

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«Тольяттинский государственный университет»

Институт химии и энергетики

(наименование института полностью)

Центр медицинской химии

(наименование)

04.03.01 Химия

(код и наименование направления подготовки)

Медицинская и фармацевтическая химия

(направленность (профиль))

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (БАКАЛАВРСКАЯ РАБОТА)

на тему Химическое пространство ковалентных ингибиторов E3-лигазы RNF4

Обучающийся

Е.О. Смиркина

(Инициалы Фамилия)

(личная подпись)

Руководитель

канд. биол. наук Д.А. Хоченков

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Консультант

А.Ю. Кудинов

(ученая степень (при наличии), ученое звание (при наличии), Инициалы Фамилия)

Тольятти 2023

Аннотация

Целью данной работы является поиск новых рекрутеров для E3-лигазы RNF4 с помощью методов ковалентного докинга и их разработка в программном пакете Schrodinger Suite. RNF4 является убиквитинлигазой E3 и ее рекрутирование важно для синтеза PROTAC. Химически индуцированное образование тройных комплексов приводит к убиквитинированию и протеасомной деградации белков-мишеней.

В литературном обзоре описана концепция PROTAC, сигнальный путь RNF4 и ее строение, роль E3 лигаз. Так же было описано молекулярное моделирование для поиска лекарств и принципы ковалентного докинга.

В экспериментальной части описана методика расчетных исследований, которая заключалась в поиске кристаллических структур, подготовке белка и структур, а так же в проведении докинга нативного лиганда и библиотеки соединений.

На основании полученных данных было найдено 9 перспективных соединений.

Работа изложена на 62 страницы, содержит 33 рисунка, 3 таблицы.

Abstract

The title of the graduation work is «Chemical space of covalent recruiters E3 ligase RNF4».

The graduation work consists of an explanatory note, an introduction, three parts on 33 pages, 11 figures, 3 tables, the list of 67 references including foreign sources.

The aim of this work is to search for new covalent types of RNF4 ligase recruiters and their development in the Schrodinger Suite software package.

The object of this thesis is the concept of covalent ligase recruitment to create PROTAC.

In the first part of the description of the PROTAC concept, the RNF4 signaling pathway and its structure, the role of the E3 ligase. Molecular modeling has also been identified to search for drugs and substantiate covalent docking.

The second part describes the methodology for the search for crystal structures, the overall design of the study, the preparation of the protein and structures, as well as the docking of the native ligand and the library of compounds.

The third part describes the methodology of computational studies.

Based on the initial data, 9 promising compounds were found that can be used in the future for the use of PROTAC.

Содержание

Определения, обозначения и сокращения	5
Введение.....	6
1 Литературный обзор	7
1.1 Роль PROTAC в убиквитин- протеосомной системе	7
1.1.1 Белковая деградация.....	9
1.1.2 Классификация убиквитин-лигаз	10
1.2 Молекулярное строение и биологические функции лигазы	12
RNF4	12
1.2.1 Сигнальный путь, контролируемый RNF4.....	13
1.3 Молекулярное моделирование	15
1.3.1 Молекулярный докинг	16
1.3.2 Скоринговые функции	17
1.3.3 Ковалентный докинг.....	21
1.4 Лиганды RNF4, подходы к синтезу.....	23
2 Результаты и их обсуждение.....	28
2.1 Аннотирование и подготовка библиотеки	29
2.3 Ковалентный докинг виртуальной библиотеки	36
3 Экспериментальная часть.....	53
3.1 Компьютерное моделирование.....	53
3.1.1 Подготовка структур белка.....	53
3.1.2 Подготовка виртуальных библиотек	53
Заключение	55
Список используемой литературы	56

Определения, обозначения и сокращения

В данной работе используются следующие термины с соответствующими определениями:

ADMET	всасывание, распределение, метаболизм, выведение, токсичность в фармакокинетике
c-Jun	белок комплекса активаторного белка-1 (AP-1)
BET	белки бромодомена и экстратерминального домена
GBSA	метод Борна сольватации площади поверхности
E1	убиквитин-активирующий фермент
E2	убиквитин-конъюгирующий фермент
E3	убиквитинлигаза, включающая убиквитин-лигирующие ферменты
HECT	домен, гомологичный карбоксильному концу E6-ассоциированного белка (E6-AP)
Myc	фактор транскрипции
PBSA	метод Пуассона-Больцмана сольватации площади поверхности
POI	белок интереса
RBR	гибридная лигаза RING-IBR-RING
RING	домен реально интересного нового гена
SUMO	малый убиквитин-подобный модификатор
TPD	направленная деградация белков
Wnt	сигнальный путь, который регулирует эмбриональное развитие и дифференцировку клеток

Введение

Целенаправленная деградация белка (TPD) – быстро развивающаяся область химической биологии и разработки лекарств. В качестве технологии, открывающей этот путь, химеры, нацеленные на протеолиз (PROTAC), состоят из двух фрагментов, соединенных линкером, и индуцируют образование тройного комплекса путем одновременного связывания лигазы E3 и целевого белка, что приводит к убиквитинированию и деградации целевого белка [1].

Современное искусство открытия и разработки новых лекарств основано на необратимых ингибиторах. Ковалентное ингибирование – это стратегия, которая используется для достижения необратимого ингибирования. Необратимые ингибиторы взаимодействуют со своими мишенями в зависимости от времени, и реакция протекает до завершения, а не до равновесия [2].

Ковалентные ингибиторы обладают некоторыми значительными преимуществами по сравнению с нековалентными ингибиторами, такими как ковалентные боеголовки, которые могут нацеливаться на редкие, неконсервативные остатки определенного целевого белка, и, таким образом, привели к разработке высокоселективных ингибиторов, ковалентные ингибиторы могут быть эффективны для нацеливания на белки с неглубоким расщеплением связывания. [2].

Цель работы: поиск новых рекрутеров для E3-лигазы RNF4 с помощью методов ковалентного докинга.

Для достижения данной цели нами были поставлены следующие задачи:

1. Отбор и подготовка белковых структур RNF4;
2. Выбор, подготовка и аннотирование коммерческой библиотеки ковалентных лигандов;
3. Проведение ковалентного докинга и анализ полученных результатов.

1 Литературный обзор

1.1 Роль PROTAC в убиквитин- протеосомной системе

PROTACs (англ. PROteolysis TArgeting Chimeras) представляют собой бифункциональные молекулы, включающие рекрутирующий лиганд убиквитинлигазы E3, лиганд интересующего белка (POI), соединенные через определенный линкер [3]. На рисунке 1 представлена схема структуры PROTAC.

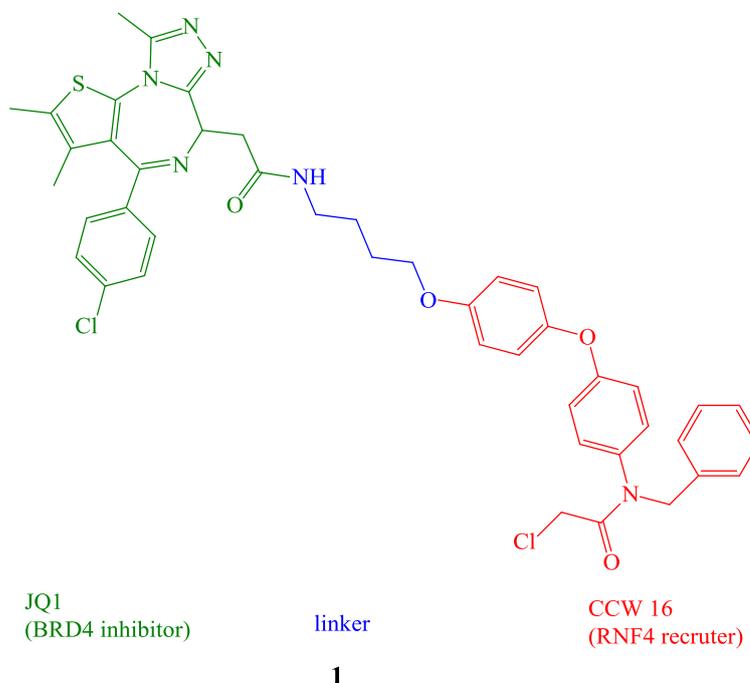


Рисунок 1 – Пример PROTAC, являющийся бифункциональным деградером

Данная многообещающая стратегия в химической биологии и открытии лекарств, как ожидается, проложит путь для усовершенствования медицины [4].

PROTAC успешно использовались для нацеливания на рецепторы эстрогена при лекарственно-устойчивом раке молочной железы и ферменты тирозинкиназы Брутона при лекарственно-устойчивых лимфомах,

демонстрируя непревзойденный потенциал PROTAC для лечения злокачественных и персистирующих опухолей [5,6]. Бифункциональные молекулы также были успешно разработаны для лечения других устойчивых к лекарствам заболеваний, таких как гепатит С, и было показано, что они работают даже против мутантных вирусов, устойчивых к ингибированию [7].

Лиганд POI (белок интереса) избирательно нацеливается и «захватывает» интересующий белок, связываясь с ним, после чего связывается с лигандом E3 для образования тройного комплекса (рисунок 2). Убиквитинлигаза E3 впоследствии опосредует перенос убиквитина от лигазы E2 к белку-мишени, так что белок-мишень метится убиквитином. Затем специфический тройной комплекс диссоциирует, и убиквитинированный белок-мишень мигрирует к протеасоме 26S для деградации. В конце концов, PROTAC вновь взаимодействует с оставшимися белками – мишенями для нового цикла [8, 9]. Образование тройного комплекса является ключевым этапом в каталитическом цикле, а «эффект крючка» потенциально вызывает снижение или даже полное отсутствие активности.

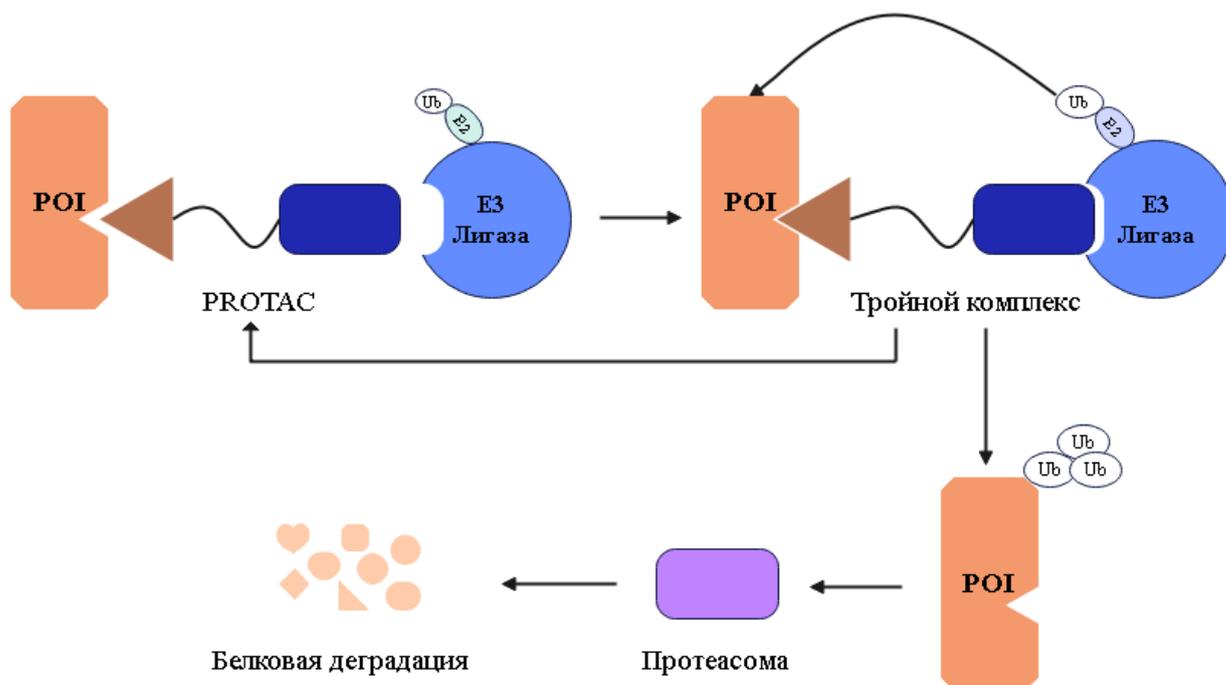


Рисунок 2 – Механизм действия классических PROTAC

При высоких концентрациях PROTAC насыщают связывание с мишенью и лигазой E3, что приводит к образованию бинарных комплексов, вместо тройного. Это предотвращает убиквитинирование и деградацию мишени. Данная ситуация наблюдалась со многими PROTAC и была названа эффектом «крючка» [10].

Эффект крючка (снижение деградации при высоких концентрациях низкомолекулярных PROTAC) часто наблюдался в исследованиях, в которых изучалась эффективность деградации низкомолекулярных PROTAC. Обычно он возникает в трехкомпонентных системах. Когда концентрация молекул PROTAC значительно выше значения DC_{50} , возникает аутоингибирование образования тройных комплексов E3:PROTAC:мишень из-за высоких концентраций бинарных комплексов PROTAC:E3 и PROTAC:мишень [11].

