

УДК: 544.165: 615.917

**Количественный анализ
«структура – противоопухолевая активность»
и рациональный молекулярный дизайн
бифункциональных VEGFR-2/HDAC-ингибиторов**

**О. В. Тиньков^{1,a}, П. Г. Полищук², Д. С. Хачатрян³, А. В. Колотаев³,
А. Н. Балаев⁴, В. Н. Осипов⁵, В. Ю. Григорьев⁶**

¹ Военный институт Министерства обороны ПМР,
Молдова, 3300, Приднестровье, г. Тирасполь, ул. Гоголя, д. 2В

² Институт молекулярной и трансляционной медицины, факультет медицины и стоматологии,
Университет Палацкого в г. Оломоуц,
Чешская Республика, 779 00, г. Оломоуц, ул. Гневотинская, д. 1333/5

³ Национальный исследовательский центр «Курчатовский институт» —
Институт химических реактивов и особо чистых химических веществ,
Россия, 107076, г. Москва, Богородский вал, д. 3

⁴ Акционерное общество «Фарм-синтез»,
Россия, 121357, г. Москва, ул. Верейская, д. 29, стр. 134

⁵ Национальный медицинский исследовательский центр онкологии имени Н. Н. Блохина
Министерства здравоохранения РФ,
Россия, 115478, г. Москва, Каширское шоссе, д. 23

⁶ Институт физиологически активных веществ Российской академии наук,
Россия, 142432, Московская область, г. Черноголовка, Северный пр., д. 1

E-mail: ^a oleg.tinkov.chem@mail.ru

*Получено 13.06.2019, после доработки — 14.08.2019.
Принято к публикации 06.09.2019.*

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (соглашение № 075-11-2018-172 от 03.12.18). Уникальный идентификатор проекта RFMEFI62418X0051.

УДК: 544.165: 615.917

Количественный анализ «структура – противоопухолевая активность» и рациональный молекулярный дизайн бифункциональных VEGFR-2/HDAC-ингибиторов

О. В. Тиньков, П. Г. Полищук, Д. С. Хачатрян, А. В. Колотаев,
А. Н. Балаев, В. Н. Осипов, В. Ю. Григорьев

Ингибиторы гистондеацетилаз (HDACi) рассматриваются в качестве перспективного класса препаратов для лечения рака из-за их влияния на рост клеток, дифференцировку и апоптоз. Ангиогенез играет важную роль в росте солидных опухолей и развитии метастазов. Фактор роста эндотелия сосудов (VEGF) является ключевым ангиогенным агентом, который секретируется злокачественными опухолями, что индуцирует пролиферацию и миграцию эндотелиальных клеток сосудов. В настоящее время наиболее перспективной стратегией в борьбе с онкологическими заболеваниями является создание гибридных лекарств, одновременно действующих на несколько физиологических мишеней. Значительный интерес с точки зрения создания бифункциональных противоопухолевых средств представляют соединения, содержащие одновременно N-фенил-4-аминохиназолин и гидроксамовую кислоту, так как данные фрагменты по отдельности присутствуют в уже успешно применяемых противоопухолевых лекарственных средствах. В этой связи в ходе литературного анализа была сформирована выборка из 42 соединений, содержащих указанные молекулярные фрагменты и обладающих экспериментальными данными по ингибированию HDAC, VEGFR-2 и росту клеток рака легкого человека MCF-7. С использованием симплексных дескрипторов и метода опорных векторов (Support Vector Machine, SVM) для указанной выборки, предварительно разделенной на обучающий и тестовый наборы, были построены удовлетворительные ($R^2_{\text{test}} = 0.64\text{--}0.87$) модели количественной связи «структура–активность» (Quantitative Structure-Activity Relationship, QSAR). Для полученных QSAR-моделей была проведена структурная интерпретация. Было оценено согласованное влияние различных молекулярных фрагментов на увеличение противоопухолевой активности исследуемых соединений. Среди заместителей N-фенильного фрагмента можно выделить положительный вклад брома в *para*-положении для всех трех видов активности. По результатам интерпретации был проведен рациональный молекулярный дизайн и предложены перспективные соединения. Для сравнительного QSAR-исследования использованы физико-химические дескрипторы, рассчитываемые программой HYBOT, метод случайного леса (Random Forest, RF), а также онлайн-версия экспертной системы OCHEM (<https://ochem.eu>). При моделировании OCHEM были выбраны PyDescriptor-дескрипторы и метод экстремального градиентного бустинга. Кроме того, полученные с помощью экспертной системы OCHEM модели были использованы для виртуального скрининга 300 соединений с целью отбора перспективных VEGFR-2/HDAC-ингибиторов для последующего синтеза и испытаний.

Ключевые слова: ингибиторы гистондеацетилазы, ангиогенез, гибридные молекулы, симплексные дескрипторы, структурная интерпретация, виртуальный скрининг

UDC: 544.165: 615.917

Quantitative analysis of “structure – anticancer activity” and rational molecular design of bi-functional VEGFR-2/HDAC-inhibitors

O. V. Tinkov^{1,a}, P. G. Polishchuk², D. S. Khachatryan³, A. V. Kolotaev³,
A. N. Balaev⁴, V. N. Osipov⁵, V. Yu. Grigorev⁶

¹ Military Institute of the Ministry of Defense,
2B Gogol st., Tiraspol, Transdnistria, 3300, Moldova

² Institute of Molecular and Translational Medicine,
Faculty of Medicine and Dentistry, Palacký University and University Hospital in Olomouc,
1333/5 Hněvotínská st., Olomouc, 779 00, Czech Republic

³ NRC «Kurchatov Institute» – IREA,
3 Bogorodsky val, Moscow, 107076, Russia

⁴ JSC «Pharm-Sintez»,
29/134 Vereyskaya st., Moscow, 121357, Russia

⁵ NMRC of oncology N. N. Blokhin of the Health Ministry of Russia,
23 Kashirskoye Shosse, Moscow, 115478, Russia

⁶ Institute of Physiologically Active Compounds, Russian Academy of Sciences,
1 Northern Passage, Chernogolovka, Moscow oblast, 142432, Russia

E-mail: ^a oleg.tinkov.chem@mail.ru

*Received 13.06.2019, after completion — 14.08.2019.
Accepted for publication 06.09.2019.*

Citation: *Computer Research and Modeling*, 2019, vol. 11, no. 5, pp. 911–930 (Russian).

The work was supported by Ministry of science and higher education of the Russian Federation (agreement No. 075-11-2018-172 of 03.12.18). Unique project ID RFMEFI62418X0051.

UDC: 544.165: 615.917

Quantitative analysis of “structure – anticancer activity” and rational molecular design of bi-functional VEGFR-2/HDAC-inhibitors

O. V. Tinkov, P. G. Polishchuk, D. S. Khachatryan, A. V. Kolotaev,
A. N. Balaev, V. N. Osipov, V. Yu. Grigorev

Inhibitors of histone deacetylases (HDACi) have considered as a promising class of drugs for the treatment of cancers because of their effects on cell growth, differentiation, and apoptosis. Angiogenesis play an important role in the growth of most solid tumors and the progression of metastasis. The vascular endothelial growth factor (VEGF) is a key angiogenic agent, which is secreted by malignant tumors, which induces the proliferation and the migration of vascular endothelial cells. Currently, the most promising strategy in the fight against cancer is the creation of hybrid drugs that simultaneously act on several physiological targets. In this work, a series of hybrids bearing N-phenylquinazolin-4-amine and hydroxamic acid moieties were studied as dual VEGFR-2/HDAC inhibitors using simplex representation of the molecular structure and Support Vector Machine (SVM). The total sample of 42 compounds was divided into training and test sets. Five-fold cross-validation (5-fold) was used for internal validation. Satisfactory quantitative structure—activity relationship (QSAR) models were constructed ($R^2_{\text{test}} = 0.64–0.87$) for inhibitors of HDAC, VEGFR-2 and human breast cancer cell line MCF-7. The interpretation of the obtained QSAR models was carried out. The coordinated effect of different molecular fragments on the increase of antitumor activity of the studied compounds was estimated. Among the substituents of the N-phenyl fragment, the positive contribution of *para* bromine for all three types of activity can be distinguished. The results of the interpretation were used for molecular design of potential dual VEGFR-2/HDAC inhibitors. For comparative QSAR research we used physico-chemical descriptors calculated by the program HYBOT, the method of Random Forest (RF), and on-line version of the expert system OCHEM (<https://ochem.eu>). In the modeling of OCHEM PyDescriptor descriptors and extreme gradient boosting was chosen. In addition, the models obtained with the help of the expert system OCHEM were used for virtual screening of 300 compounds to select promising VEGFR-2/HDAC inhibitors for further synthesis and testing.

