

УДК 51-76:577.322:539.19:004.94

## Технологии молекулярного моделирования и искусственного интеллекта для разработки потенциальных лекарственных препаратов нового поколения, исследования в области биоинформатики и вычислительной биологии

**Андрианов Александр Михайлович**

*главный научный сотрудник ИБОХ НАН Беларуси,  
доктор химических наук, профессор  
E-mail: alexande.andriano@yandex.ru*

**Тузиков Александр Васильевич**

*заведующий лабораторией математической кибернетики,  
член-корреспондент, доктор физико-математических наук, профессор  
E-mail: tuzikov@newman.bas-net.by*

**Введение.** В настоящее время компьютерный дизайн лекарств является эффективным инструментом в фармацевтических технологиях, позволяющим значительно сократить время и затраты, необходимые для разработки новых терапевтических средств [1, 2]. Важную роль в компьютерном скрининге лекарств играет молекулярный докинг, который широко используется для предсказания пространственной структуры комплекса белок-лиганд и наиболее выгодной ориентации лиганда в активном центре целевого белка, оценки энергии связывания и исследования профиля взаимодействия молекул – кандидатов в лекарственные средства с терапевтической мишенью [3]. Последние разработки полуэмпирических квантово-химических методов и методов теории функционала плотности, а также применение *ab initio* расчетов к дизайну молекул-кандидатов в контексте идентификации и оптимизации их структур показывают растущую важность квантовой химии в фармакологических исследованиях [4–6]. Молекулярная динамика (МД) также является мощным методом для скрининга потенциальных лекарств на основе структуры целевого белка [7, 8]. В отличие от молекулярного докинга МД моделирует перемещения каждого атома в силовом поле остальных атомов при явном задании растворителя и более эффективно, чем другие алгоритмы, отражает гибкость как лиганда, так и белка, что позволяет получать более корректные оценки свободной энергии связывания. Примеры успешных применений методов молекулярного докинга, квантовой химии и МД наглядно демонстрируют возможности вычислительных подходов для идентификации соединений с желаемыми свойствами и создания новых лекарств [3–8].

В последние годы при разработке лекарств все чаще используются методы машинного обучения, и в частности глубокого обучения, которые необходимы на каждом этапе этого сложного многоэтапного процесса не только для ускорения исследований, но и для оценки рисков и затрат в клинических исследованиях [9, 10]. Методы машинного обучения могут успешно применяться для решения задач перепрофилирования лекарств [11, 12], прогнозирования структуры белков [13], виртуального скрининга потенциальных лекарств и прогнозирования аффинности связывания белок-лиганд [14, 15]. В недавних работах методы глубокого обучения использовались для скрининга химических баз данных с целью выявления антибактериальных и противовирусных ингибиторов, в том числе терапевтических агентов, потенциально активных против ВИЧ-1 и SARS-CoV-2 [16–19]. Применение этих методов в сочетании с виртуальным скринингом молекулярных библиотек, содержащих одобренные Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA, США) препараты, при-

вело к открытию ряда потенциальных препаратов против SARS-CoV-2 [19]. Кроме идентификации потенциальных лекарств в химических базах данных и прогнозирования их свойств, машинное обучение используется также для создания новых соединений с заданными фармакологическими характеристиками [20–25]. В последние годы глубокие генеративные модели нашли широкое применение в исследованиях по разработке лекарств *de novo* [20–25]. Благодаря стремительному развитию методов глубокого обучения в настоящее время разработаны генеративные модели с различной архитектурой и разными алгоритмами обучения, использующие разные типы и структуры данных. Применение глубоких генеративных моделей уже показало их способность генерировать молекулы, которые могут быть синтезированы, активны *in vitro*, стабильны и проявляют активность на моделях *in vivo*, связанных с различными заболеваниями [19]. В частности, с помощью условно-сопоставительного автоэнкодера был разработан ингибитор Янус-киназы 3, представляющий новый класс иммуномодулирующих средств [23]. С применением генеративного обучения с подкреплением были разработаны активные *in vivo* ингибиторы рецепторов дискоидинового домена 1 и 2 (DDR1 и DDR2), а фармакокинетический профиль DDR1 был подтвержден экспериментами на мышах *in vivo* [25]. Однако несмотря на все большее распространение глубоких генеративных моделей в био- и хемоинформатике, их потенциал в данной области использован еще не полностью. В связи с этим разработка и применение глубоких генеративных моделей для компьютерного дизайна лекарств имеют большое научное и практическое значение.

В настоящей работе приведен обзор основных результатов, полученных группой компьютерного конструирования лекарств ОИПИ НАН Беларуси и ИБОХ НАН Беларуси в 2012–2023 гг. с использованием методов молекулярного моделирования и глубокого обучения и посвященных дизайну и идентификации малых молекул, которые перспективны для разработки противовирусных, антибактериальных и противоопухолевых препаратов нового поколения. В заключительной части обзора вкратце представлены результаты работ лаборатории математической кибернетики ОИПИ НАН Беларуси в области биоинформатики и вычислительной биологии, включающие полногеномный поиск ассоциаций, анализ структурных изменений белков при взаимодействиях и моделирование димерных белковых комплексов. В выполнении этих исследований на различных этапах участвовали А. М. Андрианов, И. В. Анищенко, А. В. Батяновский, И. П. Босько, Т. Д. Войтко, Д. А. Воробьев, А. В. Гончар, А. Д. Карпенко, И. А. Кашин, Т. В. Кирис, Ю. В. Кисель, Ю. В. Корноушенко, И. В. Лайков, Г. И. Николаев, Р. С. Сергеев, М. В. Спринджук, А. В. Тузиков, К. В. Фурс, Н. А. Шульдов, А. Ю. Хадарович, А. М. Юшкевич.

#### **Компьютерный дизайн потенциальных лекарств**

***In silico* дизайн и идентификация потенциальных ингибиторов ВИЧ-1.** В последние годы для терапии ВИЧ-1 в клинической практике используются более 25 лекарственных препаратов, которые в зависимости от механизма действия делятся на классы, включающие ингибиторы обратной транскриптазы, протеазы, интегразы и ингибиторы проникновения. Однако высокая генетическая изменчивость ВИЧ-1 приводит к выработке устойчивости к определенному препарату через некоторое время после начала его применения. С 1995 г. для лечения ВИЧ-инфекции широко используется метод высокоактивной антиретровирусной терапии, основной целью которого является преодоление резистентности вируса к отдельным антиретровирусным препаратам на основе комбинации высокоактивных лекарственных средств, обладающих различными механизмами действия. Однако необходимость пожизненного непрерывного применения нескольких терапевтических препаратов и связанные с этим токсичность и побочные эффекты требуют разработки новых, более эффективных и безопасных анти-ВИЧ-агентов. Большинство из применяемых в клинике препаратов нацелены на вирусные ферменты: обратную транскриптазу и протеазу, но они не могут предотвращать проникновение вируса в клетку-мишень, что повышает внимание к ингибиторам ВИЧ-1, которые способны вмешиваться в ранние стадии жизненного цикла вируса. В связи с этим чрезвычайно актуальным является поиск химических соединений – ингибиторов проникновения ВИЧ-1, блокирующих адсорбцию вируса на поверхности клетки хозяина и последующее слияние мембран.

Для решения данной проблемы в работах [26–28] были использованы методы компьютерного конструирования лекарств, включающие молекулярный дизайн, высокопроизводительный докинг, МД и вычислительные процедуры, предназначенные для оценки сродства лигандов к терапевтической мишени. Согласно данным литературы к соединениям, эффективно взаимодействующим с ВИЧ-1 на ранних стадиях проникновения вируса в клетку-мишень, относится гликофинголипид  $\beta$ -галактозилцерамид ( $\beta$ -GalCer). Однако применение этого гликолипида в качестве противовирусного препарата невозможно из-за его нерастворимости в воде. Тем не менее аналоги  $\beta$ -GalCer, потенциально способные образовывать устойчивые коллоидные системы выше критической концентрации мицеллообразования («водорастворимые» аналоги), могут быть использованы как перспективные базовые структуры для создания новых ингибиторов проникновения ВИЧ-1. Одно из таких соединений – деацилированный аналог  $\beta$ -GalCer  $\beta$ -галактозилфингозин (рис. 1) – авторы сконструировали и проанализировали методами молекулярного моделирования, а затем протестировали на анти-ВИЧ-активность в РНПЦ эпидемиологии и микробиологии [29, 30]. Биомедицинские испытания показали наличие ингибирующих свойств у этого соединения: химиотерапевтический индекс  $\beta$ -галактозилфингозина (отношение максимально переносимой концентрации к минимально активной концентрации соединения) составил 40,0 [30], что свидетельствует о высокой степени его анти-ВИЧ-активности. В дальнейших исследованиях были разработаны 12 производных этой молекулы, нейтрализующее действие которых оценено в компьютерных экспериментах [31]. В результате *in silico* оценки свойств этих соединений предсказано, что сконструированные гликолипиды связываются с функционально важным инвариантным участком третьего домена белка gp120 различных модификаций вируса и обладают высоким потенциалом ингибиторной активности [31–34].

Методы компьютерного моделирования были использованы еще в одном новом направлении поиска лекарственных препаратов против ВИЧ-1, основанном на применении анти-ВИЧ-антител, обладающих нейтрализующей активностью против широкого спектра вариантов вируса [35]. Несмотря на интенсивные исследования, многочисленные попытки разработать иммуноген, способный индуцировать антитела к ВИЧ-1 с широкой вирусной нейтрализацией, до сих пор не увенчались успехом [35]. Поэтому представлялся актуальным поиск низкомолекулярных соединений, способных имитировать структурно-функциональные свойства таких антител. С этой целью с помощью методов фармакофорного моделирования, высокопроизводительного докинга и МД были идентифицированы небольшие молекулы – пептидомиметики антител 3074, VRC01, 10E8, обладающих широким спектром действия против ВИЧ-1, а также малые молекулы, имитирующие первичный рецептор CD4, с помощью которого вирус адсорбируется на поверхности мембраны клетки хозяина [36–40]. Полученные результаты свидетельствуют о высокой потенциальной нейтрализующей активности идентифицированных соединений. Согласно данным компьютерного моделирования обнаруженные пептидомиметики антител VRC01, 3074, 10E8 и клеточного рецептора CD4 могут проявлять высокую аффинность связывания с функционально консервативными участками ВИЧ-1, имитируя критические взаимодействия вируса с рецепторами клеточной мембраны [36–40]. В связи с этим высказано предположение о том, что «коктейль» из низкомолекулярных соединений – пептидомиметиков этих антител, блокирующих ключевые стадии процесса проникновения ВИЧ-1 в клетку-мишень, способен подавить репликацию вируса и существенно снизить уровень вирусной нагрузки в плазме крови [40].