PROTAC являются мощными молекулами, которые могут разрушать различные внутриклеточные и трансмембранные белки в клетках мышей и, недавно, в людях. Успех PROTAC помог запустить область направленной деградации белков (TPD) [12].

1.1.1 Белковая деградация

Белковый гомеостаз, также известный как протеостаз, необходим для сбалансированного процесса модуляции внутриклеточных белков для поддержания нормальных биологических функций в организме [13]. Однако, когда происходит неправильный фолдинг белка и нерегулируемая экспрессия, естественный механизм гомеостаза будет нарушен, что приведет к возникновению и прогрессированию ряда заболеваний, не только рака, но и к сердечно-сосудистым, нейродегенеративным, иммунным заболеваниям и диабету [14]. Следовательно, очень важно, чтобы механизм контроля качества клеточного белка сохранял динамическое равновесие между правильным фолдингом белка, биологической активностью и временем жизни белка для восстановления протеостаза [15].

Протеостаз – это система поддержания протеом в функциональном состоянии. Она необходима для поддержания нормальной клеточной

метаболической функции и жизненно важна для динамических изменений, необходимых клетке для ответа на данный стимул. Обширная и сложная сеть сигнальных путей участвует в защите клеток и организмов от протеотоксического стресса. Они делают это, поддерживая тонкий баланс между производством и утилизацией белков. Сеть протеостаза включает пути, которые регулируют биогенез, фолдинг, транспортировку и деградацию белков [16].

Направленная деградация белков — новое направление в области разработки лекарственных средств, обладающее большим потенциалом для лечения сложных заболеваний. Как правило, убиквитин-протеасомная система и аутофагия являются двумя основными подходами к деградации белков и органелл в эукариотических клетках [17].

1.1.2 Классификация убиквитин-лигаз

Клетки содержат несколько протеолитических систем для осуществления процесса деградации и сложные регуляторные механизмы, обеспечивающие высокую избирательность непрерывных протеолитических процессов. Во всех тканях большинство внутриклеточных белков расщепляются убиквитин- протеасомным путем [18].

Убиквитинирование представляет собой процесс, при которой убиквитин, полипептид из 76 аминокислот, связывается с белком-субстратом, тем самым регулируя множественные гомеостатические клеточные механизмы, такие как нацеливание белка на протеасомную деградацию, каскады передачи сигнала, и врожденные иммунные реакции [19]. Процесс убиквитинирования катализируется последовательным действием трех ферментов, известных как лигазы E1, E2 и E3 (рисунок 3). Убиквитинлигазы E3 являются наиболее многочисленными по сравнению с ферментами E1 и E2 и контролируют специфичность убиквитинирования. Существует три основных класса убиквитинлигаз E3 в зависимости от их молекулярной структуры и механизма переноса убиквитина на субстраты.

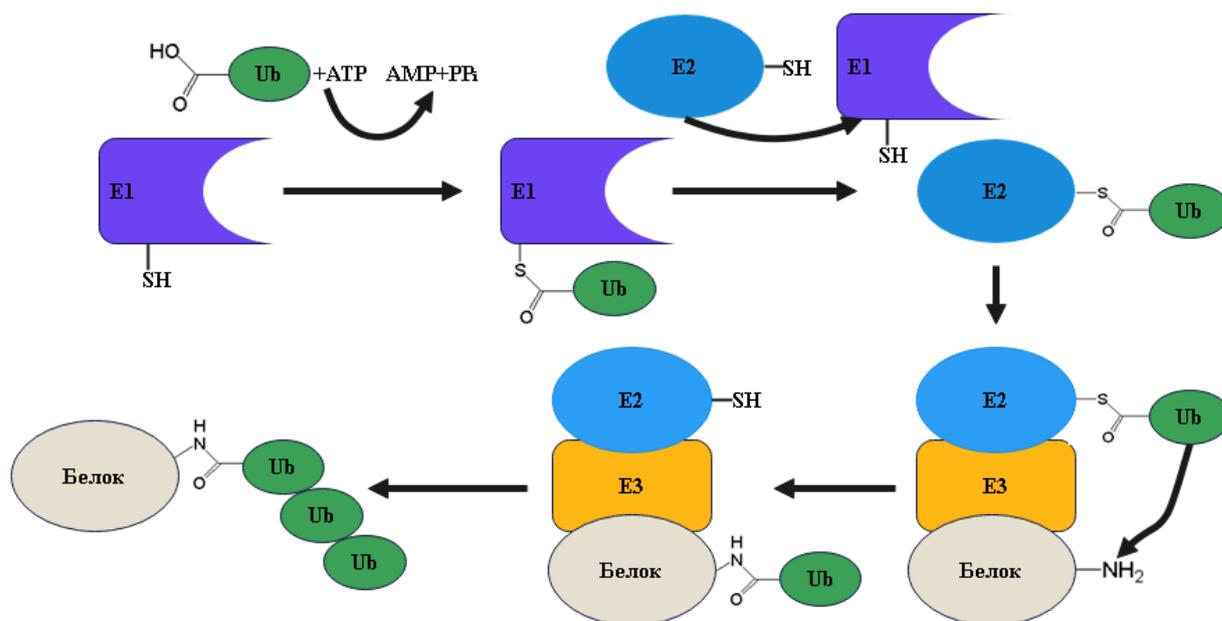


Рисунок 3 – Механизм полиубиквитинирования белка-мишени

Семейство RING содержит наибольшее количество лигаз E3, и имеет домен белка типа цинкового пальца. Домен RING или U-box служит стыковочным сайтом для связывания несущего убиквитин E2 и для облегчения переноса убиквитина непосредственно от E2 к его субстратам [20]. Семейство HECT (гомологичных карбоксильному концу E6-AP) состоит из членов с консервативным доменом HECT на C-конце. Они непосредственно катализируют убиквитинирование субстрата в двухстадийном процессе: после связывания с E2 они загружают активированный убиквитин на себя посредством образования каталитического промежуточного соединения убиквитин-тиоэфир в активном центре Cys внутри домена HECT. На втором этапе они переносят убиквитин на свой белок-мишень [21]. Наконец, семейство лигаз E3 RBR (RING-in-Between-RING) состоит из двух RING-доменов, и они катализируют перенос убиквитина на свои субстраты способом, аналогичным таковому у лигазы семейства HECT [22].

1.2 Молекулярное строение и биологические функции лигазы RNF4

RNF4 (Really Interesting New Gene Finger Protein 4) представляет собой класс убиквитинлигаз E3, которые нацелены на белки, связанные с малым убиквитин-подобным модификатором SUMO (Small Ubiquitin-like Modifier), для дальнейшего убиквитирования. На сегодняшний день регуляторная функция RNF4, по-видимому, заключается в убиквитин-опосредованной деградации сумоилированных клеточных белков [23]. На рисунке 4 представлено молекулярное строение лигазы RNF4.

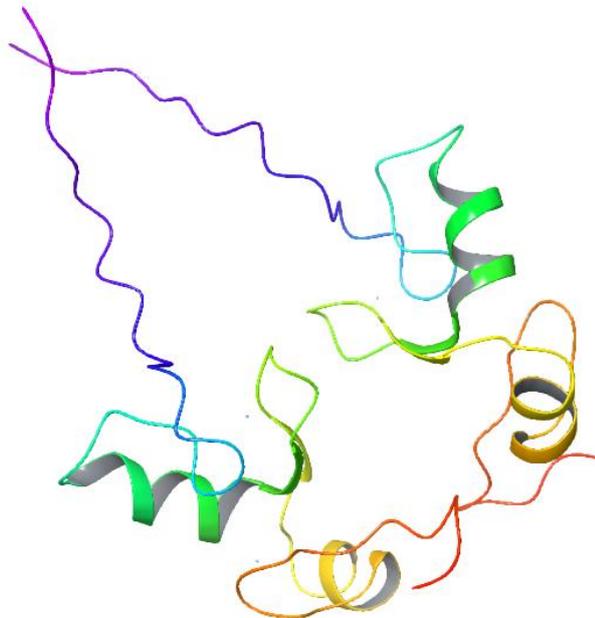


Рисунок 4 – Строение RNF4

RNF4 связывает полисумоилированные цепи, ковалентно присоединенные к белкам, и опосредует полиубиквитинирование этих субстратов, связанное с Lys-6, Lys-11, Lys-48 и Lys-63, и их последующее направление к протеасоме для деградации. Лигаза состоит из 190 остатков, содержит С-концевой домен RING, который связывает E2 и N-концевые SUMO-взаимодействующие мотивы, которые связываются с субстратом

polySUMO. Она также может связываться с ДНК/нуклеосомами и играть роль в регуляции транскрипции, например, усиливая базальную транскрипцию и активацию транскрипции, опосредованную стероидными рецепторами. Широко экспрессируется на низких уровнях во многих тканях [24].

RNF4 является преимущественно мономерным и неактивным в качестве убиквитин-Е3-лигазы. Однако в присутствии цепей SUMO RNF4 активируется путем димеризации, что приводит как к субстратному убиквитилированию, так и к аутоубиквитилированию, ответственному за деградацию RNF4 (рисунок 5). Таким образом, активность убиквитин-Е3-лигазы RNF4 напрямую связана с доступностью его polySUMO –субстратов [25].

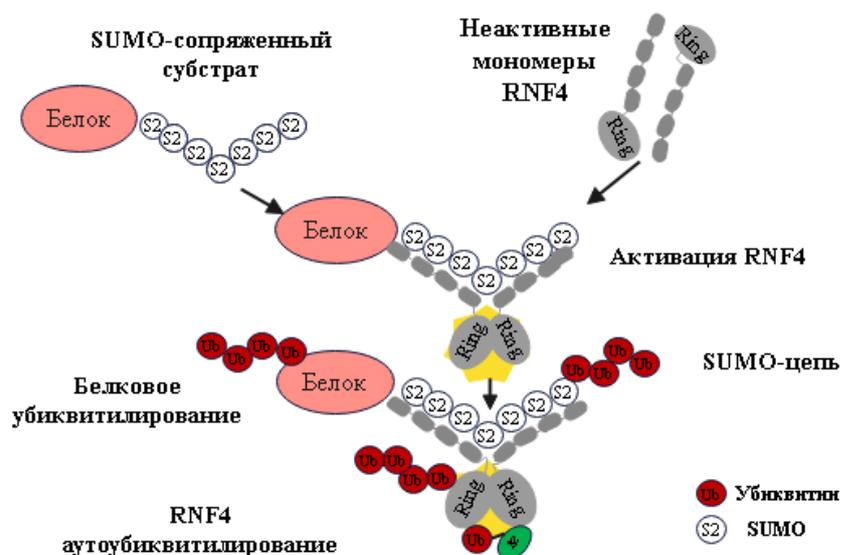


Рисунок 5 – Субстрат - опосредованная активация RNF4

1.2.1 Сигнальный путь, контролируемый RNF4

Посттранскрипционные модификации ядерных онкопротеинов играют важную роль в развитии рака. Многие онкогенные факторы транскрипции фосфорилируются митогенными сигнальными путями, что усиливает их активность. Кроме того, модификации убиквитином и убиквитин-

подобными белками регулируют критические сигнальные пути, важные для развития рака, во многих случаях отправляя белки на деградацию, а также факторы транскрипции [26]. Убиквитилирование с помощью RNF4 стабилизирует короткоживущие в остальном онкогенные факторы транскрипции, включая β -катенин, Мус, с-Jun и белок внутриклеточного домена Notch (N-ICD) (рисунок 6).

RNF4 усиливает транскрипционную активность этих факторов, а также Wnt- и Notch- зависимую экспрессию генов. В то время как RNF4 представляет собой убиквитинлигазу, нацеленную на SUMO, для стабилизации белка требуется фосфорилирование субстрата, а не SUMOилирование, и связывание с доменом RNF4, богатым аргинином [27, 28]. Стабилизация также включает образование необычного полиубиквитина и присоединение RNF4 к хроматину. Биологически RNF4 усиливает фенотип опухоли и необходима для выживания раковых клеток. Высокие уровни мРНК RNF4 коррелируют с плохой выживаемостью подгруппы пациентов с раком молочной железы, а уровни белка RNF4 повышены в 30% аденокарцином толстой кишки человека. Таким образом, RNF4-зависимое убиквитилирование переводит временный сигнал фосфорилирования в долговременную стабилизацию белка, что приводит к усиленной активации онкопротеина [29].

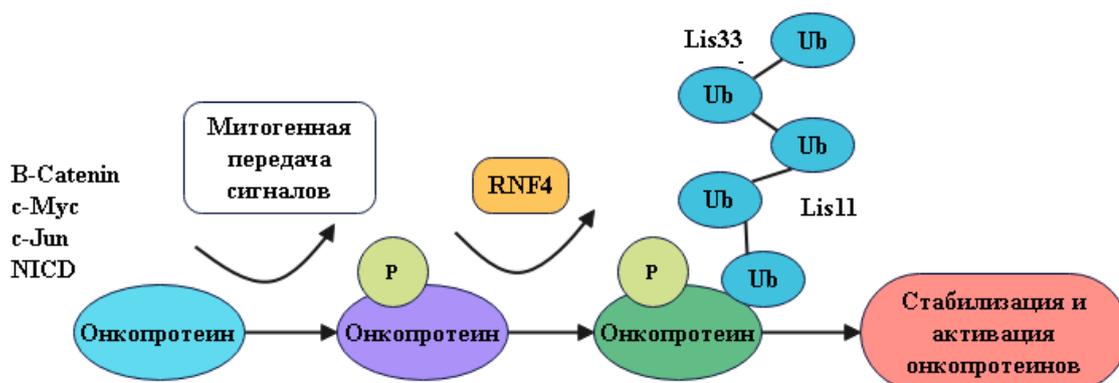


Рисунок 6 – Сигнальный путь в активации онкопротеинов

1.3 Молекулярное моделирование

Молекулярное моделирование представляет собой совокупность теоретических и вычислительных методов, используемых для исследования структуры и свойств молекул *in silico* [30].