Keywords: histone deacetylase inhibitors, angiogenesis, hybrid molecules, simplex descriptors, structural interpretation, virtual screening

1. Введение

В настоящее время, несмотря на активное развитие различных методов лечения онкологических больных, химиотерапия является основным, а при некоторых формах и стадиях распространения злокачественных опухолей — единственным методом лечения.

Наиболее серьезным препятствием к повышению эффективности химиотерапии опухолей остается фенотип их множественной лекарственной устойчивости. Утрата программы клеточной гибели дает возможность опухолевой клетке сохранять жизнеспособность в присутствии высоких концентраций противоопухолевых лекарств и формировать опухолевую ткань, абсолютно резистентную к лекарственному лечению. Достигнуть реактивации апоптоза в таких опухолевых клетках становится практически невозможным. В связи с этим актуальной остается потребность в создании новых противоопухолевых препаратов для избирательного воздействия на опухолевые клетки.

При целом ряде злокачественных опухолей различных локализаций в последние годы с успехом применяются химио- и биотерапевтические препараты направленного действия, позволяющие селективно воздействовать только на белки, участвующие в канцерогенезе и определяющие способность опухоли к прогрессии и метастазированию, — таргетные препараты. Однако через несколько месяцев монотерапии таргетным препаратом наступает резистентность. Сочетание таргетной терапии и существующих химиотерапевтических препаратов — комбинированная терапия — во многих случаях повышает эффективность лечения, но в отдельных случаях приводит к значительному увеличению токсичности, ограничивая продолжение исследований в этом направлении. Различные значения растворимости, отличия в фармакинетике и других свойствах двух или более химических веществ ограничивают применение комбинированной терапии. Кроме того, нормативные требования комбинированной терапии более сложны, поскольку профиль безопасности каждого препарата должен быть изучен перед клиническими испытаниями, затрудняя вопросы регулирования, в частности если два препарата разрабатываются разными компаниями.

В настоящее время в качестве перспективной стратегии лечения рака рассматривается проектирование и синтез гибридных соединений, состоящих из двух или более биоактивных фрагментов, объединенных в один препарат и действующих через активацию нескольких механизмов гибели клетки. В гибридных молекулах два или более фармакофора соединены ковалентной связью в одной молекуле и действуют путем ингибирования одной мишени или одновременного ингибирования двух разных мишеней. Гибридные соединения способны устранить отдельные недостатки препаратов, направленных на блокаду какого-то одного сигнального пути. Сочетание двух активных групп в одной молекуле может приводить к более выраженному терапевтическому эффекту по сравнению с индивидуальными компонентами. Концепция гибридных лекарств более подробно изложена в публикации [Meunier, 2008].

Среди перспективных терапевтических мишеней для лечения злокачественных опухолей выделяют гистондеацетилазы (Histone deacetylases, HDACs), представляющие собой ферменты, удаляющие ацетильные группы гистонов. Модифицируя гистоны и изменяя конформацию хроматина, гистондеацетилазы играют важную роль в регуляции экспрессии генов. Гистондеацетилазы представлены 18 белками, принадлежащими к 4-м классам. Внимание исследователей привлекают ингибиторы гистондеацетилаз (HDACi), индуцирующие апоптоз, арест клеточного цикла, старение, дифференцировку, иммуногенность клеток и ингибирующие ангиогенез при некоторых видах рака. Среди HDACi наиболее важной и многочисленной группой являются производные гидроксамовой кислоты (Вориностат, Ромидепсин, Белиностат). Также к HDACi относятся некоторые циклические пептиды, алифатические кислоты и бензамиды. Большинство HDACi имеют трехкомпонентную структуру, состоящую из цинк-связывающего участка (zinc binding group, ZBG), линкера, способного занимать канал фермента, и функциональной группы, взаимодействующей с аминокислотными остатками у входа в активный центр HDAC («крышка», Capping group) (рис. 1). Ингибиторы классических деацетилаз функциони-

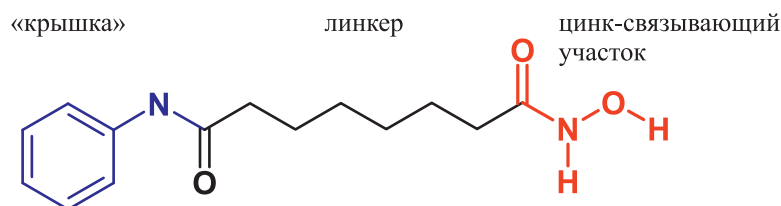


Рис. 1. Фармакофорное строение ингибиторов гистондеацетилаз на примере препарата Вориностат

руют путем связывания и вытеснения иона цинка из активного центра и таким образом инактивируют систему смены зарядов [West, Johnstone, 2014].

Одним из необходимых условий выживания опухолей разных гистологических типов является индуцирование ангиогенеза, в этой связи антиангиогенная терапия, основными мишенями которой являются фактор роста эндотелия сосудов (VEGF) и его рецепторные тирозинкиназы VEGFR-1, 2, играет важнейшую роль в борьбе с онкологическими заболеваниями. Блокирование VEGF/VEGFR-2-сигнального пути снижает кровоснабжение опухоли, приводя к гибели опухолевых клеток из-за нехватки питательных веществ и кислорода. Ингибиторы VEGFR-2, производные хиназолина, такие как Вандетаниб, Седираниб, нашли активное применение при лечении онкологических заболеваний [Wells et al., 2010].

В публикациях [Yu et al., 2013; Rajagopalan et al., 2013] успешно применено фармакофорное моделирование к различным типам VEGFR-2-ингибиторов, при этом общие признаки данных ингибиторов включают акцептор и донор водородной связи, ароматическое гидрофобное кольцо, однако дистанции между фармакофорными группами и их взаимное расположение различны и зависят от скаффолдов исследуемых соединений.

Некоторые производные хиназолина проявляют свойства ингибиторов HDAC [Lin et al., 2010]. Согласно имеющимся данным [Yang et al., 2016] хиназолиновый цикл может выступать в роли эффективного домена поверхностного взаимодействия («крышки») с активными сайтами HDAC.

С другой стороны, некоторая часть HDACi способна подавлять кровоснабжение опухоли. Исследования HDACi выявили угнетение ангиогенеза путем подавления индуцирующих ангиогенез генов, таких как VEGF, VEGFR-2, HIF-1 α , Tie-2, и эндотелиальной синтазы оксида азота (eNOS) [Zupkovitz et al., 2006]. Несомненный интерес вызывают новые публикации о снижении экспрессии обоих рецепторов VEGF в клетках нейробластомы HDACi [Deroanne et al., 2002].

Относительно недавно опубликован ряд примеров бифункциональных VEGFR-2/HDAC-ингибиторов [Peng et al., 2015; Peng et al., 2016].

С учетом вышесказанного логичным выглядит создание бифункциональных агентов, одновременно ингибирующих VEGFR-2 и HDAC, на основе двух фармакофорных групп — хиназолинов и гидроксамовых кислот.

В связи с возрастанием сложности поиска новых высокоактивных соединений значительное внимание исследователей привлекают методы количественной связи «структура – активность/свойство» (QSAR/QSPR), оказывающие значительную помощь в снижении как финансовых, так и временных затрат при разработке потенциальных лекарственных средств.

