

©Коллектив авторов

РАЗРАБОТКА ПРОТИВОВИРУСНЫХ ЛЕКАРСТВ НА ОСНОВЕ ИНГИБИТОРОВ ГЛАВНОЙ ПРОТЕАЗЫ SARS-CoV-2

*А.В. Сулимов^{1,6}, Х.С. Шихалиев², О.В. Пьянков³, Д.Н. Щербаков^{3,4}, В.Ю. Чиркова⁴,
И.С. Ильин^{1,6}, Д.К. Кутлов^{1,6}, А.С. Тацилова^{1,6}, М.Ю. Крысин², Д.В. Крыльский²,
Н.В. Столповская², Е.А. Волосникова³, С.В. Беленькая^{3,5}, В.Б. Сулимов^{1,6*}*

¹Научно-исследовательский вычислительный центр,
Московский государственный университет имени М.В. Ломоносова,
119991, Москва, ул. Ленинские горы, 1, стр. 4; *эл. почта: sulimovv@mail.ru
²Воронежский государственный университет, 394018, Воронеж, Университетская пл., 1
³Государственный научный центр вирусологии и биотехнологии "Вектор",
630559, Новосибирская область, р.п. Кольцово
⁴Алтайский государственный университет, 656049, Барнаул, пр. Ленина, 61
⁵Новосибирский Государственный Университет, 630090, Новосибирск, ул. Пирогова, 1
⁶ООО "Димонта", 117186, Москва, ул. Нагорная, 15, корп. 8

Был проведён виртуальный скрининг с использованием докинга и методов квантовой химии лекарственных-подобных соединений базы данных Воронежского университета против главной протеазы коронавируса SARS-CoV-2, вызвавшего пандемию COVID-19. Сначала был выполнен докинг более 42 тысяч трёхмерных структур лигандов с помощью программы SOL. Далее, на основании оценки свободной энергии связывания, вычисляемой с помощью программы докинга SOL, было отобрано более тысячи лигандов для вычисления энтальпии связывания белок-лиганд с помощью квантово-химического полуэмпирического метода PM7 с учётом растворителя (воды) в неявной модели COSMO. В результате для экспериментального тестирования было отобрано 20 лигандов, показавших лучшие значения функции оценки свободной энергии связывания, вычисленной программой SOL, и энтальпии связывания, вычисленной квантово-химическим методом. В ходе экспериментального тестирования измеряли ингибирующую активность соединений против главной протеазы SARS-CoV-2 и подавление репликации этого коронавируса в культуре клеток. Ингибирование определяли с использованием пептидного субстрата с флуоресцентной меткой, включающей сайт протеолиза главной протеазы. Антивирусную активность тестировали в культуре клеток Vero. В результате экспериментального тестирования 8 молекул показали активность по ингибированию главной протеазы SARS-CoV-2 в субмикромольном и микромольном диапазонах IC_{50} , и 3 из них подавляли репликацию SARS-CoV-2 в культуре клеток на уровне микромольных значений EC_{50} и имели низкую цитотоксичность. Найденные химически разнообразные ингибиторы могут быть использованы для оптимизации с целью получения соединения-лидера — основы новых антивирусных препаратов прямого действия на коронавирус SARS-CoV-2.

Ключевые слова: разработка лекарств; докинг; квантовая химия; SARS-CoV-2; главная протеаза; ингибиторы; противовирусные лекарства; COVID-19; CADD

DOI: 10.18097/PBMC20216703259

ВВЕДЕНИЕ

Одним из важнейших применений суперкомпьютеров является разработка лекарств. Рациональный подход к созданию лекарств прямого действия заключается в связывании молекулы лекарства с белком-мишенью, определяющим развитие заболевания; таким образом блокируется функционирование этого белка и изменяется течение заболевания. Разработка лекарства представляет собой непрерывный конвейер и состоит из нескольких этапов. На первом из них происходит поиск органических соединений, молекулы которых избирательно связываются с белком-мишенью, например, это могут быть ингибиторы белка-мишени, блокирующие его работу. Для успеха разработки должно быть найдено несколько десятков или даже сотен таких соединений. Чем разнообразнее химические классы найденных соединений, тем больше вероятность создания эффективного

лекарства после успешного прохождения последующих этапов разработки лекарства: доклинических испытаний (на животных) и клинических испытаний (на людях). Разработка эффективных избирательных ингибиторов заданных белков-мишеней представляет собой ключевой этап всего конвейера разработки, от успеха которого зависит сама возможность создания лекарства. Разработка новых ингибиторов часто начинается со скрининга с помощью докинга баз данных существующих соединений; это ускоряет разработку, так как найденные кандидаты в ингибиторы уже существуют. Применение молекулярного моделирования и использование для этого суперкомпьютеров может существенно увеличить эффективность разработки новых лекарств.

Основным методом молекулярного моделирования, применяемым на начальном этапе разработки лекарств, является докинг. Это стало очевидным за прошедший год, когда возникла острая необходимость разработки противовирусных лекарств,



направленных на подавление репликации коронавируса SARS-CoV-2, вызвавшего пандемию COVID-19 [1], и появились многочисленные публикации, посвященные применению докинга для поиска ингибиторов тех или иных терапевтических белков-мишеней этого коронавируса. Программы докинга позиционируют молекулы-кандидаты в ингибиторы (лиганды) в активный центр белка-мишени и оценивают свободную энергию связывания лиганда с белком. Чем ниже значение энергии, тем эффективнее будет лекарство на основе такого ингибитора, так как меньшая концентрация лекарства в организме даст нужный терапевтический эффект.

Сейчас существует несколько десятков популярных программ докинга, отличающихся друг от друга алгоритмами, нюансами описания межмолекулярных взаимодействий, учёта растворителя (воды), гибкости лиганда и подвижности атомов белка, а также сделанными упрощениями, направленными на ускорение расчётов [2, 3]. Тенденция к увеличению скорости расчётов при докинге одного лиганда на одном вычислительном узле привела к использованию крайне упрощённых моделей, описывающих взаимодействие лиганда с белком, применению грубых приближений при разработке алгоритмов докинга, что, конечно, сказалось на низкой точности большинства используемых программ докинга [4]: точность позиционирования лигандов в белках-мишенях удовлетворительная, но точность оценки свободной энергии связывания низкая. Поэтому после виртуального скрининга надо обязательно проводить экспериментальное тестирование ингибирующей активности отобранных при докинге лигандов.

