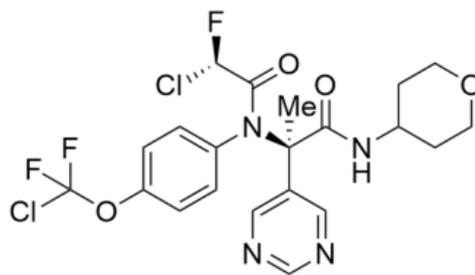
⁶ Zhu W., Liu X., Li Q., Gao F., Liu T., Chen X., Zhang M., Aliper A., Ren F., Ding X., Zhavoronkov A. Discovery of novel and selective SIK2 inhibitors by the application of AlphaFold structures and generative models // *Bioorganic & Medicinal Chemistry*. — 2023. — Т. 91. — С. 117414.

⁷ Ren F., Ding X., Zheng M., Korzinkin M., Cai X., Zhu W., Mantsyzov A., Aliper A., Aladinskiy V., Cao Z., Kong S., Long X., Man Liu B.H., Liu Y., Naumov V., Shneyderman A., Ozerov I.V., Wang J., Pun F.W., Polykovskiy D.A., Sun C., Levitt M., Aspuru-Guzik A., Zhavoronkov A. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor // *Chemical Science*. — 2023. — Т. 14. — С. 1443–1452.

идиопатического фиброза легких [8,9], а также **ISM-3312** — ингибитор главной протеазы (M^{pro}) коронавируса SARS-CoV-2 (см. Рисунок 9).



INS018_055, ингибитор TNIK
ReRSA = 2.73
MCE-18 = 74.2



ISM-3312, ингибитор M^{pro} SARS-CoV-2
ReRSA = 1.25
MCE-18 = 95.6

Рисунок 9. Клинические кандидаты Insilico Medicine, разработанные при помощи платформы генеративной химии Chemistry42.

6. Рациональный дизайн клинического кандидата **ISM-3312**, ингибитора главной протеазы SARS-CoV-2 при помощи платформы генеративной химии

Для обнаружения соединения-хита на платформе Chemistry42 был проведен виртуальный скрининг подмножества соединений библиотеки компании ChemDiv, представляющих продукты реакции Уги. В конфигурации виртуального скрининга были учтены следующие ключевые для эффективного связывания взаимодействия в виде фармакофорных точек на основе лиганда **X77** из со-кристалла PDB:6W63 (см. рис. 10A). В ходе виртуального скрининга и последующего *in vitro* скрининга виртуальных соединений-хитов была обнаружена серия из 5 активных ингибиторов M^{pro} SARS-CoV-2 с диапазоном активности IC₅₀ 0.98–8.65 мкМ, представляющих α-хлорацетамиды, получаемые по реакции Уги из α-хлоруксусной кислоты. Было высказано предположение, что серия активных α-хлорацетамидов представляет собой ковалентные ингибиторы, поскольку α-хлорацетамиды являются выраженными электрофильными агентами в реакции с остатками цистеина в активных сайтах цистеиновых протеаз. Наиболее активное соединение-хит **INSCoV-001** со значением активности 0.98 мкМ (IC₅₀) в протеазном эссе было использовано в качестве начальной

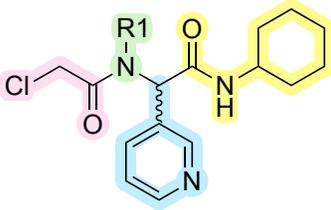
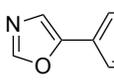
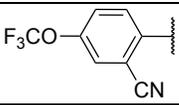
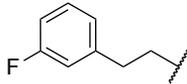
⁸ Ren F., Aliper A., Chen J., Zhao H., Rao S., Kuppe C., Ozerov I.V., Zhang M., Witte K., Kruse C., Aladinskiy V., Ivanenkov Y., Polykovskiy D., Fu Y., Babin E., Qiao J., Liang X., Mou Z., Wang H., Pun F.W., Ayuso P.T., Veviorskiy A., Song D., Liu S., Zhang B., Naumov V., Ding X., Kukharensko A., Izumchenko E., Zhavoronkov A. A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models // *Nature Biotechnology*. — 2024. — Т. 43. — С. 63–75.