Детальный анализ существующих QSAR-моделей ингибиторов HDACs приведен в публикациях [Pham et al., 2017; Tang et al., 2009], при этом необходимо отметить, что большинство опубликованных исследований основаны на методах 3D-QSAR (ComFA, CoMSIA) в сочетании с молекулярным докингом, при этом размер выборок колеблется от 16 до 1488 соединений. Полученная с их помощью разноплановая информация представляет значительный интерес для практического использования: несколько разработанных моделей применены для виртуального скрининга HDACi в таких базах данных химических соединений, как ZINC, PubChem, Maybridge [Liu et al., 2010]. В работе [Tang et al., 2009] первоначально для выборки из 59 ингибиторов гистондеацетилаз класса 1 (HDAC1) созданы QSAR-модели, которые далее использо-

вались для поиска соединений-хитов методом виртуального скрининга (virtual screening, VS) среди 9 500 000 соединений, представленных в базах данных ZINC7.0, World Drug Index (WDI), ASINEX Synergy. Среди отобранных 45 хитов находились структурно уникальные соединения, скаффолды которых отсутствовали в первоначальной выборке. Четыре соединения с новыми скаффолдами были экспериментально протестированы, и три из них подтвердили высокую активность в качестве ингибиторов HDAC1 — со значения IC_{50} , равного 1 μ M для наиболее активного из них. Четвертое соединение позднее проявило активность в качестве селективного ингибитора HDAC6. Более того, два предложенных соединения-хита — продаваемые лекарства, которые, по понятным причинам, могут быстрее, чем вновь предложенные соединения, — использованы в качестве противораковых агентов. Вышеприведенное исследование иллюстрирует возможности комбинированного метода QSAR-VS в качестве общего подхода для эффективной идентификации новых физиологически активных соединений.

Однако в ряде случаев существующие QSAR-модели ингибиторов HDACs имеют ограниченное применение. Например, в методе CoMFA [Guo et al., 2005] существуют определенные трудности при анализе соединений разнородных по структуре из-за неоднозначности трехмерного выравнивания структур. Применяемые в большинстве существующих QSAR-моделей регрессионные методы анализа данных (MLR, PLS) [Xie et al., 2004] могут давать хорошие результаты только при наличии линейной связи между структурой и активностью. Один из способов преодоления вышеуказанного недостатка регрессионного анализа заключается в использовании непараметрических методов, в частности различных методов машинного обучения (machine learning, ML). Так, в публикациях [Tang et al., 2009] применялся метод опорных векторов (Support Vector Machine, SVM), а в [Zhao et al., 2013] — метод k -ближайших соседей (k -nearest neighbor algorithm, kNN). В работе [Pham et al., 2017] приведено консенсусное QSAR-моделирование с использованием в том числе методов kNN, SVM, случайного леса (Random Forest, RF). Отметим, однако, что примененные в данной работе при QSAR-моделировании дескрипторы, рассчитываемые программой Dragon v.6.0, сильно ограничивают возможность интерпретации полученных моделей с целью получения информации, необходимой для последующего рационального молекулярного дизайна [Polishchuk, 2017].

В ряде случаев [Ragno et al., 2008] авторы не приводят важных статистических характеристик моделей, таких как коэффициент детерминации (корреляции) между наблюдаемыми и предсказанными значениями активности при кросс-валидации для соединений обучающей выборки, а также указанный параметр для соединений тестовой выборки. Хотя именно эти показатели характеризуют качество и прогнозирующую способность модели и ее возможность использования в дальнейших исследованиях.

К сожалению, большинству существующих QSAR-моделей ингибиторов VEGFR-1, 2 [Yu et al., 2007; Sun et al., 2013; Rajagopalan et al., 2013; Reid et al., 2016; Patel et al., 2009] присущи вышеуказанные проблемы и ограничения. Следует отметить, что авторами публикации [Marzago et al., 2011] в рамках вышеуказанной перспективной стратегии разработки многофункциональных противоопухолевых средств предложены мультитаргетные классификационные QSAR-модели, количественно описывающие связь «структура – антитирозинкиназная активность», в том числе в отношении VEGFR-1.

В ходе литературного анализа не найдены публикации, посвященные созданию QSAR-моделей бифункциональных VEGFR-2/HDAC-ингибиторов, содержащих гидроксамовые кислоты и фрагменты хинозалинового цикла. Поэтому в задачи данного исследования входило:

- 1) построение QSAR-моделей бифункциональных VEGFR-2/HDAC-ингибиторов, содержащих гидроксамовые кислоты и фрагменты хинозалинового цикла;
- 2) структурная интерпретация построенных QSAR-моделей с целью выявления фрагментов, наиболее влияющих на проявление VEGFR-2/HDAC-ингибиторной активности;
- 3) рациональный молекулярный дизайн и виртуальный скрининг с использованием построенных QSAR-моделей с целью выявления наиболее перспективных соединений для последующего синтеза и исследований.

2. Методы

Источником формирования выборок для QSAR-моделирования послужили публикации [Peng et al., 2015; Peng et al., 2016], в каждой из которых приведены экспериментальные значения для соединений, содержащих одновременно N-фенил-4-аминохиназолин и гидроксамовую кислоту, в отношении трех видов активностей:

- 1) ингибирование HDAC (изоформы);
- 2) ингибирование VEGFR-2;
- 3) ингибирование роста клеток рака легкого MCF-7. В настоящее время данная клеточная линия активно используется при исследовании *in vitro* цитотоксичности противоопухолевых лекарственных средств.

С целью расширения структурного пространства для QSAR-моделирования выборки из двух указанных публикаций [Peng et al., 2015; Peng et al., 2016] объединены по соответствующим видам активностей в общие выборки, что, с нашей точки зрения, допустимо, поскольку имеющиеся в публикациях результаты получены одними авторами в одинаковых условиях и значения трех видов активностей для референтных структур (Вандетаниб и Вориностат) практически совпадают.

Для выборок приведенные в первоисточниках экспериментальные значения соответствующих видов биологической активности, выраженные с помощью концентрации полумаксимального ингибирования (IC_{50} , nM или μM), были сконвертированы в отрицательный десятичный логарифм данной величины (pIC_{50}), который наиболее общепринят в QSAR-исследованиях и используется в случаях, когда линейный рост концентрации вызывает экспоненциальный рост эффекта, по формуле

$$pIC_{50} = -\log_{10}(IC_{50}). \quad (1)$$

Экспериментальные значения и структуры соединений, экспортированные из публикаций [Peng et al., 2015; Peng et al., 2016], приведены в таблице 1.

С целью оценки предсказательной способности QSAR-моделей первоначальные общие выборки разделяли на обучающие (ws) и тестовые (ts) наборы. Для этого в каждой выборке общий набор соединений упорядочивали по возрастанию активностей и каждое пятое соединение помещали в тестовый набор. Исходя из этого, размер обучающего и тестового наборов для каждого вида активности соответственно составил 80 и 20 процентов от общей выборки. Соединения, не имеющие точного значения активности, например **6ea**, **6eb**, **6ec**, **6ed** со значением $pIC_{50} < 5.0$ для VEGFR-2, при формировании выборок не учитывались.

Для стандартизации структур использовали ChemAxon Standardizer (<https://chemaxon.com/products/chemical-structure-representation-toolkit>).

При описании молекулярной структуры использованы 2D-симплексные дескрипторы, рассчитываемые в рамках симплексного представления молекулярной структуры (СПМС, Simplex representation of molecular structure, SiRMS, <https://github.com/DrrDom/sirms>). В рамках СПМС молекула рассматривается как система различных симплексов — четырехатомных молекулярных фрагментов фиксированной структуры (рис. 2).