В статьях [41, 42] разработан генеративный состязательный автоэнкодер для рационального дизайна потенциальных ингибиторов проникновения ВИЧ-1, способных блокировать участок белка gp120 оболочки вируса, критический для его связывания с клеточным рецептором CD4. Разработанная модель состоит из двух нейросетей – автоэнкодера и дискриминатора, работающих во время обучения в соревновательном режиме (рис. 2). Автоэнкодер представляет собой семислойную нейронную сеть, имеющую входной и выходной слои, латентный слой, а также четыре полносвязных слоя. На входной слой подаются молекулярные дескрипторы химических соединений, данные о которых проходят энкодер и попадают на латентный слой, где к полученному результату добавляется численная оценка энергии связывания с молекулярной мише-

нию. Далее молекулярные дескрипторы проходят декодер и попадают на выход, который представляет собой вектор молекулярного дескриптора. Латентный слой состоит из трех нейронов, два из которых получают значения от энкодера, а третий получает значение энергии связывания с молекулярной мишенью. В рабочем режиме автоэнкодера на латентный слой подаются случайные числа и пороговое значение энергии связывания, которые затем проходят через декодер, генерирующий молекулярные дескрипторы молекул с требуемыми свойствами.

Использование разработанной нейронной сети совместно с виртуальным скринингом молекулярной библиотеки ZINC15, проведенным на основе подбора молекулярных дескрипторов, и методами молекулярного моделирования позволило идентифицировать три низкомолекулярных химических соединения, проявляющих высокое сродство к белку gp120 ВИЧ-1 [42]. Согласно данным молекулярного докинга, машинного обучения, квантово-химических расчетов и МД эти соединения характеризуются низкими значениями свободной энергии связывания в комплексах с белком gp120, близкими к значениям, рассчитанным с помощью идентичных вычислительных протоколов для ингибиторов ВИЧ-1 NBD-11021 и NBD-14010, являющихся мощными антагонистами клеточного рецептора CD4 [42]. Полученные результаты показали, что разработанная нейронная сеть представляет собой эффективную математическую модель для виртуального скрининга баз данных химических соединений, направленного на поиск малых молекул с высоким сродством к белку gp120 и разработку на их основе новых анти-ВИЧ-препаратов широкого спектра действия [41, 42].

В работах [43, 44] предложен комбинированный алгоритм компьютерного конструирования лекарственных препаратов против ВИЧ-1, состоящий из базовых блоков – шагов, включающих:

- 1) формирование базы данных химических соединений, предназначенных для использования в качестве исходных структурных блоков в реакции азид-алкинового циклоприсоединения – основной реакции клик-химии, которая позволяет получать гибридные молекулы, способные обеспечить их связывание с белком gp120 ВИЧ-1 и его нейтрализацию (библиотеки 1 и 2). Формирование базы данных выполняется из набора химических соединений, депонированных в молекулярной библиотеке ZINC15 (<http://zinc.docking.org>), на основе соответствия их физико-химических свойств заданным фильтром параметров;

- 2) компьютерный дизайн молекул-кандидатов с привлечением в качестве исходных структурных блоков соединений, определенных на предыдущем шаге, на основе методов комбинаторики и математического подбора функциональных групп, обладающих потенциалом вступать в реакцию азид-алкинового циклоприсоединения;

- 3) отбор молекул, удовлетворяющих правилу пяти Липинского, накладываемому на молекулу, взаимодействующую с заданной молекулярной мишенью, условия подобия лекарству [45]. На данном шаге практически осуществляется скрининг соединений, полученных на шаге 2; результатом является уменьшенный набор соединений, согласующихся с указанными в правиле пяти Липинского ограничениями;

- 4) построение структурных комплексов сконструированных молекул с белком gp120 ВИЧ-1 с помощью методов молекулярного докинга. На данном шаге задается мишень для отобранных соединений и на основе анализа их геометрического сходства и оценки энергии формируется весовая функция для идентификации потенциальных лекарств. Результатом шага 4 является набор молекул-кандидатов, отобранных по оценочным функциям;

- 5) построение молекулярно-динамической траектории комплексов отобранных соединений с белком gp120 ВИЧ-1 и расчет свободной энергии их образования. На данном шаге выполняется математическое моделирование возможных изменений комплексов во времени для отобранных по результатам докинга соединений. Шаг 5 служит дополнительным к молекулярному докингу фильтром, позволяющим обнаруживать молекулы, формирующие устойчивые комплексы с белковой мишенью.

Таким образом, алгоритм представляет собой двойной каскад фильтров, позволяющий сначала сконструировать из простых соединений гибридные молекулы для формирования возможных комплексов с белком gp120 ВИЧ, а затем отобрать среди них лучшие для связывания и нейтрализации ВИЧ-1. Особенностью алгоритма является метод формирования базы данных

химических соединений на основе молекулярных моделей с потенциальной биологической активностью и методологии клик-химии, которая унифицирует некоторые химические реакции и позволяет получать их результаты прямо на компьютере. Такой подход позволил сформировать с помощью методов компьютерного моделирования базу перспективных соединений – потенциальных ингибиторов ВИЧ, содержащую более миллиона молекул. Реализация данного алгоритма позволила идентифицировать пять химических соединений, перспективных для дальнейших исследований биологами, химиками и медиками на стадиях доклинического и клинического тестирования [44].

В работах [46–49] представлена комплексная методика предсказания потенциальных противовирусных средств, ориентированная на поиск пептидомиметиков анти-ВИЧ-антител и включающая большие вычислительные блоки, которые выполняются в разных программных комплексах:

– блок построения фармакофорной модели антитела с помощью программно-алгоритмического обеспечения веб-сервера Pharmit на основе анализа его комплекса с молекулярной мишенью, полученного, как правило, экспериментальным путем, и последующего виртуального скрининга баз данных химических соединений с использованием заданного набора пространственных и электронных признаков, необходимых для обеспечения оптимальных межмолекулярных взаимодействий;

– блок отбора перспективных химических соединений и построения комплексов лиганд-мишень методами молекулярного докинга (программный пакет AutoDock Vina);

– блок квантово-химического уточнения построенных комплексов (MOPAC2016);

– блок молекулярной динамики (AMBER18).

Особенностью предложенной методики является использование на одном из ее этапов методов квантовой химии, которые до недавнего времени не применялись для моделирования структур белков и их комплексов с лигандами из-за чрезвычайной сложности расчетов, требующих огромных затрат временных и компьютерных ресурсов. Реализация этого блока выполнена с помощью полуэмпирического квантово-химического метода PM7, позволяющего проводить уточнение и оценку результатов молекулярного докинга, что дало возможность существенно повысить качество отбора потенциальных кандидатов в противовирусные средства. В настоящее время разработаны различные программные пакеты, реализующие этапы, которые входят в методику и доступны пользователям в виде клиентских приложений, свободных кодов или коммерческих продуктов. При этом пользователь должен обеспечить согласование форматов данных при переходе от одного этапа моделирования к другому.

Эффективность методики подтверждена в процессе поиска потенциальных пептидомиметиков нейтрализующего ВИЧ-1 антитела N6. В результате поиска были обнаружены девять малых молекул, которые связываются с белком gp120 сильнее, чем лучшие из известных в настоящее время ингибиторов проникновения вируса [46–49]. Разработанная методика универсальна и может быть использована для идентификации малых молекул, перспективных для разработки новых эффективных лекарственных препаратов против других болезней.

**Идентификация потенциальных ингибиторов SARS-CoV-2 методами виртуального скрининга и глубокого обучения.** Вспышка коронавирусной инфекции в Китае, зарегистрированная в конце 2019 г. и вызванная вирусом SARS-CoV-2 (этиологическим агентом COVID-19), стала причиной серьезной обеспокоенности мирового сообщества, так как число инфицированных людей постоянно увеличивалось со значительным географическим распространением. В связи с этим были предприняты многочисленные попытки разработать эффективную противовирусную вакцину и найти новые терапевтические средства против COVID-19. Определение в 2020 г. пространственной структуры основной протеазы ( $M^{Pro}$ ) SARS-CoV-2 – фермента, играющего чрезвычайно важную роль в жизненном цикле вируса, – создало предпосылки не только для понимания его функции и механизма действия, но и для разработки новых эффективных ингибиторов на основе прямых методов компьютерного конструирования лекарств, использующих данные о структуре молекулярной мишени. В частности, 25 марта 2020 г. в Банк данных белков была депонирована структура  $M^{Pro}$  SARS-CoV-2 в комплексе с высокоаффин-

ным лигандом X77 (Банк данных белков, код 6w63, <https://doi.org/10.2210/pdb6W63/pdb>, <http://www.rcsb.org/structure/6W63>), представляющим собой мощный нековалентный ингибитор как SARS-CoV, MERS-CoV, так и SARS-CoV-2.

В работах [50, 51] осуществлен виртуальный скрининг низкомолекулярных химических соединений, имитирующих фармакофорные свойства ингибитора X77, выполнена *in silico* оценка их потенциальной анти-SARS-CoV-2-активности и идентифицированы молекулы, перспективные для создания новых эффективных препаратов для терапии COVID-19.

Проведенные исследования включали следующие этапы (рис. 3):

1) построение модели фармакофора, описывающей совокупность структурно-функциональных свойств ингибитора X77, которые обеспечивают специфичность его взаимодействий с активным центром M<sup>Pro</sup> SARS-CoV-2;

2) фармакофорный анализ молекулярных библиотек веб-сервера Pharmit (<http://pharmit.csb.pitt.edu>), позволяющего проводить интерактивное исследование химического пространства с целью поиска потенциальных лекарств на основе сходства фармакофорных моделей с высокоаффинными лигандами белка-мишени;

3) отбор соединений, удовлетворяющих правилу пяти Липинского;

4) молекулярный докинг отобранных соединений с M<sup>Pro</sup> SARS-CoV-2;

5) расчет величин свободной энергии связывания с последующей идентификацией молекул, перспективных для разработки эффективных противовирусных препаратов.

Виртуальный скрининг осуществляли в девяти молекулярных библиотеках веб-сервера Pharmit, содержащих информацию о химических структурах более 213,5 миллионов молекул (<http://pharmit.csb.pitt.edu>). В результате был идентифицирован набор соединений, удовлетворяющих заданной модели фармакофора и правилу пяти Липинского, и с помощью методов молекулярного моделирования проведена оценка потенциала их ингибиторной активности против M<sup>Pro</sup>, в результате которой обнаружены пять соединений-лидеров, перспективных для синтеза и тестирования на противовирусную активность [51].