⁹ Xu Z., Ren F., Wang P., Cao J., Tan C., Ma D., Zhao L., Dai J., Ding Y., Fang H., Li H., Liu H., Luo F., Meng Y., Pan P., Xiang P., Xiao Z., Rao S., Satler C., Liu S., Lv Y., Zhao H., Chen S., Cui H., Korzinkin M., Gennert D., Zhavoronkov A. A generative AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis: a randomized phase 2a trial // *Nature Medicine*. — 2025. — Т. 31. — № 8. — С. 2602–2610.

точки для дальнейшей оптимизации хемотипа. В ходе наработки и оптимизации химического пространства вокруг соединения-хита была улучшена активность в протеазном эссе и ряд *in vitro* ADME свойств для представителей хемотипа. Моделирование структур на данном этапе осуществлялось посредством генеративных экспериментов, в ходе которых использовалась логика оптимизонов. В общей сложности рассматривалось 4 оптимизона, каждый из которых соответствовал отдельному компоненту из 4-компонентной реакции Уги: **изонитрилу**, **альдегиду**, **амину** и **кислоте**.

В ходе первого генеративного цикла по расширению-оптимизации хемотипа производилось исследование **аминного** оптимизона, исходя из модели, полученной на основе докинг-позы соединения-хита **INSCoV-001**, в то время как остальная часть молекулы закреплялась при помощи якорных точек. В ходе данного генеративного цикла удалось установить, что наиболее благоприятными аминными компонентами в рамках соответствующего оптимизона с точки зрения активности в протеазном эссе (Mpro SARS-CoV-2, IC₅₀) являются 5-трифторметокси-,2-аминобензонитрил (**INSCoV-517**), 4-(оксазол-5-ил)анилин (**INSCoV-501**) и 2-(3-фторфенил)этан-1-амин (**INSCoV-558**) (см. табл. 2).

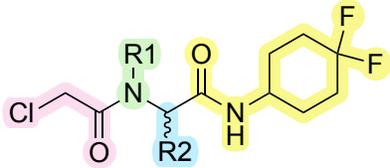
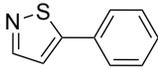
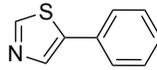
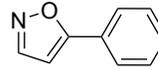
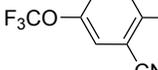
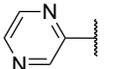
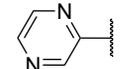
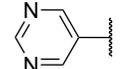
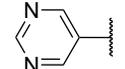
Таблица 2. Наиболее активные в протеазном эссе соединения, полученные в ходе первого оптимизационного цикла

ID	INSCoV-001	INSCoV-501	INSCoV-517	INSCoV-558
Свойства \ R1				
M ^{pro} SARS-CoV-2, IC ₅₀ (нМ)	980	680	242	399
ЧПМ/МПМ Cl _{int} (мкМ/мин/мг)	---	205.0/212.8	---	665.8/420.1
Сасо-2 проницаемость, (10 ⁻⁶ см/с)	---	0.15	---	0.23
Растворимость в воде, рН 7.4 (мкМ)	---	2.88	---	8.51
СУР Р450 1A2/2C9/2C19/2D6/3A4, IC ₅₀ (мкМ)	---	>50/7.59/5.79 /6.75/0.59	---	>50/9.22/7.50 / >50/1.49

Последующие генеративные циклы в рамках других оптимизонов позволили значительно улучшить профиль активности в протеазном эссе для хемотипа и основные *in vitro* ADME-параметры (см. табл 3):