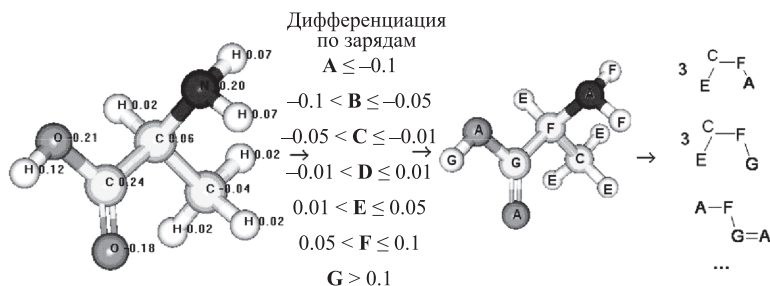


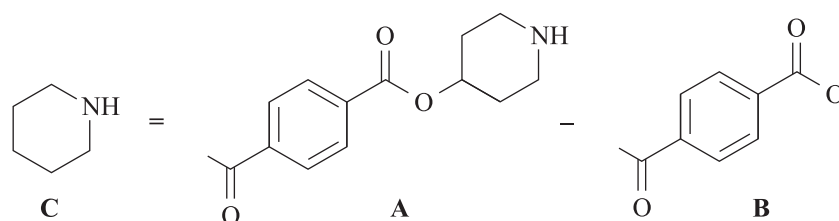
Рис. 2. Пример 2D-генерации симплексных дескрипторов для аланина на 2D-уровне при использовании дифференциации атомов по их парциальным зарядам

Таблица 1. Структуры соединений и их экспериментальные значения активностей

№	Соединение	R ¹	R ²	R ³	R ⁴	n	pIC ₅₀			Источник
							VEGFR-2	HDAC	MCF-7	
1	6aa	F	H	H	H	2	6.06	6.17	<4	[Peng et al., 2016]
2	6ab	F	H	H	H	3	6.08	6.26	<4	
3	6ac	F	H	H	H	4	6.04	6.84	4.08	
4	6ad	F	H	H	H	5	6.12	7.49	4.82	
5	6ba	H	H	F	H	2	6.42	5.87	<4	
6	6bb	H	H	F	H	3	6.34	6.86	4.13	
7	6bc	H	H	F	H	4	6.28	7.54	4.74	
8	6bd	H	H	F	H	5	6.57	8.07	5.62	
9	6ca	Cl	H	H	H	2	6.07	5.90	<4	
10	6cb	Cl	H	H	H	3	6.08	6.89	4.11	
11	6cc	Cl	H	H	H	4	6.16	7.54	4.82	
12	6cd	Cl	H	H	H	5	6.44	7.74	5.11	
13	6da	H	H	Cl	H	2	6.74	6.16	4.07	
14	6db	H	H	Cl	H	3	6.66	6.88	4.21	
15	6dc	H	H	Cl	H	4	6.50	7.96	5.40	
16	6dd	H	H	Cl	H	5	6.67	8.49	5.77	
17	6ea	Br	H	H	H	2	<5.0	6.08	<4	
18	6eb	Br	H	H	H	3	<5.0	7.13	4.48	
19	6ec	Br	H	H	H	4	<5.0	7.51	4.77	
20	6ed	Br	H	H	H	5	<5.0	8.06	5.59	
21	6fa	H	H	Br	H	2	7.01	6.26	4.28	
22	6fb	H	H	Br	H	3	7.23	7.17	4.44	
23	6fc	H	H	Br	H	4	6.82	7.92	5.46	
24	6fd	H	H	Br	H	5	7.13	8.66	6.07	
25	Вандетаниб	–	–	–	–	–	7.27	<5.0	4.73	
26	Вориностат	–	–	–	–	–	<5.0	7.82	5.38	
27	6a	H	Cl	H	H	2	6.32	6.19	<4	[Peng et al., 2015]
28	6b	H	Cl	H	H	3	6.45	6.92	4.19	
29	6c	H	Cl	H	H	4	6.21	7.77	5.07	
30	6d	H	Cl	H	H	5	6.30	8.27	5.49	
31	6e	H	Br	H	H	2	6.20	6.27	<4	
32	6f	H	Br	H	H	3	6.36	7.01	4.16	
33	6g	H	Br	H	H	4	6.38	7.82	5.07	
34	6h	H	Br	H	H	5	6.34	8.35	5.55	
35	6i	Cl	H	Cl	H	2	6.43	6.67	4.02	
36	6j	Cl	H	Cl	H	3	6.86	7.07	4.25	
37	6k	Cl	H	Cl	H	4	6.94	7.85	5.19	
38	6l	Cl	H	Cl	H	5	7.08	8.55	5.92	
39	6m	Cl	H	H	Cl	2	5.90	6.46	<4	
40	6n	Cl	H	H	Cl	3	5.67	7.24	4.35	
41	6o	Cl	H	H	Cl	4	5.28	7.68	4.86	
42	6p	Cl	H	H	Cl	5	5.61	8.08	5.40	
25	Вандетаниб	–	–	–	–	–	7.21	<5.0	4.73	
26	Вориностат	–	–	–	–	–	<5.0	7.92	5.35	

Дескриптором в этом случае служит количество симплексов одного вида. На 2D-уровне атомы (вершины симплекса) дифференцируют не только по природе «метки» атома, но и с учетом различных физико-химических свойств — частичного заряда на атоме, липофильности, рефракции, способности выступать донором или акцептором водорода при образовании водородной связи.

Одним из принципов проведения QSAR-исследований, предложенных экспертной группой OECD (Organization for Economic Cooperation and Development), является интерпретация полученной модели. Симплексные дескрипторы позволяют интерпретировать получаемые модели без существенных ограничений, присущих другим дескрипторам. В данном исследовании структурная интерпретация проводилась в соответствии с подходом [Polishchuk et al., 2013], при котором вклад исследуемого фрагмента (C) рассчитывался путем нахождения разницы между рассчитанными значениями активности/свойства для родительской структуры (A) и гипотетической структуры (B), полученной удалением исследуемого фрагмента (C) из родительской структуры (A) (рис. 3).



$$W(C) = X(A) - X(B)$$

Рис. 3. Использованный принцип структурной интерпретации: $W(C)$ — вклад фрагмента (C), $X(A)$ — предсказанное значение активности/свойства родительской структуры (A), $X(B)$ — предсказанное значение активности/свойства для гипотетической структуры (B)

При моделировании с использованием симплексных дескрипторов применяли пакет Scikit-learn для языка программирования Python, в котором реализован метод опорных векторов (Support Vector Machine, SVM) [Pedregosa et al., 2011]. Для внутренней валидации использовали процедуру пятикратной кросс-валидации (5-folds). При QSAR-моделировании для тестовых выборок учитывали вхождение соединений в область применимости (applicability domain, AD) [Jaworska et al., 2005], при этом если значение хотя бы одного дескриптора выходило за пределы его минимального или максимального значения для обучающей выборки, то соединение тестовой выборки, содержащее данный дескриптор, не входило в область применимости (bounding box). Вышеописанный подход для QSAR-моделирования с использованием симплексных дескрипторов реализован в виде программного обеспечения SPCI, которое находится в свободном доступе (http://qsar4u.com/pages/sirms_qsar.php).

Для сравнительного анализа предсказательной способности QSAR-моделей, построенных с помощью симплексных дескрипторов, при моделировании также использовались физико-химические дескрипторы, рассчитываемые программой HYBOT, в основе которых лежит оценка Н-донорной и Н-акцепторной способности молекул [Raevsky et al., 1992]. Всего было рассчитано 39 дескрипторов, из них в финальные QSAR-модели вошло 5 дескрипторов: ΣE_{ad} — сумма энтальпийных Н-акцепторных и Н-донорных дескрипторов; ΣC_a — сумма свободноэнергетических Н-акцепторных дескрипторов; $\Sigma E_{a/\alpha}$ — сумма энтальпийных Н-акцепторных дескрипторов, деленная на молекулярную поляризуемость; $\Sigma E_{ad/\alpha}$ — сумма энтальпийных Н-акцепторных и Н-донорных дескрипторов, деленная на молекулярную поляризуемость; Csp^3ac — число ациклических атомов углерода в состоянии sp^3 -гибридизации.

