

# БИОИНФОРМАТИКА

## BIOINFORMATICS



УДК 51-76:577.322:539.19:004.94  
<https://doi.org/10.37661/1816-0301-2024-21-3-80-93>

Оригинальная статья  
Original Article

## Адаптация архитектуры нейронной сети REINVENT для генерации потенциальных ингибиторов проникновения ВИЧ-1

Д. А. Воробьев<sup>1✉</sup>, А. Д. Карпенко<sup>1</sup>, А. В. Тузиков<sup>1</sup>, А. М. Андрианов<sup>2</sup>

<sup>1</sup>Объединенный институт проблем информатики  
Национальной академии наук Беларуси,  
ул. Сурганова, 6, Минск, 220012, Беларусь  
✉E-mail: [daniel.vorobiov.2002@yandex.ru](mailto:daniel.vorobiov.2002@yandex.ru)

<sup>2</sup>Институт биоорганической химии  
Национальной академии наук Беларуси,  
ул. академика В. Ф. Купревича, 5, корп. 2, Минск, 220141, Беларусь

### Аннотация.

**Цели.** Основной целью настоящей работы является адаптация архитектуры нейронной сети REINVENT для генерации потенциальных ингибиторов белка gp120 оболочки ВИЧ-1 с использованием в процессе обучения с подкреплением молекулярного докинга на графических процессорах.

**Методы.** Для модификации исходной модели сети использован внедренный в процессе обучения с подкреплением молекулярный докинг на графических процессорах и разработан алгоритм, позволяющий преобразовывать генерируемые сетью SMILES представления соединений в необходимый для выполнения докинга формат PDBQT. С целью ускорения обучения нейронной сети в модифицированной версии модели REINVENT использована программа докинга AutoDock-Vina-GPU-2.1, а для уточнения результатов ее работы – процедура переоценки сродства соединений к мишени с помощью оценочной функции RFScore-4.

**Результаты.** С помощью модифицированной версии модели REINVENT получено более 60 000 соединений, из которых около 52 000 молекул имеют величину энергии связывания с белком gp120 ВИЧ-1, сопоставимую со значением, рассчитанным для ингибитора ВИЧ-1 NBD-14204, использованного в расчетах в качестве позитивного контроля. Из отобранных 52 000 соединений около 34 000 молекул удовлетворяют ограничениям, налагаемым на потенциальное лекарство для обеспечения его биодоступности при пероральном приеме.

**Заключение.** Полученные результаты позволяют продемонстрировать эффективность адаптированной нейронной сети на примере конструирования новых потенциальных ингибиторов белка gp120 ВИЧ-1, способных блокировать CD4-связывающий сайт белка gp120 оболочки вируса и предотвращать его проникновение в клетки хозяина.

**Ключевые слова:** генеративные модели ИИ, обучение с подкреплением, компьютерный дизайн лекарств, молекулярный докинг, ВИЧ-1, белок gp120, анти-ВИЧ-препараты

**Благодарности.** Работа выполнена при поддержке Государственной программы научных исследований «Конвергенция 2025» (подпрограмма «Междисциплинарные исследования и новые технологии», задание 3.04.1).

**Для цитирования.** Адаптация архитектуры нейронной сети REINVENT для генерации потенциальных ингибиторов проникновения ВИЧ-1 / Д. А. Воробьев [и др.] // Информатика. – 2024. – Т. 21, № 3. – С. 80–93. <https://doi.org/10.37661/1816-0301-2024-21-3-80-93>

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию | Received 01.07.2024

Подписана в печать | Accepted 28.08.2024

Опубликована | Published 30.09.2024

## Adaptation of the REINVENT neural network architecture to generate potential HIV-1 entry inhibitors

D. A. Varabyeu<sup>1✉</sup>, A. D. Karpenko<sup>1</sup>, A. V. Tuzikov<sup>1</sup>, A. M. Andrianov<sup>2</sup>

<sup>1</sup>*The United Institute of Informatics Problems  
of the National Academy of Sciences of Belarus,  
st. Surganova, 6, Minsk, 220012, Belarus*

✉E-mail: [daniel.vorobiov.2002@yandex.ru](mailto:daniel.vorobiov.2002@yandex.ru)

<sup>2</sup>*Institute of Bioorganic Chemistry  
of the National Academy of Sciences of Belarus,  
st. Academician V. F. Kuprevich, 5, bldg. 2, Minsk, 220141, Belarus*

### Abstract

**Objectives.** The main purpose of this work is to adapt the architecture of the REINVENT neural network to generate potential inhibitors of the HIV-1 envelope protein gp120 using in the learning process with reinforcement of molecular docking on GPUs.

**Methods.** To modify the initial network model, molecular docking on GPUs implemented in the learning process with reinforcement was used, and an algorithm was developed that allows converting the representations of connections generated by the SMILES network into the PDBQT format necessary for docking. To accelerate the learning of the neural network in the modified version of the REINVENT model, the AutoDock-Vina-GPU-2.1 docking program was used, and to clarify the results of its work, the procedure for revaluing the affinity of compounds to the target using the RFScore-4 evaluation function was used.

**Results.** Using a modified version of the REINVENT model, more than 60,000 compounds were obtained, of which about 52,000 molecules have a binding energy value to the HIV-1 gp120 protein comparable to the value calculated for the HIV-1 inhibitor NBD-14204, used in calculations as a positive control. Of the 52,000 compounds selected, about 34,000 molecules satisfy the restrictions imposed on a potential drug to ensure its bioavailability when taken orally.

**Conclusion.** The results obtained allow us to demonstrate the effectiveness of an adapted neural network by the example of designing new potential inhibitors of the gp120 HIV-1 protein capable of blocking the CD4-binding site of the gp120 virus envelope protein and preventing its penetration into host cells.

**Keywords:** generative AI, reinforcement learning, computer-aided drug design, molecular docking, HIV-1, gp120 protein, anti-HIV drugs

**Acknowledgments.** The work was carried out with the support of the State Scientific Research Program "Convergence 2025" (subprogram "Interdisciplinary Research and New Technologies", task 3.04.1).

**For citation.** Varabyeu D. A., Karpenko A. D., Tuzikov A. V., Andrianov A. M. *Adaptation of the REINVENT neural network architecture to generate potential HIV-1 entry inhibitors*. Informatika [Informatics], 2024, vol. 21, no. 3, pp. 80–93 (In Russ.). <https://doi.org/10.37661/1816-0301-2024-21-3-80-93>

**Conflict of interest.** The authors declare of no conflict of interest.

**Введение.** Разработка эффективных и безопасных лекарственных средств является чрезвычайно важной и востребованной задачей мирового здравоохранения. Многие заболевания, которые ранее считались неизлечимыми, а причины и условия их возникновения оставались неясными, сегодня успешно исследуются с помощью компьютерных технологий, позволяющих значительно сократить временные и финансовые затраты, необходимые для создания новых лекарственных препаратов. В последние годы эти технологии все чаще применяются на всех этапах разработки лекарственных средств, включая виртуальный скрининг потенциальных лекарств [1–6], перепрофилирование лекарств [7, 8], предсказание аффинности связывания малых молекул с белковой мишенью и профиля межмолекулярных взаимодействий [9–12].