1. Метаболическую стабильность соединения в микросомальной фракции печеночных микросом человека (ЧПМ) и мыши (МПМ), определенную в терминах клиренса;
2. Проникающую способность соединения через монослой клеток Сасо-2;
3. Растворимость соединения в воде;
4. Профиль ингибирования цитохромов Р450, ключевых для метаболизма ксенобиотиков;

Таблица 3. Избранные соединения из оптимизационной кампании и их свойства

				
ID	INSCoV-601I	INSCoV-601G	INSCoV-600J	INSCoV-517A
R1				
R2				
M ^{pro} SARS-CoV-2, IC ₅₀ (нМ)	62	81	106	86
ЧПМ/МПМ Cl _{int} (мкМ/мин/мг)	123.7/629.9	46.2/437.6	17.0/27.8	23.6/22.4
Сасо-2 проницаемость, (10 ⁻⁶ см/с)	0.04	0.04	0.002	0.31
Растворимость в воде, рН 7.4 (мкМ)	<1.6	<1.6	34.5	5.31
СУР Р450 1A2/2C9/2C19/2D6/3A4, IC ₅₀ (мкМ)	>50/11.3/9.5 /5.5/0.3	>50/20.8/4.6 /3.5/0.6	>50/>50/49.9 /9.4/43.2	>50/>50/>50 />50/>50

Во время разработки характеризовались преимущественно рацематы, однако для тех соединений, активность рацематов которых попадала в диапазон значений IC₅₀ менее 100 нМ, проводилось разделение стереоизомеров. Так, для рацемического соединения **INCoV-601I** был получен активный стереоизомер **INCoV-601I-R** со значением IC₅₀ 50 нМ и предварительно установленной абсолютной R-конфигурацией стереоцентра по результатам молекулярного моделирования на платформе Chemistry42. Отметим, что R-конфигурация присваивалась всем молекулярным структурам, поскольку таковая была более благоприятна с точки зрения моделирования связывания в активном сайте протеазы. Для подтверждения абсолютной конфигурации стереоцентра в соединении **INCoV-601I-R** и подтверждении ковалентного характера связывания оно было со-кристаллизовано с протеазой, а со-кристалл охарактеризован при помощи кристаллографического метода.

Рентгеноструктурное исследование на примере соединения **INCoV-601I-R** (активный энантиомер, выделенный из рацемата **INCoV-601I**) подтвердило ковалентное связывание описываемых α-хлорацетамидов с остатком Cys145 (см. рис 10Б); однако фактический режим связывания (PDB: 9WHE) частично отличался от смоделированного, который имитировал режим связывания соединения X77 (см. рис. 10А). Ожидаемое занятие S4-подкармана 4,4-дифторметилциклогексильным заместителем в соединении **INCoV-601I-R**, аналогичное тому, как циклогексильная группа занимает S4 в случае соединения X77, не наблюдается. Вместо этого вся область R3-оптимизона соединения **INCoV-601I-R** экспонируется в область растворителя. Амидный фрагмент R3-оптимизона соединения **INCoV-601I-R** непосредственно связывается с остатком Asn142 через водородную связь (жёлтый пунктир), тогда как у соединения X77 связывание с Asn142 происходит через водяной мостик, что позволяет ему образовывать

водородную связь с остовом Glu166. Прямое связывание с Asn142 у соединения **INCoV-601I-R** также приводит к потере водородной связи с Glu166. Также нам удалось наблюдать более выраженный, чем в случае соединения **X77**, параллельный π -стэкинг (синий пунктир) между боковой цепью His41 и биарильным фрагментом соединения **INCoV-601I-R**. Наличие второго атома азота в пиразиновом кольце R2-оптимизона соединения **INCoV-601I-R** позволило ему связываться с цепочкой из молекул воды, что не наблюдается у соединения **X77**. Наконец, ковалентно связанное соединение **INCoV-601I-R** прочно фиксируется в S1'-области сайта связывания протеазы благодаря водородным связям с остовами Gly143 и Ser144, тогда как соединение **INCoV-601I-R** связывается там только остовом остатка Gly143.