Создание QSAR-моделей при использовании вышеуказанных физико-химических дескрипторов проводили с использованием метода случайного леса (Random forest, RF). При выполнении регрессии в модели RF применяли ансамбль решающих деревьев [Breiman, 2001]. Каждое

дерево формировали на основе бутстреп-процедуры. Для конструирования деревьев на стадии расщепления использовали фиксированное число случайно выбранных переменных. Деревья решений создавали без обрезания. RF-моделирование проводили с помощью компьютерной программы RRforest (http://www.stat.berkeley.edu/~breiman/RandomForests/reg_examples/RFR.f) с использованием следующих параметров: jbt (число деревьев) = 500, mtry (число случайно выбираемых переменных на каждом узле) = $m/3$ (m — общее число переменных) и ndsize (число точек в узле, ниже которого дерево не расщепляли) = 5.

Отбор физико-химических дескрипторов, полученных с помощью программы HYBOT, и конструирование на их основе QSAR-моделей осуществляли путем перебора всех возможных комбинаций из 1-й, 2-х и 3-х переменных. Для оценки качества моделей использовали величину FIT (модифицированный критерий Фишера), которую рассчитывали по формуле [Kubinyi, 1994]

$$\text{FIT} = \frac{R^2(n-k-1)}{((n+k^2) \times (1-R^2))}, \quad (2)$$

где R^2 — квадрат коэффициента линейной корреляции, n — число соединений, k — число независимых переменных.

В качестве лучших моделей выбирали те из них, которые имели максимальные значения FIT в условиях кросс-валидации с размером блока, равным 5 (5-fold).

Область применимости оценивали интервальным методом путем анализа независимых (X-AD) и зависимых (Y-AD) переменных.

Обучающая выборка:

HDAC: X-AD: $\Sigma E_{ad}/\alpha$ (0.347÷0.427); Csp^3ac (4÷7); Y-AD: pIC_{50} (5.87÷8.66)

VEGFR-2: X-AD: ΣE_{ad} (14.06÷17.07); Y-AD: pIC_{50} (5.28÷7.27)

MCF-7: X-AD: ΣC_a (6.03÷9.10); $\Sigma E_a/\alpha$ (0.153÷0.273); $\Sigma E_{ad}/\alpha$ (0.307÷0.404); Y-AD: pIC_{50} (4.02÷6.07)

Тестовая выборка:

HDAC: X-AD: $\Sigma E_{ad}/\alpha$ (0.356÷0.407); Csp^3ac (4÷7); Y-AD: pIC_{50} (5.90÷8.27)

VEGFR-2: X-AD: ΣE_{ad} (16.69÷17.07); Y-AD: pIC_{50} (5.67÷7.01)

MCF-7: X-AD: ΣC_a (8.99÷9.09); $\Sigma E_a/\alpha$ (0.238÷0.276); $\Sigma E_{ad}/\alpha$ (0.350÷0.408); Y-AD: pIC_{50} (4.13÷5.59)

Помимо этого, для сравнительного QSAR-анализа нами была использована экспертная система OCHEM (<https://ochem.eu>). Для моделирования были выбраны дескрипторы PyDescriptor [Masand et al., 2017] и метод экстремального градиентного бустинга [Sheridan et al., 2016], при этом область применимости оценивалась с помощью концепции «расстояние до модели», в частности подхода CLASS-LAG [Manallack et al., 2003].

В интернет-ресурсе OCHEM реализован метод молекулярных пар [Sushko et al., 2014], который также позволяет проводить интерпретацию моделей, построенных на любых дескрипторах. Однако для интерпретации в нашем случае были использованы только модели, построенные с помощью симплексных дескрипторов, поскольку экспертная система OCHEM для построенных ею моделей не идентифицировала молекулярные пары, что, вероятно, связано с небольшим размером выборок, использованных для QSAR-моделирования.

Для расчета статистических характеристик QSAR-моделей использовали экспериментальные (наблюдаемые) и рассчитанные (предсказанные) значения активности для соединений обучающей выборки (в условиях кросс-валидации) и тестовой выборки.

Качество прогнозов на основе QSAR-моделей оценивалось коэффициентом детерминации (R^2):

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - y_{\text{mean}})^2}, \quad (3)$$

а также среднеквадратичной ошибкой (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m-1}}, \quad (4)$$

где y_i — наблюдаемое значение активности для i -го соединения, \hat{y}_i — предсказанное значение активности для i -го соединения, mean — среднее значение наблюдаемой активности, m — количество соединений в выборке.

3. Результаты и обсуждение

Результаты QSAR-моделирования приведены в таблице 2, из которой видно, что получены адекватные QSAR-модели, обладающие сопоставимой предсказательной способностью. Рассчитанные значения активностей для исследуемых соединений приведены в дополнительных материалах, в таблицах 1S–3S.

Таблица 2. Статистические характеристики QSAR-моделей. Созданные с помощью OCHEM модели доступны по ссылкам, приведенным в скобках

Вид активности	Программа для построения QSAR-модели (дескрипторы)	R^2_{cv}	RMSE _{cv}	R^2_{test}	RMSE _{test}
HDAC	SPCI	0.91	0.24	0.69	0.43
	OCHEM (https://ochem.eu/model/11031947)	0.81	0.35	0.65	0.45
	HYBOT ($\Sigma E_{ad}/\alpha$; Csp ³ ac). RRforest	0.94	0.21	0.70	0.45
VEGFR-2	SPCI	0.53	0.32	0.64	0.26
	OCHEM (https://ochem.eu/model/44455912)	0.55	0.31	0.86	0.16
	HYBOT (ΣE_{ad}). RRforest	0.53	0.33	0.92	0.12
MCF-7	SPCI	0.66	0.37	0.87	0.22
	OCHEM (https://ochem.eu/model/19379008)	0.70	0.35	0.82	0.26
	HYBOT (ΣC_a , $\Sigma E_a/\alpha$, $\Sigma E_{ad}/\alpha$). RRforest	0.80	0.31	0.86	0.22

Для интерпретации были выбраны молекулярные фрагменты (заместители), отмеченные в публикациях [Peng et al., 2015; Peng et al., 2016] и использованные в данных исследованиях в качестве вариативного параметра с сохранением молекулярного остова, представленного одновременно N-фенил-4-аминохиназолином и гидроксамовой кислотой, при синтезе с последующим испытанием потенциальных бифункциональных VEGFR-2/HDAC-ингибиторов. Краткое описание фрагментов представлено в таблице 3.

Результаты интерпретации приведены на рис. 4. Наибольшее положительное влияние на ингибирующую активность HDAC и ингибирование роста клеток рака легкого MCF-7 оказывает длина линкера, которая снижается в ряду (CH₂)₆ (фрагмент CH2_5) > (CH₂)₅ (фрагмент CH2_4) > (CH₂)₄ (фрагмент CH2_3). Эта закономерность подтверждается проведенным ранее SAR-анализом в публикации [Ding et al., 2017]. Для проявления VEGFR-2-ингибирующей активности длина линкера не является определяющей. Оптимальная длина линкера в этом случае составляет 4 метиленовых звена (фрагмент CH2_3).

Среди заместителей N-фенильного фрагмента можно выделить согласованный положительный вклад брома в *para*-положении (фрагмент R3_Br) для всех трех видов активности.

Согласно полученным результатам варьирование положения и типа атома галогена в фенильном фрагменте оказывает заметное влияние на VEGFR-2-ингибирующую активность соединений.