Определение фармакокинетических свойств молекул может быть выполнено компьютерным путем. Текущие программы молекулярного моделирования и веб-инструменты включают модули для исследования свойств ADMET на основе правил Липински. Таким образом, можно получить представление о молекулах-предшественниках без необходимости проведения экспериментальных исследований [31].

Компьютерный поиск лекарств (CADD) — популярный подход к ускорению процесса. CADD на основе лиганда использует информацию об известных низкомолекулярных связующих, чтобы предсказать, какие дополнительные молекулы также могут быть фармакологически активными. Принимая во внимание физико-химические и структурные свойства известных биоактивных соединений, можно разработать родственные молекулы с улучшенным сродством или другими свойствами, даже если конкретная мишень неизвестна. Однако методы на основе лигандов имеют ряд недостатков. Например, они могут не идентифицировать новые лиганды, которые существенно отличаются от известных, и они неэффективны против мишеней без известных лигандов [32].

Напротив, основанный на рецепторе (т.е. на структуре) CADD использует трехмерную структуру известной мишени для идентификации новых лигандов. CADD на основе рецепторов не зависит от уже известных лигандов и поэтому может идентифицировать первые в своем классе биологически активные молекулы. Однако подходы, основанные на структуре, требуют известной лекарственной мишени с известной структурой с атомарным разрешением.

1.3.1 Молекулярный докинг

В области молекулярного моделирования докинг представляет собой метод, предсказывающий предпочтительную ориентацию одной молекулы по отношению к другой, когда лиганд и мишень связываются друг с другом с образованием стабильного комплекса [33]. Молекулярные взаимодействия очень часто приводят к образованию устойчивых комплексов белок-белок или белок-лиганд.

На рисунке 7 показаны ключевые этапы докинга, общие для всех протоколов. Стыковка включает в себя поиск наиболее благоприятного режима связывания лиганда с интересующей мишенью. Способ связывания лиганда с рецептором может быть однозначно определен его переменными состояниями. Они состоят из его положения, ориентации и, если лиганд является гибким, его конформация (торсионные углы для каждой вращающейся связи). Каждая из этих переменных состояния описывает одну степень свободы в многомерном пространстве поиска, а их границы описывают степень поиска [34].

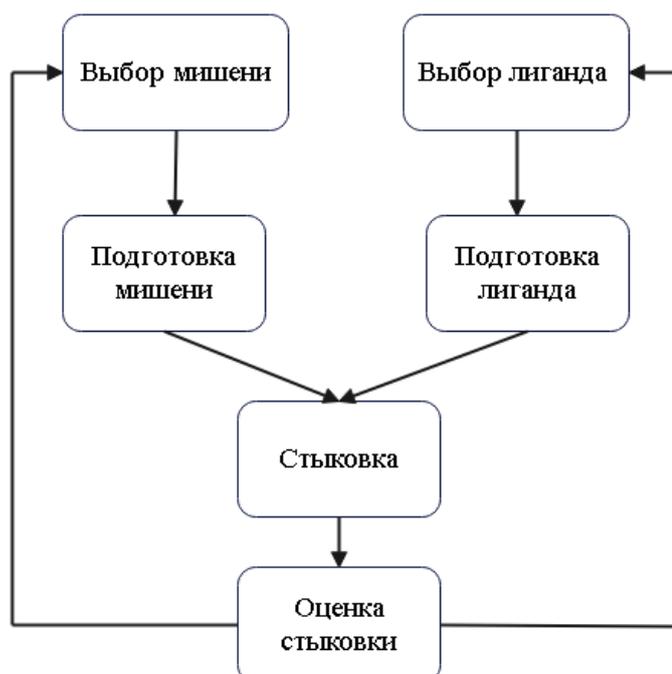


Рисунок 7 – Основные этапы докинга

Процедура докинга основана на образовании комплекса белок-лиганд (рисунок 8). Первоначально лиганд в растворе принимает множество конформаций. Затем белок распознает одну или несколько из этих конформаций. При распознавании лиганд, белок и растворитель следуют определенному энергетическому механизму, образуя окончательный комплекс [35].

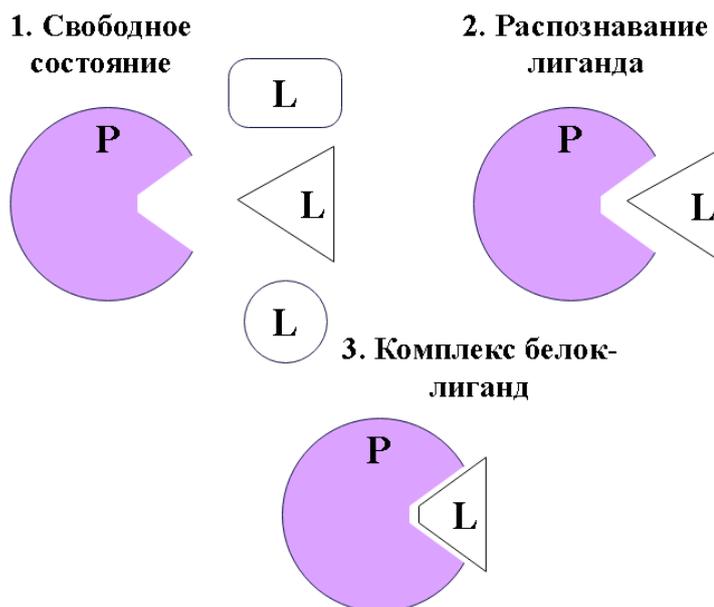


Рисунок 8 – Схема модели образования комплекса белок-лиганд

1.3.2 Скоринговые функции

В области вычислительной химии и молекулярного моделирования скоринговые функции представляют собой математические выражения, используемые для приблизительного предсказания аффинности связывания между двумя молекулами после их стыковки [36].

Можно рассмотреть несколько важных аспектов оценочных функций для стыковки белок-лиганд (рисунок 9).



Рисунок 9 – Иллюстрация категорий и оценок функций оценки для стыковки белок-лиганд

Существуют три класса оценочных функций:

Функция оценки силового поля – состоит из суммы энергетических членов классического силового поля (межмолекулярные силы Ван-дер-Ваальса и электростатические взаимодействия), обычно учитывающих энергии взаимодействия комплекса белок-лиганд и внутреннюю энергию лиганда, тогда как энергия сольватации может быть рассчитана с помощью континуальных моделей сольватации таких как GBSA или PBSA [37, 38].

Одной из типичных функций оценки силового поля при молекулярном докинге является функция оценки DOCK, энергетические параметры которой взяты из силовых полей Amber [39, 40]. Функция подсчета очков состоит из двух энергетических компонентов vdW (Ван-дер-Ваальса) Леннарда-Джонса и электростатического элемента определяется по формуле 1 [41].

$$E = \sum_i \sum_j \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right), \quad (1)$$

где r_{ij} - расстояние между атомами белка i и атомом лиганда j ;
 A_{ij} и B_{ij} – коэффициенты притяжения и отталкивания;
 q_i и q_j – атомные заряды;
 $\varepsilon(r_{ij})$ – диэлектрическая проницаемость.

Второй вид скоринговых функций — это эмпирические скоринговые функции, основанные на подсчете количества различных типов взаимодействий между двумя партнерами по связыванию [37]. Условия взаимодействия функции могут включать гидрофобные (благоприятные), гидрофобно- гидрофильные (неблагоприятные), количество водородных связей, число вращающихся связей. Примером эмпирических оценочных функций является Glide [42].

Подсчет по формуле 2 основан на количестве атомов лиганда и рецептора, находящихся друг с другом в контакте, или путем расчета изменения площади, доступной для контакта с растворителем ($\Delta SASA$) в комплексе по сравнению с лигандом и белком без комплекса [41].

$$\Delta G = \sum_i W_i \Delta G_i, \quad (2)$$

где ΔG – различные энергетические термины (энергия VDW, электростатика, водородная связь, десольватация, энтропия, гидрофобность);
 W_i – коэффициент аффинности связывания.

Третий тип оценочных функций - это оценочные функции, основанные на знаниях основаны на статистическом анализе взаимодействующих пар атомов из комплексов белок-лиганд с доступными трехмерными структурами. Эти данные о парных атомах преобразуются в псевдопотенциал, также известный как потенциал

средней силы, который описывает предпочтительную геометрию парных атомов белок-лиганд [43].

Принцип, лежащий в основе оценочных функций, основанных на знаниях, прост: парные потенциалы получаются непосредственно из частоты встречаемости пар атомов в базе данных с использованием обратной зависимости Больцмана [41]. Для исследований белок-лиганд потенциалы рассчитываются по формулам 3 и 4:

$$w(r) = -k_B T \ln[g(r)], \quad (3)$$

$$g(r) = \rho(r) / \rho^*(r). \quad (4)$$

где $w(r)$ – потенциал средней силы;

k_B – постоянная Больцмана;

T – абсолютная температура системы;

$g(r)$ – функция распределения в простой жидкостной системе;

$\rho(r)$ – числовая плотность пары атомов белок– лиганд на расстоянии r ;

$\rho^*(r)$ – плотность пар в эталонном состоянии, когда межатомные взаимодействия равны нулю.

Несмотря на большое количество разработанных скоринговых функций, ни одна из них не является идеальной с точки зрения точности и общей применимости. Каждая скоринговая функция имеет свои преимущества и ограничения. Чтобы воспользоваться преимуществами и сбалансировать недостатки различных функций оценки, был введен метод оценки консенсуса, чтобы повысить вероятность нахождения правильных решений путем объединения оценок от нескольких функций оценки [44]. Важнейшим шагом в оценке консенсуса является разработка соответствующей стратегии оценки консенсуса для отдельных баллов, чтобы можно было соответствующим образом отличить истинные режимы/связующие вещества от других [45].

1.3.3 Ковалентный докинг

Ковалентное ингибирование обычно представляет собой механизм, при котором низкомолекулярные соединения обратимо или необратимо инактивируют свои белковые рецепторы-мишени. Как правило, для ковалентного ингибирования требуется двухстадийный зависимый процесс (рисунок 10). Прежде всего, ингибитор образует обратимую связь с целевым ферментом, приближая химическую боеголовку ингибитора к определенному реакционноспособному аминокислотному остатку фермента. На втором этапе происходит образование ковалентной связи между ингибитором и реактивным компонентом фермента [46, 47].

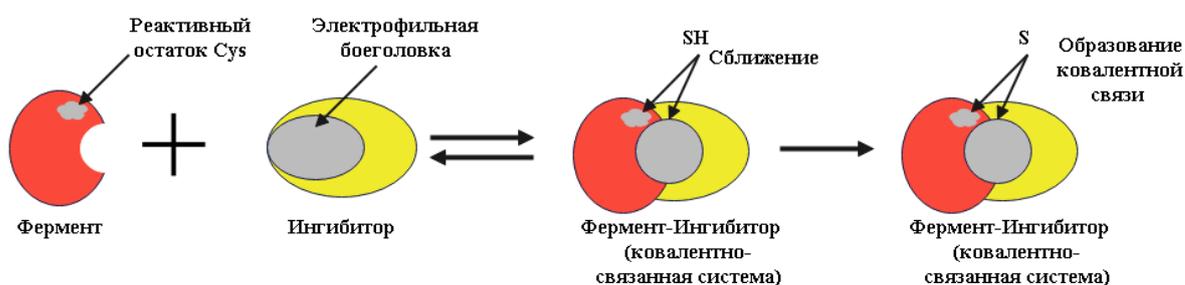


Рисунок 10 – Две стадии, необходимые для ковалентного ингибирования фермента

Обратимые ингибиторы отличаются от ковалентных ингибиторов отсутствием второй стадии [48]. Через некоторое время ковалентно-конъюгированный ингибитор может подвергаться дополнительным химическим превращениям, чтобы высвободиться из целевого фермента. Кроме того, он может образовывать необратимую реакцию с мишенью, тем самым ограничивая фермент в неактивном состоянии.

Обычно требуется три этапа для создания ковалентного ингибитора для данного фермента-мишени. Во-первых, структурный анализ мишени показывает, какой нуклеофил (например, цистеин, серин) присутствует в потенциальном кармане связывания или рядом с ним. Нуклеофил должен быть уникальным в пределах этого семейства белков, в противном случае

селективность будет низкой. Во-вторых, идентифицируется обратимый ингибитор с некоторой эффективностью (значения IC_{50} в диапазоне от моль/л до микромоль/л), и выясняется его способ связывания и взаимодействия. Этот ингибитор может быть получен из ранее разработанных ингибиторов родственных ферментов. Наконец, электрофильная «боеголовка» помещается в представляющий интерес обратимый ингибитор для точного взаимодействия с выбранным нуклеофилом фермента-мишени. Как правило, изостерическое замещение и аналоговый синтез используются для создания кандидатов в активные ковалентные ингибиторы [49]. На рисунке 11 представлен типичный рабочий процесс для проведения моделирования ковалентной стыковки.