Использование алгоритмов глубокого обучения может существенно повысить эффективность проводимых исследований и, кроме того, сконструировать соединения, еще не зарегистрированные в существующих химических базах данных [5, 6], открывая таким образом многообещающие перспективы для создания новых низкомолекулярных соединений, которые могли бы служить более мощными ингибиторами терапевтических мишеней, чем доступные в настоящее время лекарственные препараты. Однако при создании новой модели нейронной сети для дизайна потенциальных лекарств исследователи неизменно сталкиваются с одной и той же проблемой, связанной с необходимостью оперировать большими массивами обучающих данных, что возможно в полной мере лишь для очень ограниченного числа белковых мишеней. В связи с этим перспективным методом разработки лекарств становится обучение с подкреплением [13–15].

Концепция обучения с подкреплением использует два основных понятия – агента и среды. Агенту обучения, имитирующему некоторую сущность, приходится сталкиваться с определенной средой, имитирующей ряд реальных процессов, и по мере взаимодействия с ней достигать большего вознаграждения, развивая те действия, которые дают положительное подкрепление, и избегая действий, влекущих за собой штрафы. При этом в качестве «действия» агента может выступать упорядоченный набор символов, в том числе и представления молекул. Основным преимуществом обучения с подкреплением служит то, что такая модель часто не требует специализированного набора обучающих данных (этот факт является серьезной проблемой применения технологий глубокого обучения), а нуждается только в специфической среде, позволяющей выделить среди совершенных агентом действий те, которые наиболее подходят для реальных условий.

На сегодняшний день одной из наиболее перспективных моделей нейронной сети для дизайна потенциальных лекарств с помощью обучения с подкреплением является серия платформ REINVENT [14–16]. В данной модели реализовано множество способов задания обучающей среды, в том числе с помощью оценки сродства низкомолекулярного соединения к белковой мишени, общепризнанным критерием которого является минимизация значения свободной энергии связывания ( $\Delta G$ ). Большинство потенциальных лекарственных средств, разрабатываемых с помощью методов компьютерного моделирования, проходят процедуру прогнозирования величины  $\Delta G$ , которая обычно выполняется с использованием алгоритмов молекулярного докинга [17].

В модели нейронной сети Reinvent для проведения молекулярного докинга применяется платформа DockStream [18], объединяющая несколько методов молекулярного докинга в один вычислительный инструмент для обучения с подкреплением. Однако все используемые в DockStream методы молекулярного докинга реализованы на центральных процессорах, что с учетом трудоемкости применяемых ими алгоритмов вынуждает пользователей либо прибегать к использованию больших вычислительных мощностей (что не всегда возможно), либо понижать точность предсказания энергии связывания для сокращения времени вычислений,

а это может значительно повлиять на результат обучения агента. На сегодняшний день разработаны алгоритмы молекулярного докинга с использованием графических процессоров, позволяющие значительно ускорить время расчетов, не снижая точности прогнозирования значения  $\Delta G$ .

Цель настоящей работы заключается в адаптации архитектуры нейронной сети REINVENT для генерации потенциальных ингибиторов белка gp120 оболочки ВИЧ-1 с использованием в процессе обучения с подкреплением молекулярного докинга на графических процессорах. Для достижения этой цели в работе проведены исследования, которые включали:

- разработку модифицированной версии REINVENT, использующей для ускорения обучения с подкреплением программу AutoDock-Vina-GPU-2.1 [19];
- создание программного конвейера для конвертации генерируемых нейронной сетью SMILES-строк молекул в формат PDBQT;
- обучение модели с привлечением в качестве терапевтической мишени структуры белка gp120 ВИЧ-1 – гликопротеина оболочки вируса, критически важного для его проникновения в клетки хозяина;
- тестирование нейронной сети и анализ результатов ее работы.

## 1. Материалы и методы

**1.1. Архитектура нейронной сети.** В качестве базовой модели использовали нейронную сеть SMILES-RL [20], входящую в состав архитектуры REINVENT [15] и представляющую различные реализации обучения с подкреплением. Модель SMILES-RL является рекуррентной нейронной сетью со слоями LSTM, в которой реализованы 29 комбинаций алгоритмов обучения с подкреплением и буферов воспроизведения. Как и в большинстве предложенных в литературе моделей обучения с подкреплением, процесс обучения сети SMILES-RL происходил циклично (здесь и далее один цикл обучения будем называть шагом). Для каждого шага генерировали 128 строк SMILES (Simplified Molecular Input Line Entry System) [21] – формата представления низкомолекулярных соединений в виде строк, состоящих из букв, обозначающих химические элементы, и специальных символов.

Для достижения цели исследования авторами был выбран алгоритм обучения с подкреплением Advantage Actor-Critic (A2C) [22] стохастического метода Policy Gradient в сочетании с буфером воспроизведения AC, который использует для обучения агента все сгенерированные на текущем шаге последовательности и учитывает только опыт текущего шага. Выбор алгоритма A2C и буфера воспроизведения AC обусловлен тем, что их комбинация обеспечивала генерацию одного из наибольших наборов потенциально активных молекул, полученных с помощью базовой модели SMILES-RL [20]. Авторы работы [20] для оценки сгенерированных нейронной сетью SMILES-RL соединений на каждом шаге обучения применяли алгоритм случайного леса (Random Forest, RF), с помощью которого каждому полученному соединению ставили в соответствие число-вознаграждение, варьирующее от 0 до 1. В модели нейронной сети, реализованной в настоящей работе, этап оценки был полностью изменен. Так, в предлагаемой версии модели REINVENT вместо алгоритма Random Forest был реализован другой алгоритм оценки сгенерированных соединений (см. подразд. 1.2). При этом все остальные этапы обучения нейронной сети проводили по схеме, реализованной в базовой модели SMILES-RL [20].

**1.2. Обучение нейронной сети.** Процесс обучения длился 10 полных эр и одну нулевую эру, каждая из полных эр включала 50 шагов. Нулевая эра состояла всего из одного шага, и ее вводили для проверки корректности работы алгоритма генерации и оценки соединений. Это было необходимо, так как обучение полной эры, т. е. проведение 50 шагов обучения, занимало в среднем около 6 ч. В конце каждой эры веса для агента сохраняли и в начале следующей эры загружали в качестве исходных данных для агента. В начале нулевой эры использовали случайные веса для неспецифической модели, предварительно обученной генерировать правильные представления SMILES.

На каждом шаге обучения сгенерированные последовательности SMILES конвертировали в формат PDBQT (Protein Data Bank, Partial Charge (Q) & Atom Type (T)), необходимый для проведения молекулярного докинга (см. подразд. 1.3). Далее с помощью программы AutoDock

Vina-GPU-2.1, работающей на графических процессорах, для всех корректно конвертированных соединений проводили молекулярный докинг с белком gp120 ВИЧ-1 (см. подразд. 1.5). Затем результаты докинга для каждого соединения подвергали переоценке с помощью оценочной функции (ОФ) машинного обучения RFScore-4 для того, чтобы уточнить прогнозируемое значение  $\Delta G$  (см. подразд. 1.6). Исходя из значения  $\Delta G$ , каждому соединению ставили в соответствие число-вознаграждение от 0 до 1 (см. подразд. 1.7). Если сгенерированная последовательность SMILES-строк не могла быть конвертирована в формат PDBQT, то этой последовательности ставили в соответствие вознаграждение 0. Все соединения проверяли также на соответствие заданным ограничениям на физико-химические параметры, традиционно используемые в виртуальном скрининге наиболее вероятных кандидатов на лекарственные средства для предсказания молекул с большим потенциалом биодоступности при пероральном введении (см. подразд. 1.8). Если соединение не удовлетворяло ограничениям какого-либо физико-химического параметра, то вознаграждение для этого соединения умножали на соответствующий штраф – число от 0 до 1.