Усилия, направленные на разработку лекарств прямого действия на коронавирусы, делались задолго до начала пандемии COVID-19. Коронавирусы связаны с большим разнообразием респираторных заболеваний, включая обычную простуду и острую пневмонию, называемую Severe Acquired Respiratory Syndrome (SARS) и вызываемую коронавирусом SARS-CoV [5]. В 2012 году был идентифицирован новый коронавирус, вызвавший тяжелое заболевание похожее на SARS, получивший название Middle East Respiratory Syndrome (MERS-CoV) [6]. Вскоре были определены основные этапы размножения этих коронавирусов и выявлены два белка-мишени: papain-like протеаза (PL^{pro}) и главная протеаза 3-chymotrypsin-like protease (M^{pro} или 3CL^{pro}) [7], играющие важную роль при репликации коронавируса. Была получена трёхмерная структура главной протеазы SARS-CoV [8] и установлено, что каталитическая диада состоит из аминокислотных остатков CYS145 и HIS41. Было разработано много ингибиторов M^{pro} первого коронавируса SARS-CoV [2], но за все эти годы ни один ингибитор не перешёл на стадию клинических испытаний. Однако благодаря этим исследованиям при вспышке COVID-19 оказалось возможным быстро выявить вызвавший его коронавирус SARS-CoV-2 и определить терапевтические мишени для создания противовирусных препаратов прямого действия.

Для некоторых из этих мишеней была определена трёхмерная структура, в том числе и для главной протеазы M^{pro} [9], что даёт возможность применять докинг и другие методы для разработки ингибиторов. За прошедший год поиски ингибиторов M^{pro} проводились во многих базах органических соединений, прежде всего среди существующих лекарств [2], и в некоторых случаях были обнаружены экспериментально подтвержденные ингибиторы микромолярного и наномолярного диапазонов. Так, в работе [9] при экспериментальном скрининге около 10000 соединений, в которые входили одобренные к применению лекарства, лекарства, находящиеся на стадии доклинических и клинических испытаний, а также природные соединения (natural products), было обнаружено четыре наиболее интересных ингибитора главной протеазы SARS-CoV-2. Среди них два одобренных к применению лекарства — Дисульфирам (Disulfiram, IC₅₀=9,35 мкМ) и Кармофур (Carmofur, IC₅₀=1,82 мкМ), а также ещё два лекарства, проходящих доклинические и клинические испытания — Эбселен (Ebselen, IC₅₀=0,67 мкМ) и Тидеглусиб (Tideglusib, IC₅₀=1,55 мкМ). Из этих четырёх ингибиторов Эбселен и Кармофур представляют собой ковалентные необратимые ингибиторы [10], при этом Эбселен показал неплохую эффективную концентрацию EC₅₀=4,67 мкМ по подавлению репликации SARS-CoV-2 в культуре клеток. В работе [11] был проведён поиск ингибиторов главной протеазы SARS-CoV-2 среди около 2000 разрешённых к применению лекарств. При осуществлении докинга четырьмя различными программами (Glide SP, AutoDock Vina, и двумя протоколами программы AutoDock 4.2) для виртуального скрининга использовали протокол консенсуса. Консенсус заключался в том, что в лучшие лиганды отбирались только те, которые одновременно оказались среди 200 лучших лигандов при докинге по трём или по четырём протоколам. Так было отобрано 42 лиганда, анализ положений которых в белке-мишени и наблюдение стабильности связанных комплексов белок-лиганд при молекулярно-динамическом моделировании в течение 70 нс позволил отобрать для тестирования только 17 соединений, из которых в экспериментах ингибирующая активность главной протеазы SARS-CoV-2 на уровне IC₅₀<10 мкМ была подтверждена только для двух лекарств: Manidipine (IC₅₀=4,81 мкМ) и Восепреви́г (IC₅₀=5,40 мкМ). О двух новых ковалентных ингибиторах SARS-CoV-2 M^{pro}, подавляющих репликацию коронавируса в культуре клеток с эффективной концентрацией EC₅₀ равной 0,53 мкМ и 0,72 мкМ сообщается в работе [12]. В настоящее время довольно часто появляются публикации, посвященные ингибиторам главной протеазы SARS-CoV-2 (см., например, обзор [13]), и эти исследования отражают острую необходимость систематической разработки эффективной терапии против SARS-CoV-2.

В данной работе с помощью методов суперкомпьютерного молекулярного моделирования проведён поиск ингибиторов главной протеазы M^{pro} коронавируса SARS-CoV-2 в базе данных

существующих лекарственных соединений. На основании проведённых расчётов отобрано 20 соединений, которые наиболее сильно связываются с активным центром главной протеазы коронавируса и блокируют работу её каталитического центра. Из 20 отобранных соединений 18 соединений были проверены экспериментально на их способность ингибировать M^{pro} коронавируса SARS-CoV-2. Для нескольких из этих соединений получено экспериментальное подтверждение их активности по ингибированию главной протеазы M^{pro} коронавируса SARS-CoV-2 на уровне нескольких микромолей. Все отобранные соединения были экспериментально проверены на способность подавлять репликацию коронавируса SARS-CoV-2 в культуре клеток, и некоторые из них действительно подавляли репликацию коронавируса, а также обладали при этом низкой цитотоксичностью, так что они представляются перспективными для дальнейших исследований на животных.

МЕТОДИКА

Модель белка-мишени

К началу наших исследований в банке белковых структур (Protein Data Bank, PDB) существовало несколько комплексов белка M^{pro} коронавируса SARS-CoV-2 с хорошим разрешением. В результате детального рассмотрения в качестве основы модели для проведения докинга была выбрана структура PDB ID 6W63, в которой главная протеаза закристаллизована с нековалентным ингибитором, имеющим 7 внутренних вращательных степеней свободы. Детальная информация о причинах выбора именно этой структуры PDB и методика приготовления из неё модели для докинга представлены в работе [14]. Вкратце, из всех существовавших в PDB к началу данной работы структур M^{pro} , закристаллизованных с нековалентными ингибиторами, высокое разрешение ($<2,2$ Å) и хорошее качество, то есть без пропусков атомов и/или аминокислотных остатков, имели только 3 структуры: 5R7Z, 5R83 и 6W63. Приготовление моделей белка на основе этих трёх структур заключалось в удалении из них всех атомов, ионов и молекул, не относящихся к белку, и добавлении атомов водорода с помощью программы APLITE [15] для pH=7,4. Для двух моделей M^{pro} , построенных на основе структур 5R83 и 6W63, докинг нативных лигандов с помощью программы докинга SOL [16] оказался успешным со среднеквадратичными отклонениями RMSD = 1,19 Å и RMSD = 1,31 Å между соответствующими атомами лиганда в лучшей полученной при докинге конформации и в закристаллизованном положении. Для виртуального скрининга нами была выбрана модель на основе структуры 6W63, поскольку её нативный лиганд X77 был достаточно большим и гибким — имел 7 внутренних вращательных степеней свободы, так что активный центр главной протеазы был достаточно открытым для связывания самых разнообразных лигандов.