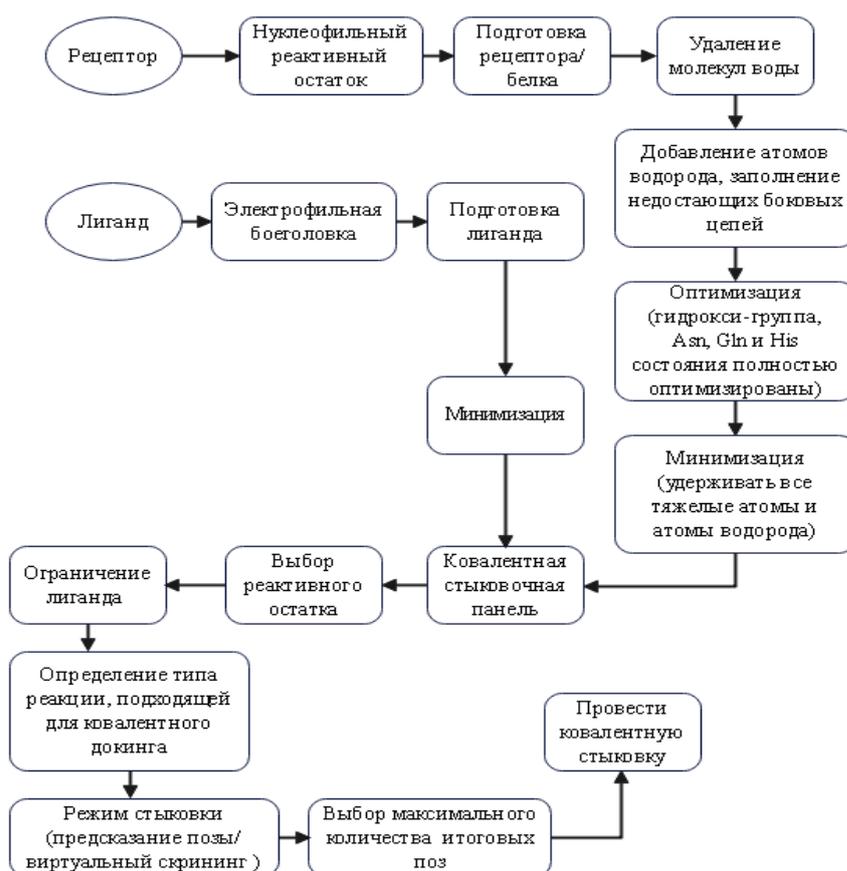


Рисунок 11 – Моделирование ковалентной стыковки

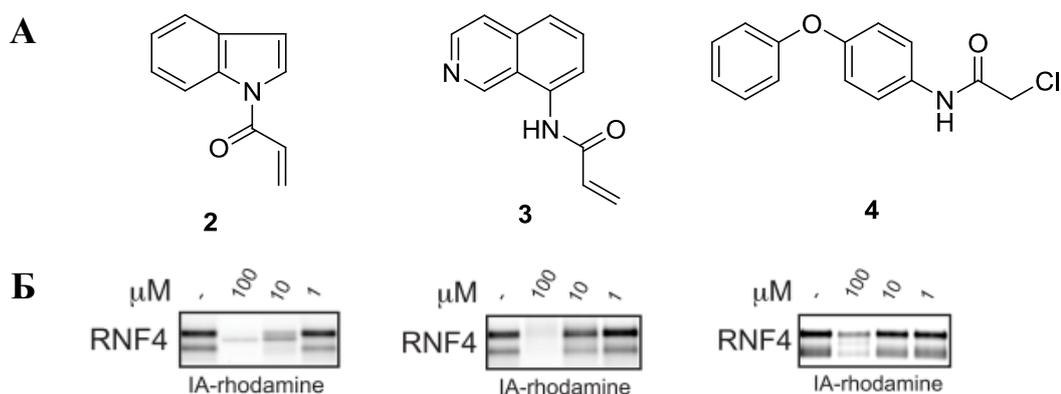
Надежный метод ковалентной стыковки должен быть способен взвесить все важные вклады в связывающие взаимодействия, в том числе в результате образования ковалентной связи. К сожалению, существенным недостатком является невозможность эффективного подсчета энергии, связанной с образованием связи. Моделирование химической реакции требует трудоемких квантово-механических расчетов, которые еще не подходят для методов быстрой стыковки. Оценочная функция систем ковалентной стыковки недостаточна с точки зрения образования связи, что препятствует оценке лигандов с различной степенью реакционной способности и, следовательно, может отрицательно сказаться на успешности виртуального скрининга при идентификации ковалентных ингибиторов[50].

1.4 Лиганды RNF4, подходы к синтезу

В целевой деградации белков чаще всего используют гетеробифункциональные молекулы деградации, состоящие из лиганда, нацеливающего на субстрат, связанного с рекрутером лигазы E3. Эти деструкторы способны рекрутировать лигазы E3 к специфическим белковым мишеням для убиквитинирования и маркировки мишеней для деградации протеосомозависимым образом. Поскольку функциональное ингибирование мишени не является необходимым для эффективности деградации, эта стратегия может нацеливаться и разрушать любой белок в протеоме, для которого существует лиганд [51].

Профилирование белка на основе активности (ABPP) стало мощной платформой для обнаружения лигандов против интересующих мишеней, включая белки, которые обычно считаются не поддающимися лечению. В этом подходе используются низкомолекулярные химические зонды для понимания механизмов взаимодействия между соединениями и мишенями [52].

В поисках лигандов, нацеленных на RNF4 в 2019 году группа исследователей под руководством Ward провели скрининг библиотеки ковалентных лигандов, реагирующих с цистеином, против мечения IA-родамином чистого человеческого RNF4 с использованием ABPP на основе геля [53]. Соединения YP 1-44 **2**, TRH 1-163 **3**, и TRH 1-23 **4** продемонстрировали воспроизводимое и дозозависимое ингибирование мечения RNF4 иодородамином (IA-родамин) (рисунок 12).



А – Ковалентные лиганды; Б – мечение IA-родамином RNF4

Рисунок 12 – Химические структуры ковалентных лигандов

TRH 1-23 **4** оказался наиболее многообещающим хитом RNF4. Он ковалентно модифицировал любой из двух цинккоординирующих цистеинов Cys 132 и Cys135 в RING домене RNF4. Были синтезированы несколько аналогов TRH 1-23 **4**. Среди них был найден CCW 16 **5**, самый мощный среди аналогов (рисунок 13).

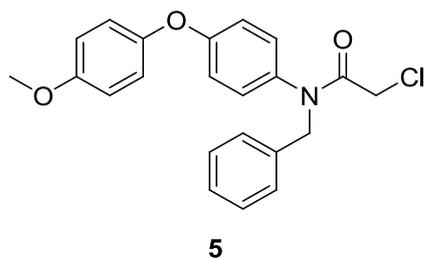


Рисунок 13 – CCW-16, ведущий лиганд RNF4

На рисунке 14 показана модель ковалентного связывания CCW-16 с на общедоступной кристаллической структуре RNF4 (PDB ID: 4PPE [54]).

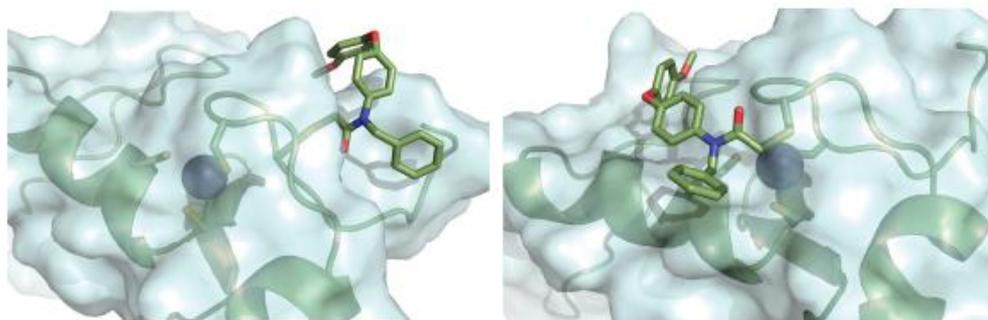
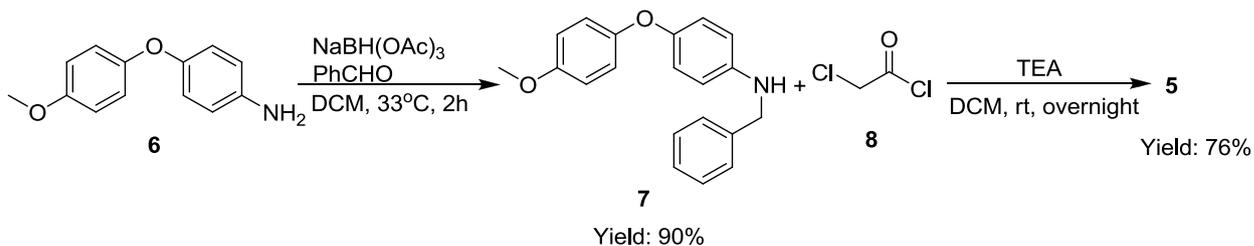


Рисунок 14 – Моделирование связывания CCW-16 в RNF4 (PDB ID: 4PPE)

Синтез соединения CCW-16 представлен на схеме 1. *N*-бензил-4-(4-метоксифенокси)анилин **6** обрабатывали бензальдегидом в присутствии триацетоксиборгидрида натрия с получением соединения **7**. Последующее алкилирование дизамещенного амина с **8** давало продукт **5**.

Схема 1



4-метоксигруппа на CCW -16 представляла собой идеальное положение для удлинения линкера с получением деструктора на основе RNF4. Чтобы продемонстрировать, использование CCW-16 в качестве потенциального рекрутера RNF4 для целевой деградации белка, авторы синтезировали CCW 28-3, бифункциональный деградер, связывающий CCW-16 с ингибитором семейства бромодоменов BET JQ1 (рисунок 15).

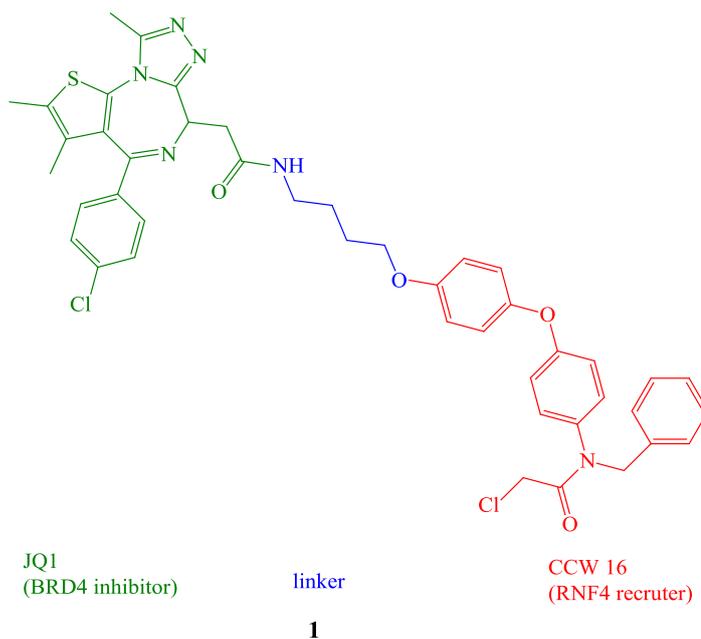
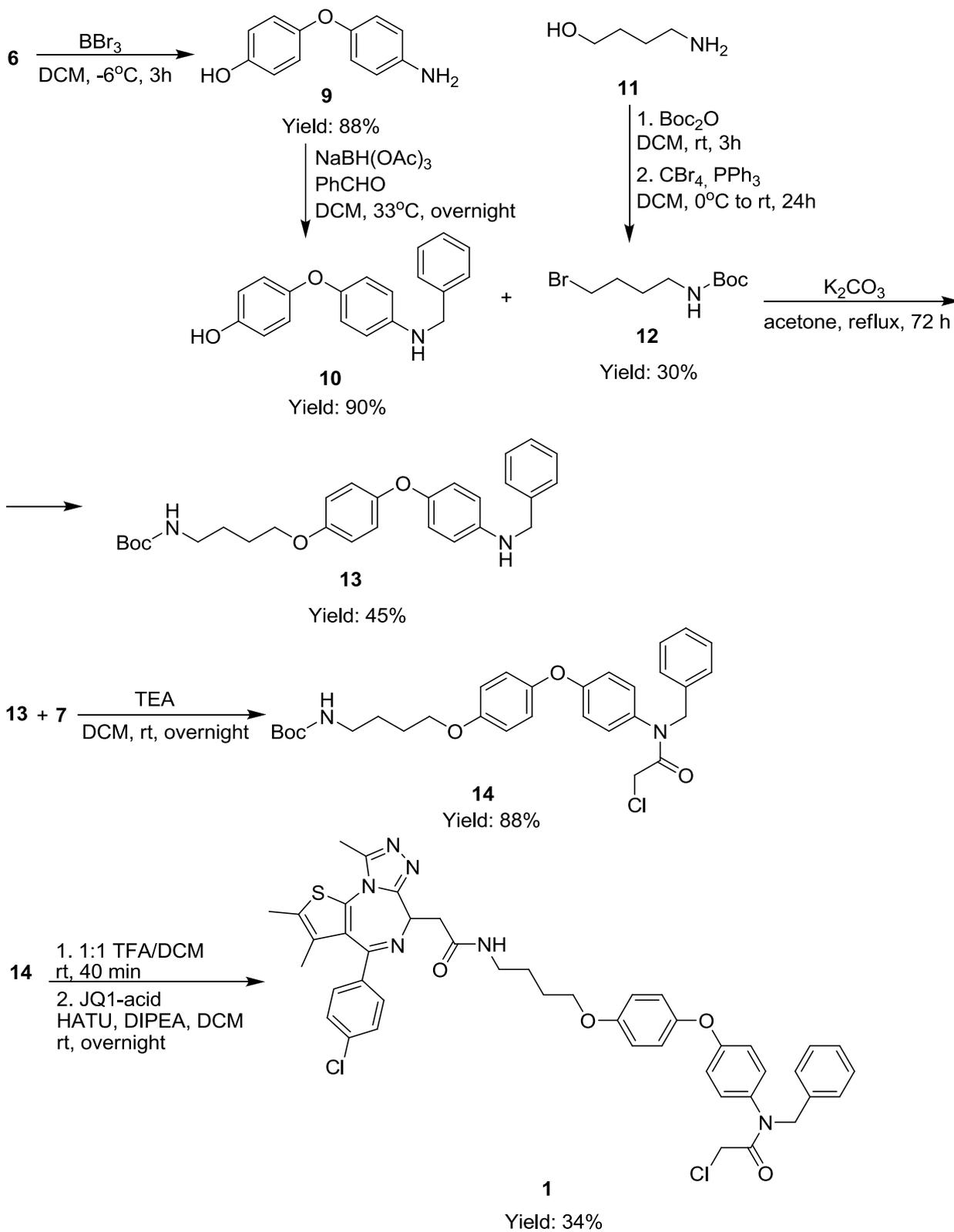