К сожалению, у авторов не было возможности провести синтез и тестирование обнаруженных соединений на анти-SARS-CoV-2-активность за относительно небольшой промежуток времени. В связи с этим было принято решение использовать стратегию перепрофилирования лекарств, под которой понимается процесс определения новых терапевтических показаний для зарегистрированных ранее и доступных препаратов. Так как профили безопасности этих препаратов хорошо задокументированы, виртуальный скрининг молекулярных библиотек лекарственных веществ может существенно облегчить и ускорить процесс обнаружения молекул с клиническим потенциалом с целью их перепрофилирования для лечения пациентов, инфицированных коронавирусом нового типа. Для реализации этого подхода была сформирована библиотека биологически активных молекул, состоящая из 28 860 молекул и включающая одобренные для применения в клинике препараты и лекарственные вещества, находящиеся на различных стадиях доклинических и клинических испытаний, и методами молекулярного моделирования проведен ее виртуальный скрининг, направленный на идентификацию потенциальных ингибиторов M<sup>Pro</sup> SARS-CoV-2 [52].

Выполненные расчеты позволили идентифицировать шесть молекул, проявляющих высокое сродство к активному центру M<sup>Pro</sup> SARS-CoV-2 [52]. Об этом свидетельствуют низкие значения свободной энергии образования комплексов этих соединений с ферментом, сопоставимые с величиной, предсказанной для мощного ингибитора SARS-CoV-2 X77 с использованием идентичного вычислительного протокола. На основе полученных данных был сделан вывод о том, что найденные соединения обладают хорошим терапевтическим потенциалом для ингибирования каталитической активности фермента и формируют перспективные базовые структуры для разработки новых эффективных препаратов против COVID-19 [52]. В настоящее время совместно с партнерами из Шанхайского института лекарственных соединений Китайской академии наук выполняются исследования по тестированию идентифицированных соединений на моделях *in vitro* и оптимизации их структур.

В работах [53–55] с использованием методов глубокого обучения разработаны две генеративные модели автоэнкодера для создания новых лекарственных веществ, способных ингибировать каталитическую активность  $M^{Pro}$  SARS-CoV-2. Построена архитектура нейронной сети (рис. 4), сформирована виртуальная библиотека потенциальных ингибиторов фермента и выполнен виртуальный скрининг соединений из этой библиотеки с последующими расчетами значений свободной энергии связывания. Обучение нейронной сети и ее тестирование показали, что созданные модели автоэнкодера позволяют генерировать малые молекулы с высокой противовирусной активностью и приемлемыми фармакологическими свойствами [53]. С помощью разработанной нейронной сети осуществлен *de novo* дизайн 95 775 потенциальных ингибиторов  $M^{Pro}$  SARS-CoV-2 с последующей *in silico* оценкой их ингибиторной активности [54, 55]. В результате проведенных исследований отобраны семь соединений-лидеров, которые характеризуются низкими значениями свободной энергии связывания, близкими к величинам, полученным с помощью идентичного вычислительного протокола для двух мощных нековалентных ингибиторов  $M^{Pro}$ , использованных в расчетах в качестве позитивного контроля [55]. Полученные результаты свидетельствуют о перспективности применения идентифицированных соединений в работах по созданию новых препаратов против COVID-19, терапевтическое действие которых основано на ингибировании ферментативной активности  $M^{Pro}$  SARS-CoV-2 [55].

Технология перепрофилирования лекарств была также использована еще в одном исследовании, в котором с помощью комплексного вычислительного подхода проведен виртуальный скрининг сформированной ранее молекулярной библиотеки [52], направленный на идентификацию потенциальных ингибиторов домена HR1 белка S SARS-CoV-2 – консервативного участка вируса, критически важного для слияния мембран и его проникновения в клетку-мишень [56, 57]. Методами молекулярного моделирования идентифицированы 12 молекул, характеризующихся согласно расчетным данным высоким сродством к этому функционально значимому домену оболочки вируса [56]. В результате биомедицинского тестирования девяти из двенадцати идентифицированных молекул, проведенного в Университете Фудань (Шанхай, Китай), обнаружено соединение-лидер – противоопухолевый препарат Навитоклак (рис. 5) [57]. Показано, что Навитоклак проявляет высокую противовирусную активность по отношению к различным штаммам SARS-CoV-2, в том числе практически ко всем субвариантам штаммов Дельта и Омикрон, со значениями константы ингибирования  $IC_{50}$ , варьирующими в диапазоне от 0,5 до 3,7 мкМ, а также к родственным коронавирусам SARS-CoV и MERS-CoV [57]. На моделях *in vitro* показано, что Навитоклак селективно связывается с терапевтической мишенью, блокируя проникновение вируса в клетки хозяина [57]. Полученные результаты свидетельствуют о том, что Навитоклак формирует многообещающую основу для создания эффективного и безопасного перорального препарата широкого спектра действия против SARS-CoV-2 и других известных коронавирусов человека, а также их новых вариантов, которые могут появиться в ближайшем будущем.

***In silico* скрининг потенциальных ингибиторов микобактерии туберкулеза.** В октябре 2018 г. Организация Объединенных Наций провела беспрецедентное совещание высокого уровня с целью привлечь внимание к эпидемии туберкулеза (ТБ) и разработать стратегию, направленную на его ликвидацию к 2030 г. (A/RES/73/3. 2018; <https://www.stoptb.org/sites/default/files/UN%2520Declaration%2520on%2520TB.pdf>). Однако пандемия COVID-19 существенно снизила прогресс, достигнутый в борьбе с ТБ во всем мире, а уровень смертности от этого инфекционного заболевания увеличился впервые за более чем 10 лет. На сегодняшний день ТБ продолжает оставаться одной из 10 основных причин смертности во всем мире и ведущей причиной смерти пациентов с ВИЧ и сахарным диабетом прежде всего из-за резистентности к противомикробным препаратам, используемым в клинической практике. Возрастающая распространенность лекарственной устойчивости представляет собой серьезную проблему для эффективной борьбы с ТБ. В связи с этим одной из ключевых задач, стоящих перед научным сообществом, является лечение ТБ с множественной и широкой лекарственной устойчивостью. Это определяет актуальность, важность и практическую значимость исследований по разработ-

ке новых эффективных антибактериальных препаратов широкого спектра действия, воздействующих на функционально важные участки микобактерии туберкулеза (МБТ).

Среди белковых мишеней, играющих важную роль в биосинтезе клеточной стенки МБТ, особое внимание следует уделить бета-кетоацил-(ацил-белок-носитель)-синтазе I (KasA) – одному из ключевых ферментов пути FAS-II синтеза жирных кислот. Потеря активности KasA приводит к лизису клеток бактерии. Это указывает на то, что данный белок может иметь ключевое значение для жизненного цикла МБТ и, следовательно, является важной мишенью при разработке новых эффективных препаратов для терапии лекарственно-устойчивых форм ТБ.

С целью идентификации соединений, способных блокировать фермент KasA МБТ, в работе [58] был выполнен виртуальный скрининг библиотеки [52], который включал: молекулярный докинг белка C171Q KasA с соединениями из библиотеки биоактивных молекул; оценку сродства этих соединений к активному центру фермента с помощью трех оценочных функций с последующим расчетом экспоненциального консенсусного ранга (ECR) для каждого соединения и ранжированием лигандов по значениям ECR; МД комплексов лиганд-C171Q KasA и расчеты свободной энергии их образования для лучших по значениям ECR-соединений; анализ результатов МД для идентификации наиболее перспективных соединений – кандидатов в лекарственные средства, способных ингибировать каталитическую активность фермента KasA МБТ.

В результате проведенных исследований были идентифицированы шесть перспективных соединений (рис. 6), обладающих высоким сродством к каталитическому центру фермента [58]. При этом значения свободной энергии связывания этих соединений с белком C171Q KasA оказались значительно ниже по сравнению с величинами, рассчитанными с использованием идентичных вычислительных протоколов для известного ингибитора KasA TLM5 [58]. Эти результаты позволяют предположить, что идентифицированные соединения формируют перспективные базовые структуры для разработки новых противотуберкулезных агентов, имеющих клиническое значение и обладающих высокой активностью против белка KasA МБТ. Поскольку идентифицированные соединения были найдены в библиотеке одобренных FDA лекарств и экспериментальных препаратов-кандидатов, их фармакологические свойства известны или находятся на стадии изучения. Очевидно, что это может существенно сократить время и затраты, связанные с длительным процессом создания новых эффективных лекарств, обладающих высокой антибактериальной активностью и приемлемыми фармакологическими свойствами.

Анализ терапевтических профилей обнаруженных соединений показывает, что соединение I (рис. 6) – Лумакафтор – является одобренным FDA препаратом, соединения II и III имеют статус экспериментальных препаратов, а соединения IV–VI представляют собой исследуемые кандидаты в лекарственные средства. В частности, Лумакафтор используется в клинике в комбинации с Ивакафтором для лечения пациентов с муковисцидозом. Поскольку Лумакафтор имеет хорошую пероральную биодоступность, зрелую технологию производства и хорошо охарактеризованную безопасность *in vivo*, этот препарат может быть быстро перепрофилирован и разработан как противотуберкулезное средство для клинического использования в ближайшем будущем.

С целью предварительной оценки антимикобактериального эффекта Лумакафтора исследователи из Университета Фудань (Шанхай, Китай) измерили минимальную ингибиторную концентрацию (МИК50) этой молекулы против вакцинного штамма *Mycobacterium bovis* BCG – облигатного аэробного медленно растущего штамма микобактерий, тесно связанного с МБТ. Полученные результаты показали, что эта молекула имеет значение МИК50, равное 62,5 мкМ и совпадающее с величиной, опубликованной для ингибитора KasA МБТ тиолактомицина [58]. В совокупности данные результаты указывают на то, что Лумакафтор является многообещающим кандидатом в лекарственное средство для лечения ТБ.

В настоящее время китайские партнеры из Университета Фудань проводят биомедицинские испытания найденных *in silico* соединений на *in vitro* моделях штаммов МБТ с множественной лекарственной устойчивостью.



Исследования белорусской группы проводились в рамках международного проекта по разработке информационного портала по лекарственно-устойчивому туберкулезу (<https://tbportals.niaid.nih.gov>).