База данных органических соединений для скрининга

Для виртуального скрининга использовали базу данных лекарственно-подобных органических соединений, имеющихся в наличии в лаборатории кафедры органической химии Воронежского государственного университета (ВГУ). Всего было рассмотрено более 16000 молекул различных соединений, включающих широкий спектр азот-, кислород- и серу-содержащих гетероциклических соединений. Среди них были производные гидрохинолина, аминопиримидины и пирроло[3,2,1-*ij*]хинолин-2-оны, часто встречающиеся в качестве биологически активных соединений. Были подготовлены трёхмерные структуры всех этих молекул, и с учётом различных конформаций неароматических колец и макроциклов мы получили более 42 тысяч молекулярных структур для проведения докинга.

Программа докинга

Для проведения виртуального скрининга была использована программа докинга SOL [16]. Работа этой программы основана на парадигме докинга, согласно которой лучшая поза лиганда в активном центре белка-мишени соответствует глобальному минимуму энергии системы белок-лиганд, вычисляемой обычно в рамках какого-либо силового поля — набора потенциалов, описывающих внутримолекулярные и межмолекулярные взаимодействия между атомами. Для описания взаимодействия лиганда с белком в программе SOL используется силовое поле MMFF94 [17] и генетический алгоритм глобальной оптимизации. Суть этого алгоритма заключается в следующем. Случайным образом создается начальная популяция особей, представляющих собой различные возможные положения лиганда в области докинга. Затем рассматривается эволюция этой популяции через заданное число поколений. В процессе эволюции отбираются самые сильные особи, то есть позы лиганда, в которых целевая функция оптимизации имеет наиболее отрицательные значения. Целевая функция представляет собой сумму потенциальной энергии системы белок-лиганд с учётом растворителя и внутренней энергии лиганда, каждая из которых вычисляется в рамках силового поля MMFF94. При переходе от поколения к поколению размер популяции особей сохраняется, а особи в каждом последующем поколении создаются на основе прямого переноса самых “сильных” особей путём кроссинговера и мутаций. Решением задачи оптимизации является поза лиганда (особь) с минимальным значением целевой функции. Для ускорения проведения глобальной оптимизации с помощью модуля SOLGRID заранее создается трёхмерная сетка, в узлах которой записываются потенциалы кулоновских и ван-дер-ваальсовых взаимодействий для различных типов пробных атомов лиганда со всеми атомами белка-мишени. Кроме этих потенциалов в узлах сетки записываются и потенциалы десольватации, полученные с помощью упрощённого варианта обобщённой модели Борна [18].

В процессе глобальной оптимизации с помощью модуля SOL белок остается жёстким, то есть все атомы белка неподвижны, а лиганд может свободно перемещаться и вращаться в пространстве как жёсткое целое и изменять свою конформацию благодаря внутренним вращательным степеням свободы в пределах выделенного для докинга пространства, охватывающего активный центр белка-мишени. Таким образом, в программе SOL реализуется докинг гибкого лиганда в жёсткий белок. Поскольку генетический алгоритм глобальной оптимизации является эвристическим, для повышения надежности нахождения глобального минимума целевой функции генетический алгоритм глобальной оптимизации выполняли 50 независимых раз для каждого лиганда, а найденные лучшие 50 поз лиганда кластеризовали, то есть в каждый кластер отбирали решения докинга (так называемые “позы”), отличающиеся друг от друга не более чем на 1 Å по среднеквадратичному отклонению координат соответствующих атомов друг от друга в двух разных позах. При этом докинг считался успешным, только если кластер, содержащий позы лигандов с наиболее отрицательными значениями целевой энергетической функции, содержал не менее 10 поз лиганда. Для лучшей позы лиганда программа SOL рассчитывает так называемую функцию оценки (score), которая даёт оценку свободной энергии связывания белок-лиганд, в которой в простейшем виде учитывается и энтропийный вклад в свободную энергию связывания: энтропийная составляющая пропорциональна числу внутренних вращательных степеней свободы. Такой способ учёта изменения энтропии при связывании лиганда с белком широко используется в большинстве программ докинга [2, 3]. Параметризация и программы докинга SOL и используемого в ней силового поля MMFF94 сделаны для комнатной температуры, так что в параметрах модели температура учитывается неявно.

Программа SOL участвовала в международных соревнованиях по докингу CSAR 2011-2012 вместе с такими программами докинга, как Gold, AutoDock, AutoDock Vina, ICMVLS, Glide и другими, и для двух из трёх использовавшихся в этом соревновании белков-мишеней заняла одно из первых мест [16]. Программа докинга SOL адаптирована для проведения виртуально скрининга больших баз данных лигандов на суперкомпьютере Московского государственного университета (МГУ) “Ломоносов-2” [19] и успешно применялась за последние 15 лет для разработки ингибиторов ряда белков-мишеней.

При поиске ингибиторов M^{pro} с помощью программы SOL был проведён докинг трёхмерных молекулярных структур (более 42000) всех имеющихся в наличии соединений базы данных BГУ. Докинг проводили со стандартными параметрами генетического алгоритма, используемыми по умолчанию, из которых самые важные для качества и времени глобальной оптимизации — размер популяции особей (положений лиганда в белке) — равен 30000, число поколений — 1000. Ранжирование этих соединений

по значениям функции оценки SOL выявило 1045 соединений, имеющих более отрицательные значения, чем значение при докинге нативного лиганда X77 комплекса 6W63 (-6,3 ккал/моль).

Отбор кандидатов в ингибиторы методом квантовой химии

Для лучших отобранных в результате докинга лигандов с помощью программы MOPAC [20] квантово-химическим полуэмпирическим методом PM7 [21] с учётом растворителя в континуальной модели COSMO [22] был проведён расчёт энтальпии связывания лигандов с белком-мишенью. В модели COSMO растворитель (вода) представляет собой однородную среду, окружающую молекулу. Учитывая, что диэлектрическая проницаемость воды при комнатной температуре высока (78,5), уравнения для поляризационных зарядов на поверхности растворителя, наведённых зарядами атомов, решают для случая бесконечной диэлектрической проницаемости, а её конечную величину учитывают в результирующем выражении для поляризационной энергии.