**1.3. Программный конвейер для конвертации строк SMILES в формат PDBQT.** Форматом выходных данных для выбранной модели нейронной сети являются строки SMILES [21], однако программный пакет AutoDock-Vina-GPU-2.1 оперирует с данными в формате PDBQT. Для создания представлений сгенерированных соединений в формате PDBQT авторами был разработан программный конвейер на языке программирования python (рис. 1).

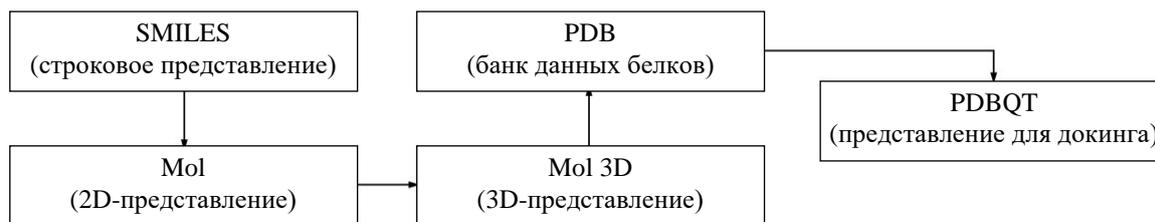


Рис. 1. Общая схема программного конвейера для подготовки соединений к графически ускоренному молекулярному докингу

Fig. 1. The software pipeline general scheme to prepare compounds for graphically accelerated molecular docking

На первом этапе с помощью программного пакета RDKit (URL: <http://www.rdkit.org/>) генерировали двухмерные (2D) структуры молекул с сохранением их топологии, состава и типов химических связей в файлах с расширением .mol, которые преобразовывали в трехмерные (3D) структуры в файлах .mol3d (рис. 1). Затем с привлечением программных средств MGLtools полученные файлы .mol3d конвертировали в формат PDB, используемый для хранения соединений в банке данных белков (Protein Data Bank [23]), и переводили их в формат PDBQT.

**1.4. Молекулярный докинг.** Кристаллическую структуру белка gp120 ВИЧ-1 в комплексе с контрольным ингибитором NBD-14204 заимствовали из Банка данных белков (PDB ID: 8F9Z) [24]. Подготовку белка-мишени к молекулярному докингу осуществляли с помощью пакета UCSF Chimera [25] и программных средств веб-сервера открытого доступа SWISS-MODEL [26]. Атомы водорода добавляли к структуре белка gp120 с использованием программы UCSF Chimera [25], а исходный файл преобразовывали из формата .pdb в формат .pdbqt с помощью программного пакета AutoDockTools-1.5.7 (<https://ccsb.scripps.edu/mgltools/1-5-7/>). Молекулярный докинг проводили с использованием программы AutoDock-Vina-GPU-2.1 [19] в приближении жесткого рецептора и гибких лигандов. Ячейка для докинга включала CD4-связывающий сайт белка gp120 ВИЧ-1 [24] и имела следующие параметры:  $\Delta X = 21,5 \text{ \AA}$ ,  $\Delta Y = 21 \text{ \AA}$ ,  $\Delta Z = 25 \text{ \AA}$  с центром в точках  $X = 2,4 \text{ \AA}$ ,  $Y = 16,57 \text{ \AA}$ ,  $Z = 11,95 \text{ \AA}$ . Параметр, характеризующий полноту поиска (охват конформационного пространства), был задан равным 30. В качестве позитивного контроля в расчетах использовали ингибитор проникновения ВИЧ-1 NBD-14204, блокирующий связывание вируса с первичным рецептором CD4 клеточной мембраны и его проникновение в клетки-мишени [24].

**1.5. Сравнение работы программ AutoDock-Vina и AutoDock-Vina-GPU-2.1.** Для реализации молекулярного докинга было крайне важно сократить время его выполнения, поскольку даже при графическом ускорении расчетов обучение на 50 шагах занимало около 6 ч. Более длительные вычисления или меньший охват конформационного пространства сделали бы использование этой процедуры на этапе обучения нейронной сети нецелесообразным. Для решения данной проблемы может быть использован молекулярный докинг на графических процессорах. Однако в этом случае необходимо показать, что его применение как минимум не уменьшает точность прогнозирования энергии связывания по сравнению с традиционно используемой программой молекулярного докинга AutoDock-Vina 1.2. Поэтому, для того чтобы оценить эффективность докинга на графических процессорах, на наборе данных с известными экспериментальными значениями констант диссоциации ( $K_d$ ) комплексов лиганд-белок был проведен сравнительный анализ точности предсказания аффинности связывания малых молекул с белками-мишенями с использованием программ AutoDock-Vina 1.2.5 и AutoDock-Vina-GPU-2.1 [19]. AutoDock-Vina-GPU-2.1 базируется на алгоритмах программы AutoDock-Vina 1.0, адаптированных для проведения расчетов на графических процессорах [19]. Для сравнения точности прогнозирования величины  $K_d$  (моль/л) с помощью данных программ использовалась база данных Binding MOAD [27], из которой были отобраны кристаллические структуры комплексов лиганд-белок с разрешением  $< 2,5 \text{ \AA}$ , рекомендованным для корректного выполнения молекулярного докинга. Затем для отобранных из этой базы данных 1967 комплексов с помощью ОФ AutoDock-Vina [17] и AutoDockVina-GPU-2.1 [19] были предсказаны значения  $pK_d = -\log(K_d)$  и выполнен повторный молекулярный докинг с использованием программы AutoDock-Vina-GPU-2.1 для оценки разброса данных, полученных одним и тем же методом при изменении случайного начального значения. Результаты проведенных расчетов представлены на рис. 2.

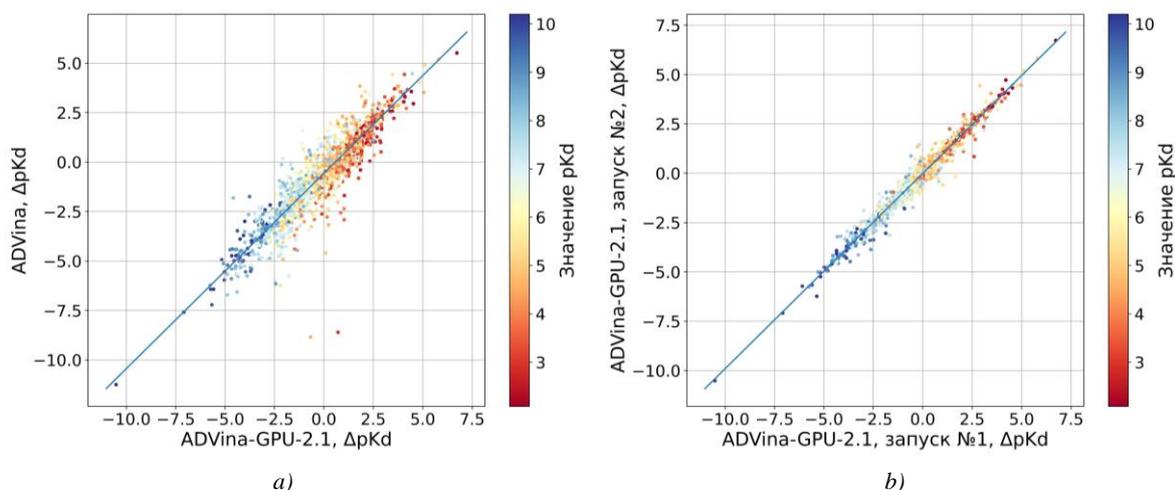


Рис. 2. Отклонения значений  $pK_d$  ( $\Delta pK_d$ ), рассчитанных программами AutoDock-Vina и AutoDock-Vina-GPU-2.1, от известных экспериментальных значений ( $pK_d(\text{рассчит.}) - pK_d(\text{эксперим.})$ ) (a) и отклонения значений  $pK_d$  для того же набора данных, предсказанные AutoDock-Vina-GPU-2.1 в результате двух расчетов с разными случайными начальными значениями (b)