Рисунок 15 – Деградер на основе рекрутера RNF4

Деметилирование **5** дает соединение **9**, которое подвергается восстановительному аминированию бензальдегидом с использованием $\text{NaBH}(\text{OAc})_3$ в качестве восстанавливающего агента с образованием **10**. Алкилирование фенольного гидроксила бромидом **12** и последующая реакция с 2-хлорацетилхлоридом **7** позволили функционализировать рекрутер RNF4 линкером, содержащим Вос-защищенную аминогруппу. Удаление Вос-защиты с помощью трифторуксусной кислоты восстанавливало функциональный первичный амин, который затем можно было реагировать с гидролизованной свободной кислотной формой JQ1 посредством сочетания амида с получением **8** (схема 2).



2 Результаты и их обсуждение

Направленная деградация белка представляет собой мощную стратегию для открытия лекарств, позволяющая истощать белки, не поддающиеся лечению, с помощью механизмов протеосомной деградации. Для этих целей часто используют гетеробифункциональные деструкторы PROTAC, состоящие из лиганда, нацеленного на белок, связанного линкером с рекрутером лигазы E3.

Одна из проблем концепции химерных молекул заключается в том, что в настоящее время существует лишь ограниченное количество рекрутеров лигаз, которые могут быть использованы несмотря на сотни известных E3 в геноме человека [53]. Соответственно, разработка новых лигандов, равно как и поиск новых рекрутируемых убиквитинлигаз является важной и перспективной областью исследований.

За последние 30 лет дизайн ковалентных лекарств привлек большой интерес, и ковалентное нацеливание на неконсервативные аминокислоты для повышения селективности стало обычным явлением [55]. Между тем, в поиске новых ковалентных рекрутеров E3 сейчас наблюдается большой прогресс, и не последнее роль в этом сыграли методы *in silico* моделирования.

Целью данной работы стал поиск новых рекрутеров для E3-лигазы RNF4 с помощью методов ковалентного докинга.

На рисунке 16 представлен общий дизайн исследования. На первом этапе нами был осуществлен выбор коммерческой библиотеки ковалентных фрагментов ChemDiv [56], а так же отбор X-Ray структур из PDB [57]. Затем мы выполнили ковалентный докинг с дополнительным расчетом аффинности MM/GBSA [58].

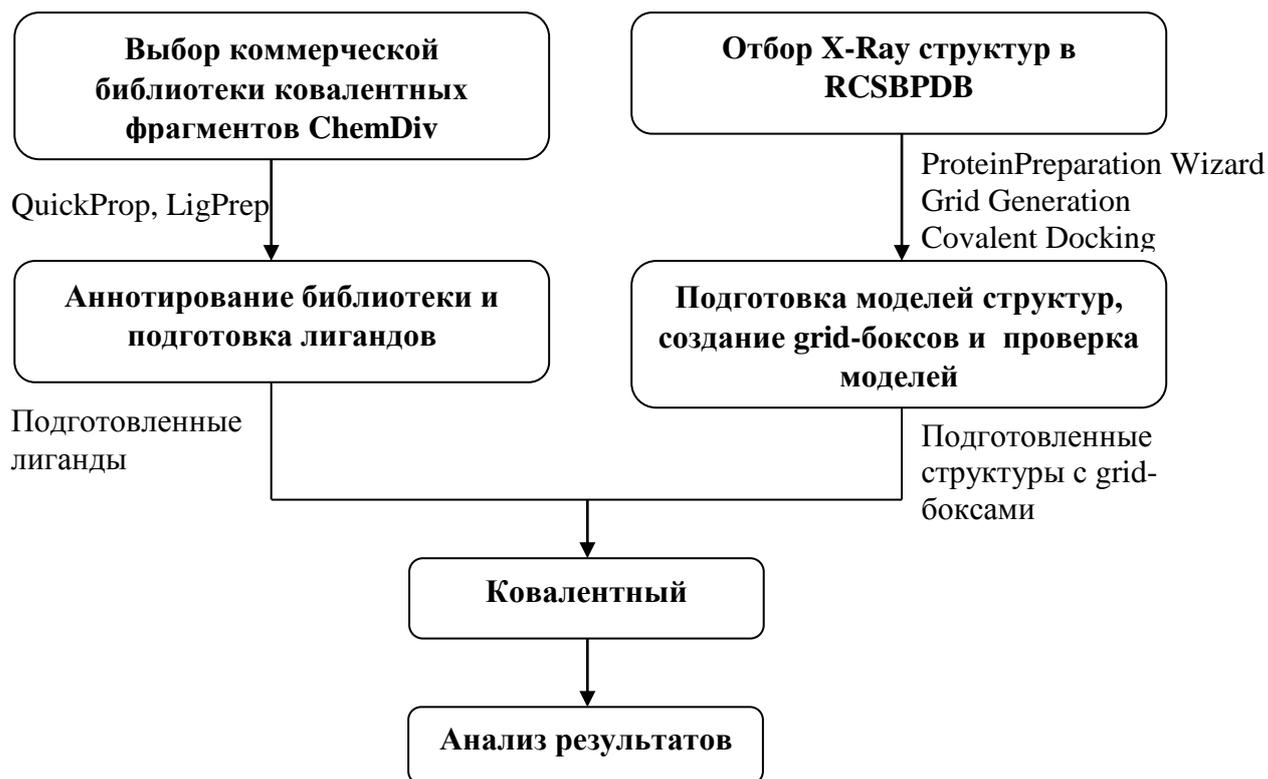


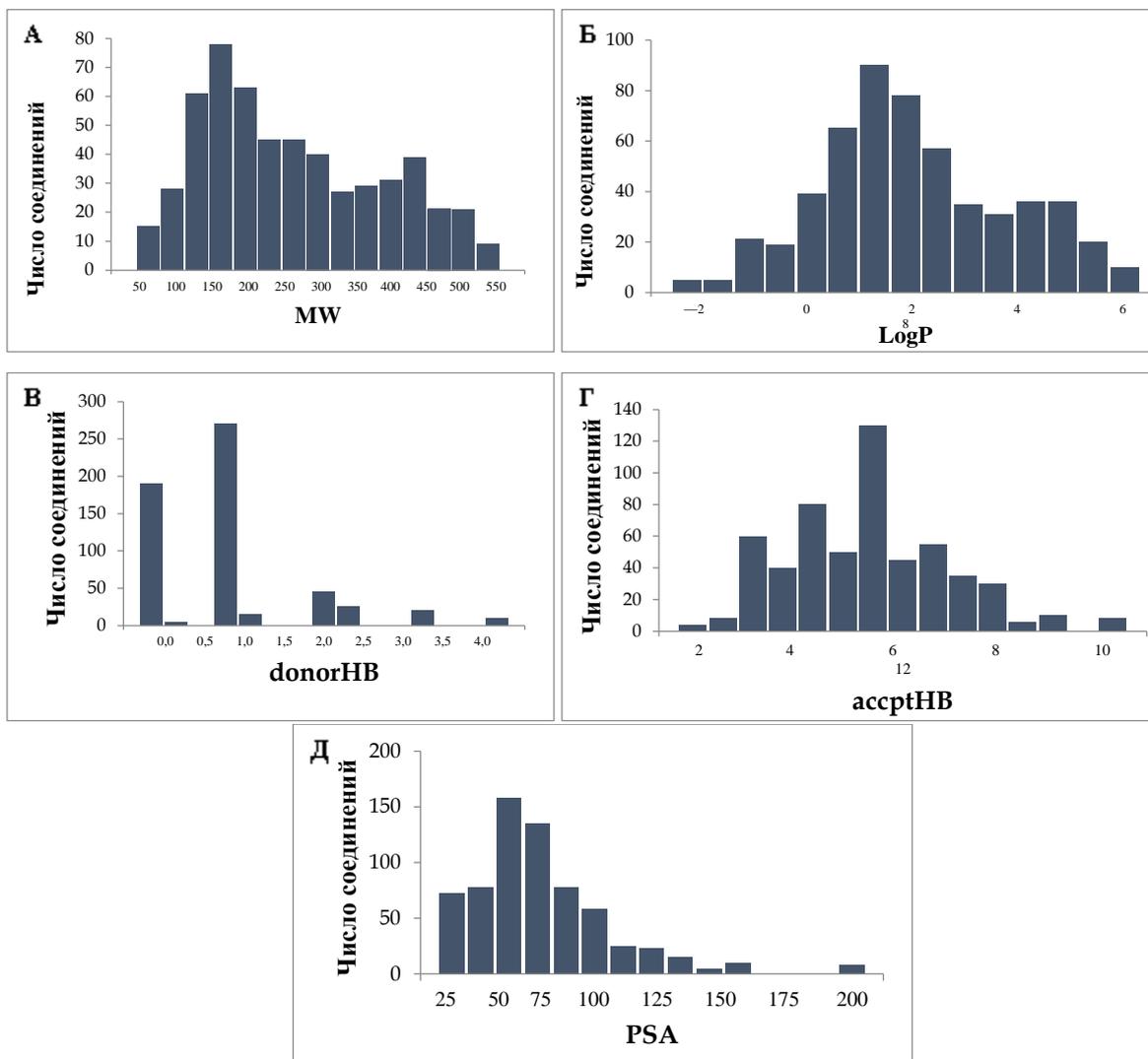
Рисунок 16 – общий дизайн исследования

2.1 Аннотирование и подготовка библиотеки

Был загружен набор соединений Covalent Fragments Smart Library [67], которые были подвергнуты фильтрации с целью оставить только те лиганды, которые имеют хлорацетильный фрагмент в качестве ковалентно-реактивной группы. Трехмерные структуры и состояния протонирования затем определялись с помощью модуля LigPrep [59], после чего с использованием модуля QuickProp [60] были рассчитаны различные физико-химические свойства представленных 542 соединений.

Диаграммы распространения свойств, определенных правилами Липински [61] представлены на рисунке 17, где ось ординат на диаграммах показывает число соединений библиотеки, а ось абсцисс – диапазон значений свойства. Как можно увидеть, объединенная библиотека целиком охватывает весь диапазон свойств (рисунок 17).

Также все соединения библиотеки были разделены на кластеры согласно их структурам: в одном кластере находятся лиганды, имеющие схожий скаффолд (рисунки 18 и 19).



А – Молекулярная масса, Да; Б – липофильность; В и Г – число доноров и акцепторов водородной связи, шт; \square^2 ; Д – площадь полярной поверхности, \square^2 .

Рисунок 17 – Распределение свойств среди соединений библиотеки.