Метод PM7 разработан сравнительно недавно, если сравнивать, например, с возрастом популярных полуэмпирических методов AM1 и PM3, разработанных в восьмидесяти годах прошлого века. Методы AM1 и PM3 плохо описывают и дисперсионные взаимодействия, и водородные связи. В методе PM7 эти недостатки устранены: ещё на стадии параметризации учтены поправки, позволяющие достаточно аккуратно учесть дисперсионные взаимодействия, водородные и галогенные связи, то есть именно те взаимодействия, которые играют важную роль в связывании лигандов с белками. Параметризация этого метода выполнена на беспрецедентно широком наборе молекулярных систем, опираясь не только на известные экспериментальные данные, но и на квантово-химические расчёты из первых принципов с учётом электронных корреляций. В результате точность расчётов методом PM7 часто не уступает точности расчётов с помощью методов функционала плотности [21], и среди нескольких недавно разработанных полуэмпирических методов PM7 является одним из самых надёжных [23].

Для получения энтальпии связывания лиганда с белком рассчитывали (без учёта колебательных степеней свободы) полную энергию комплекса белок-лиганд с использованием найденной при докинге лучшей позы лиганда, и также вычисляли энергию свободного белка в конформации, использовавшейся для докинга, и энергию свободного лиганда. Для последнего находили конформацию, соответствующую глобальному минимуму энергии, вычисляемой в методе PM7 с учётом растворителя в модели COSMO. Энтальпию связывания вычисляли как разность этих трёх энергий: из энергии комплекса вычитали энергии белка и лиганда. Детали этих вычислений приведены в работе [14]. Лиганды-кандидаты для экспериментального

тестирования их ингибирующей активности главной протеазы коронавируса SARS-CoV-2 отбирали с учётом выполнения в основном двух требований: минимального значения функции оценки докинга программы SOL, дающей свободную энергию связывания белок-лиганд, и энтальпии связывания лиганда с белком. Реперными значениями этих величин были соответствующие значения, полученные при докинге нативного лиганда X77, закристаллизованного вместе с главной протеазой в комплексе PDB ID 6W63. Значение функции оценки программы SOL для этого лиганда было равно -6,3 ккал/моль, а энтальпии связывания — -58,5 ккал/моль. Дополнительным критерием отбора служил визуальный анализ положения лиганда в активном центре белка. В отдельных случаях при прочих равных условиях предпочтение отдавали тем лигандам, которые в пространстве сильнее перекрывали каталитическую диаду CYS145 и HIS41. Все расчёты — докинг и последующие квантово-химические расчёты — проводились на суперкомпьютере МГУ “Ломоносов-2” [19].

Измерение ингибирования главной протеазы SARS-CoV-2

Последовательность, кодирующую основную протеазу SARS-CoV-2 (M^{pro} или 3CL pro) изолята Wuhan-Hu-1 (GenBank:MN908947), синтезировали на заказ (“ДНК-синтез”, Россия). Перед заказом проводили оптимизацию кодонного состава последовательности с помощью инструмента GeneOptimizer для оптимального синтеза протеазы в клетках *E. coli*. Чтобы избежать аккумуляцию целевого белка в виде плохо растворимых телец включения и обеспечить транспорт M^{pro} в периплазматическое пространство было проведено клонирование генов GST (глутатион-S-трансфераза) и M^{pro} в единой рамке считывания для получения слитого белка GST- M^{pro} . Для этого были рассчитаны олигонуклеотидные праймеры GST-F 5'-aaaaacatatgtccctatactaggttatt-3' и GST3CL-R 5'-aaaaaaggatccttttgaggatggtcgccac-3' (умножение копий GST в ПЦР) и 3CLGST-F 5'-aaaaaaggatccacctcagctgttttcagagcggttttcgtaaatggcat-3' и 3CL-R 5'-aaaaaaggcgccgcttaaccgctaccaccgctctgaaaggt-aacaccgctacactg-3' (умножение копий 3CL в ПЦР). Для обеспечения получения активной формы протеазы M^{pro} были заложены фланкирующие последовательности, содержащие гомологичные сайты протеолиза SAVLQSGFR на N-конце и TFQSG на C-конце, которые впоследствии помогают вырезать M^{pro} из синтезированного слитого белка. После умножения количества копий молекул ДНК два ПЦР продукта были одновременно перенесены в составе вектора pET21. Полученной плазмидой pET21-GST-3CL трансформировали клетки *E. coli* BL21(DE3). Индивидуальные колонии *E. coli*, содержащие рекомбинантные плазмиды, культивировали в течение ночи на орбитальном шейкере (“Biosan”, Латвия) в среде LB (“AppliChem”, США), содержащей 100 мкг/мл ампициллина (для селекции клеток, которые получили плазмиду), при 37°C и 180 об./мин.

После культивации раствор с трансформированными бактериями в соотношении 1/100 переносили в колбу Эрленмейера, содержащую среду LB, и растили до поглощения 0,8 (при $\lambda=600$ нм). Добавляли индуктор изопропил- β -D-1-тиогалактопиранозид (ИПТГ) (“Anatrace Products”, США) до конечной концентрации 1 mM для запуска синтеза белка. Культуру дополнительно культивировали на шейкере в течение 5 ч при 37°C и 180 об/мин. Биомассу осаждали центрифугированием (“Beckman Coulter”, США) в течение 20 мин при 5000 g и 4°C. Полученный осадок растворяли в буфере: 12 mM Трис; 120 mM NaCl; 0,1 mM EDTA; 1 mM DTT. Клетки разрушали с использованием ультразвукового гомогенизатора Soniprep 150 Plus (“MSE”, Китай). Очистку проводили последовательной хроматографией с промежуточным диализом в 20 mM Трис-HCl, pH 7,6 на колонках с катионообменными сорбентами (SP-сефароза и Nuvia HR-S), уравновешенными 20 mM Трис-HCl, pH 7,6. Целевую фракцию элюировали в линейном градиенте концентрации NaCl от 0 M до 1 M в 20 mM Трис-HCl, pH 7,6 и анализировали при помощи электрофореза в денатурирующих условиях в 15% ПААГ. Конечный препарат диализовали против 20 mM Трис-HCl, pH 7,6 и подвергали стерилизующей фильтрации через фильтры с размером пор 0,22 мкм.