Fig. 2.  $pK_d$  values deviations ( $\Delta pK_d$ ) calculated by AutoDock-Vina and AutoDock-Vina-GPU-2.1 from known experimental values ( $pK_d(\text{calc}) - pK_d(\text{exp})$ ) (a) and  $pK_d$  values deviations for the same data set predicted by AutoDock-Vina-GPU-2.1 as a result of two calculations with different random initial values (b)

**1.6. Сравнение оценочных функций AutoDock-Vina, AutoDock-Vina-GPU-2.1 и RFScore-4.** После проведения молекулярного докинга для уточнения полученных результатов и переоценки  $K_d$  была использована ОФ машинного обучения RFScore-4 [28], основанная на алгоритме случайного леса и предназначенная для виртуального скрининга химических баз данных. Отклонения значений  $pK_d$ , предсказанных с помощью ОФ RFScore-4 и AutoDock-Vina, от экспериментальных величин ( $\Delta pK_d$ ) представлены на рис. 3.

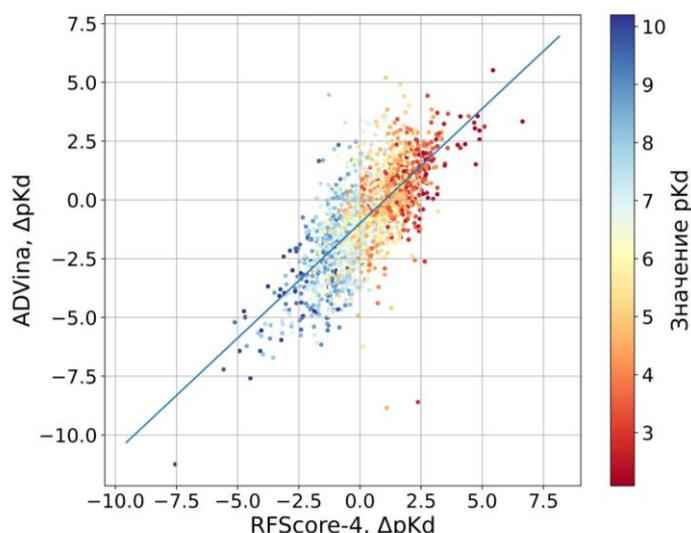


Рис. 3. Отклонения значений  $pK_d$ , рассчитанных с помощью ОФ AutoDock-Vina и RFScore-4, от экспериментальных величин ( $\Delta pK_d$ )

Fig. 3.  $pK_d$  values deviations calculated by the SF AutoDock-Vina and RFScore-4 from the experimental data ( $\Delta pK_d$ )

Средние значения отклонений  $pK_d$  между предсказанными и экспериментальными величинами и соответствующие среднеквадратические отклонения свидетельствуют о том (табл. 1), что графически ускоренный молекулярный докинг AutoDock-Vina-GPU-2.1 показал на использованном наборе экспериментальных данных результаты, сопоставимые с результатами, полученными с помощью программы AutoDock-Vina. Несмотря на то что для ОФ RFScore-4 наблюдается большее среднее отклонение предсказанных значений, чем для ОФ AutoDock-Vina-GPU-2.1, ОФ RFScore-4 характеризуется меньшим среднеквадратическим отклонением между экспериментальными и предсказанными величинами аффинности связывания, а также большим значением коэффициента корреляции Пирсона, что чрезвычайно важно для успешного решения задачи обучения нейронной сети (табл. 1).

Таблица 1  
Сравнение оценочных функций

Table 1  
Comparison of scoring functions

Оценочная функция <i>Evaluation function</i>	Среднее отклонение, $\Delta pK_d$ <i>Average deviation, <math>\Delta pK_d</math></i>	Среднеквадратическое отклонение, $\Delta pK_d$ <i>Standard deviation, <math>\Delta pK_d</math></i>	Коэффициент корреляции Пирсона <i>Pearson's correlation coefficient</i>	$p$ -уровень значимости для коэффициента корреляции Пирсона <i>p-level of significance for Pearson correlation coefficient</i>
AutoDock-Vina	-0,736	1,897	0,469	$1,69 \cdot 10^{-108}$
AutoDock-Vina-GPU-2.1	-0,169	1,769	0,476	$6,13 \cdot 10^{-112}$
RFScore-4	0,266	1,411	0,661	$7,37 \cdot 10^{-248}$

**1.7. Функция вознаграждения.** Для применения молекулярного докинга в процессе обучения агента необходимо было преобразовать его выходные данные, а именно значения  $\Delta G$ , в величины, с которыми оперирует нейронная сеть. Так как базовая модель [20] использует в качестве вознаграждения за каждое сгенерированное соединение реальное значение в интервале от

0 до 1, было решено конвертировать величины  $\Delta G$  в ту же форму. Для этой цели было удобно использовать асимптотические функции, которые достигают значения 0 или 1 только при бесконечно высоких или низких величинах аргумента соответственно, поскольку в таком случае отсутствует вероятность ошибки при получении неожиданно высокого или низкого значения  $\Delta G$ .

Для решения поставленной задачи была выбрана модифицированная кумулятивная функция нормального распределения (1) и (2) в предположении, что значения  $\Delta G$  для сгенерированных соединений имеют нормальное распределение. Функция (1) отличается от кумулятивной функции нормального распределения тем, что для нее большие значения  $F(x)$  соответствуют меньшему значению  $x$ . Это изменение связано с тем, что большему значению аффинности связывания, на которое в данной работе было ориентировано обучение модели, соответствует меньшее значение  $\Delta G$ :

$$F(x) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{-x + \mu}{\sigma\sqrt{2}} \right) \right], \quad (1)$$

где  $x$  – значение  $\Delta G$  для определенного соединения;  $\mu$  – среднее значение  $\Delta G$  (значение, при котором вознаграждение  $F(x) = 0,5$ ), выбранное таким образом, чтобы величина вознаграждения для контрольного ингибитора находилась в диапазоне от 0,5 до 0,6;  $\sigma$  – стандартное отклонение, подобранное эмпирически так, чтобы «энергетическая ценность» вознаграждения  $F(x) = 0,95$  находилась в диапазоне от 10,5 до 11,0 ккал/моль.