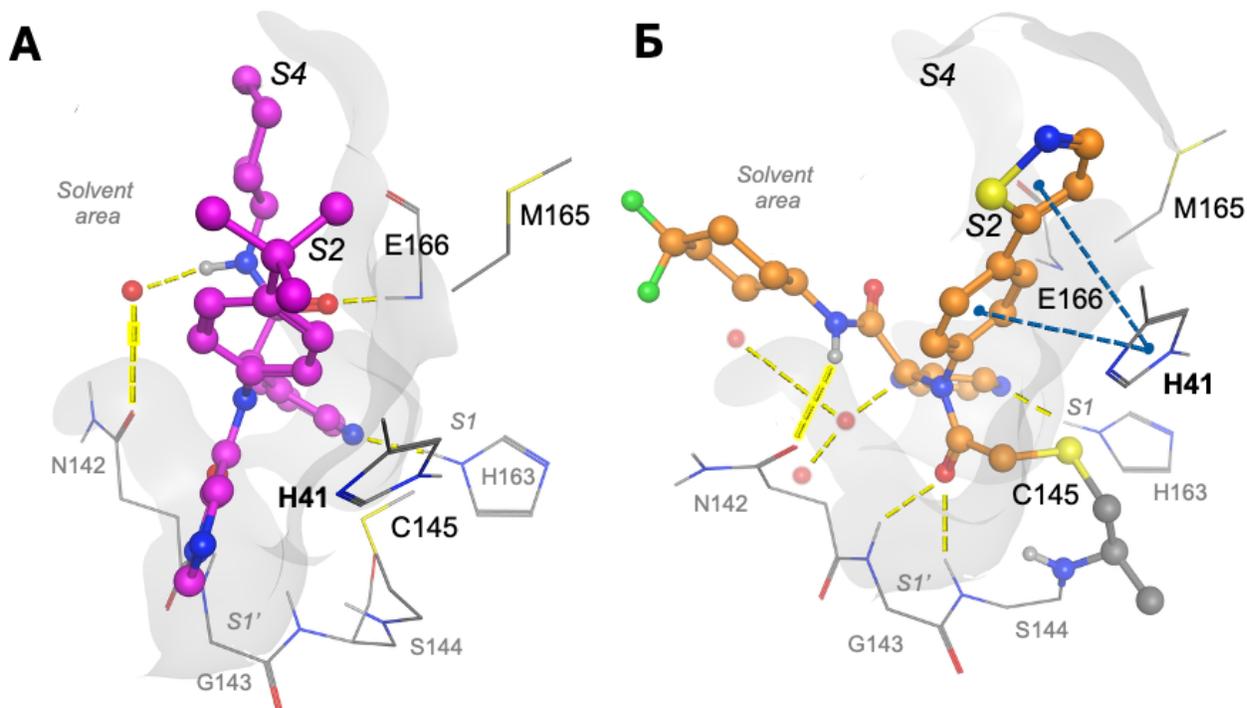


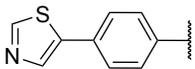
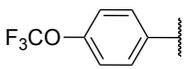
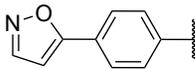
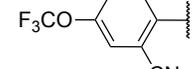
Рисунок 10. Режим связывания нековалентного ингибитора **X77** (пурпурный), PDB: 6W63 (А). Режим связывания ковалентно связанного соединения **INCoV-601I-R** (оранжевый), PDB: 9WHE (Б).

Несмотря на то, что удалось оптимизировать ряд свойств, таких как активность в протеазном эссе, метаболическая стабильность в микросомальной фракции печени и профиль ингибирования цитохромов P450, другие свойств: растворимость в воде, проникающая способность не были должным образом улучшены. Более того, низкая стабильность в плазме крови, препятствовала дальнейшему развитию серии в целом. Так, например, период полураспада соединения **INCoV-601I-R** в плазме составлял всего 4.2 мин. В целом, такая нестабильность α -хлорацетамидной серии соединений в нуклеофильных условиях плазмы крови объясняется крайне высокой электрофильностью связывающего фрагмента молекулы (СФМ).