Анализ ингибиторов проводили с использованием синтетического флуоресцентно меченого пептидного субстрата, включающего сайт протеолиза основной протеазы коронавируса состава DabcyL-VNSTLQSGLRK(FAM)MA (чистота более 95%, “CPC Scientific Inc”, Китай). Интенсивность флуоресценции контролировали с помощью прибора CLARIOstar Plus (“BMG Labtech”, Германия), используя длины волн 495 нм и 520 нм для возбуждения и излучения соответственно. При взаимодействии фермента с пептидными субстратом происходит расщепление сайта протеолиза основной протеазы коронавируса, что приводит к увеличению интенсивности флуоресценции из-за физического удаления флуорофора от тушителя. При добавлении в реакционную смесь тестируемых соединений наблюдается снижения уровня свечения. Реакционные смеси готовили в 384-луночном планшете, после чего инкубировали при комнатной температуре в течение 5 мин. Все измерения проводили при 25°C. Реакционные смеси содержали 5 мкл Трис-HCl буфера (pH 8,0; 50 mM Трис; 50 mM NaCl), 2 мкл субстрата (C = 33 мкг/мл) и ингибитор различной концентрации (0,000058594–0,03 мг/мл). Реакцию запускали добавлением 3 мкл M^{pro} (C = 0,025 мг/мл). За величину IC_{50} принимали такую концентрацию тестируемого вещества, которая снижает уровень флуоресценции на 50% от максимального, наблюдаемого без добавления ингибитора. В качестве стандартных, то есть реперных, соединений взяты два известных лекарственных соединения — Дисульфирам и Эбселен, для которых ранее [9, 10] была экспериментально продемонстрирована активность $IC_{50}=9,35$ мкМ и $IC_{50}=0,67$ мкМ по ингибированию главной протеазы SARS-CoV-2 соответственно.

Измерение противовирусной активности

Противовирусную активность соединений измеряли по “профилактической” схеме, в рамках которой разные концентрации препарата вносили в клетки за 2 ч до заражения и оставляли на 96 ч на весь период наблюдения до учёта результата в отношении вируса SARS-CoV2 в клетках Vero E6. Был использован штамм коронавируса hCoV-19/Russia/Vologda-171613-1208/2020; множественность инфицирования (MOI) 0,1; 5 последовательных разведений с шагом 5; начальная концентрация 100 мкМ.

Количество жизнеспособных клеток (защищённых от цитопатического действия вируса) определяли с помощью 3-(4,5-диметилтиазол-2-ил)-2,5-дифенил-2Н-тетразолий бромид (метод МТТ). В основе теста лежит реакция превращения бледно-жёлтого 3-(4,5-диметилтиазолин-2)-2,5-дифенилтетразолий бромид (“Sigma-Aldrich”, США) в формазан фиолетового цвета под действием фермента сукцинатдегидрогеназы. Для этого к супернатанту клеток добавляли раствор МТТ (50 мкл, 5 мг/мл). Планшет с клетками инкубировали в течение 90 мин при температуре 37°C. Супернатант удаляли, а клетки фиксировали 4% формальдегидом в течение 30 мин. Кристаллы тетразолия растворяли в 1 мл 96% этанола в течение 10 мин.

Значения оптической плотности (ОП), которая коррелирует с количеством жизнеспособных клеток в монослое, защищённых от цитопатического действия (ЦПД) вируса или цитотоксического действия исследуемого соединения, измеряли на многофункциональном спектрофотометре xMark (“Bio-Rad”, США) в 96-луночных планшетах при длине волны 450 нм.

По показателям ОП определяли 50% цитотоксическую концентрацию (CC₅₀) и 50% ингибирующую (эффективную) концентрацию (EC₅₀) препарата при помощи ПО SoftMaxPro-4.0. Значение CC₅₀ — это концентрация вещества, при которой разрушаются (теряют жизнеспособность) 50% клеток в неинфицированном монослое. Значение EC₅₀, являющееся показателем эффективности подавления репликации вируса, — это концентрация вещества, при которой не разрушаются (остаются жизнеспособными) 50% клеток в инфицированном монослое.

Таблица. Характеристики ингибирования SARS-CoV-2 M^{pro} выявленными соединениями. Номера в первой колонке соответствуют номерам соединений в базе ВГУ

Лиганд	IC ₅₀ , мкМ	Score SOL, ккал/моль	ΔH, ккал/моль	EC ₅₀ , мкМ	SI
16826	0,20±0,01	-6,6	-47,8	—	—
2425	1,77±0,01	-7,1	-54,4	—	—
2900	4,05±0,02	-6,7	-44,6	—	—
15689	6,25±0,03	-6,9	-55,5	18,19±0,51	11,9
6276	8,59±0,03	-6,6	-44,0	—	—
223350	24,50±0,12	-7,1	-49,5	6,53±0,42	27,3
214225	41,54±0,15	-6,8	-46,1	—	—
18302	41,62±0,12	-6,9	-50,8	4,77±0,33	45,1
Дисульфирам	6,10±0,04	—	—	—	—
Эбселен	1,39±0,03	—	—	—	—
Ремдесивир	—	—	—	2,94±0,67	56,5

На основании этих показателей рассчитывали индекс селективности (SI) исследуемого образца: $SI = CC_{50} : EC_{50}$. В качестве реперного для подавления репликации коронавируса SARS-CoV-2 было использовано существующее лекарственное средство Ремдесивир (Remdesivir), которое должно связываться, как предсказано на основе докинга, не с M^{pro}, а с RdRp [24].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В результате проведения докинга молекул-лигандов всех имеющихся в наличии соединений базы данных ВГУ и последующих квантово-химических расчётов энтальпии связывания белок-лиганд для лучших по результатам докинга лигандов были отобраны 20 лучших соединений для проведения их экспериментального тестирования, которое проводилось как путём измерения величины IC₅₀ для рекомбинантного белка главной протеазы коронавируса SARS-CoV-2, так и путём измерения подавления репликации этого вируса в культуре клеток и цитотоксичности этих соединений. У 20 этих лучших соединений либо значения функции оценки находились в диапазоне до -7,8 ккал/моль, либо энтальпия связывания находилась в диапазоне до -69 ккал/моль [14]. Столь отрицательные величины энтальпии связывания — характерная черта рассматриваемого метода вычисления, что видно из тщательно проведённого исследования на сравнительно большом наборе разнообразных белков и лигандов [25]. В качестве дополнительного критерия отбора лучших лигандов для тестирования, при прочих равных условиях, было принято их расположение в активном центре белка-мишени: блокирование каталитической диады и заполнение карманов связывания. Измерения были проведены для 18 из 20 отобранных соединений. В результате были выявлены новые ингибиторы главной протеазы коронавируса SARS-CoV-2 субмикромольного и микромольного диапазонов.