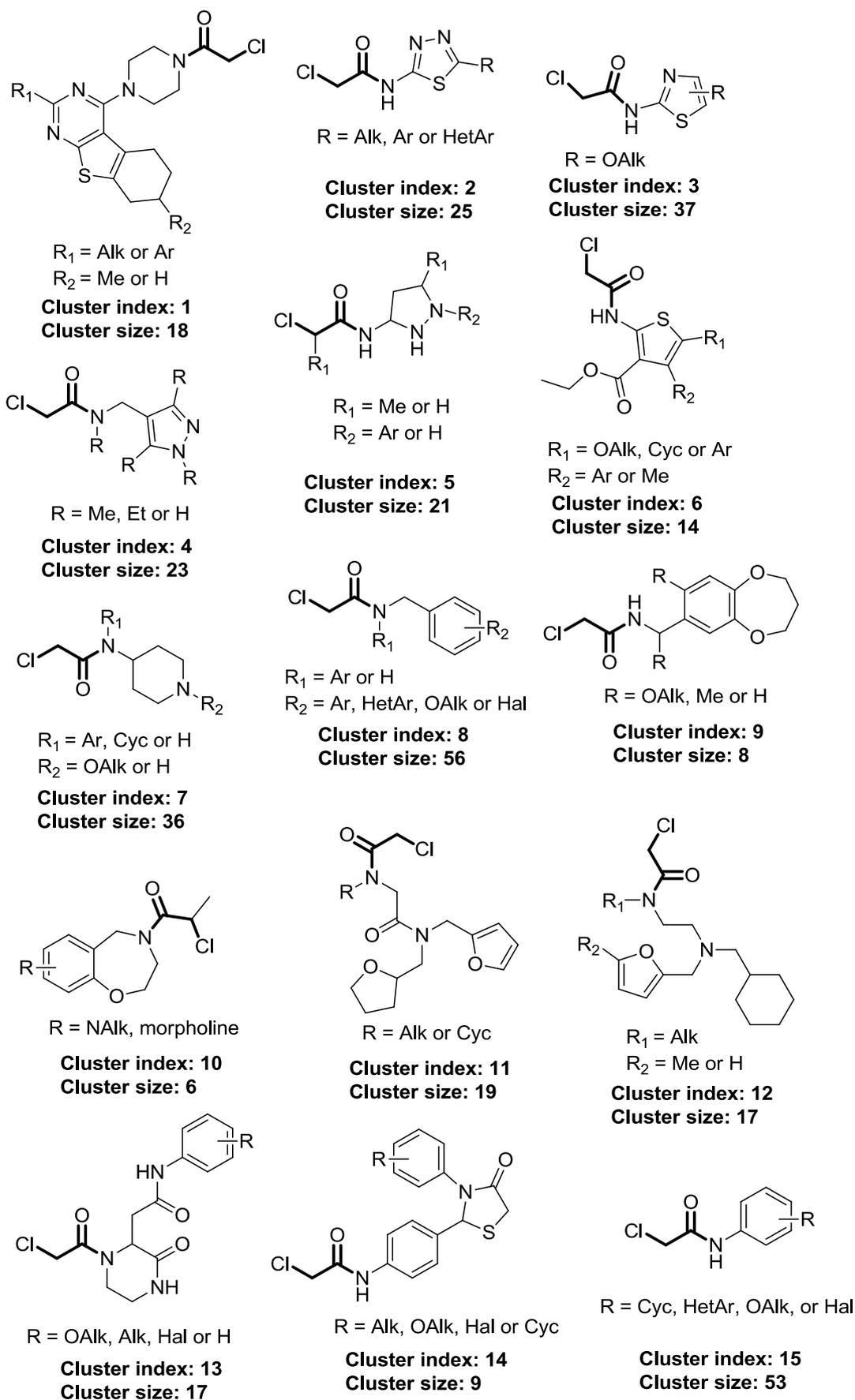
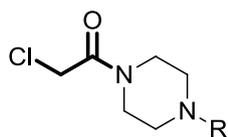
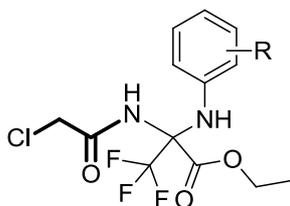


Рисунок 18 – Представленные в библиотеке структурные кластеры



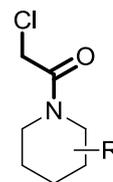
R = Ar or HetAr

Cluster index: 16
Cluster size: 34



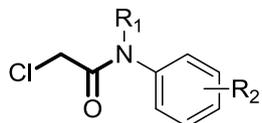
R = Alk, Hal or H

Cluster index: 17
Cluster size: 7



R = Ar, Alk, Cyc

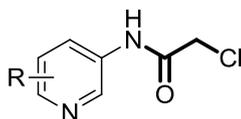
Cluster index: 18
Cluster size: 19



R₁ = Cyc, Me or H

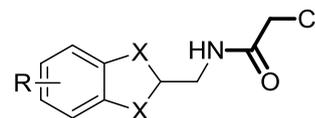
R₂ = NAlk, Ar, Hal or Cyc

Cluster index: 19
Cluster size: 23



R = OAlk, Ar, Cyc

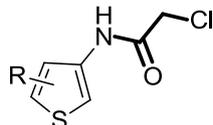
Cluster index: 20
Cluster size: 16



R = Me or H

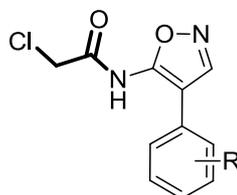
X = NH or CH₂

Cluster index: 21
Cluster size: 25



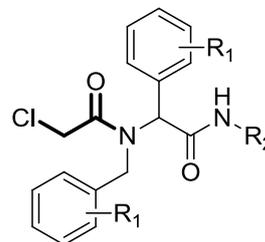
R = HetAr, OAlk

Cluster index: 22
Cluster size: 13



R = Hal, Me or H

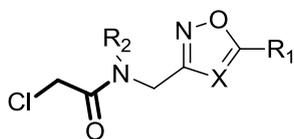
Cluster index: 23
Cluster size: 9



R₁ = OAlk, Hal or H

R₂ = Cyc

Cluster index: 24
Cluster size: 6

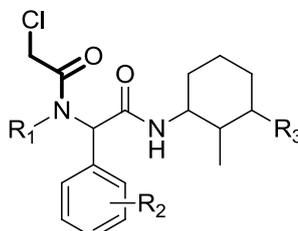


R₁ = Ar or HetAr

R₂ = Me or H

X = N or CH

Cluster index: 25
Cluster size: 15

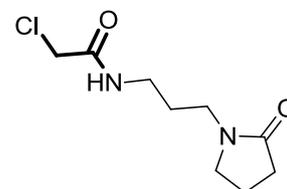


R₁ = OAlk, Cyc, Alk or Ar

R₂ = OAlk, Me or Hal

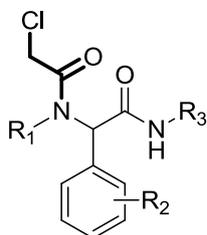
R₃ = Me or H

Cluster index: 26
Cluster size: 4



Cluster index: 27

Cluster size: 1

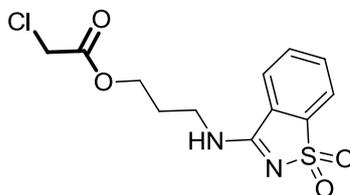


R₁ = OAlk or Cyc

R₂ = OAlk, Me or Hal

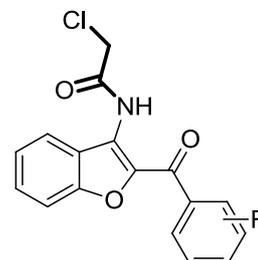
R₃ = Cyc

Cluster index: 28
Cluster size: 5



Cluster index: 29

Cluster size: 1



R = Hal, Me or H

Cluster index: 30
Cluster size: 3

Рисунок 19 – Представленные в библиотеке структурные кластеры

2.2 Подготовка структур белков и нативный докинг

В Protein Data Bank [57] находятся 3 кристаллических комплекса RNF4 (таблица 1). Все структуры имели кристаллографическое разрешение в диапазоне 2.00-3.40 Å.

Таблица 1 – Количество комплексов RNF4 в RCSB PDB

Комплекс	PDB ID	X-Ray resolution, Å
RNF4	4PPE	2.00
	5AIT	3.40
	5AIU	2.21

Выбранный комплекс 4PPE [54] был загружен из Protein Data Bank, их трехмерные структуры были загружены и подготовлены с использованием панели сборки Maestro в пакете Schrödinger.

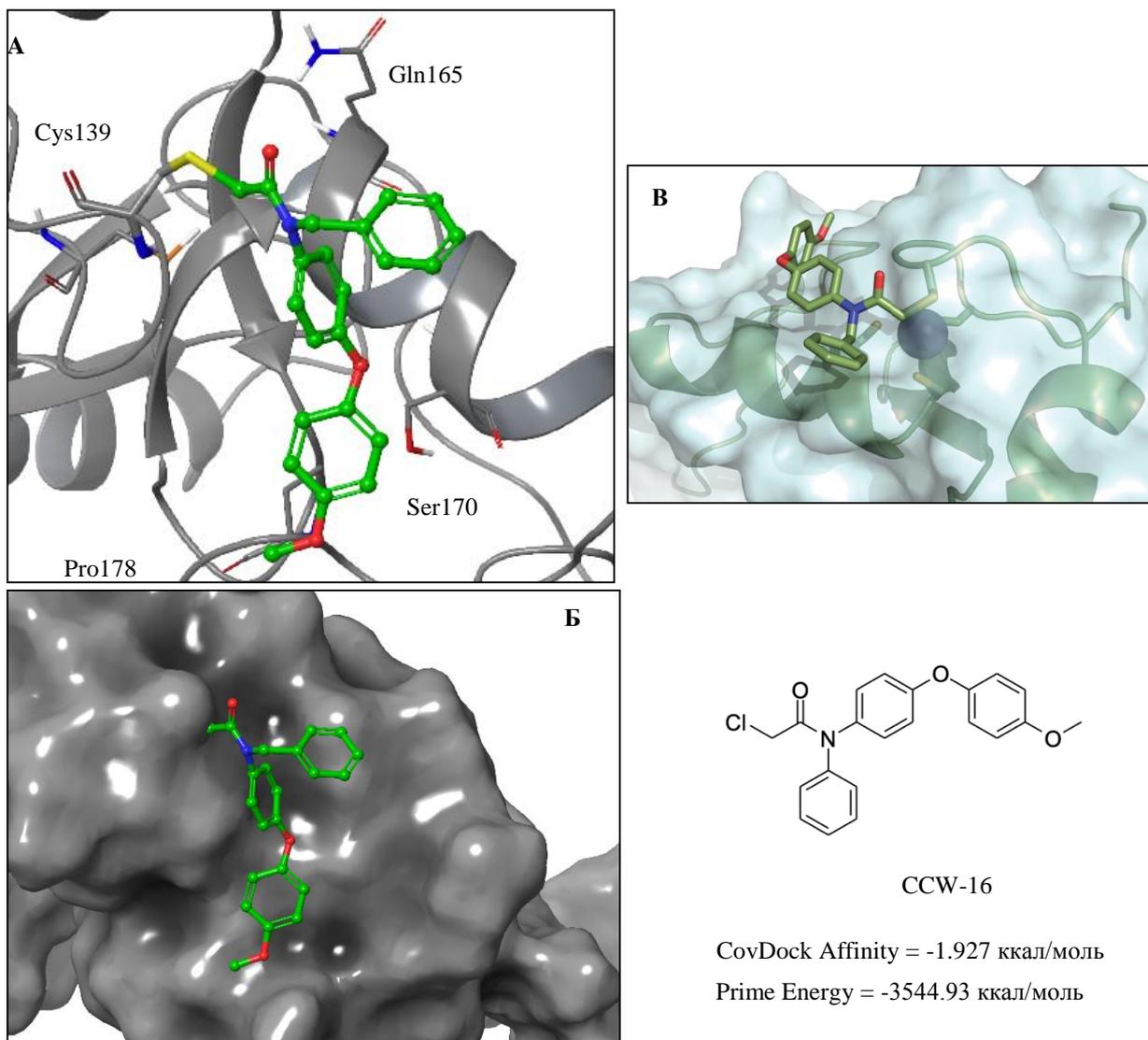
С помощью модуля подготовки ProteinPreparationWizard [62] была проведена оптимизация: добавлены атомы водорода, исправлена валентность некоторых атомов, достроены отсутствующие боковые цепи. Лигаза RNF4 является димером с реактивными остатками Cys 136 и Cys 139. Атом цинка вблизи этих остатков был удален.

В качестве референсного лиганда нами была выбрана структура CCW-16, опубликованная в работе [53] и являющаяся аффинным ковалентным лигандом RNF4.

Поскольку для дальнейшего ранжирования результатов виртуального скрининга выгруженных нами библиотек нам были необходимы данные об энергии и геометрии позы связывания данной структуры, а на данный момент в RCSB PDB не депонировано ни одного кристаллического комплекса с этим лигандом (5), мы провели ковалентный нативный докинг. Стыковка осуществлялась с помощью модуля CovalentDocking [63], реактивным остатком выбран Cys в позиции 139, область докинга определялась как куб с ребром в 20 Å с центром, совпадающим с позицией Cys.

Тип реакции был установлен как «нуклеофильное замещение» с атомами F, Cl, Br и I в качестве возможных реакционных атомных паттернов. Стыковка осуществлялась на самом точном протоколе «Pose Prediction», доступном в режиме ковалентной стыковки с расчетом энергии связывания модулем Prime MM GBSA [58].