Посредством анализа литературных источников нам удалось найти эффективную альтернативу α -хлорацетамидному СФМ в виде (R)- α -фтор, α -хлорацетамидного фрагмента (ФХА), который может быть лаконичным образом инкорпорирован в хемотип путем замены α -хлоруксусной кислоты на (R)- α -фтор, α -

хлоруксусную кислоту. В результате проблема нестабильности соединений в плазме крови была нивелирована, при том, что негативного эффекта на активность в протеазном эссе не наблюдалось. Стабильность ФХА-фрагмента в физиологических условиях объяснима с точки зрения SN2-механизма, который делает α -атом углерода СФМ стерически менее доступным для нуклеофилов в плазме крови. Более того, путем замены СФМ удалось улучшить растворимость соединений, их метаболическую стабильность и проникающую способность. Последнее, по-видимому, объясняется несколько более высокой липофильностью ФХА-фрагмента. Таким образом, заметное улучшение профиля свойств хемотипа посредством внедрения нового СФМ позволило номинировать серию соединений-лидеров, представленных в табл. 4.

Таблица 4. Избранные представители лидирующей серии ингибиторов M^{pro} SARS-CoV-2

ID	INSCoV-614-R	INSCoV-517D-R	INSCoV-614A-R	INSCoV-517C-R
R1				
M ^{pro} SARS-CoV-2, IC ₅₀ (нМ)	61±1	121±7	128±3	69±3
ЧПМ/МПМ Cl _{int} (мкМ/мин/мг)	29.6/38.9	<9.6/<9.6	<9.6/<9.6	50.3/39.6
Сасо-2 проницаемость, (10 ⁻⁶ см/с)	0.5	1.6	0.4	0.3
Растворимость в воде, рН 7.4 (мкМ)	164	156	158	55
СУР Р450 1A2/2C9/2C19/2D6/3A4, IC ₅₀ (мкМ)	24.3/15.6/9.5 /6.5/1.45	>50/36.6/40.4 /44.5/37.8	>50/>50/>50 />50/>50	0.08/24.0/13.2 /22.7/5.8
Стабильность в плазме крови человека, (t _{1/2} , мин)	>289.1	>289.1	>289.1	9
Связывание с белками плазмы крови, % (человек/мышь)	29.4/16.2	18.0/10.4	29.8/26.0	63.1/30.7
Противовирусная активность, SARS-CoV-2, Корейский вариант, клеточная линия Vero, EC ₅₀ (мкМ)	2.23	1.08	4.25	—
Цитотоксичность в клеточной линии Vero, CC ₅₀ (мкМ)	>50	>50	>50	—

Часть соединений-лидеров была охарактеризована в клеточном эссе на предмет наличия противовирусной активности. Для всех протестированных соединений (INSCoV-614-R, INSCoV-517D-R и INSCoV-614A-R) наблюдалась заметная способность подавлять SARS-CoV-2 (см. табл. 4), сопоставимая с

другими известными противовирусными агентами, такими как нирматрелвир ($EC_{50} = 4.00$ мкМ) и ремдесивир ($EC_{50} = 7.86$ мкМ), испытанных в тех же условиях. Важно отметить, что ни одно из соединений не проявляло выраженной цитотоксичности (CC_{50}) в клеточной линии Vero.

Два лучших соединения-лидера были охарактеризованы с фармакокинетической точки зрения в мышах (см. табл. 5). Пероральная биодоступность (F, %) соединения **INSCoV-517D-R** была близка к 100%, что указывало на большой потенциал этого соединения для дальнейшей оптимизации соединения-лидера. Стоит заметить, что с точки зрения структуры соединение **INSCoV-517D-R** не претерпело значительных структурных изменений на пути к клиническому кандидату **ISM-3312**.