Результаты для выявленных и для реперных ингибиторов представлены в таблице. Структуры выявленных ингибиторов приведены на рисунке. Учитывая, что реперные ингибиторы SARS-CoV-2 M^{pro} (Дисульфирам и Эбселен) связываются с этим белком неспецифично и ковалентно [9, 10], мы не приводим

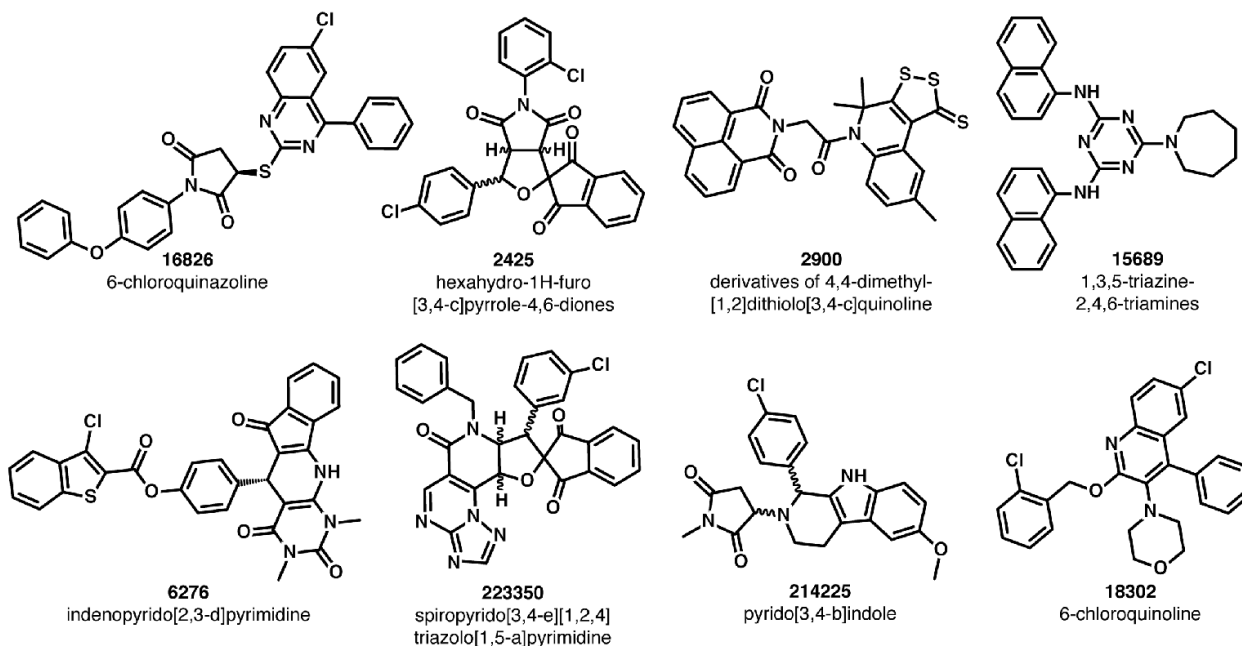


Рисунок. Структуры выявленных ингибиторов SARS-CoV-2 M^{pro}.

для них ни значения функции оценки докинга, ни величину энтальпии связывания белок-лиганд, так как для этих соединений такие расчёты не имеют большого смысла.

Как видим, прогноз ингибирования лигандами мишени SARS-CoV-2 M^{pro} с использованием докинга с помощью программы SOL и последующего постпроцессинга с помощью квантово-химического полумпирического метода PM7 с учётом растворителя в модели COSMO оказался достаточно надёжным: из 18 измеренных соединений 8 оказались ингибиторами. Детали механизма ингибирования исследуются, но о результатах этих исследований пока говорить преждевременно. Три лучших соединения (16826, 2425 и 2900) не уступают по значениям IC₅₀ реперным соединениям, а одно из них (16826) даже существенно превосходит по своей активности Эбселен. Кроме того, по сравнению с Эбселеном выявленные соединения не содержат атом селена, который обеспечивает неспецифическую модификацию цистеиновых остатков ряда белков, в том числе главной протеазы SARS-CoV-2 M^{pro}. По сравнению с Дисульфирамом данные ингибиторы характеризуются более высокой сложностью с точки зрения органической химии, а также в большинстве своём не содержат дисульфидную реактивную группировку, которая ответственна за неспецифический характер действия данного соединения против широкого круга мишеней. Отметим и большое химическое разнообразие наших ингибиторов: восемь ингибиторов относятся к семи различным химическим классам, которые обозначены на рисунке.

Эти 8 ингибиторов SARS-CoV-2 M^{pro} были исследованы на их способность подавлять репликацию этого коронавируса и на цитотоксичность в культуре клеток, и соответствующие результаты

также представлены в таблице. Как видим, только три из восьми ингибиторов подавляют репликацию коронавируса в культуре клеток. Похожий эффект наблюдался в работе [26], где было обнаружено 6 ингибиторов главной протеазы SARS-CoV-2, но ни один из них не показал противовирусной активности в культуре клеток.

В нашей работе из трёх соединений, подавлявших репликацию коронавируса, у двух (223350 и 18302) подавление репликации коронавируса оказалось достаточно сильным, характеризующимся величинами EC₅₀ в несколько мкМ, а индекс селективности оказался достаточно высоким — несколько десятков. Это указывает, что, по крайней мере, эти два соединения, а, возможно, и их аналоги, могут быть перспективными для дальнейших исследований на животных в качестве основы для создания новых противовирусных препаратов. За прошедший с начала пандемии COVID-19 год было открыто несколько нековалентных ингибиторов SARS-CoV-2 M^{pro}, однако большинство из них имеют активность от единиц до десятков мкМ, и далеко не все из них подавляют репликацию коронавируса SARS-CoV-2. Для многих из них был установлен неспецифический характер действия, что ограничивает их дальнейшую разработку [10]. Таким образом, найденные в данной работе нековалентные ингибиторы главной протеазы SARS-CoV-2 не уступают по ингибирующей активности лучшим известным нековалентным ингибиторам, а в некоторых случаях и превосходят их.