В качестве средств оценки нами использовались такие параметры как CovDock affinity и Prime MM GBSA. Первый параметр является кажущейся оценкой аффинности связывания и представляет собой усредненное значение скоринга Glide, проводимого для пред- и послереакционной позы связывания лиганда. В публикации [64], посвященной данному алгоритму авторы отмечают, что именно этот параметр является ключевым средством оценки эффективности связывания. На рисунке ниже представлена лучшая поза связывания для лиганда CCW-16, обладающая значениями CovDock affinity = -1.927 ккал/моль и Prime Energy = -3544.93 ккал/моль (рисунок 20).



А – связывание CCW-16 с Cys139 на поверхности RNF4; Б – Укладка нативного лиганда на поверхности RNF4 (с демонстрацией поверхности белка); В – Связывание CCW-16 на поверхности RNF4 (рисунок из статьи [53])

Рисунок 20 – Режим связывания нативного лиганда CCW-16

Нам удалось воспроизвести режим связывания лиганда, приведенный в статье (сравнение на рисунке 20): укладка дифенилэфирного фрагмента осуществляется в неглубокую ложбинку на поверхности RNF4, образованную остатками Pro178 и Ser170, где образует многочисленные локальные гидрофобные взаимодействия (рисунок 21); перпендикулярно ему бензильный фрагмент простирается близ Gln 165.

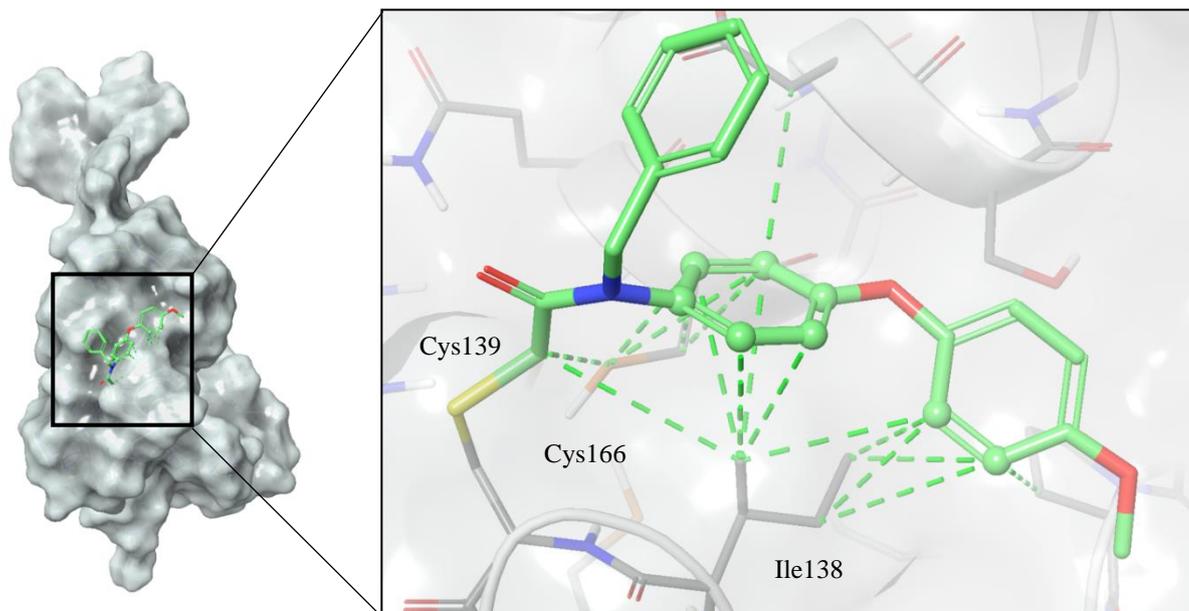


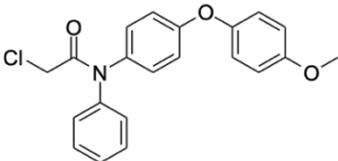
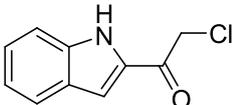
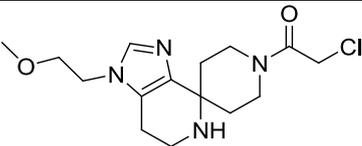
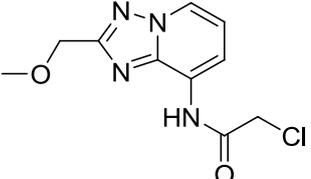
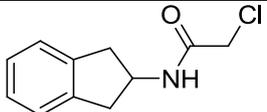
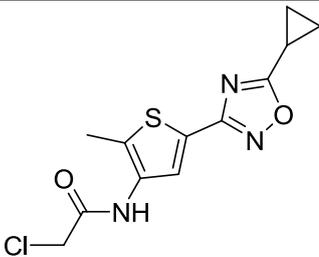
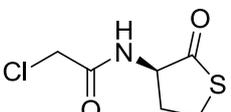
Рисунок 21 – Гидрофобные контакты лиганда; зеленый пунктир – гидрофобные взаимодействия

Нераскрытым остается вклад в связывание бензильного фрагмента, который не простирается вдоль поверхности RNF4, а ориентируется четко в область растворителя.

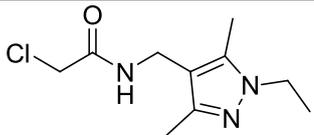
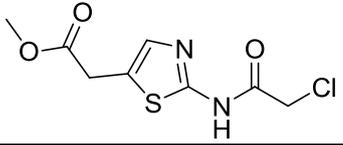
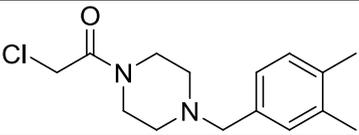
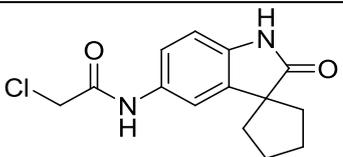
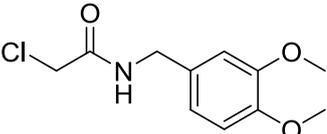
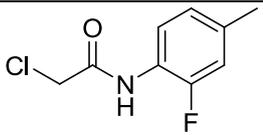
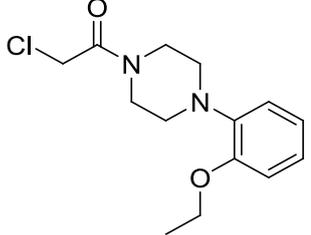
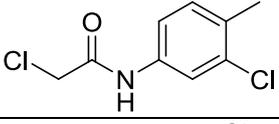
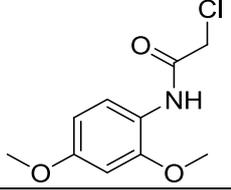
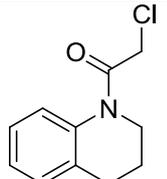
2.3 Ковалентный докинг виртуальной библиотеки

Подготовленные и отфильтрованные структуры лигандов библиотеки были подвергнуты ковалентному докингу по тому же протоколу, что и для нативного CCW-16. Результаты ковалентной стыковки 35 соединений, доступных для покупки, представлены в таблице ниже (таблица 2).

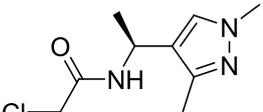
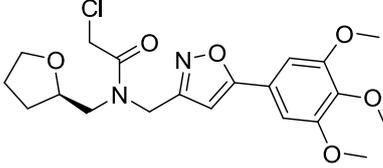
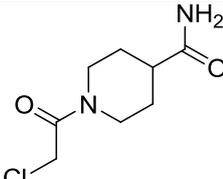
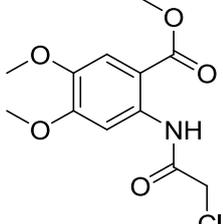
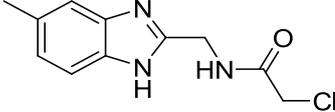
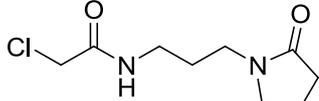
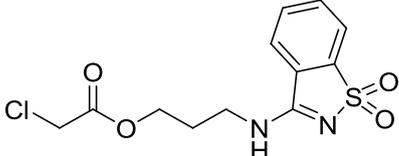
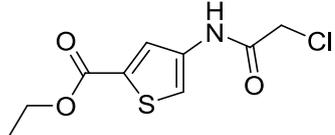
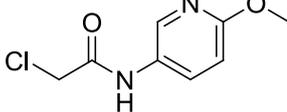
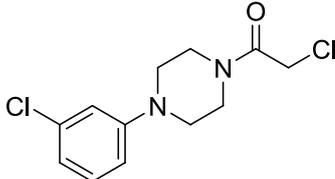
Таблица 2 – Результаты ковалентного докинга виртуальной библиотеки

Структура	Код соединения	Индекс кластера	MM GBSA ΔG_{Bind} , ккал/моль	Prime Energy, ккал/моль	CovDock affinity, ккал/моль
1	2	3	4	5	6
	CCW-16	-	-25.53	-3544.93	-1.927
	6292-0250	21	-19.53	-4420.09	-2.528
	ZE09-1535	18	-16.41	-4446.48	-2.499
	ZE09-1192	20	-14.55	-4385.14	-2.411
	8018-4716	21	-16.81	-4483.60	-2.355
	ZE09-1329	22	-26.46	-4455.49	-2.319
	8019-7122	22	-17.24	-4462.04	-2.310

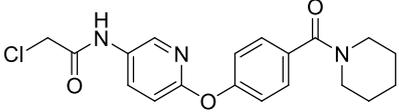
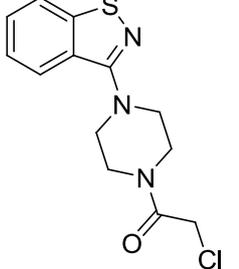
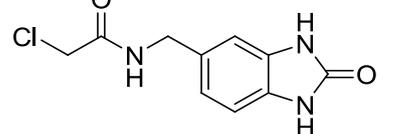
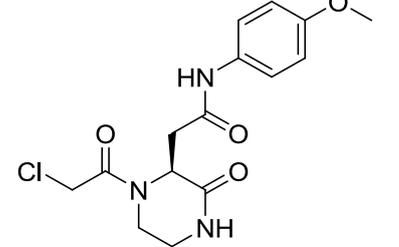
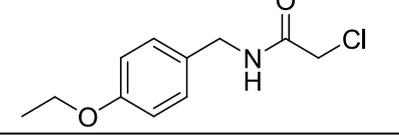
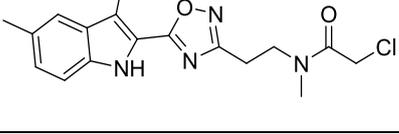
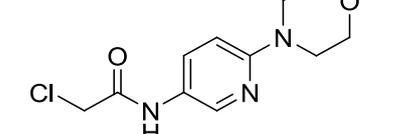
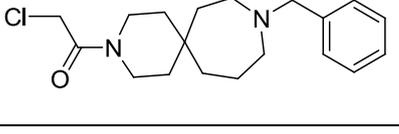
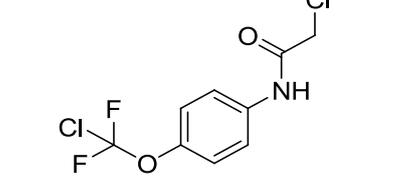
Продолжение таблицы 2

1	2	3	4	5	6
	T001-1392	14	-13.01	-4478.35	-2.279
	8018-6498	3	-17.49	-4460.77	-2.267
	T001-1615	16	-20.67	-4452.56	-2.252
	ZE09-1274	15	-19.70	-4446.30	-2.216
	T001-1599	8	-17.56	-4462.39	-2.196
	R152-1570	15	-24.09	-4455.21	-2.191
	T001-1017	16	-20.08	-4390.22	-2.186
	R152-1954	18	-24.14	-4471.06	-2.151
	R052-1725	18	-23.10	-4460.27	-2.121
	R152-0359	18	-12.48	-4411.70	-2.121

Продолжение таблицы 2

1	2	3	4	5	6
	T001-1400	14	-13.00	-4490.45	-2.121
	V002-0745	25	-27.84	-4425.49	-2.103
	Y020-5255	18	-9.89	-4476.57	-2.096
	Y020-5248	18	-19.53	-4459.17	-2.079
	T001-1322	21	-27.09	-4480.68	-2.075
	T001-0382	26	-19.03	-4496.68	-2.072
	4816-0003	29	-24.84	-4468.95	-2.063
	ZE09-1286	22	-23.33	-4483.33	-2.058
	T001-0952	20	-22.14	-4470.98	-2.053
	T001-0450	16	-11.37	-4420.13	-2.038

Продолжение таблицы 2

1	2	3	4	5	6
	ZE09-0733	20	-36.32	-4448.75	-2.037
	ZE09-1379	16	-16.93	-4393.04	-2.026
	8018-7135	8	-17.94	-4517.10	-2.016
	Y020-9951	16	-25.91	-4481.56	-2.014
	T001-1882	8	-18.92	-4473.94	-2.011
	ZE09-1290	25	-11.65	-4418.27	-2.003
	ZE09-1182	20	-24.02	-4443.45	-1.996
	ZE09-1936	18	-20.32	-4429.13	-1.993
	8011-6560	15	-24.82	-4471.24	-1.816

Хотелось бы отметить, что все соединения демонстрируют примерно одинаковые значения как по CovDock affinity (диапазон 1.816-2.528 ккал/моль), так и по Prime Energy (4517.10-4385.14 ккал/моль), при этом показывая скоринг лучше, чем у нативного лиганда.