Таблица 5. Фармакокинетический профиль избранных соединений-лидеров

Соединение	INSCoV-614-R	INSCoV-517D-R
Доза (iv/по, мг/кг)	5/20	1/20
$t_{1/2}$ [iv, ч]	0.21 ± 0.03	0.64 ± 0.15
$t_{1/2}$ [по, ч]	0.80 ± 0.09	1.07 ± 0.18
AUC_{INF} (iv, нг*ч/мл)	646 ± 120	97.8 ± 4.5
AUC_{INF} (по, нг*ч/мл)	232 ± 103	1952 ± 269
F [%]	9	99.8

ЗАКЛЮЧЕНИЕ

1. На примере модельных экспериментов продемонстрирована способность разработанной платформы генеративной химии проводить *in silico* эксперименты согласно базовым сценариям молекулярного моделирования потенциальных лекарственных веществ.

2. Платформа генеративной химии была многократно валидирована экспериментально в рамках биологических тестирований и клинических испытаний. Польза использования платформы в разработке потенциальных лекарственных подтверждена успешным прохождением полноценных доклинических испытаний для ряда малых лекарственных молекул, а также II фазы клинических исследований для ингибитора TNK киназы и вовлечением ингибитора главной протеазы коронавируса SARS-CoV-2 в I фазу клинических испытаний.

3. Для платформы генеративной химии разработан метод моделирования синтетической доступности ReRSA, сочетающий лучшие практики оценки синтезируемости молекулярных структур. *In silico* валидация алгоритма ReRSA выявила значимую корреляцию значений ReRSA с ретросинтетической сходимостью, получаемой в ходе работы автоматизированного ретросинтетического движка.

4. Разработан дескриптор MCE-18, эффективно моделирующий соответствие генерируемых платформой молекулярных структур структурным трендам медицинской химии на основе данных из патентов крупнейших фармацевтических компаний.

Основное содержание работы изложено в следующих публикациях

Научные статьи, опубликованные в журналах Scopus, Web of Science, RSCI, рекомендованных для защиты в диссертационном совете МГУ по специальностям 1.4.16. Медицинская химия и 1.4.3. Органическая химия:

1) Ivanenkov Y. A., **Zagribelnyy B.**, Aladinskiy V. Are we Opening the Door to a New Era of Medicinal Chemistry or Being Collapsed to a Chemical Singularity? // *Journal of Medicinal Chemistry* — 2019. — V. 62. — No. 22. — pp. 10026-10043, — DOI: 10.1021/acs.jmedchem.9b00004. — EDN DFNVNW. **Импакт-фактор 6.8 (JIF)**, 2.08 п.л., доля вклада **60%**

2) Zhavoronkov A., Ivanenkov Y.A., Aliper A., Veselov M.S., Aladinskiy V.A., Aladinskaya A.V., Terentiev V.A., Polykovskiy D.A., Kuznetsov M.D., Asadulaev A., Volkov Y., Zholus A., Shayakhmetov R.R., Zhebrak A., Minaeva L.I., **Zagribelnyy B.**, Lee L.H., Soll R., Madge D., Xing L., Guo T., Aspuru-Guzik A. Deep learning enables rapid identification of potent DDR1 kinase inhibitors // *Nature Biotechnology* — 2019. — V. 37. — No. 9. — pp. 1038-1040, — DOI: 10.1038/s41587-019-0224-x. — EDN YKXOEF. **Импакт-фактор 41.7 (JIF)**, 0.35 п.л., доля вклада **10%**