Таблица 3. Молекулярные фрагменты (заместители), использованные при интерпретации QSAR-моделей. Нумерация заместителей соответствует таблице 1. Визуализация молекулярных фрагментов посредством SMARTS возможна с помощью ресурса <https://smarts.plus/>

Название фрагмента	Состав фрагмента	Окружение фрагмента	SMARTS
R1_F	R ¹ = F	R ² = R ³ = R ⁴ = H	<chem>[\$(Fc1ccccc1(N))]</chem>
R1_Br	R ¹ = Br	R ² = R ³ = R ⁴ = H	<chem>[\$(Br1ccccc1(N))]</chem>
R1_Cl	R ¹ = Cl	R ² = R ³ = R ⁴ = H	<chem>[\$(Clc1[cH][cH][cH]c1(N))]</chem>
R2_Br	R ² = Br	R ¹ = R ³ = R ⁴ = H	<chem>[\$(Br1cccc(N)c1)]</chem>
R2_Cl	R ² = Cl	R ¹ = R ³ = R ⁴ = H	<chem>[\$(Clc1[cH][cH][cH]c(N)[cH]1)]</chem>
R3_F	R ³ = F	R ¹ = R ² = R ⁴ = H	<chem>[\$(Fc1ccc(N)cc1)]</chem>
R3_Br	R ³ = Br	R ¹ = R ² = R ⁴ = H	<chem>[\$(Br1ccc(N)cc1)]</chem>
R3_Cl	R ³ = Cl	R ¹ = R ² = R ⁴ = H	<chem>[\$(Clc1[cH][cH]c(N)[cH][cH]1)]</chem>
R1_Cl(R3_Cl)	R ¹ = Cl	R ³ = Cl; R ² = R ⁴ = H	<chem>[\$(Clc1cc(Cl)ccc1(N))]</chem>
R3_Cl(R1_Cl)	R ³ = Cl	R ¹ = Cl; R ² = R ⁴ = H	<chem>[\$(Clc1ccc(N)c(Cl)c1)]</chem>
R1_Cl(R4_Cl)	R ¹ = Cl	R ⁴ = Cl; R ² = R ³ = H	<chem>[\$(Clc1ccc(Cl)cc1(N))]</chem>
R4_Cl(R1_Cl)	R ⁴ = Cl	R ¹ = Cl; R ² = R ³ = H	<chem>[\$(Clc1ccc(Cl)c(N)c1)]</chem>
CH2_2	(CH ₂) _n , n = 2		<chem>[\$([CH2](Oc)[CH2][CH2]C(=O)[NH][OH])[CH2][CH2]</chem>
CH2_3	(CH ₂) _n , n = 3		<chem>[\$([CH2](Oc)[CH2][CH2][CH2]C(=O)[NH][OH])[CH2][CH2][CH2]</chem>
CH2_4	(CH ₂) _n , n = 4		<chem>[\$([CH2](Oc)[CH2][CH2][CH2][CH2]C(=O)[NH][OH])[CH2][CH2][CH2][CH2]</chem>
CH2_5	(CH ₂) _n , n = 5		<chem>[\$([CH2](Oc)[CH2][CH2][CH2][CH2][CH2]C(=O)[NH][OH])[CH2][CH2][CH2][CH2][CH2]</chem>

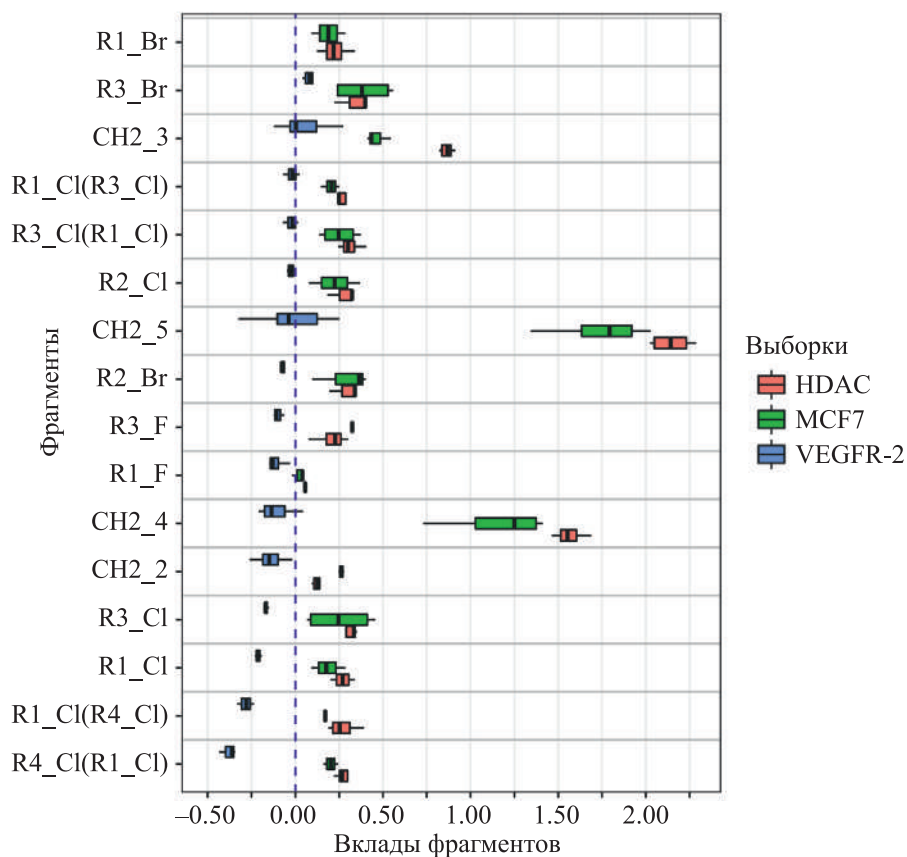


Рис. 4. Вклады молекулярных фрагментов в значения ингибирующей активности по отношению к HDAC. Фрагменты упорядочены по величине вклада в ингибирующую активность VEGFR-2. М-количество соединений, содержащих фрагмент. N-количество обнаружений соответствующего фрагмента в выборке

Построенные нами адекватные QSAR-модели с использованием физико-химических дескрипторов, представляющих те или иные количественные характеристики Н-связи, указывают на высокую значимость водородной связи в проявлении противоопухолевых свойств, что подтверждается в обзорной публикации [Patel et al., 2009]. В данном исследовании по результатам молекулярного докинга (PDB код: 1YWN) и фармакофорного анализа предложена консенсусная модель ингибиторов VEGFR-2, согласно которой для проявления высокой активности вышеуказанные соединения должны обладать следующими структурными особенностями: 1) обладать донором и акцептором водородной связи для взаимодействия соответственно с аминокислотами Glu883 и Cys919; 2) иметь липофильную группу, которая будет располагаться в «кармане», образованном аминокислотами Ile886, Leu887, Ile890, Val896, Leu1017; 3) содержать ароматическое кольцо.

Результаты проведенной нами интерпретации также подтверждают, что с ростом липофильности заместителей в фенильном фрагменте увеличивается активность ингибиторов VEGFR-2. Однако необходимо отметить, что положительным вкладом в VEGFR-2-ингибирующую активность обладает только вышеуказанный бром в *para*-положении относительно атома азота в N-фенильном фрагменте (фрагмент R3_Br). Остальные исследуемые фрагменты оказывают негативный вклад в повышение активности ингибиторов VEGFR-2. Выявленная тенденция может быть объяснена гипотезой, предложенной в публикации [Peng et al., 2016], согласно которой при введении объемного заместителя в *ortho*-положение возникают стерические препятствия.

Таким образом, обобщая анализ, проведенный в публикациях [Peng et al., 2016; Patel et al., 2009], а также результаты нашей интерпретации, можно полагать, что для данной серии соединений значительная ингибирующая активность VEGFR-2 проявляется при наличии высоколипофильных заместителей в фенильном фрагменте, не создающих стерических препятствий, располагающихся, например, в *para*-положении.

В то же время для HDAC- и MCF-7-ингибирующей активности вклады заместителей в бензольном кольце почти не меняются — атомы галогенов, за исключением фтора, вносят небольшой положительный вклад, независимо от их положения в цикле.