Функция  $\operatorname{erf}(z)$  определяется следующим образом:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt. \quad (2)$$

Следует отметить, что чем ниже значение  $\Delta G$ , тем сильнее средство, поскольку свободная энергия связывания представляет собой разницу между внутренней энергией комплекса лиганда с белком и суммой их энергий в свободном состоянии. Используемая для оценки  $\Delta G$  функция вознаграждения представлена на рис. 4.

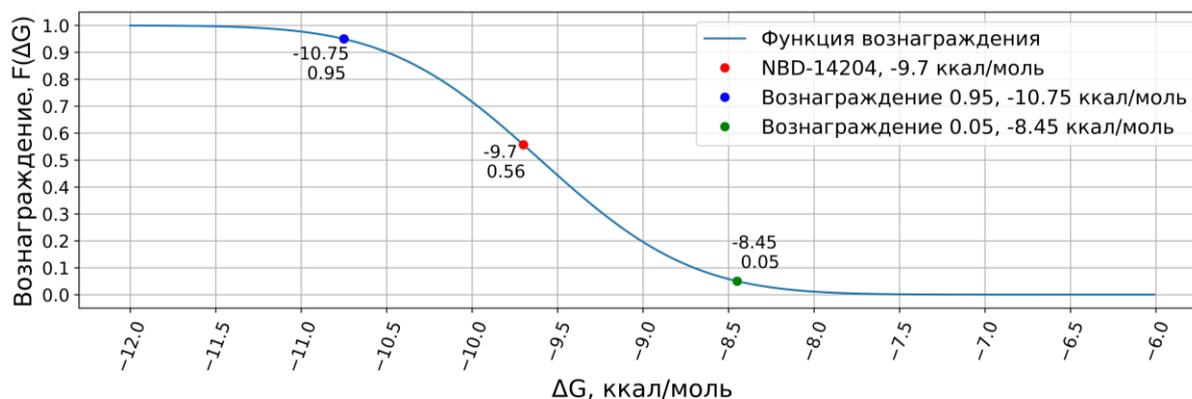


Рис. 4. Функция, использованная в процессе обучения с подкреплением для преобразования значений  $\Delta G$  в вознаграждение

Fig. 4. Function, used for converting the values of binding free energy into reward during the reinforcement learning

**1.8. Дополнительные ограничения.** Кроме ограничений на значения  $\Delta G$ , на генерируемые соединения налагались штрафы за отклонение физико-химических параметров молекул от требований, предъявляемых к ним правилом пяти Липинского [29], которое используется для предсказания молекул с большим потенциалом биодоступности при пероральном введении.

Правило пяти принимает в рассмотрение такие важные характеристики потенциального лекарства, как всасывание, распределение, метаболизм и экскреция. Кроме того, в функцию вознаграждения вводили штрафы за нарушение ограничений на топологическую площадь полярной поверхности, число вращательных связей, логарифм молярной растворимости соединений в воде и их реакционную способность (степень ненасыщенности), которые также оказывают существенное влияние на оральную биодоступность кандидатов в лекарства [29–31]. Полный список ограничений на физико-химические параметры, использованных в процессе обучения нейронной сети, представлен в табл. 2.

Таблица 2

Штрафные коэффициенты, введенные в функцию вознаграждения за нарушение ограничений на физико-химические параметры соединений

Table 2

Penalty coefficients introduced into the reward function for violating restrictions on the physicochemical parameters of compounds

Физико-химические параметры соединений <i>Physicochemical parameters of compounds</i>	Ограничения на физико-химические параметры <i>Limitations on physical and chemical parameters</i>		Штрафной коэффициент <i>Penalty coefficient</i>
Правило пяти Липински	Молекулярная масса (Да)	$\leq 500$	0,8
	Число доноров водородной связи	$\leq 5$	
	Число акцепторов водородной связи	$\leq 10$	
	Липофильность соединения	$-0,7 < XLOGP3 < 5,0$	
	При этом допускалось нарушение не более одного условия		
Топологическая площадь полярной поверхности (TPSA)	$20 \text{ \AA}^2 < \text{TPSA} < 130 \text{ \AA}^2$		0,8
Число вращательных связей (N)	$N \leq 10$		0,8
Логарифм молярной растворимости соединения в воде (LogS)	$-6 < \text{Log S} < 0$		0,8
Ненасыщенность соединения <sup>1,2</sup> (Csp3)	$0,25 < \text{Csp3} < 1$		$\frac{< \text{Csp3} >}{0,25}$

<sup>1</sup>Ненасыщенные соединения (непредельные соединения) – органические молекулы, скелет которых содержит кратные (C=C, C≡C) связи. Ненасыщенные соединения характеризуются высокой реакционной способностью.

<sup>2</sup>Штраф за нарушение ограничений, относящихся к ненасыщенности соединений, отличается от остальных. Он может быть более строгим, так как при обучении сети без наложения штрафных санкций, свыше 85 % сгенерированных соединений не соответствовали этим ограничениям, в то время как по остальным параметрам не удовлетворяло требованиям от десятой части до двух третей соединений.

<sup>1</sup>Unsaturated compounds (unsaturated compounds) are organic molecules whose skeleton contains multiple (C=C, C≡C) bonds. Unsaturated compounds are characterized by high reactivity.

<sup>2</sup>The penalty for violating the restrictions related to the unsaturation of compounds differs from the others. It can be more stringent, since when training the network without imposing penalties, over 85 % of the generated compounds did not meet these restrictions, while for other parameters, from a tenth to two thirds of the compounds did not meet the requirements.

Итоговое вознаграждение для каждого соединения рассчитывали путем умножения вознаграждения, полученного на основе данных молекулярного докинга, на штрафы, наложенные за отклонения физико-химических параметров от установленных ограничений:

$$\text{Reward} = F(\Delta G) \prod_i x_i, \quad (3)$$

где  $x_i$  – штрафной коэффициент для  $i$ -го параметра, равный значению в табл. 2, если ограничение нарушено, и единице, если соединение удовлетворяет ограничениям  $i$ -го параметра.  $F(\Delta G)$  вычисляется по формуле (1) для величины  $\Delta G$ .

**2. Результаты и обсуждение.** Процесс обучения нейронной сети длился около 60 ч с использованием графического процессора RTX 3080 TI и включал 500 шагов, разделенных на 10 эр, в течение которых нейронная сеть генерировала в среднем от 120 до 128 оригинальных (без учета дубликатов) соединений за шаг. Потери в количестве соединений связаны с тем, что не все сгенерированные нейронной сетью последовательности символов из словаря SMILES удалось преобразовать по правилам SMILES в структуры химических соединений. Соответственно, такие последовательности символов не могли быть корректно конвертированы программным конвейером в PDBQT-формат и их пришлось отбросить. Средняя величина итоговой функции вознаграждения за каждый шаг показана на рис. 5.