**Дизайн, *in silico* и *in vitro* оценка противоопухолевой активности новых потенциальных ингибиторов тирозинкиназы Bcr-Abl.** Тирозинкиназа Bcr-Abl играет ключевую роль в патогенезе хронического миелоидного лейкоза (ХМЛ), характеризующегося быстрым неконтролируемым ростом миелоидных клеток в периферической крови и костном мозге, и в 20–50 % случаев является причиной острого В-лимфобластного лейкоза взрослых. За последние два десятилетия для лечения ХМЛ разработано достаточно много ингибиторов тирозинкиназы Bcr-Abl, проявляющих высокую ингибиторную активность и эффективных во многих случаях, когда возникает резистентность к этому препарату. Их применение позволило значительно увеличить продолжительность жизни пациентов с ХМЛ, а также некоторыми стромальными опухолями желудочно-кишечного тракта. Однако наряду с терапевтическим эффектом лечение этими препаратами может сопровождаться рядом побочных эффектов, связанных с их относительно высокой токсичностью, а мутации, возникающие в каталитическом центре фермента, могут вызывать резистентность к используемым препаратам, оставляя пациентам ограниченные возможности лечения. На сегодняшний день мутация T315I в каталитическом центре фермента является главной причиной развития первичной и вторичной резистентности к терапии ингибиторами тирозинкиназы Bcr-Abl у пациентов с хронической фазой ХМЛ. В связи с этим в настоящее время проводятся многочисленные исследования по разработке ингибиторов тирозинкиназы Bcr-Abl, эффективных для лечения больных ХМЛ с непереносимостью предшествующей терапии.

Совместно с сотрудниками ИХНМ НАН Беларуси и ИФОХ НАН Беларуси в работе [59] авторами был осуществлен дизайн малых молекул, содержащих фармакофорные группы, способные обеспечить высокое сродство этих соединений к активному центру нативной и мутантной (T315I) тирозинкиназы Bcr-Abl. Выполнена *in silico* оценка их ингибиторного потенциала с последующим определением противоопухолевой активности на моделях *in vitro*. Для решения данной задачи был использован комплексный подход, который включал следующие этапы: компьютерное конструирование молекул; молекулярный докинг этих соединений с каталитическим центром нативной и мутантной тирозинкиназы Bcr-Abl; МД комплексов лиганд-Bcr-Abl и расчет свободной энергии их образования; синтез соединений и оценку их ингибиторной активности на моделях опухолевых клеток линий K562 (хронический миелоидный лейкоз), HL-60 (острый промиелоцитарный лейкоз) и HeLa (карцинома шейки матки). В результате совместного анализа расчетных и экспериментальных данных выявлены три соединения (рис. 7), проявляющие противоопухолевую активность по отношению к клеткам линий K562 и HL-60 [59]. Обнаружено соединение-лидер (соединение I на рис. 7), демонстрирующее эффективное ингибирование роста этих клеток, что подтверждается низкими значениями IC<sub>50</sub>, равными 2,80 ± 0,76 мкМ (K562) и 3,51 ± 0,23 мкМ (HL-60). Все идентифицированные соединения удовлетворяют правилу пяти Липинского и согласно данным компьютерного прогнозирования обладают приемлемой токсичностью. Полученные данные позволяют предположить, что идентифицированные соединения могут служить основой для разработки новых эффективных лекарственных препаратов, способных ингибировать каталитическую активность тирозинкиназы Bcr-Abl путем блокирования АТР-связывающей полости фермента. Кроме того, обнаруженные соединения могут быть использованы для создания многоцелевых ингибиторов протеинкиназ, что подтверждается полученными данными об их противоопухолевой активности по отношению к линии клеток острого промиелоцитарного лейкоза HL-60, характеризующегося образованием аномального онкогенного фузионного белка PML-RARalpha. В связи с этим одно из дальнейших направлений развития настоящей работы предполагает исследование механизма действия идентифицированных соединений на *in vitro* моделях потенциальных терапевтических мишеней.

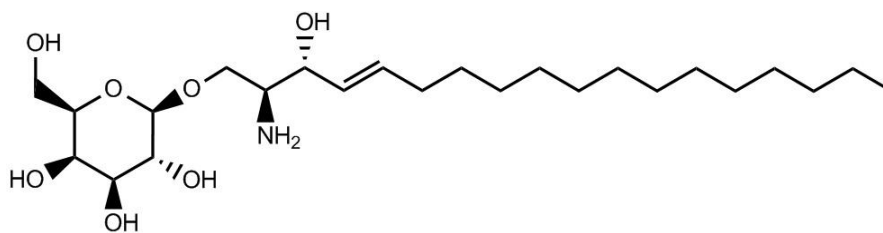


Рис. 1. Химическая структура β-галактозилсфингозина – ингибитора белка gp120 оболочки ВИЧ-1

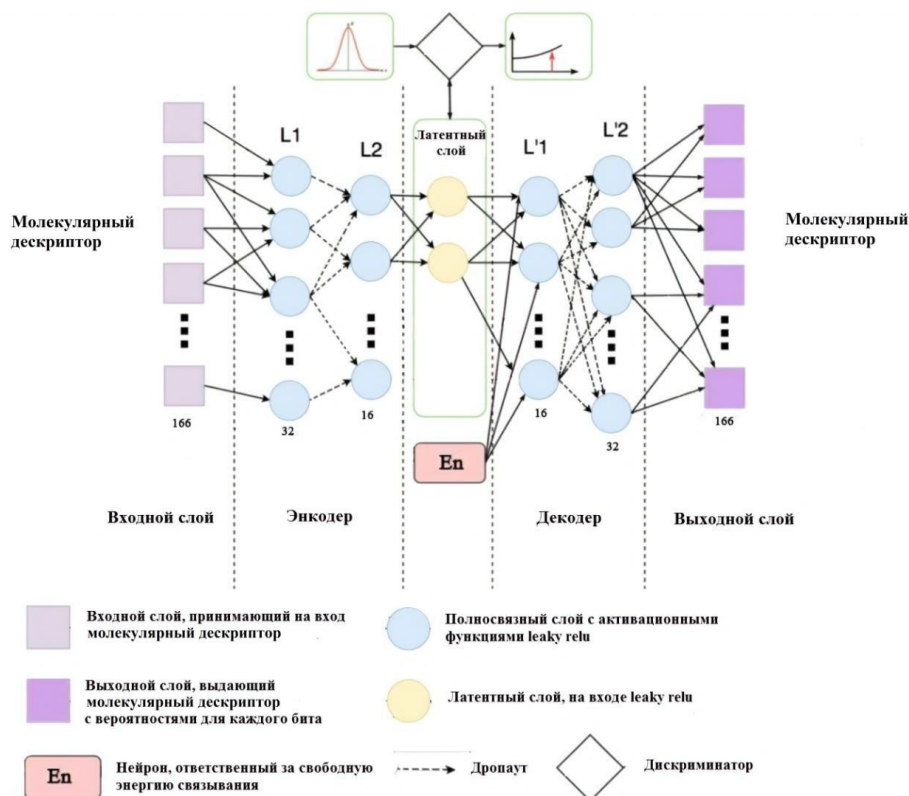


Рис. 2. Архитектура нейронной сети для генерации потенциальных ингибиторов ВИЧ-1, блокирующих CD4-связывающий сайт белка gp120 оболочки вируса [41, 42]



Рис. 3. Схема алгоритма виртуального скрининга потенциальных ингибиторов M<sup>Pro</sup> SARS-CoV-2 [50, 51]

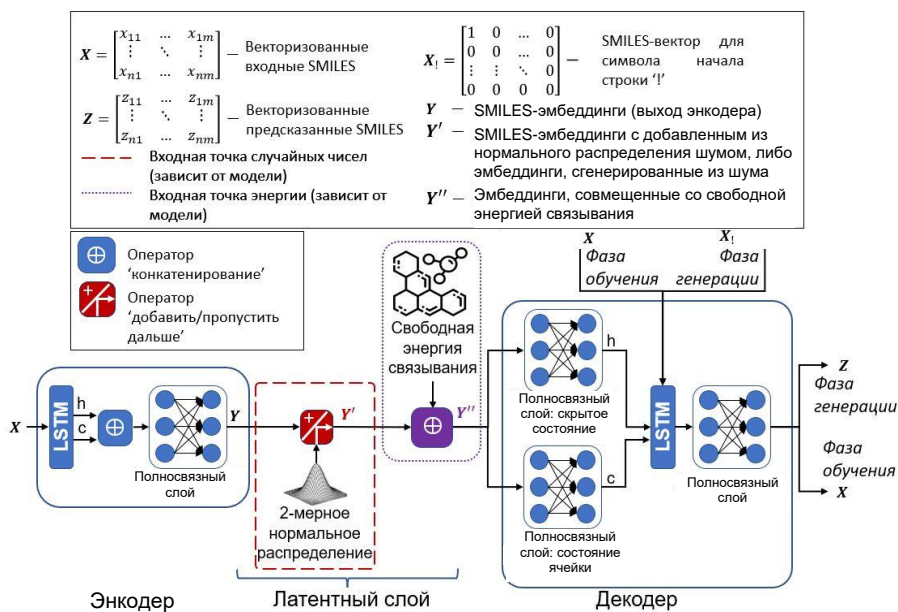


Рис. 4. Высокоуровневая архитектура моделей разработанной нейронной сети [53, 55]

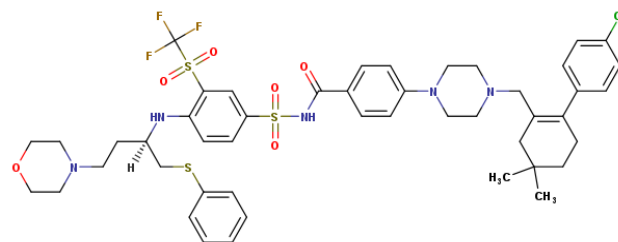


Рис. 5. Химическая структура Навитоклакса [57]

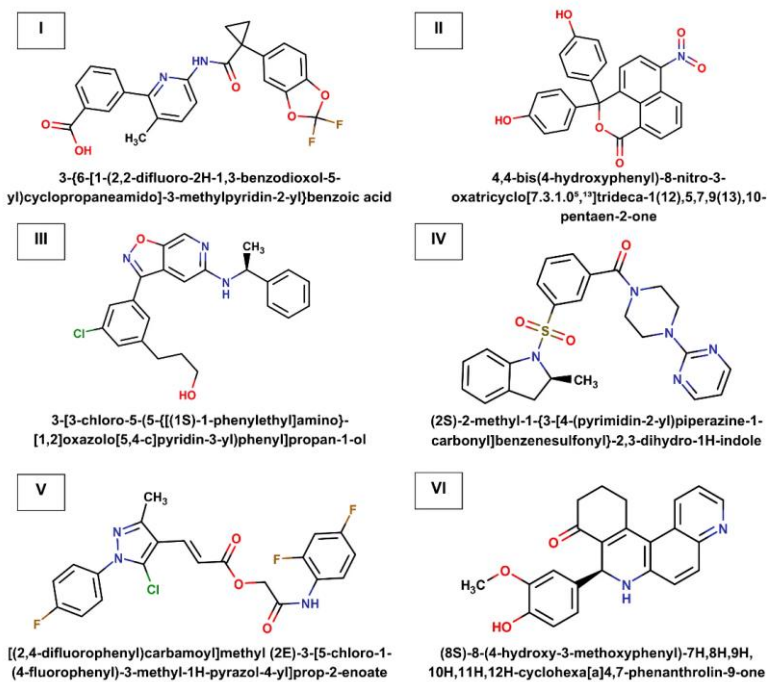


Рис. 6. Химические структуры идентифицированных соединений (приведены систематические названия молекул)