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

В данной работе приведены результаты поиска ингибиторов главной протеазы коронавируса SARS-CoV-2. Поиск проводили в базе данных

имеющихся в наличии лекарственно-подобных соединений кафедры органического синтеза Воронежского государственного университета. Для поиска ингибиторов использовали методику с комбинированным применением классического докинга с использованием программы докинга SOL и квантово-химических расчётов энтальпии связывания белок-лиганд с помощью полуэмпирического метода PM7 и континуальной модели растворителя COSMO. Используемая методика показала сравнительно высокую эффективность идентификации ингибиторов главной протеазы (8 из 18 лигандов, переданных для экспериментального тестирования) и ингибирующую активность субмикромольного и микромольного диапазонов величины IC₅₀, причём найденные ингибиторы относятся к 7 различным химическим классам. Три соединения из этих ингибиторов продемонстрировали микромольную активность по подавлению репликации коронавируса SARS-CoV-2 в культуре клеток Vero E6 с низкой цитотоксичностью. Полученные результаты показывают, что наиболее активные ингибиторы главной протеазы коронавируса (**2425**, **2900**, **16826**) не блокируют репликацию коронавируса в культуре клеток. Причины этого могут быть самые разные, поскольку культура клеток с находящимися в ней вирусами гораздо более сложная система, чем система белок-субстрат для измерения ингибирования главной протеазы. Одной из этих причин может быть плохое проникновение ингибиторов через мембрану клетки. Кроме того, некоторые соединения могут содержать реактивные фрагменты, неселективно модифицирующие разные белки-мишени. Примером служит соединение **2900**, содержащее дисульфидную реактивную группу, схожую с той, что есть в Дисульфираме. По-видимому, по аналогии с Дисульфирамом данное соединение необратимо связывается с несколькими типами клеточных белков по остаткам цистеина, что проявляется в отсутствии активности против SARS-CoV-2 в тесте на культуре клеток.

В данной работе были найдены менее активные ингибиторы главной протеазы коронавируса — **15689**, **18302**, **223350**, которые блокируют репликацию коронавируса SARS-CoV-2 и, в свою очередь, открывают возможность для разработки новых противовирусных препаратов прямого действия. У найденных ингибиторов есть потенциал для дальнейшей оптимизации и получения соединений с более высокой ингибирующей активностью как для ингибирования главной протеазы, так и для создания противовирусного препарата против коронавируса SARS-CoV-2.

ФИНАНСИРОВАНИЕ

Работа выполнена при поддержке гранта РНФ № 21-71-20031 с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит описания проведённых авторами исследований с участием людей или использованием животных в качестве объектов.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов

ЛИТЕРАТУРА

1. Wu F, Zhao S, Yu B, Chen Y.-M., Wang W, Song Z.-G., Hu Y, Tao Z.-W., Tian J.-H., Pei Y.-Y., Yuan M.-L., Zhang Y.-L. Y.-Z., Dai F.-H., Liu Y., Wang Q.-M., Zheng J.-J., Xu L., Holmes E.C., Zhang Y.-L. Y.-Z. (2020) *Nature*, **579**(7798), 265-269.
2. Sulimov V.B., Kutov D.C., Taschilova A.S., Ilin I.S., Tyrtysnikov E.E., Sulimov A.V. (2021) *Curr. Top. Med. Chem.*, **21**(6), 507-546.
3. Sulimov V.B., Kutov D.C., Sulimov A.V. (2019) *Curr. Med. Chem.*, **26**(42), 7555-7580.
4. Сулимов А.В., Кутов Д.К., Каткова Е.В., Кондакова О.А., Сулимов В.Б. (2017) *Изв. АН, Сер. Хим.*, **10**, 1913-1924. [Sulimov A.V., Kutov D.C., Katkova E.V., Kondakova O.A., Sulimov V.B. (2017) *Russ. Chem. Bull.*, **66**, 1913-1924.
5. Ksiazek T.G., Erdman D., Goldsmith C.S., Zaki S.R., Peret T., Emery S., Tong S., Urbani C., Comer J.A., Lim W., Rollin P.E., Dowell S.F., Ling A.-E., Humphrey C.D., Shieh W.-J., Guarner J., Paddock C.D., Rota P., Fields B., DeRisi J., Yang J.-Y., Cox N., Hughes J.M., LeDuc J.W., Bellini W.J., Anderson L.J. (2003) *N. Engl. J. Med.*, **348**(20), 1953-1966.
6. Zaki A., Boheemen S., Bestebroer T., Osterhaus A., Fouchier R. (2012) *N. Engl. J. Med.*, **367**, 1814-1820.
7. Barretto N., Jukneliene D., Ratia K., Chen Z., Mesecar A.D., Baker S.C. (2005) *J. Virol.*, **79**(24), 15189-15198.
8. Yang H., Yang M., Ding Y., Liu Y., Lou Z., Zhou Z., Sun L., Mo L., Ye S., Pang H., Gao G.F., Anand K., Bartlam M., Hilgenfeld R., Rao Z. (2003) *Proc. Natl. Acad. Sci. USA*, **100**(23), 13190-13195.
9. Jin Z., Du X., Xu Y., Deng Y., Liu M., Zhao Y., Zhang B., Li X., Zhang L., Peng C., Duan Y., Yu J., Wang L., Yang K., Liu F., Jiang R., Yang X., You T., Liu X., Yang X., Bai F., Liu H., Liu X., Guddat L.W., Xu W., Xiao G., Qin C., Shi Z., Jiang H., Rao Z., Yang H. (2020) *Nature*, **582**(7811), 289-293.
10. Ma C., Hu Y., Townsend J.A., Lagarias P.I., Marty M.T., Kolocouris A., Wang J. (2020) *ACS Pharmacol. Transl. Sci.*, **3**(6), 1265-1277.
11. Ghahremanpour M.M., Tirado-Rives J., Deshmukh M., Ippolito J.A., Zhang C.-H., Cabeza de Vaca I., Liosi M.-E., Anderson K.S., Jorgensen W.L. (2020) *ACS Med. Chem. Lett.*, **11**(12), 2526-2533.
12. Dai W., Zhang B., Jiang X.-M., Su H., Li J., Zhao Y., Xie X., Jin Z., Peng J., Liu F., Li C., Li Y., Bai F., Wang H., Cheng X., Cen X., Hu S., Yang X., Wang J., Liu X., Xiao G., Jiang H., Rao Z., Zhang L.-K., Xu Y., Yang H., Liu H. (2020) *Science*, **368**(6497), 1331-1335.
13. Wu Y., Li Z., Zhao Y.S., Huang Y.Y., Jiang M.Y., Luo H.B. (2021) *Med. Res. Rev.*, **41**(Jan 3), 1775-1797.
14. Sulimov A.V., Kutov D.C., Taschilova A.S., Ilin I.S., Solpovskaya N.V., Shikhaliev K.S., Sulimov V.B. (2020) *Supercomput. Front. Innov.*, **7**(3), 41-56.
15. Kutov D.C., Katkova E.V., Kondakova O.A., Sulimov A.V., Sulimov V.B. (2017) *Bull. South Ural State Univ. Ser. Math. Model. Program. Comput. Softw.*, **10**(3), 94-107.