Принимая во внимание выявленную тенденцию и учет существующих фармакофорных моделей, проведен рациональный молекулярный дизайн и предложен ряд структур (**7a-f**), обладающих, согласно спрогнозированным значениям, сопоставимой активностью с наилучшими соединениями (**6fd**, **6l** — таблица 1), приведенными в публикациях [Peng et al., 2015; Peng et al., 2016]. Предложенные соединения **7a**, **7b**, **7c**, **7e**, **7f**, исходя из рассчитанных значений, обладают большей ингибирующей активностью в отношении HDAC и MCF-7 по сравнению с Вандетанибом и Вориностатом и достаточно высокой ингибирующей активностью по отношению к киназе VEGFR-2, что позволяет их рассматривать в качестве перспективных бифункциональных VEGFR-2/HDAC-ингибиторов. Результаты молекулярного дизайна приведены в таблице 4.

Помимо этого, с целью выявления перспективных бифункциональных ингибиторов среди соединений с различными скаффолдами, содержащими гидроксамовые кислоты и фрагменты хинозалинового цикла, проведен скрининг 300 виртуальных соединений, сгенерированных комбинаторно из структур известных и перспективных противоопухолевых лекарственных средств, например Эрлотиниб, Ралтитрексид, Седираниб. Для скрининга использовали модели, построенные экспертной системой OCHEM. QSAR-модели, полученные с помощью SPCI, ввиду вышеописанного метода определения области применимости, достаточно сильно ограничивающего структурное пространство модели, не использовались.

В ходе проведения виртуального скрининга отобраны соединения, входящие в области применимости всех трех QSAR-моделей, построенных экспертной системой OCHEM, и удовлетворяющие следующим условиям:

- 1) рассчитанные значения pIC_{50} для MCF-7 — больше 5.4, для HDAC — больше 7.9. Указанный минимальный порог обусловлен соответствующими экспериментальными значениями активностей лекарственного средства Вориностат, которые были определены в публикации [Peng et al., 2016] и приведены в таблице 1 данного исследования;
- 2) 10 соединений с наилучшим значением pIC_{50} для VEGFR-2.

Таблица 4. Результаты рационального молекулярного дизайна. С целью сравнительного анализа для соединений **6fd** и **6l** приведены экспериментальные значения согласно [Peng et al., 2015; Peng et al., 2016]

Название	Структура	Рассчитанные значения pIC_{50}					
		HDAC		VEGFR-2		MCF-7	
		SPCI	OCHEM	SPCI	OCHEM	SPCI	OCHEM
7a		8.20	8.23	5.97	6.56	5.71	5.92
7b		8.38	8.25	6.52	6.27	5.70	5.92
7c		8.29	8.25	6.51	6.26	5.74	5.92
7d		7.98	8.54	6.31	6.3	5.03	5.87
7e		8.27	8.15	6.29	6.27	5.46	5.92
7f		8.27	8.15	6.29	6.23	5.46	5.92
6fd		8.66		7.23		6.07	
6l		8.55		7.08		5.92	

Отобранные с учетом указанных условий перспективные соединения приведены в таблице 5.

Таблица 5. Наиболее перспективные соединения по результатам виртуального скрининга

Название	Структура	Рассчитанные значения pIC_{50}		
		HDAC	VEGFR-2	MCF-7
OVFV-173		7.97	6.59	5.48
OVFV-115		8.01	6.58	5.43
OVFV-184		8.09	6.56	5.54
OVFV-037		7.95	6.52	5.8
OVFV-282		8.02	6.49	5.43

Таблица 5 (продолжение). Наиболее перспективные соединения по результатам виртуального скрининга

Название	Структура	Рассчитанные значения pIC_{50}		
		HDAC	VEGFR-2	MCF-7
OVFV-028		8.02	6.48	5.54
OVFV-112		8.18	6.47	5.48
OVFV-125		8.37	6.45	5.53
OVFV-017		7.95	6.44	5.78
OVFV-122		8.48	6.44	5.53

Для всех приведенных в таблицах 4 и 5 перспективных бифункциональных VEGFR-2/HDAC-ингибиторов была проведена проверка наличия структур в базах данных Reaxys (<https://www.reaxys.com>), PubChem (<https://pubchem.ncbi.nlm.nih.gov/>), в том числе среди патентов, в ходе которой указанные вещества не были обнаружены.

4. Заключение и выводы

Проведенные исследования позволили:

- получить ряд адекватных регрессионных QSAR-моделей, обладающих удовлетворительными статистическими показателями и предсказательной способностью;
- проведенная интерпретация позволила количественно описать и ранжировать структурные фрагменты, устойчиво повышающие активность ингибиторов HDAC, VEGFR-2, MCF-7;
- в результате рационального молекулярного дизайна и виртуального скрининга предложен ряд перспективных бифункциональных VEGFR-2/HDAC-ингибиторов.

В последующем планируется синтезировать вышеуказанные бифункциональные VEGFR-2/HDAC-ингибиторы для оценки их биологической активности и сопоставить полученные результаты с расчетными.

Полученные результаты будут способствовать снижению финансовых, временных и трудовых затрат при синтезе и испытании перспективных противоопухолевых препаратов. Результаты структурной интерпретации могут оказать помощь в установлении механизмов действия новых бифункциональных гибридных лекарственных средств, направленных на лечение онкологических заболеваний.

Список литературы (References)

- Breiman L.* Random Forests // *Machine Learning*. — 2001. — Vol. 45. — P. 5–32.
- Deroanne C. F., Bonjean K., Servotte S., Devy L., Colige A., Clausse N., Blacher S., Verdin E., Foidart J. M., Nusgens B. V., Castronovo V.* Histone deacetylases inhibitors as anti-angiogenic agents altering vascular endothelial growth factor signaling // *Oncogene*. — 2002. — Vol. 21. — P. 427–436.
- Ding C., Chen S., Zhang C., Hu G., Zhang W., Li L., Chen Y. Z., Tan C., Jiang Y.* Synthesis and investigation of novel 6-(1,2,3-triazol-4-yl)-4-aminoquinazolin derivatives possessing hydroxamic acid moiety for cancer therapy // *Bioorg. Med. Chem.* — 2017. — Vol. 25, No. 1. — P. 27–37.
- Guo Y., Xiao J., Guo Z., Chu F., Cheng Y., Wu S.* Exploration of a binding mode of indole amide analogues as potent histone deacetylase inhibitors and 3D-QSAR analyses // *Bioorg. Med. Chem.* — 2005. — Vol. 13. — P. 5424–5434.
- Jaworska J., Nikolova-Jeliazkova N., Aldenberg T.* QSAR applicability domain estimation by projection of the training set descriptor space: a review // *Altern. Lab. Anim.* — 2005. — Vol. 33, No. 5. — P. 445–459.
- Kubinyi H.* Variable Selection in QSAR Studies. II. A Highly Efficient Combination of Systematic Search and Evolution // *Quant. Struct.-Act. Relat.* — 1994. — Vol. 13. — P. 393–401.
- Lin Z., Murray P. M., Ding Y., Denny W. A., Ferguson L. R.* Quinazolines as novel anti-inflammatory histone deacetylase inhibitors // *Mutation Research*. — 2010. — Vol. 690. — P. 81–88.
- Liu X. H., Song H. Y., Zhang J. X., Han B. C., Wei X. N., Ma X. H., Cui W. K., Chen Y. Z.* Identifying novel type ZBGs and nonhydroxamate HDAC inhibitors through a SVM based virtual screening approach // *Mol. Inf.* — 2010. — Vol. 29. — P. 407–420.