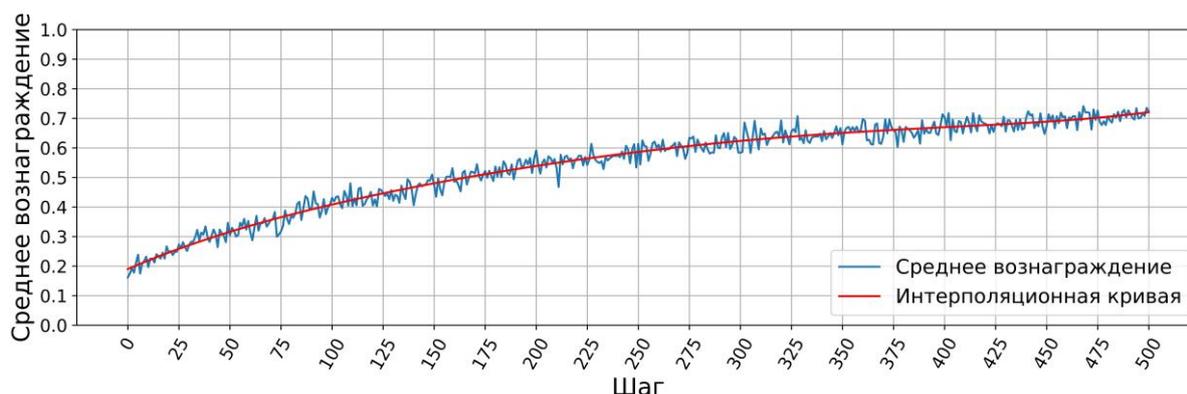


Рис. 5. Средние значения функции вознаграждения для сгенерированных на каждом шаге обучения SMILES-строк

Fig. 5. Mean values of reward function for generated SMILES strings per step of training

В результате нейронная сеть сгенерировала более 62 000 соединений, из которых свыше 52 000 молекул имели значения  $\Delta G$  ниже  $-9,2$  ккал/моль, сопоставимые с величиной  $-9,7$  ккал/моль, полученной для контрольного ингибитора ВИЧ-1 NDB-14204. На рис. 6 изображен график зависимости числа генерируемых соединений за шаг, которые удовлетворяли ограничениям, наложенным на их физико-химические параметры, и имели величину  $\Delta G$  ниже  $-9,2$  ккал/моль. Итоговое число таких соединений превысило 34 000.

Следует отметить, что в работе также было проведено обучение модели с использованием ОФ AutoDock-Vina-GPU-2.1, но без учета ограничений на физико-химические параметры молекул. Когда полученные таким путем молекулы были отобраны по критерию соответствия введенным ограничениям, то их количество составило менее 3000. Это связано с тем, что при обучении сети без дополнительных ограничений штрафование соединений, которые не удовлетворяют этим ограничениям, отсутствует и единственным ориентиром для сети становятся соединения с наименьшей величиной  $\Delta G$ . Однако при наложении штрафов за нарушения сеть обучается избегать молекулы с отклонениями от заданных параметров и учится ориентироваться на соединения, удовлетворяющие ограничениям, из-за чего с каждой эрой их количество возрастает.

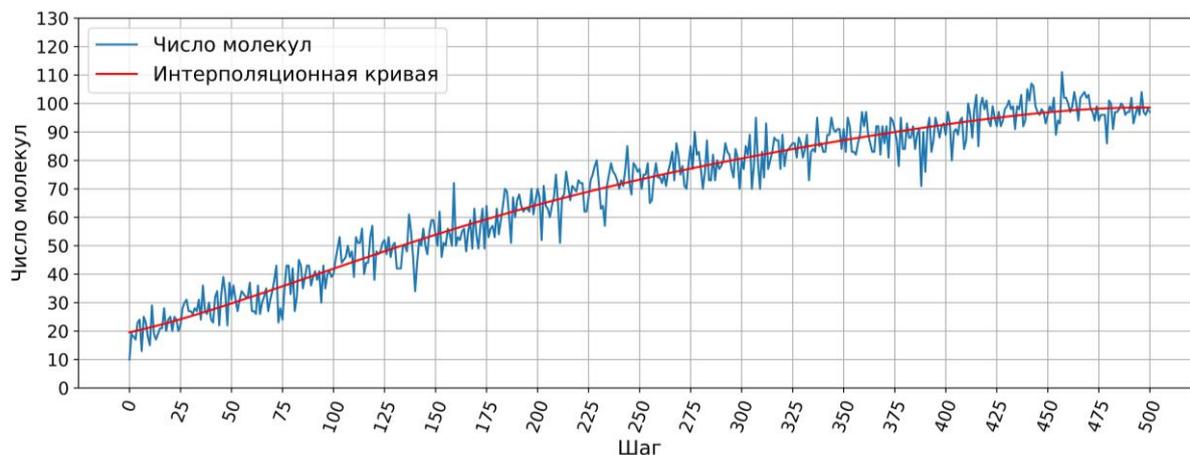


Рис. 6. Число сгенерированных за шаг обучения нейронной сети молекул, удовлетворяющих всем заданным ограничениям на физико-химические параметры и имеющих значения  $\Delta G$ , более низкие или сопоставимые с величиной, рассчитанной для контрольного ингибитора NBD-14204

*Fig. 6. Number of molecules that meet all requirements imposed on physicochemical parameters and exhibit values of  $\Delta G$  comparable or lower than that calculated for the reference inhibitor NBD-14204*

Также важно отметить, что уменьшение свободной энергии связывания для генерируемых соединений во время обучения слабо связано с тенденцией к росту их молекулярной массы. Это может существенно влиять на величину молекулярной массы по крайней мере из-за того, что с увеличением числа атомов возрастает количество возможных межмолекулярных взаимодействий между лигандом и белком-мишенью. Большинство соединений как с очень большими, так и с очень малыми молекулярными массами были получены в эры 1 и 2, т. е. когда агент был еще слабо специфичен к целевому белку gp120. Кроме того, вклад каждой эры в конечное распределение молекулярной массы скорее одинаков, с небольшим перевесом большей массы для поздних эр, в то время как для конечного распределения свободной энергии связывания ее более низкие значения явно соответствуют более поздним эрам (рис. 7).

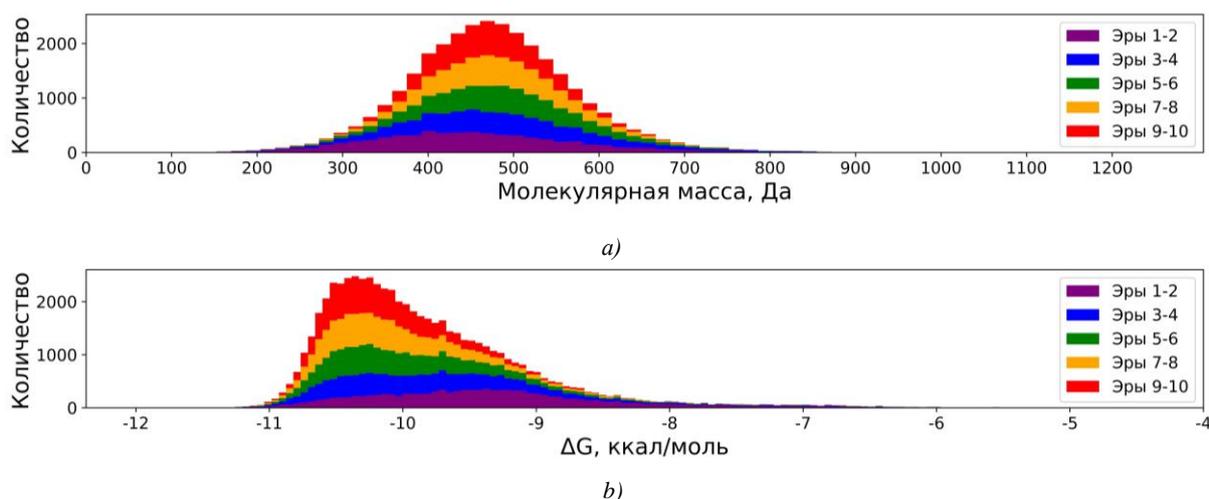


Рис. 7. Распределение молекулярных масс (a) и свободной энергии связывания (b) с указанием вклада соответствующих эр обучения