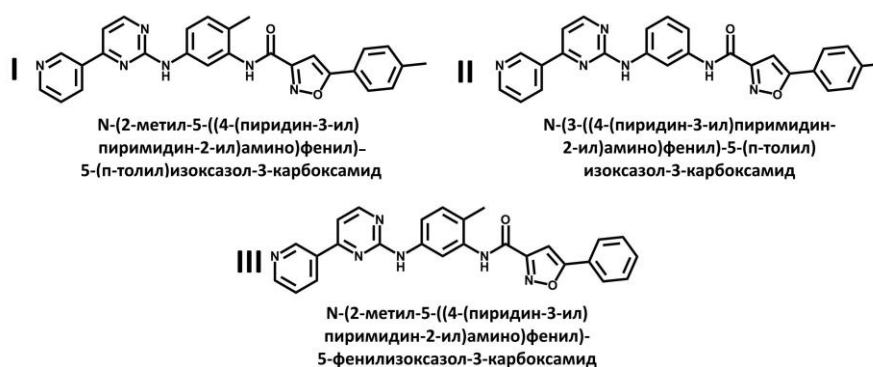


Рис. 7. Химические структуры ингибиторов роста опухолевых клеток K562 и HL-60 [59]

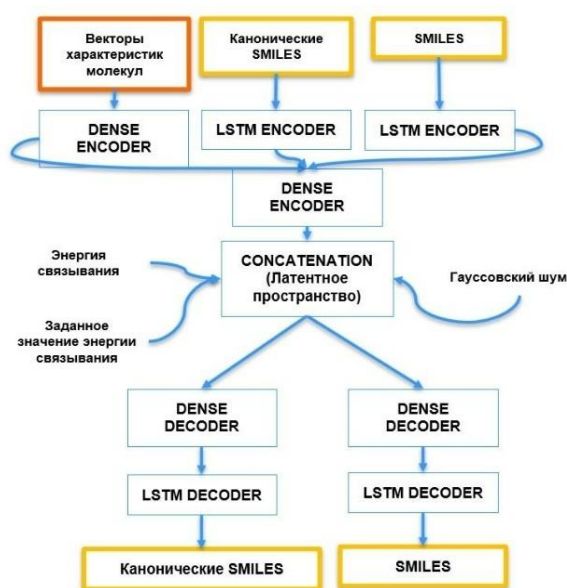


Рис. 8. Архитектура разработанной модели гетероэнкодера [60]

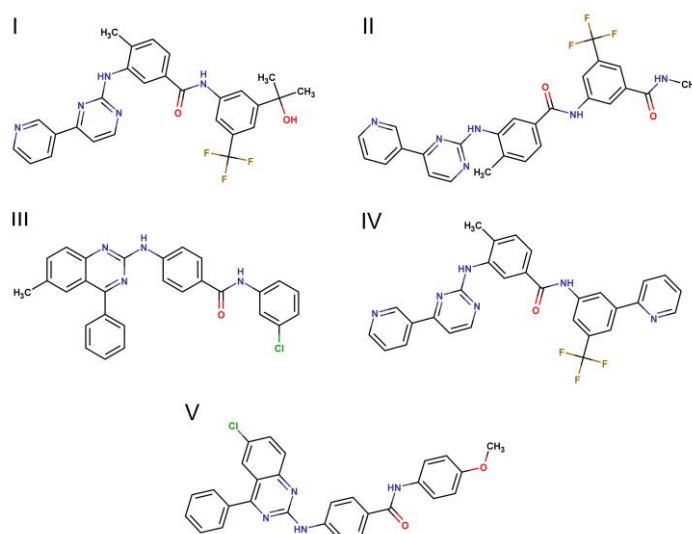


Рис. 9. Химические структуры идентифицированных соединений – потенциальных ингибиторов Vsrc-Abl тирозинкиназы и ее мутантной формы T3151 (названия соединений соответствуют систематической номенклатуре ИЮПАК [61])

*De novo* дизайн потенциальных ингибиторов Vcr-Abl тирозинкиназы методами глубокого обучения и виртуального скрининга. В работе [60] на основе рекуррентных и полносвязных нейронных сетей прямого распространения создана модель генеративной нейронной сети для компьютерного дизайна потенциальных ингибиторов тирозинкиназы Vcr-Abl, проведены обучение и тестирование этой модели на наборе химических соединений, которые содержат 2-ариламинопиримид, присутствующий в качестве основного фармакофора в структурах многих низкомолекулярных ингибиторов протеинкиназ. Разработанная модель нейронной сети (рис. 8) базируется на архитектуре гетероэнкодера – автоэнкодера, предназначенного для решения задач, в которых входные данные представлены в нескольких разных форматах. Такая архитектура нейронной сети позволяет получать более информативное латентное пространство за счет большего числа начальных признаков, что расширяет возможности поиска зависимостей между ними в процессе ее обучения. В исследовании [60] была реализована модель гетероэнкодера с тремя энкодерами и двумя декодерами, которая использует открытую библиотеку Keras (<https://keras.io>), обеспечивающую работу с искусственными нейронными сетями. В этой модели входные данные задаются в строковых форматах SMILES (Simplified Molecular Input Line Entry System) и канонический SMILES, а также числовым вектором характеристики молекулы (<https://www.rdkit.org/docs/source/rdkit.Chem.Descriptors.htm>) (рис. 8).

С помощью разработанной модели нейронной сети в работе [61] был получен набор из 1083 новых молекул-кандидатов и методами молекулярного докинга и МД выполнен анализ их сродства к активным центрам тирозинкиназы Vcr-Abl и ее модификации T315I. В результате проведенных исследований идентифицированы пять соединений-лидеров (рис. 9), обладающих согласно расчетным данным лекарственными свойствами и характеризующихся низкими значениями свободной энергии связывания с ферментом, которые сопоставимы с величинами, предсказанными для иматиниба, нилотиниба и понатиниба – противоопухолевых препаратов, широко используемых в клинике для лечения пациентов с ХМЛ.

Совместный анализ данных компьютерного моделирования позволяет предположить, что сконструированные с помощью технологий искусственного интеллекта соединения представляют значительный интерес для проведения дальнейших экспериментальных и теоретических исследований, включающих синтез, биомедицинские испытания *in vitro* и оптимизацию их структур, направленную на получение аналогов с улучшенной противоопухолевой активностью и приемлемыми фармакокинетическими и токсикологическими параметрами. Сгенерированные *de novo* соединения I, II и IV, содержащие 2-ариламинопиримидин (рис. 9), могут также быть использованы для разработки ингибиторов протеинкиназ многоцелевого действия. Это подтверждается многочисленными исследованиями, свидетельствующими о том, что производные 2-ариламинопиримидина обладают большим потенциалом в качестве кандидатов в лекарственные средства для противоопухолевой терапии.

#### **Биоинформатика и вычислительная биология**

*Полногеномный поиск ассоциаций [62–64].* Полногеномный поиск ассоциаций (Genome-Wide Association Study, GWAS) является мощным инструментом, позволяющим выполнять поиск мутаций в геномах живых организмов, связанных с фенотипическими признаками, под которыми могут пониматься как отдельные биологические свойства и характеристики, присущие организмам, так и их совокупность. Задача полногеномного поиска ассоциаций является частным случаем более общей задачи отбора признаков, где в качестве признаков выступают геномные мутации, представляющие собой отличия в ДНК-последовательностях организмов. При этом переменная, которая характеризует измеряемый фенотипический признак, может быть категориальной или количественной и принимать дискретные или непрерывные значения.

Исследование генетических или геномных маркеров имеет большое практическое значение для своевременной диагностики и понимания биологических механизмов становления лекарственной устойчивости микроорганизмов. В отличие от комплексных заболеваний человека, где гены могут играть определенную роль, молекулярные основы лекарственной устойчивости, как правило, напрямую связаны с мутационными изменениями в последовательностях ДНК

микроорганизмов, возникающими в ходе естественного отбора или под воздействием применяемых препаратов при ненадлежащем лечении.

За прошедшее десятилетие был достигнут значительный прогресс в разработке методов полногеномного анализа ассоциаций. Вместе с тем существует ряд сложностей, связанных с их применением на практике. Прежде всего эти сложности обусловлены большим числом подлежащих проверке геномных мутаций, наличием взаимозависимостей между ними, слабыми ассоциациями между геномными мутациями и фенотипическими признаками. Решение возникших задач особенно актуально для исследований, где фенотип описывается принадлежностью к одному из двух классов, что обычно имеет место при анализе лекарственной устойчивости, поскольку в случае дискретных или непрерывных значений фенотипа статистические тесты, как правило, позволяют иметь большую мощность и лучшую интерпретируемость результатов.

В лаборатории математической кибернетики ОИПИ НАН Беларуси разработан метод, позволяющий автоматизировать поиск алгоритма машинного обучения, который с точки зрения используемого критерия качества (F-мера, доля верных ответов) эффективно находит статистические взаимосвязи в пространстве бинарных переменных и может применяться для решения задачи поиска ассоциаций между геномными мутациями и бинарным откликом [64].

Разработан алгоритм, позволяющий выполнять анализ первичных последовательностей белков ВИЧ-1 с целью определения мутаций, связанных с проведением противовирусной терапии. Был предложен способ оценки коррелированности мутаций в позициях белков ВИЧ-1, основанный на расчете коэффициента ранговой корреляции и учитывающий эволюционную взаимосвязанность исследуемых образцов вируса [62].

Разработан алгоритм поиска мутаций в полных геномах микобактерий туберкулеза, связанных с развитием устойчивости к лекарственным препаратам. Алгоритм основан на применении методов однофакторного и многофакторного анализа данных и использует графическую модель для окончательного отбора значимых признаков, что позволяет оценивать вклад мутаций в совокупности и корректировать результаты в случае их противоречивости [63].