- Manallack D. T., Tehan B. G., Gancia E., Hudson B. D., Ford M. G., Livingstone D. J., Whitley D. C., Pitt W. R.* A Consensus neural network-based technique for discriminating soluble and poorly soluble compounds // *J. Chem. Inf. Comput. Sci.* — 2003. — Vol. 43. — P. 674–679.
- Marzaro G., Chilin A., Guiotto A., Uriarte E., Brun P., Castagliuolo I., Tonus F., González-Díaz H.* Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors // *Eur. J. Med. Chem.* — 2011. — Vol. 46, No. 6. — P. 2185–2192.
- Masand V., Rastija V.* PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors // *Chemometrics and Intelligent Laboratory Systems.* — 2017. — Vol. 169. — P. 12–18.
- Meunier B.* Hybrid molecules with a dual mode of action: dream or reality? // *Acc. Chem. Res.* — 2008. — Vol. 41. — P. 69–77.
- Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information.* URL: <https://ochem.eu> (accessed: 12.05.2019).
- Patel A., Karthikeyan C., Moorthy N. S., Trivedi P.* QSAR study on N-(aryl)-4-(azolyethyl) thiazole-5-carboxamides: novel potent inhibitors of VEGF receptors I and II // *Med. Chem.* — 2009. — Vol. 5, No. 5. — P. 455–461.
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E.* Scikit-learn: Machine Learning in Python // *J. Mach. Learn. Res.* — 2011. — Vol. 12. — P. 2825–2830.
- Peng F. W., Wu T. T., Zeng Z. W., Xue J. Y., Shi L.* Hybrids from 4-anilinoquinazoline and hydroxamic acid as dual inhibitors of vascular endothelial growth factor receptor-2 and histone deacetylase // *Bioorg. Med. Chem. Lett.* — 2015. — Vol. 25, No 22. — P. 5137–5141.
- Peng F. W., Xuan J., Wu T. T., Xue J. Y., Ren Z. W., Liu D. K., Wang X. Q., Chen X. H., Zhang J. W., Xu Y. G., Shi L.* Design, synthesis and biological evaluation of N-phenylquinazolin-4-amine hybrids as dual inhibitors of VEGFR-2 and HDAC // *Eur. J. Med. Chem.* — 2016. — Vol. 109. — P. 1–12.
- Pham-The H., Casañola-Martin G., Diéguez-Santana K., Nguyen-Hai N., Ngoc N. T., Vu-Duc L., Le-Thi-Thu H.* Quantitative structure-activity relationship analysis and virtual screening studies for identifying HDAC2 inhibitors from known HDAC bioactive chemical libraries // *SAR and QSAR in Environmental Research.* — 2017. — Vol. 28, No. 3. — P. 199–220.
- Polishchuk P.* Interpretation of quantitative structure-activity relationship models: past, present, and future // *J. Chem. Inf. Model.* — 2017. — Vol. 57. — P. 2618–2639.
- Polishchuk P. G., Kuz'min V. E., Artemenko A. G., Muratov E. N.* Universal approach for structural interpretation of QSAR/QSPR models // *Mol. Inform.* — 2013. — Vol. 32. — P. 843–853.
- Program RRforest. — URL: http://www.stat.berkeley.edu/~breiman/RandomForests/reg_examples/RFR.f (accessed: 12.05.2019).
- Raevsky O. A., Grigor'ev V. Yu., Kireev D. B., Zefirov N. S.* Complete thermodynamic description of H-bonding in the framework of multiplicative approach // *Quant. Struct.-Act. Relat.* — 1992. — Vol. 11. — P. 49–63.
- Ragno R., Simeoni S., Rotili D., Caroli A., Botta G., Brosch G., Massa S., Mai A.* Class II-selective histone deacetylase inhibitors. Part 2: alignment-independent GRIND 3D-QSAR, homology and docking studies // *Eur. J. Med. Chem.* — 2008. — Vol. 43. — P. 621–632.
- Rajagopalan M., Balasubramanian S., Ramaswamy A., Mathur P. P.* Pharmacophore based 3D-QSAR modeling and free energy analysis of VEGFR-2 inhibitors // *J. Enzyme Inhib. Med. Chem.* — 2013. — Vol. 28, No. 6. — P. 1236–1246.
- Reaxys — Chemistry data and chemical literature. — URL: <https://www.reaxys.com> (accessed: 08.08.2019).

- Reid T. E., Fortunak J. M., Wutoh A., Simon Wang X. Cheminformatic-based drug discovery of human tyrosine kinase inhibitors // *Curr. Top. Med. Chem.* — 2016. — Vol. 16, No. 13. — P. 1452–1462.
- Sheridan R. P., Wang W. M., Liaw A., Ma J., Gifford E. M. Extreme gradient boosting as a method for quantitative structure–activity relationships // *J. Chem. Inf. Model.* — 2016. — Vol. 56, No. 12. — P. 2353–2360.
- Simplex representation of molecular structure — a chemoinformatic tool for calculation of simplex descriptors. — URL: <https://github.com/DrrDom/sirms> (accessed: 12.05.2019).
- SPCI — Tool for mining structure-property relationships from chemical datasets. — URL: http://qsar4u.com/pages/sirms_qsar.php (accessed: 12.05.2019).
- Standardizer, version 5.4; ChemAxon: Budapest, Hungary. — URL: <https://chemaxon.com/products/chemical-structure-representation-toolkit> (accessed: 12.05.2019).
- Sun J., Li D. D., Li J. R., Fang F., Du Q. R., Qian Y., Zhu H. L. Design, synthesis, biological evaluation, and molecular modeling study of 4-alkoxyquinazoline derivatives as potential VEGFR2 kinase inhibitors // *Org. Biomol. Chem.* — 2013. — Vol. 11, No. 44. — P. 7676–7686.
- Sushko Y., Novotarskyi S., Korner R., Vogt J., Abdelaziz A., Tetko I. Prediction-driven matched molecular pairs to interpret QSARs and aid the molecular optimization process // *J. Cheminf.* — 2014. — Vol. 6. — P. 48
- Tang H., Wang X. S., Huang X. P., Roth B. L., Butler K. V., Kozikowski A. P., Jung M., Tropsha A. Novel inhibitors of human histone deacetylase (HDAC) identified by QSAR modeling of known inhibitors, virtual screening, and experimental validation // *J. Chem. Inf. Model.* — 2009. — Vol. 49, No. 2. — P. 461–476.
- U.S. EPA. “User’s Guide for T.E.S.T. (version 4.2) (Toxicity Estimation Software Tool): A Program to Estimate Toxicity from Molecular Structure.” — URL: <https://www.epa.gov/chemical-research/users-guide-test-version-42-toxicity-estimation-software-tool-program-estimate> (accessed: 10.08.2019).
- Wells S. A., Gosnell J. E., Gagel R. F., Moley J., Pfister D., Sosa J. A., Skinner M., Krebs A., Vasselli J., Schlumberger M. Vandetanib for the treatment of patients with locally advanced or metastatic hereditary medullary thyroid cancer // *J. Clin. Oncol.* — 2010. — Vol. 28. — P. 767–772.
- West A. C., Johnstone R. W. New and emerging HDAC inhibitors for cancer treatment // *J. Clin. Invest.* — 2014. — Vol. 124, No. 1 — P. 30–39.
- Xie A., Liao C., Li Z., Ning Z., Hu W., Lu X., Shi L., Zhou J. Quantitative structure-activity relationship study of histone deacetylase inhibitors // *Curr. Med. Chem. Anticancer Agents.* — 2004. — Vol. 4. — P. 273–299.
- Yang Z., Wang T., Wang F., Niu T., Liu Z., Chen X., Long C., Tang M., Cao D., Wang X., Xiang W., Yi Y., Ma L., You J., Chen L. Discovery of selective histone deacetylase 6 inhibitors using the quinazoline as the cap for the treatment of cancer // *J. Med. Chem.* — 2016. — Vol. 59, No. 4. — P. 1455–1470.
- Yu H., Wang Z., Zhang L., Zhang J., Huang Q. Pharmacophore modeling and in silico screening for new KDR kinase inhibitors // *Bioorg. Med. Chem. Lett.* — 2007. — Vol. 17, No. 8. — P. 2126–2133.
- Zhao L., Xiang Y., Song J., Zhang Z. A novel two-step QSAR modeling work flow to predict selectivity and activity of HDAC inhibitors // *Bioorg. Med. Chem. Lett.* — 2013. — Vol. 23. — P. 929–933.
- Zupkovitz G., Tischler J., Posch M., Sadzak I., Ramsauer K., Egger G., Grausenburger R., Schweifer N., Chiocca S., Decker T., Seiser C. Negative and positive regulation of gene expression by mouse histone deacetylase 1 // *Mol. Cell. Biol.* — 2006. — Vol. 26. — P. 7913–7928.