16. Sulimov A.V., Kutov D.C., Oferkin I.V., Katkova E.V., Sulimov V.B. (2013) J. Chem. Inf. Model., **53**(8), 1946-1956.
17. Halgren T.A. (1996) J. Comput. Chem., **17**(5-6), 490-641.
18. Купервассер О.Ю., Жабин С.Н., Мартынов Я.Б., Федюлов К.М., Офёркин И.В., Сулимов А.В., Сулимов В.Б. (2011) Вычислительные методы и программирование, **12**, 246-261. [Kupervasser O.Y., Zhabin S.N., Martynov Y.B., Fedulov K.M., Oferkin I.V., Sulimov A.V., Sulimov V.B. (2011) Numer. Methods Program., **12**, 246-261.]
19. Voevodin V.V., Antonov A.S., Nikitenko D.A., Shvets P.A., Sobolev S.I., Sidorov I.Y., Stefanov K.S., Voevodin V.V., Zhumatiy S.A. (2019) Supercomput. Front. Innov., **6**(2), 4-11.
20. Stewart J.J.P. (2016) Stewart Computational Chemistry. MOPAC2016, Colorado Springs, CO, USA, URL: <http://openmopac.net/MOPAC2016.html>
21. Stewart J.J. (2013) J. Mol. Model., **19**(1), 1-32.
22. Klamt A., Schuurmann G (1993) J. Chem. Soc. Perkin Trans., **2**(5), 799-805.
23. Hostaš J., Řezáč J., Hobza P. (2013) Chem. Phys. Lett., **568-569**(Supplement C), 161-166.
24. Elfiky A.A. (2020) Life Sci., **253**, 117592. DOI: 10.1016/j.lfs.2020.117592
25. Sulimov A.V., Kutov D.C., Gribkova A.K., Ilin I.S., Tashchilova A.S., Sulimov V.B. (2019) Commun. Comput. Inf. Sci., Cham: Springer, **1129**, 363-378.
26. Kuzikov M., Costanzi E., Reinshagen J., Esposito F., Vangeel L., Wolf M., Ellinger B., Claussen C., Geisslinger G., Corona A., Iaconis D., Talarico C., Manelfi C., Cannalire R., Rossetti G., Gossen J., Albani S., Musiani F., Herzog K., Ye Y., Giabbai B., Demitri N., Jochmans D., de Jonghe S., Rymenants J., Summa V., Tramontano E., Beccari A.R., Leyssen P., Storici P., Neyts J., Gribbon P., Zaliani A. (2021) ACS Pharmacol. Transl. Sci., **2021**, DOI: 10.1021/acspstsci.0c00216

Поступила в редакцию: 15. 04. 2021.
 После доработки: 14. 05. 2021.
 Принята к печати: 15. 05. 2021.

DEVELOPMENT OF ANTIVIRAL DRUGS BASED ON INHIBITORS OF THE SARS-CoV-2 MAIN PROTEASE

A.V. Sulimov^{1,6}, Kh.S. Shikhaliev², O.V. Pyankov³, D.N. Shcherbakov^{3,4}, V.Yu. Chirkova⁴, I.S. Ilin^{1,6}, D.C. Kutov^{1,6}, A.S. Tashchilova^{1,6}, M.Yu. Krysin², D.V. Krylskiy², N.V. Stolpovskaya², E.A. Volosnikova³, S.V. Belenkaya^{3,5}, V.B. Sulimov^{1,6*}

¹Research Computer Center of Lomonosov Moscow State University,
 1 bldg 4, Leninskie Gory, Moscow, 119991 Russia; *e-mail: sulimovv@mail.ru

²Department of Organic Chemistry, Faculty of Chemistry, Voronezh State University,
 1 Universitetskaya sq., Voronezh, 394018 Russia

³State Research Centre of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk Region, 630559 Russia

⁴Altai State University, 61 Lenin ave., Barnaul, 656049 Russia

⁵Novosibirsk State University, 1 Pirogova str., Novosibirsk, 630090 Russia

⁶Dimonta Ltd., 15 bldg 8, Nagornaya str., Moscow, 117186 Russia

Docking and quantum-chemical methods have been used for screening of drug-like compounds from the own database of the Voronezh State University to find inhibitors the SARS-CoV-2 main protease, an important enzyme of the coronavirus responsible for the COVID-19 pandemic. Using the SOL program more than 42000 3D molecular structures were docked into the active site of the main protease, and more than 1000 ligands with most negative values of the SOL score were selected for further processing. For all these top ligands, the protein-ligand binding enthalpy has been calculated using the PM7 semiempirical quantum-chemical method with the COSMO implicit solvent model. 20 ligands with the most negative SOL scores and the most negative binding enthalpies have been selected for further experimental testing. The latter has been made by measurements of the inhibitory activity against the main protease and suppression of SARS-CoV-2 replication in a cell culture. The inhibitory activity of the compounds was determined using a synthetic fluorescently labeled peptide substrate including the proteolysis site of the main protease. The antiviral activity was tested against SARS-CoV-2 virus in the Vero cell culture. Eight compounds showed inhibitory activity against the main protease of SARS-CoV-2 in the submicromolar and micromolar ranges of the IC₅₀ values. Three compounds suppressed coronavirus replication in the cell culture at the micromolar range of EC₅₀ values and had low cytotoxicity. The found chemically diverse inhibitors can be used for optimization in order to obtain a leader compound, the basis of new direct-acting antiviral drugs against the SARS-CoV-2 coronavirus.

Key words: drug discovery; docking; quantum chemistry; SARS-CoV-2; M^{pro}; inhibitors; antiviral drugs; COVID-19; CADD

Funding. The work was financially supported by the Russian Science Foundation (project no. 21-71-20031). The research was carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University, including the Lomonosov supercomputer.

Received: 15.04.2021, revised: 14.05.2021, accepted: 15.05.2021.