Разработан алгоритм поиска информативно-значимых комбинаций бинарных переменных, позволяющий без полного перебора всевозможных вариантов определять комбинации геномных мутаций, ассоциированные с исследуемым бинарным откликом. Предложенный алгоритм может использоваться как шаг предварительного отбора переменных при построении моделей, допускающих наличие взаимодействующих переменных [64].

#### **Исследования в области структурной и вычислительной биологии**

Лаборатория математической кибернетики ОИПИ НАН Беларуси работала в области структурной и вычислительной биологии в тесном контакте с Центром биоинформатики Университета Канзаса (США), который возглавлял профессор Илья Ваксер (<https://molecularbiosciences.ku.edu/people/ilya-vakser>).

*Анализ структурных изменений белков при взаимодействиях [65–68].* Понимание процессов взаимодействия белков является необходимым шагом для моделирования биохимических взаимодействий в живой клетке, что служит основой для создания лекарств. Информация о схожести пространственных структур белков в значительной степени облегчает предсказание функций неизвестных белков и определение отдаленных эволюционных связей между белками.

Несмотря на разнообразие алгоритмов прогнозирования взаимодействия белков (белок-белкового докинга), задача предсказания комплекса белков по их свободным структурам на в период 2012–2018 гг. интенсивно исследовалась многими научными группами. Поскольку часто при взаимодействии пространственные структуры белков изменяются, другим важным аспектом докинга является учет гибкости белка. Включение гибкости структур белка в алгоритмы докинга значительно усложняет их из-за увеличения числа степеней свободы, что приводит к существенному увеличению времени работы алгоритмов и большому числу ложноположительных результатов.

Для дальнейшего улучшения протоколов докинга необходимо было понять связь между энергетическими ландшафтами и изменениями пространственных структур белков при взаимодействии. Преимуществом современных компьютерных технологий является то, что они поз-

воляют извлекать дополнительные знания из больших объемов данных, накопленных экспериментальными методами.

В лаборатории математической кибернетики ОИПИ НАН Беларуси разработан алгоритм предсказания взаимодействия белков, базирующийся на использовании структурной схожести с белок-белковым интерфейсом из базы данных. Алгоритм основан на выравнивании пространственных структур пары свободных белков с пространственной структурой белок-белкового интерфейса из базы данных. Выравнивание реализуется с помощью метода динамического программирования путем максимизации корреляции между матрицами расстояний интерфейса и белка и выполнения преобразования движения в трехмерном пространстве. Для каждой пары свободных белков строится 10 наиболее предпочтительных моделей комплекса на основе интерфейсов с наибольшей схожестью.

Для построения библиотек ротамеров разработан иерархический алгоритм кластеризации с изменяемым радиусом сфер, покрывающих пространство, который позволяет более точно аппроксимировать форму распределения двугранных углов аминокислот. На основе библиотек были построены карты и матрицы переходов боковых цепей аминокислот из свободного состояния в связанное. Построенные библиотеки, карты и матрицы переходов являются важным инструментом при переборе возможных структур боковых цепей аминокислот в алгоритмах предсказания взаимодействия белков, поскольку позволяют значительно уменьшить вычислительную сложность алгоритмов.

Для моделирования белок-белковых взаимодействий разработан способ подтверждения гипотезы механизма выбора связанной структуры белка при взаимодействии, основанный на моделировании энергетического ландшафта и структурной схожести интерфейса набора связанных и свободных структур белков; определены энергетические характеристики набора свободных и связанных структур белков; подтвержден механизм конформационного выбора для четырех из шести рассмотренных белков.

**Моделирование димерных белковых комплексов [69–72].** Белок-белковые взаимодействия определяют большинство процессов в клетке и базируются на их специфическом контакте, для которого необходима определенная пространственная структура. Поэтому биологическая функция белков определяется их уникальной трехмерной структурой, даже небольшие изменения которой часто ведут к утере или резкому изменению активности.

Вступая во взаимодействие, белки образуют белок-белковый комплекс. Задача нахождения трехмерной структуры комплекса, образованного при взаимодействии белков, называется белковым докингем. Так же как и в случае индивидуальных белков, экспериментальные методы для определения структуры комплекса могут использоваться в ограниченном количестве случаев и требуют длительного времени, а в массовом порядке и вовсе неприменимы. Поэтому ускоренное развитие получили вычислительные методы, которые позволяют получать трехмерные структуры комплексов быстро, используя в качестве входных данных трехмерные структуры белков, составляющих комплекс. Наибольшее значение при этом имеет область связывания белков, или интерфейс белкового комплекса, поскольку именно в данной области находятся аминокислоты, являющиеся необходимыми для образования комплекса, который, в свою очередь, выполняет соответствующую функцию в организме.

В лаборатории математической кибернетики ОИПИ НАН Беларуси была рассмотрена задача моделирования димерных белковых комплексов, т. е. состоящих из двух белков, разработаны алгоритмы, позволяющие решать эту задачу, а также общий подход к моделированию димерных белковых комплексов.

Получены следующие новые результаты:

1. Алгоритм моделирования структур димерных белковых комплексов с выбором шаблона на основе генной онтологии, позволяющий эффективно строить и ранжировать их модели в случае, когда структуры белков, которые поступают на вход алгоритма, неточно соответствуют экспериментальным данным, и показавший на тестовом множестве более высокую точность по сравнению с другими алгоритмами [69, 70].

2. Алгоритм генерации структур внутрибелковых мотивов на основе представления структуры белка в виде графа, с помощью которого было проведено структурное и функциональное сравнение межмолекулярных интерфейсов и внутримолекулярных мотивов, показавшее, что их свойства, включающие структурный состав и укладку, в значительной степени различаются. Алгоритм позволил определить структуры, которые могут использоваться в качестве дополнительной информации при моделировании белковых комплексов и включают междоменные интерфейсы, а также три класса структурных мотивов внутри белка [72].

3. Алгоритм предсказания области связывания гомодимерных белковых комплексов на основе нейронной сети глубокого обучения, который для 53 % протестированных комплексов правильно идентифицировал большую часть интерфейсных элементов и обеспечил существенное сокращение количества возможных ориентаций структур белков относительно друг друга при моделировании комплекса [71].

4. Алгоритм моделирования структур димерных белковых комплексов на основе матрицы контактов, позволивший получать на тестовом множестве белковых комплексов корректные трехмерные модели для 94 % гетеродимеров и 96 % гомодимеров [71].

**Заключение.** Как было отмечено выше, в настоящее время в разработке новых лекарственных препаратов продолжает увеличиваться роль молекулярного моделирования и машинного обучения. Однако современный уровень развития компьютерных методик не позволяет создавать новые лекарства, используя только компьютеры, но существенно сокращает время их выпуска и снижает стоимость разработки. Современные технологии конструирования лекарств включают *in silico* идентификацию низкомолекулярных соединений с высокой потенциальной активностью и приемлемыми фармакологическими характеристиками с последующим их синтезом, анализом терапевтической эффективности, оптимизацией соединений-лидеров, направленной на улучшение их ингибиторной активности и фармакологических свойств, синтезом оптимизированных соединений и их детальными биомедицинскими испытаниями. Исследования *in silico* были реализованы во всех представленных в настоящем обзоре работах, а в ряде из них проведено тестирование идентифицированных в компьютерных экспериментах соединений на моделях *in vitro*. Поэтому дальнейшее развитие наших исследований предполагает реализацию последующих этапов многостадийного процесса создания новых лекарственных препаратов. Продолжаются также работы в области биоинформатики и вычислительной биологии.

**Защиты диссертаций.** При выполнении исследований были защищены следующие диссертации:

Кирис Т. В. – диссертация на соискание ученой степени кандидата технических наук «Анализ и моделирование структурных изменений белков при взаимодействиях» по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (05.03.2013).

Анищенко И. В. – диссертация на соискание ученой степени кандидата технических наук «Компьютерное моделирование новых потенциальных анти-ВИЧ агентов, блокирующих белок gp120 оболочки вируса» по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (07.10.2014).

Корноушенко Ю. В. – диссертация на соискание ученой степени кандидата химических наук «Конструирование потенциальных ингибиторов ВИЧ-1 на основе гликофинголипидов методами молекулярного моделирования и химического синтеза» по специальности 02.00.10 – Биоорганическая химия (28.02.2017).

Кашин И. А. – диссертация на соискание ученой степени кандидата химических наук «Компьютерный скрининг и идентификация потенциальных ингибиторов ВИЧ-1 на основе высокоаффинных лигандов белков оболочки вируса» по специальности 02.00.10 – Биоорганическая химия (18.01.2018).

Сергеев Р. С. – диссертация на соискание ученой степени кандидата технических наук «Алгоритмы анализа и поиска ассоциаций в генетических данных» по специальности 05.13.01 – Системный анализ, управление и обработка информации (17.09.2019).



Хадарович А. Ю. – диссертация на соискание ученой степени кандидата технических наук «Алгоритмы моделирования димерных белковых комплексов» по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (02.02.2020).

Николаев Г. И. – диссертация на соискание ученой степени кандидата технических наук «Разработка алгоритмов, основанных на технологии глубокого обучения и молекулярном моделировании, для конструирования потенциальных ингибиторов проникновения ВИЧ-1» по специальности 05.13.18 – Математическое моделирование, численные методы и комплексы программ (23.06.2022).

### Конкурсы и премии

2015 г.: ТОП 10 результатов НАН Беларуси за 2015 г. (А. В. Тузиков, А. М. Андрианов, И. А. Кашин, Ю. В. Корноушенко) за обнаружение на основе компьютерного скрининга и моделирования новых низкомолекулярных химических соединений с ароматическими фрагментами, формирующих перспективные базовые структуры для создания эффективных лекарственных препаратов против ВИЧ/СПИД с широким спектром нейтрализующего действия.

2017 г.: премия Национальной академии наук Беларуси «Компьютерный дизайн потенциальных ингибиторов ВИЧ-1, перспективных для создания противовирусных препаратов нового поколения» (А. М. Андрианов, И. А. Кашин, А. В. Тузиков).

2020 г.: ТОП 10 результатов НАН Беларуси за 2020 г. (А. М. Андрианов, И. П. Босько, А. Д. Карпенко, Ю. В. Корноушенко, А. В. Тузиков) за идентификацию методами компьютерного скрининга и молекулярного моделирования потенциальных ингибиторов коронавируса SARS-CoV-2.