*Fig. 7. The molecular weights distribution (a) and binding free energy (b) indicating the contribution of the corresponding eras of training*

**Заключение.** В настоящем исследовании адаптирована модель глубокой генеративной нейронной сети на основе архитектуры REINVENT, использующая в процессе обучения с подкреплением графически ускоренный молекулярный докинг и учитывающая ограничения на физико-химические параметры генерируемых соединений, введенные для идентификации молекул с большим потенциалом оральной биодоступности. Для решения поставленной задачи создан программный конвейер, предназначенный для преобразования SMILES-строк молекул в формат PDBQT. Для преобразования результатов молекулярного докинга в данные, приемлемые для обучения с подкреплением, использована функция вознаграждения, основанная на кумулятивной функции нормального распределения. На наборе данных для комплексов лиганд-белок с известными экспериментальными значениями  $K_d$  проведен сравнительный анализ точности предсказания значений  $\Delta G$  с помощью программ AutoDock-Vina и AutoDock-Vina-GPU-2.1 и ОФ Vina, Vina-GPU-2.1 и RFScore-4. В результате показано, что использование графически ускоренного метода AutoDock-Vina-GPU-2.1 в сочетании с ОФ RFScore-4 является наиболее продуктивным для успешного обучения предлагаемой модели. Полученная модификация нейронной сети не требует набора обучающих данных и способна генерировать большое количество соединений с высокой прогнозируемой активностью по отношению к заданной терапевтической мишени и приемлемыми фармакологическими свойствами.

Эффективность адаптированной нейронной сети продемонстрирована на примере конструирования новых потенциальных ингибиторов белка gp120 ВИЧ-1, способных блокировать CD4-связывающий сайт белка gp120 оболочки вируса и предотвращать его проникновение в клетки хозяина. В результате проведенных расчетов идентифицировано более 34 000 высокоаффинных лигандов этого гликопротеина, потенциально активных при пероральном приеме. В связи с этим развитие настоящей работы предполагает проведение дальнейших теоретических исследований, включающих молекулярную динамику комплексов лиганд-gp120, отбор соединений, перспективных для тестирования *in vitro*, и оптимизацию структур соединений-лидеров, направленную на получение их аналогов с улучшенной противовирусной активностью.

В заключение отметим, что предлагаемая модель нейронной сети может быть использована для компьютерного дизайна потенциальных кандидатов в лекарственные средства, способных ингибировать разные терапевтические мишени, играющие важную роль в развитии опасных заболеваний человека.

**Вклад авторов.** Д. А. Воробьев и А. Д. Карпенко разработали и реализовали модификацию модели генеративной нейронной сети, обучили, протестировали и апробировали нейронную сеть. Д. А. Воробьев провел молекулярный докинг сгенерированных соединений с белком gp120 ВИЧ-1. А. М. Андрианов и А. В. Тузиков руководили проектом и написали рукопись. Все авторы проанализировали данные расчетов, обсудили полученные результаты и внесли свой вклад в окончательную версию рукописи.

## References

1. Li H., Sze K. H., Lu G., Ballester P. Machine-learning scoring functions for structure-based virtual screening. *Wiley interdisciplinary reviews: Computational Molecular Science*, 2020, vol. 11. Available at: <https://wires.onlinelibrary.wiley.com/doi/10.1002/wcms.1478> (accessed 20.05.2024). <https://doi.org/10.1002/wcms.1478>
2. Xiong G. L., Ye W. L., Shen C., Lu A. P., Hou T. J., Cao D. S. Improving structure-based virtual screening performance via learning from scoring function components. *Briefings in Bioinformatics*, 2020. Available at: <https://academic.oup.com/bib/article/22/3/bbaa094/5851268> (accessed 20.05.2024). <https://doi.org/10.1093/bib/bbaa094>
3. Stokes J. M., Yang K., Swanson K., Jin W., Cubillos-Ruiz A. A deep learning approach to antibiotic discovery. *Cell*, 2020, vol. 180, pp. 688–702. <https://doi.org/10.1016/j.cell.2020.01.021>
4. Timmons P. B., Hewage C. M. ENNAVIA is a novel method which employs neural networks for antiviral and anti-coronavirus activity prediction for therapeutic peptides. *Briefings in Bioinformatics*, 2021. Available at: <https://academic.oup.com/bib/article/22/6/bbab258/6326528> (accessed 20.05.2024). <https://doi.org/10.1093/bib/bbab258>