3) Ivanenkov Y. A., Polykovskiy D., Bezrukov D., **Zagribelnyy B.**, Aladinskiy V., Kamya P., Aliper A., Ren F., Zhavoronkov A. Chemistry42: An AI-Driven Platform for Molecular Design and Optimization // *Journal of Chemical Information and Modeling* — 2023. — V. 63. — No. 3. — pp. 695-701, — DOI 10.1021/acs.jcim.2c01191. — EDN DOBCMI. **Импакт-фактор 5.3 (JIF)**, 0.81 п.л., доля вклада **50%**

4) Ivanenkov Y., **Zagribelnyy B.**, Malyshev A., Evteev S., Terentiev V., Kamya P., Bezrukov D., Aliper A., Ren F., Zhavoronkov A. The Hitchhiker's Guide to Deep Learning Driven Generative Chemistry // *ACS Medicinal Chemistry Letters* — 2023. — V. 14. — No. 7. — pp. 901-915, — DOI 10.1021/acsmchemlett.3c00041. — EDN VWDMBT. **Импакт-фактор 4.0 (JIF)**, 1.73 п.л., доля вклада **50%**

5) Бондарев Н., **Загрибельный Б.**, Федорченко С.А., Иваненков Я.А., Палюлин В.А. Моделирование синтетической доступности потенциальных лекарственных веществ, содержащих пятичленные ароматические гетероциклы // *Известия Академии наук. Серия химическая*. — 2025. — Т. 74. — №. 6. — С. 1687-1703, — EDN MOBTKX. **Импакт-фактор 1.035 (РИНЦ)**, 0.98 п.л., доля вклада **60%**

(Переводная версия: Bondarev N., **Zagribelnyy B.**, Fedorchenko S.A., Ivanenkov Y.A., Paliulin V.A. Modeling of synthetic accessibility of potential drug molecules containing five-membered aromatic heterocycles // *Russian Chemical Bulletin*. — 2024. — Vol. 74. — pp. 1687-1703. — DOI 10.1007/s11172-024-4652-5. — EDN JWIXSA. **Импакт-фактор 0.305 (SJR)**, 0.98 п.л., доля вклада **60%**)

6) Sun J., Sun D., Yang Q., Wang D., Peng J., Guo H., Ding X., Chen Zh., Yuan B., Ivanenkov Y.A., Yuan J., **Zagribelnyy B.**, He Y., Su J., Wang L., Tang J., Li Zh., Li R., Li T., Hu X., Liang X., Zhu A., Wei P., Fan Y., Liu S., Zheng J., Guan X., Aliper A., Yang M., Bezrukov D.S., Xie Zh., Terentiev V.A., Peng G., Polykovskiy D.A., Malyshev A.S., Malkov M.N., Zhu Q., Aspuru-Guzik A., Ding X., Cai X., Zhang Man, Zhao J., Zhong N., Ren F., Chen X., Zhavoronkov A., Zhao J. A novel, covalent broad-spectrum inhibitor targeting human coronavirus Mpro // *Nature Communications* — 2025. — V. 16. — P. 4546, – DOI 10.1038/s41467-025-59870-4. – EDN DTDLMX. **Импакт-фактор 15.7 (JIF)**, 1.15 п.л., **доля вклада 25%**

7) Патент № US20220172802A1. Retrosynthesis systems and methods: опубл. 02.06.2022 / Konstantinov A., Putin E.O., **Zagribelnyy B.**, Ivanenkov Y.A., Zhavoronkovs A., 1.96 п.л., **доля вклада 30%**

8) Патент № US20230154572A1. Retrosynthesis Related Synthetic Accessibility: опубл. 18.05.2023 / **Zagribelnyy B.**, Putin E.O., Fedorchenko S.A., Ivanenkov Y.A., Zhavoronkovs A., 1.5 п.л., **доля вклада 50%**

9) Патент № US20230174488A1. Sars-cov-2 inhibitors having covalent modifications for treating coronavirus infections: опубл. 08.06.2023 / Zhavoronkovs A., Ivanenkov Y.A., **Zagribelnyy B.**, 9.03 п.л., **доля вклада 40%**