ТОП 10 результатов НАН Беларуси за 2015 г.,  
В. Г. Гусаков, А. В. Тузиков, А. М. Андрианов, И. А. Кашин, Ю. В. Корноушенко



На симпозиуме в Шэньчжэне, Китай, 27–30 октября 2018 г.,  
Г. И. Николаев, А. В. Тузиков, проф. Shibo Jiang, А. М. Андрианов, Ю. В. Корноушенко



Семинар по проекту ANSO-CR-PP-2021-04, Шанхай (Китай), 18 апреля 2023 г.,  
А. В. Тузиков, проф. Hong Liu

### Публикации

1. Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review / V. T. Sabe, T. Ntombela, L. A. Jhamba [et al.] // *European Journal of Medicinal Chemistry*. – 2021. – Vol. 224. – P. 113705.
2. Current perspectives and trend of computer-aided drug design: a review and bibliometric analysis / Z. Wu, S. Chen, Y. Wang [et al.] // *International Journal of Surgery*. – 2024. – Vol. 110, no. 6. – P. 3848–3878.
3. Fan, J. Progress in molecular docking / J. Fan, A. Fu, L. Zhang // *Quantitative Biology*. – 2019. – Vol. 7. – P. 83–89.
4. Cavasotto, C. N. Quantum chemical approaches in structure-based virtual screening and lead optimization / C. N. Cavasotto, N. S. Adler, M. G. Aucar // *Frontiers in Chemistry*. – 2018. – Vol. 6. – P. 188.

5. Ryde, U. Ligand-binding affinity estimates supported by quantum-mechanical methods / U. Ryde, P. Söderhjelm // *Chemical Reviews*. – 2016. – Vol. 116. – P. 5520–5566.
6. Yilmazer, N. D. Prospects of applying enhanced semi-empirical QM methods for 2101 virtual drug design / N. D. Yilmazer, M. Korth // *Current Medicinal Chemistry*. – 2016. – Vol. 23. – P. 2101–2111.
7. Childers, M. C. Insights from molecular dynamics simulations for computational protein design / M. C. Childers, V. Daggett // *Molecular Systems Design & Engineering*. – 2017. – Vol. 2, no. 1. – P. 9–33.
8. Hollingsworth, S. A. Molecular dynamics simulation for all / S. A. Hollingsworth, R. O. Dror // *Neuron*. – 2018. – Vol. 99. – P. 1129–1143.
9. Applications of machine learning in drug discovery and development / J. Vamathevan, D. Clark, P. Czodrowski [et al.] // *Nature Reviews Drug Discovery*. – 2019. – Vol. 18(6). – P. 463–477.
10. Advances and perspectives in applying deep learning for drug design and discovery / C. F. Lipinski, V. G. Maltarollo, P. R. Oliveira [et al.] // *Frontiers in Robotics and AI*. – 2019. – Vol. 6. – P. 108.
11. A Machine learning-based method to improve docking scoring functions and its application to drug repurposing / S. L. Kinnings, N. Liu, P. J. Tonge [et al.] // *Journal of Chemical Information and Modeling*. – 2011. – Vol. 51. – P. 408–419.
12. Agastheeswaramoorthy, K. Drug REpurposing using AI/ML tools – for Rare Diseases (DREAM-RD): A case study with Fragile X Syndrome (FXS) / K. Agastheeswaramoorthy, A. Sevilimedu // *bioRxiv*. – 2020. – DOI: 10.1101/2020.09.25.311142.
13. Improved protein structure prediction using potentials from deep learning / A. W. Senior, R. Evans, J. Jumper [et al.] // *Nature*. – 2020. – Vol. 577. – P. 706–710.
14. Machine-learning scoring functions for structure-based virtual screening / H. Li, K.-H. Sze, G. Lu, P. J. Ballester // *WIREs Computational Molecular Science*. – 2020. – Vol. 11. – P. e1478.
15. Improving structure-based virtual screening performance via learning from scoring function components / G.-L. Xiong, W.-L. Ye, C. Shen [et al.] // *Briefings in Bioinformatics*. – 2020. – Vol. 22, iss. 3. – P. bbaa094. – DOI: 10.1093/bib/bbaa094.
16. A deep learning approach to antibiotic discovery / J. M. Stokes, K. Yang, K. Swanson [et al.] // *Cell*. – 2020. – Vol. 180. – P. 688–702.
17. Timmons, P. B. ENNAVIA is a novel method which employs neural networks for antiviral and anti-coronavirus activity prediction for therapeutic peptides / P. B. Timmons, C. M. Hewage // *Briefings in Bioinformatics*. – 2021. – Vol. 22, iss. 6. – P. bbab258. – DOI: 10.1093/bib/bbab258.
18. Classification of HIV-1 protease inhibitors by machine learning methods / Y. Li, Y. Tian, Z. Qin, A. Yan // *ACS Omega*. – 2018. – Vol. 3, no. 11. – P. 15837–15849.
19. Deep learning driven drug discovery: Tackling Severe Acute Respiratory Syndrome Coronavirus 2 / Y. Zhang, T. Ye, H. Xi [et al.] // *Frontiers in Microbiology*. – 2021. – Vol. 12. – P. 739684.
20. Practical notes on building molecular graph generative models / R. Mercado, T. Rastemo, E. Lindelöf [et al.] // *Applied AI Letters*. – 2020. – Vol. 1, no. 2. – DOI: 10.1002/ail2.18.
21. Exploring the GDB-13 chemical space using deep generative models / J. Arús-Pous, T. Blaschke, S. Ulander [et al.] // *Journal of Cheminformatics*. – 2019. – Vol. 11: Article 20. – DOI: 10.1186/s13321-019-0341-z.
22. A de novo molecular generation method using latent vector based generative adversarial network / O. Prykhodko, S. V. Johansson, P.-C. Kotsias [et al.] // *Journal of Cheminformatics*. – 2019. – Vol. 11: Article 74. – DOI: 10.1186/s13321-019-0397-9.
23. Entangled conditional adversarial autoencoder for de novo drug discovery / D. Polykovskiy, A. Zhebrak, D. Vetrov [et al.] // *Molecular Pharmaceutics*. – 2018. – Vol. 15. – P. 4398–4405.
24. Comparative study of deep generative models on chemical space coverage / J. Zhang, R. Mercado, O. Engkvist, H. Chen // *Journal of Chemical Information and Modeling*. – 2021. – Vol. 61. – P. 2572–2581.
25. Deep learning enables rapid identification of potent DDR1 kinase inhibitors / A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper [et al.] // *Nature Biotechnology*. – 2019. – Vol. 37. – P. 1038–1040.
26. Анищенко, И. В. Компьютерный дизайн потенциальных лекарственных препаратов для терапии СПИДа: β-галактозилцерамид и петля V3 белка gp120 ВИЧ-1 / И. В. Анищенко, А. В. Тузиков, А. М. Андрианов // *Математическая биология и биоинформатика*. – 2011. – Т. 6, № 2. – С. 161–172.
27. Computer-aided design of novel HIV-1 entry inhibitors targeting the envelope gp120 V3 loop / A. M. Andrianov, I. V. Anishchenko, M. A. Kisel [et al.] // *Biopolymers and Cell*. – 2012. – Vol. 28, no. 6. – P. 468–476.
28. Компьютерное конструирование новых ингибиторов проникновения ВИЧ-1 на основе гликофинголипидов / А. М. Андрианов, Ю. В. Корноушенко, И. А. Кашин, А. В. Тузиков // *Математическая биология и биоинформатика*. – 2013. – Т. 8, № 1. – С. 88–105.

29. Конструирование ингибиторов репликации ВИЧ-1 на основе  $\beta$ -галактозилцерамида методами молекулярного моделирования и химического синтеза / А. М. Андрианов, И. В. Анищенко, М. А. Кисель [и др.] // Доклады Национальной академии наук Беларуси. – 2011. – Т. 55, № 3. – С. 70–78.
30. Получение и анти-ВИЧ активность  $\beta$ -галактозилсфингозина / Ю. В. Корноушенко, В. А. Николаевич, М. А. Кисель [и др.] // Весці Нацыянальнай акадэміі навук Беларусі. Серыя хімічных навук. – 2015. – № 1. – С. 85–88.
31. In silico design of novel broad anti-HIV-1 agents based on glycosphingolipid  $\beta$ -galactosylceramide, a high affinity receptor for the envelope gp120 V3 loop / A. M. Andrianov, Y. V. Kornoushenko, I. A. Kashyn [et al.] // Journal of Biomolecular Structure and Dynamics. – 2015. – Vol. 33, no. 5. – P. 1051–1066.
32. Andrianov, A. M. HIV-1 gp120 V3 loop for anti-AIDS drug discovery: computer-aided approaches to the problem solving / A. M. Andrianov // Expert Opinion on Drug Discovery. – 2011. – Vol. 6, no. 4. – P. 419–435.
33. Andrianov, A. M. Discovery of novel promising targets for anti-AIDS drug developments by computer modeling: application to the HIV-1 gp120 V3 loop / A. M. Andrianov, I. V. Anishchenko, A. V. Tuzikov // Journal of Chemical Information and Modeling. – 2011. – Vol. 51, no. 10. – P. 2760–2767.
34. Андрианов, А. М. Конформационный анализ белков. Теория и приложения / А. М. Андрианов. – Минск : Беларус. навука, 2013. – 518 с.
35. Jiang, S. Small-molecule HIV-1 entry inhibitors targeting the epitopes of broadly neutralizing antibodies / S. Jiang, A. V. Tuzikov, A. M. Andrianov // Cell Chemical Biology. – 2022. – Vol. 29, no. 5. – P. 757–773.
36. Andrianov, A. M. Discovery of novel anti-HIV-1 agents based on a broadly neutralizing antibody against the envelope gp120 V3 loop: a computational study / A. M. Andrianov, I. A. Kashyn, A. V. Tuzikov // Journal of Biomolecular Structure and Dynamics. – 2014. – Vol. 32, no. 12. – P. 1993–2004.
37. Andrianov, A. Computational discovery of novel HIV-1 entry inhibitors based on potent and broad neutralizing antibody VRC01 / A. Andrianov, I. Kashyn, A. Tuzikov // Journal of Molecular Graphics and Modelling. – 2015. – Vol. 61. – P. 262–271.
38. Andrianov, A. M. Computer-based technologies for virtual screening and analysis of chemical compounds promising for anti-HIV-1 drug design / A. M. Andrianov, I. A. Kashyn, A. V. Tuzikov // Pattern Recognition and Information Processing (PRIP 2016) / ed.: V. Krasnoproshin, S. Ablameyko. – Springer, 2017. – P. 14–23. – (Communications in Computer and Information Science ; vol. 673).
39. Andrianov, A. M. Identification of novel HIV-1 fusion inhibitor scaffolds by virtual screening, high-throughput docking and molecular dynamics simulations / A. M. Andrianov, I. A. Kashyn, A. V. Tuzikov // JSM Chemistry. – 2016. – Vol. 4, no. 2. – P. 1022.
40. Andrianov, A. M. Computational identification of novel entry inhibitor scaffolds mimicking primary receptor CD4 of HIV-1 gp120 / A. M. Andrianov, I. A. Kashyn, A. V. Tuzikov // Journal of Molecular Modeling. – 2017. – Vol. 23, no. 1. – P. 1–13.
41. Разработка генеративной состязательной нейронной сети для идентификации потенциальных ингибиторов ВИЧ-1 методами глубокого обучения / Г. И. Николаев, Н. А. Шульдов, А. И. Анищенко [и др.] // Информатика. – 2020. – Т. 17, № 1. – С. 7–17.
42. Application of deep learning and molecular modeling to identify small drug-like compounds as potential HIV-1 entry inhibitors / A. M. Andrianov, G. I. Nikolaev, N. A. Shuldov [et al.] // Journal of Biomolecular Structure and Dynamics. – 2022. – Vol. 40, no. 16. – P. 7555–7573.
43. Разработка потенциальных ингибиторов ВИЧ-1 методами in silico клик-химии и молекулярного моделирования / А. М. Андрианов, Г. И. Николаев, И. А. Кашин, А. В. Тузиков // Математическая биология и биоинформатика. – 2018. – Т. 13, № 2. – С. 507–525.
44. In silico identification of novel aromatic compounds as potential HIV-1 entry inhibitors mimicking cellular receptor CD4 / A. M. Andrianov, G. I. Nikolaev, Y. V. Kornoushenko [et al.] // Viruses. – 2019. – Vol. 11, no. 8. – P. 746.
45. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings / C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney // Advanced Drug Delivery Reviews. – 2001. – Vol. 46. – P. 3–26.
46. Виртуальный скрининг и идентификация потенциальных ингибиторов ВИЧ-1 на основе кросс-реактивного нейтрализующего антитела N6 / А. М. Андрианов, Г. И. Николаев, Ю. В. Корноушенко [и др.] // Доклады Национальной академии наук Беларуси. – 2019. – Т. 63, № 4. – С. 445–456.
47. In silico идентификация высокоаффинных лигандов белка gp120 ВИЧ-1 – потенциальных пептидомиметиков нейтрализующего антитела N6 / А. М. Андрианов, Г. И. Николаев, Ю. В. Корноушенко [и др.] // Математическая биология и биоинформатика. – 2019. – Т. 14, № 2. – С. 430–449.
48. Идентификация функциональных миметиков нейтрализующего анти-ВИЧ антитела N6 методами виртуального скрининга и молекулярного моделирования / А. М. Андрианов, Г. И. Николаев, Ю. В. Корноушенко [и др.] // Доклады Национальной академии наук Беларуси. – 2019. – Т. 63, № 5. – С. 561–571.