5. Andrianov A. M., Nikolaev G. I., Shuldov N. A., Bosko I. P., Anischenko A. I., Tuzikov A. V. Application of deep learning and molecular modeling to identify small drug-like compounds as potential HIV-1 entry inhibitors. *Journal of Biomolecular Structure and Dynamics*, 2022, vol. 40, pp. 7555–7573. <https://doi.org/10.1080/07391102.2021.1905559>
6. Zhang Y., Ye T., Xi H., Juhas M., Li J. Deep learning driven drug discovery: Tackling Severe Acute Respiratory Syndrome Coronavirus 2. *Frontiers in Microbiology*, 2021. Available at: <https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.739684/full> (accessed 20.05.2024). <https://doi.org/10.3389/fmicb.2021.739684>
7. Kinnings S. L., Liu N., Tonge P. J., Jackson R. M., Xie L., Bourne P. E. A Machine learning-based method to improve docking scoring functions and its application to drug repurposing. *Journal of Chemical Information and Modeling*, 2011, vol. 51, no. 5, pp. 1195–1197. <https://doi.org/10.1021/ci2001346>
8. Agastheeswaramoorthy K., Sevilimedu A. Drug REpurposing using AI/ML tools – for Rare Diseases (DREAM-RD): A case study with Fragile X Syndrome (FXS). *bioRxiv*, 2020. Available at: <https://www.biorxiv.org/content/10.1101/2020.09.25.311142v1.full.pdf> (accessed 20.05.2024). <https://doi.org/10.1101/2020.09.25.311142>
9. Heinzlmann G., Gilson M. K. Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation. *Scientific Reports*, 2021, vol. 11, no. 1, p. 1116. <https://doi.org/10.1038/s41598-020-80769-1>
10. Meli R., Morris G. M., Biggin P. C. Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: A review. *Frontiers in Bioinformatics*, 2022, vol. 2. Available at: <https://www.frontiersin.org/journals/bioinformatics/articles/10.3389/fbinf.2022.885983/full> (accessed 20.05.2024). <https://doi.org/10.3389/fbinf.2022.885983>
11. Metcalf D. P., Glick Z. L., Bortolato A., Jiang A., Cheney D. L., Sherrill C. D. Directional  $\Delta G$  neural network (Dr  $\Delta G$ -Net): A modular neural network approach to binding free energy prediction. *Journal of Chemical Information and Modeling*, 2024, vol. 64, no. 6, pp. 1907–1918. <https://doi.org/10.1021/acs.jcim.3c02054>
12. Li Y., Fan Z., Rao J., Chen Z., Chu Q., ..., Li X. An overview of recent advances and challenges in predicting compound-protein interaction (CPI). *Medical Review*, 2023, vol. 3, no. 6, pp. 465–486. <https://doi.org/10.1515/mr-2023-0030>
13. Zhavoronkov A., Ivanenkov Y. A., Aliper A., Veselov M. S., Aladinskiy V. A., ..., Aspuru-Guzik A. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology*, 2019, vol. 37, pp. 1038–1040. <https://doi.org/10.1038/s41587-019-0224-x>
14. Olivecrona M., Blaschke T., Engkvist O., Chen H. J. Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics*, 2017, vol. 9, p. 48. <https://doi.org/10.1186/s13321-017-0235-x>
15. Blaschke T., Arús-Pous J., Chen H., Margreitter C., Tyrchan C., ..., Patronov A. Reinvent 2.0: an AI tool for de novo drug design. *Journal of Chemical Information and Modeling*, 2020, vol. 60, no. 12, pp. 5918–5922. <https://doi.org/10.1021/acs.jcim.0c00915>
16. Loeffler H. H., He J., Tibo A., Janet J. P., Voronov A., ..., Engkvist O. Reinvent 4: Modern AI-driven generative molecule design. *Journal of Cheminformatics*, 2024, vol. 16, no. 20. Available at: <https://jcheminf.biomedcentral.com/articles/10.1186/s13321-024-00812-5> (accessed 20.05.2024). <https://doi.org/10.1186/s13321-024-00812-5>
17. Trott O., Olson A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 2010, vol. 31, no. 2, pp. 455–461. <https://doi.org/10.1002/jcc.21334>
18. Guo J., Janet J. P., Bauer M. R., Voronov A., Mervin L. H., Engkvist O. DockStream: a docking wrapper to enhance de novo molecular design. *Journal of Cheminformatics*, 2021, vol. 13. Available at: <https://jcheminf.biomedcentral.com/articles/10.1186/s13321-021-00563-7> (accessed 20.05.2024). <https://doi.org/10.1186/s13321-021-00563-7>
19. Tang S., Ding J., Zhu X., Wang Z., Zhao H., Wu J. Vina-GPU 2.1: towards further optimizing docking speed and precision of AutoDock Vina and its derivatives. *bioRxiv*, 2023. Available at: <https://www.biorxiv.org/content/10.1101/2023.11.04.565429v1.full.pdf> (accessed 20.05.2024). <https://doi.org/10.1101/2023.11.04.565429>
20. Svensson H., Tyrchan C., Engkvist O., Chehreghani M. H. Utilizing Reinforcement learning for de novo drug design. *Machine-Mediated Learning*, 2024, vol. 113, no. 2, pp. 4811–4843. <https://doi.org/10.1007/s10994-024-06519-w>
21. Weininger D. SMILES, a chemical language and information system. *Journal of Chemical Information and Computer Sciences*, 1998, vol. 28, pp. 31–36. <https://doi.org/10.1021/ci00057a005>

22. Mnih V., Badia A. P., Mirza M., Graves A., Lillicrap T. P., ..., Kavukcuoglu K. Asynchronous methods for deep reinforcement learning. *Proceedings of the 33rd International Conference on Machine Learning, New York, USA, 20–22 June 2016*, vol. 48, pp. 1928–1937. <https://proceedings.mlr.press/v48/mniha16.html>
23. Berman H. M., Battistuz T., Bhat T. N., Bluhm W. F., Bourne P., ..., Zardecki C. The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, 2002, vol. 58, no. 6, pp. 899–907. <https://doi.org/10.1093/nar/28.1.235>
24. Curreli F., Kwon Y. D., Nicolau I., Burgos G., Altieri A., ..., Debnath A.K. Antiviral activity and crystal structures of HIV-1 gp120 antagonists. *International Journal of Molecular Sciences*, 2022, vol. 23, no. 24, p. 15999. <https://doi.org/10.3390/ijms232415999>
25. Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., ..., Ferrin T. E. UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 2004, vol. 25, no. 13, pp. 1605–1612. <https://doi.org/10.1002/jcc.20084>
26. Kiefer F., Arnold K., Künzli M., Bordoli L., Schwede T. The Swiss-MODEL Repository and associated resources. *Nucleic Acids Research*, 2009, vol. 37, pp. 387–392. <https://doi.org/10.1093/nar/gkn750>
27. Benson M. L., Smith R. D., Khazanov N. A., Dimcheff B., Beaver J., ..., Carlson H. A. Binding MOAD, a high-quality protein–ligand database. *Nucleic Acids Research*, 2007, vol. 36, pp. 674–678. <https://doi.org/10.1093/nar/gkm911>
28. Wójcikowski M., Ballester P. J., Siedlecki P. Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, 2017, vol. 7, no. 1, pp. 1–10. <https://doi.org/10.1038/srep46710>
29. Lipinski C. A., Lombardo F., Dominy B. W., Feeney P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 2001, vol. 46, no. 1–3, pp. 3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0)
30. Daina A., Michielin O., Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 2017, vol. 7. Available at: <https://www.nature.com/articles/srep42717> (accessed 20.05.2024). <https://doi.org/10.1038/srep42717>
31. Veber D. F., Johnson S. R., Cheng H. Y., Smith B. R., Ward K. W., Kopple K. D. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 2002, vol. 45, no. 12, pp. 2615–2623. <https://doi.org/10.1021/jm020017n>

### Информация об авторах

*Воробьев Данила Александрович*, стажер младшего научного сотрудника, Объединенный институт проблем информатики Национальной академии наук Беларуси.

E-mail: [daniel.vorobiov.2002@yandex.ru](mailto:daniel.vorobiov.2002@yandex.ru)

*Карпенко Анна Дмитриевна*, научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси.

E-mail: [rfe.karpenko@gmail.com](mailto:rfe.karpenko@gmail.com)

*Тузиков Александр Васильевич*, член-корреспондент, доктор физико-математических наук, профессор, заведующий лабораторией математической кибернетики, Объединенный институт проблем информатики Национальной академии наук Беларуси.

E-mail: [tuzikov@newman.bas-net.by](mailto:tuzikov@newman.bas-net.by)

*Андрианов Александр Михайлович*, доктор химических наук, профессор, главный научный сотрудник, Институт биоорганической химии Национальной академии наук Беларуси.

E-mail: [alexande.andriano@yandex.ru](mailto:alexande.andriano@yandex.ru)

### Information about the authors

*Danila A. Varabyeu*, Trainee Junior Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.

E-mail: [daniel.vorobiov.2002@yandex.ru](mailto:daniel.vorobiov.2002@yandex.ru)

*Anna D. Karpenko*, Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.

E-mail: [rfe.karpenko@gmail.com](mailto:rfe.karpenko@gmail.com)

*Alexander V. Tuzikov*, Corresponding Member, D. Sc. (Phys.-Math.), Prof., Head of the Laboratory of Mathematical Cybernetics, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus.

E-mail: [tuzikov@newman.bas-net.by](mailto:tuzikov@newman.bas-net.by)

*Alexander M. Andrianov*, D. Sc. (Chem.), Prof., Principal Researcher, Institute of Bioorganic Chemistry of the National Academy of Sciences of Belarus.

E-mail: [alexande.andriano@yandex.ru](mailto:alexande.andriano@yandex.ru)