49. In silico-guided discovery of potential HIV-1 entry inhibitors mimicking bNAb N6: Virtual screening, docking, molecular dynamics, and post-modeling analysis / A. M. Andrianov, G. I. Nikolaev, Y. V. Kornoushenko [et al.] // *Bioinformatics Research and Applications* / ed.: Z. Cai [et al.]. – Springer, 2020. – P. 243–249. – (Lecture Notes in Computer Science ; vol. 12304).
50. Идентификация потенциальных ингибиторов коронавируса SARS-CoV-2 методами виртуального скрининга и молекулярного моделирования / А. М. Андрианов, Ю. В. Корноушенко, А. Д. Карпенко, А. В. Тузиков // Доклады Национальной академии наук Беларуси. – 2020. – Т. 64, № 3. – С. 308–316.
51. Computational discovery of small drug-like compounds as potential inhibitors of SARS-CoV-2 main protease / A. M. Andrianov, Y. V. Kornoushenko, A. D. Karpenko [et al.] // *Journal of Biomolecular Structure and Dynamics*. – 2021. – Vol. 39, no. 15. – P. 5779–5791.
52. Применение технологий виртуального скрининга и молекулярного моделирования для идентификации потенциальных ингибиторов основной протеазы коронавируса SARS-CoV-2 / А. М. Андрианов, К. В. Фурс, А. В. Гончар [и др.] // *Математическая биология и биоинформатика*. – 2023. – Т. 18, № 1. – С. 15–32.
53. Разработка генеративной нейронной сети глубокого обучения для компьютерного дизайна потенциальных ингибиторов коронавируса SARS-CoV-2 / Н. А. Шульдов, А. М. Юшкевич, К. В. Фурс [и др.] // *Математическая биология и биоинформатика*. – 2022. – Т. 17, № 2. – С. 188–207.
54. De novo дизайн потенциальных ингибиторов основной протеазы коронавируса SARS-CoV-2 с помощью технологий искусственного интеллекта и молекулярного моделирования / А. М. Андрианов, К. В. Фурс, Н. А. Шульдов, А. В. Тузиков // Доклады Национальной академии наук Беларуси. – 2023. – Т. 67, № 3. – С. 197–206.
55. AI-driven de novo design and molecular modeling for discovery of small-molecule compounds as potential drug candidates targeting SARS-CoV-2 main protease / A. M. Andrianov, M. A. Shuldau, K. V. Furs [et al.] // *International Journal of Molecular Sciences*. – 2023. – Vol. 24, no. 9. – P. 8083.
56. In silico скрининг потенциальных ингибиторов SARS-CoV-2, блокирующих тример HR1 белка S коронавируса / А. М. Андрианов, К. В. Фурс, А. М. Юшкевич [и др.] // Доклады Национальной академии наук Беларуси. – 2022. – Т. 66, № 2. – С. 156–166. – DOI: 10.29235/1561-8323-2022-66-2-156-166.
57. Repurposing Navitoclax to block SARS-CoV-2 fusion and entry by targeting heptapeptide repeat sequence 1 in S2 protein / F. Jiao, A. M. Andrianov, L. Wang [et al.] // *Journal of Medical Virology*. – 2023. – Vol. 95. – P. e29145.
58. Virtual screening and identification of promising therapeutic compounds against drug-resistant *Mycobacterium tuberculosis*  $\beta$ -ketoacyl-acyl carrier protein synthase I (KasA) / A. M. Andrianov, K. V. Furs, A. V. Gonchar [et al.] // *Journal of Biomolecular Structure and Dynamics*. – 2023. – DOI: 10.1080/07391102.2023.2293276.
59. Design, in silico evaluation, and determination of antitumor activity of potential inhibitors against protein kinases: Application to BCR-ABL tyrosine kinase / E. V. Koroleva, A. L. Ermolinskaya, Z. V. Ignatovich [et al.] // *Biochemistry (Moscow)*. – 2024. – Vol. 89, no. 6. – P. 1094–1108.
60. Генеративная нейронная сеть на основе модели гетероэнкодера для de novo дизайна потенциальных противоопухолевых препаратов: применение к Bcr-Abl тирозинкиназе / А. Д. Карпенко Т. Д. Войтко, А. В. Тузиков, А. М. Андрианов // *Информатика*. – 2023. – Т. 20, № 3. – С. 7–20.
61. De novo дизайн и виртуальный скрининг потенциальных ингибиторов тирозинкиназы Bcr-Abl с помощью технологий глубокого обучения и молекулярного моделирования / А. М. Андрианов, К. В. Фурс, А. Д. Карпенко и [др.] // Доклады Национальной академии наук Беларуси. – 2024. – Т. 68, № 3. – С. 196–206.
62. Сергеев, Р. С. Алгоритмы анализа мутаций в первичных последовательностях белков ВИЧ-1 субтипа А / Р. С. Сергеев, А. В. Тузиков, В. Ф. Еремин // *Информатика*. – 2011. – № 3(31). – С. 88–97.
63. Алгоритмы поиска мутаций лекарственной устойчивости в геномах микобактерий туберкулеза / Р. С. Сергеев, И. С. Ковалев, А. В. Тузиков [и др.] // *Информатика*. – 2016. – № 1(49). – С. 75–91.
64. Genome-wide analysis of MDR and XDR tuberculosis from Belarus: machine-learning approach / R. S. Sergeev, I. S. Kavaliou, U. V. Sataneuski [et al.] // *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. – 2019. – Vol. 16, iss. 4. – P. 1398–1408. – DOI: 10.1109/TCBB.2017.2720669.
65. Side-chain conformational changes upon protein-protein association / A. M. Ruvinsky, T. Kirys, A. V. Tuzikov, I. A. Vakser // *Journal of Molecular Biology*. – 2011. – Vol. 408. – P. 356–365.
66. Rotamer libraries and probabilities of transition between rotamers for the side chains in protein-protein binding / T. Kirys, A. M. Ruvinsky, A. V. Tuzikov, I. A. Vakser // *Proteins: Structure, Function and Bioinformatics*. – 2012. – Vol. 80. – P. 2089–2098.

67. Structure fluctuations and conformational changes in protein binding / A. Ruvinsky, T. Kirys, A. V. Tuzikov, I. A. Vakser // *Journal of Bioinformatics and Computational Biology*. – 2012. – Vol. 10, no. 2. – P. 1241002.

68. Correlation analysis of the side-chains conformational distribution in bound and unbound proteins / T. Kirys, A. M. Ruvinsky, A. V. Tuzikov, I. A. Vakser // *BMC Bioinformatics*. – 2012. – Vol. 13. – P. 236–244.

69. Gene ontology improves template selection in comparative protein docking / A. Y. Hadarovich, I. Anishchenko, A. V. Tuzikov [et al.] // *Proteins: Structure, Function, Bioinformatics*. – 2019. – Vol. 87(3). – P. 245–253.

70. Алгоритм предсказания структур белковых комплексов на основе генной онтологии / А. Ю. Хадарович, И. В. Анищенко, П. Кундротас [и др.] // *Доклады Национальной академии наук Беларуси*. – 2020. – № 64(2). – С. 150–158.

71. Хадарович, А. Ю. Предсказание структуры гомодимерных белковых комплексов на основе глубокой нейронной сети / А. Ю. Хадарович, А. А. Калиновский, А. В. Тузиков // *Информатика*. – 2020. – № 17(2). – С. 44–53.

72. Structural motifs in protein cores and at protein-protein interfaces are different / A. Y. Hadarovich, D. Chakravarty, A. V. Tuzikov [et al.] // *Protein Science*. – 2021. – Vol. 30, iss. 2. – P. 381–390. – DOI: 10.1002/pro.3996.