В первую очередь мы пытались проанализировать сходства и различия режимов связывания соединений библиотеки относительно CCW-16. Ниже представлены структуры, которые демонстрируют отличный от нативного лиганда формат укладки на поверхности RNF4: ZE09-1192, 8018-6498, R152-1570, R152-0359 (рисунок 22).

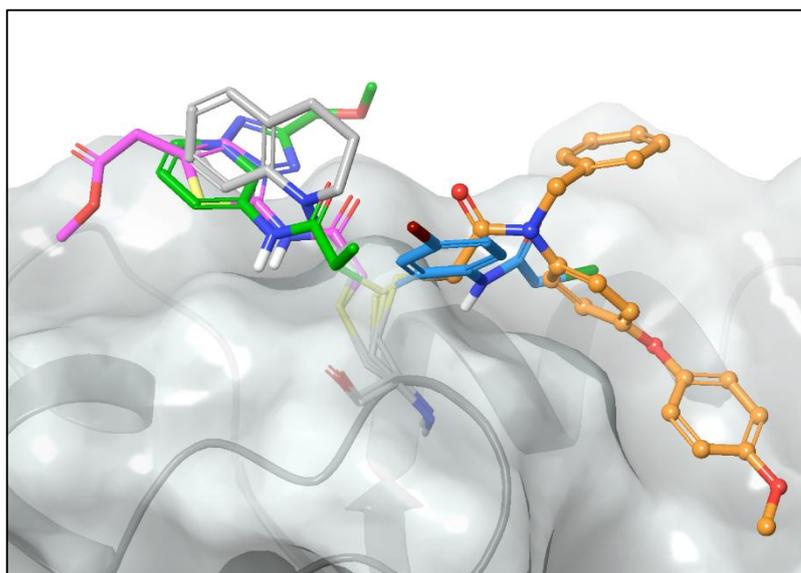
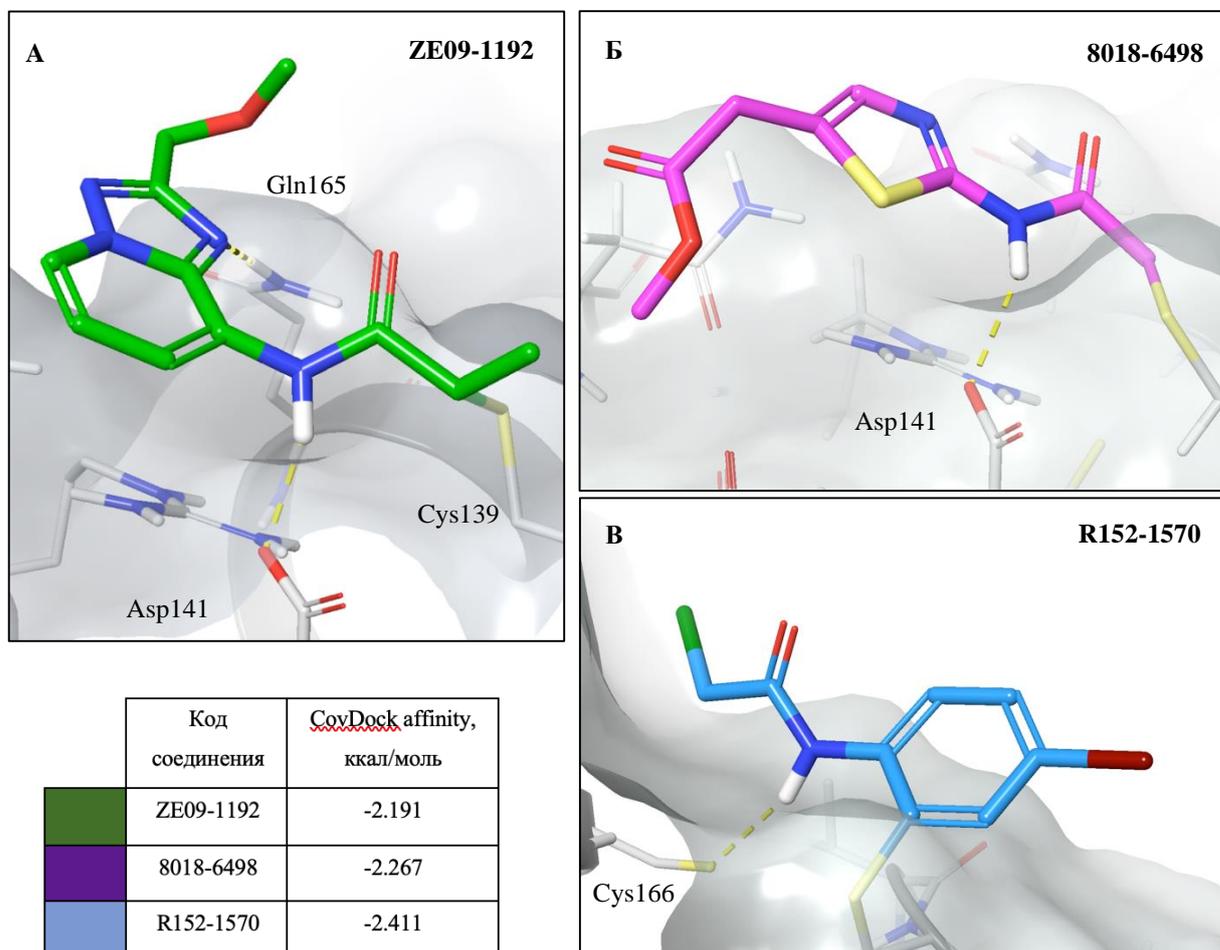


Рисунок 22 – Связывание соединений ZE09-1192, 8018-6498, R152-1570, R152-0359 и CCW-16 на поверхности RNF4

ZE09-1192, имеющий бензотриазоловый скаффолд, ориентируется в противоположную от CCW-16 полость, образованную остатками Arg151, Asn149, Asp141 (как и тиазол 8018-6498, тетрагидрохиолин R152-0359), где образует связь с Asp141 за счет амидной группы, а атом азота триазольного фрагмента служит акцептором связи для Gln165. Схожее взаимодействие с Asp показывает и 2-аминотиазол 8018-6498. Соединение R152-1570 выходит в полость нативного лиганда, но занимает ортогональную CCW-16 конформацию, дополнительно демонстрируя связь между своим амидным

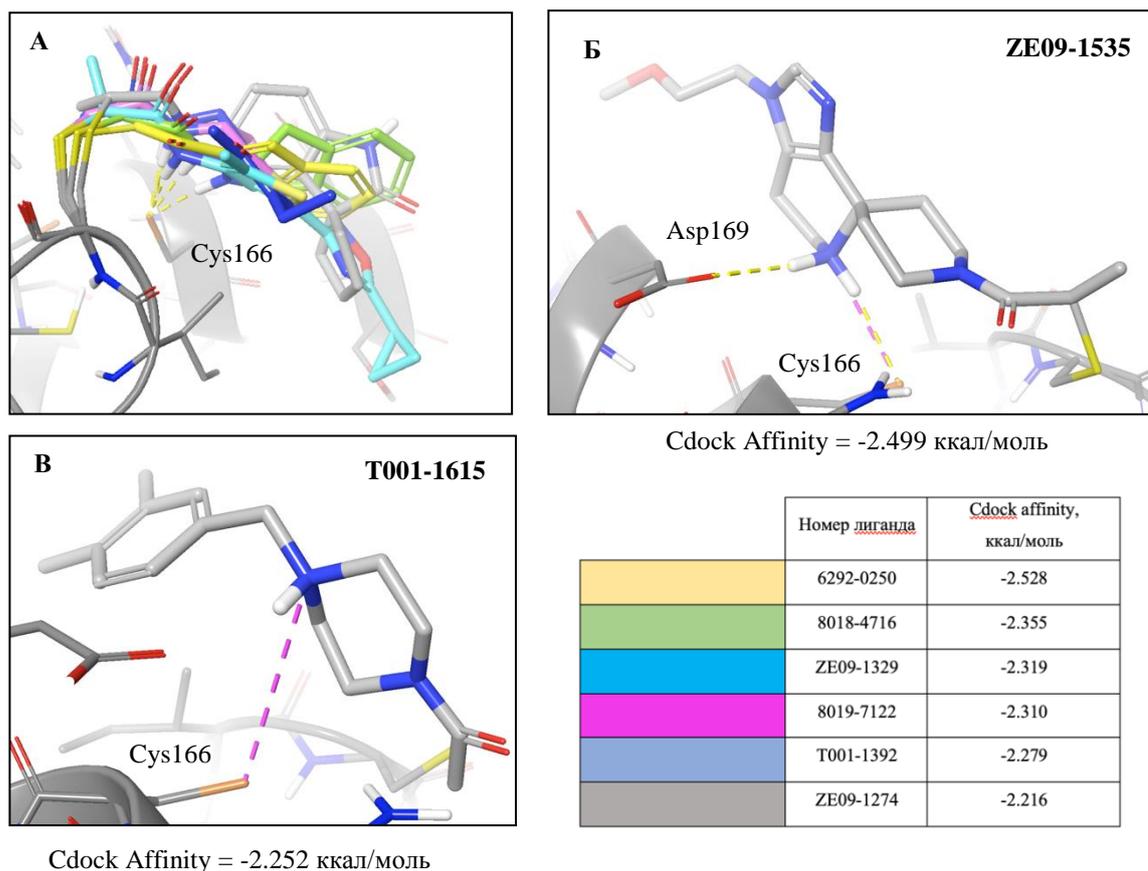
азотом и тиолом Cys166. Интересно, что все эти лиганды обладают хорошими значениями скоринга (рисунок 23).



А, Б, В – Режимы связывания **ZE09-1192** (зеленый), **8018-6498** (фуксия) и **R152-1570** (синий), соответственно

Рисунок 23 – Соединения, демонстрирующие ориентацию на поверхности RNF4, отличную от CCW-16

Подавляющее большинство соединений, показывающих схожую с CCW-16 ориентацию, образуют водородную связь с остатком Cys в позиции 166 в основном за счет азота амидной группы (рисунок 24).



А – Наложение структур **6292-0250** (желтый), **8018-4716** (зеленый), **ZE09-1329** (голубой), **8019-7122** (фуксия), **T001-1932** (синий), **ZE09-1274** (серый) на поверхности RNF4 в связывающих конформациях; Б – Связывание **ZE09-1535**; В – Связывание **T001-1615**

Рисунок 24 – Некоторые структуры, демонстрирующие схожий режим связывания с Cys166; желтый пунктир – водородные связи; фиолетовый пунктир – ионные взаимодействия

Причем данный паттерн характерен для структур как с низким, так и с высоким значением CovDock. На рисунке ниже приведено несколько примеров таких структур. Spirocyclic compound on the basis of tetrahydroimidazopyridine **ZE09-1535**, simultaneously with the formation of hydrogen bonds with Asp169 and Cys166 residues, demonstrates also ionic interaction with the thiol of the latter. Binding with the same residue due to the salt bridge is realized by **T001-1615** thanks to the protonated nitrogen of the piperazine scaffold (Figure 24).

Представленные на рисунке 24 соединения в основном ранжируются в верхних строчках рейтинга CovDock. Однако мы не можем не отметить, что в нативной структуре RNF4 Cys166 участвует в координации атома цинка, который был удален перед проведением докинга (рисунок 25).

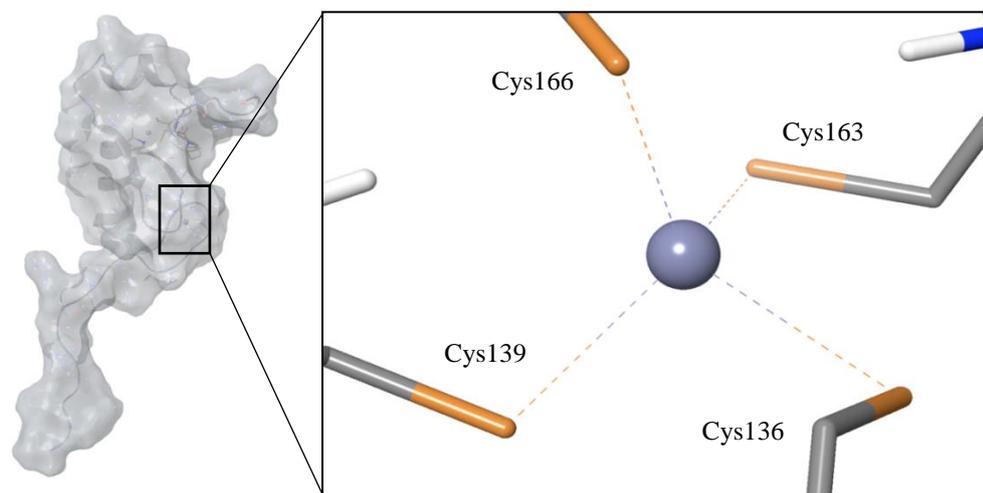
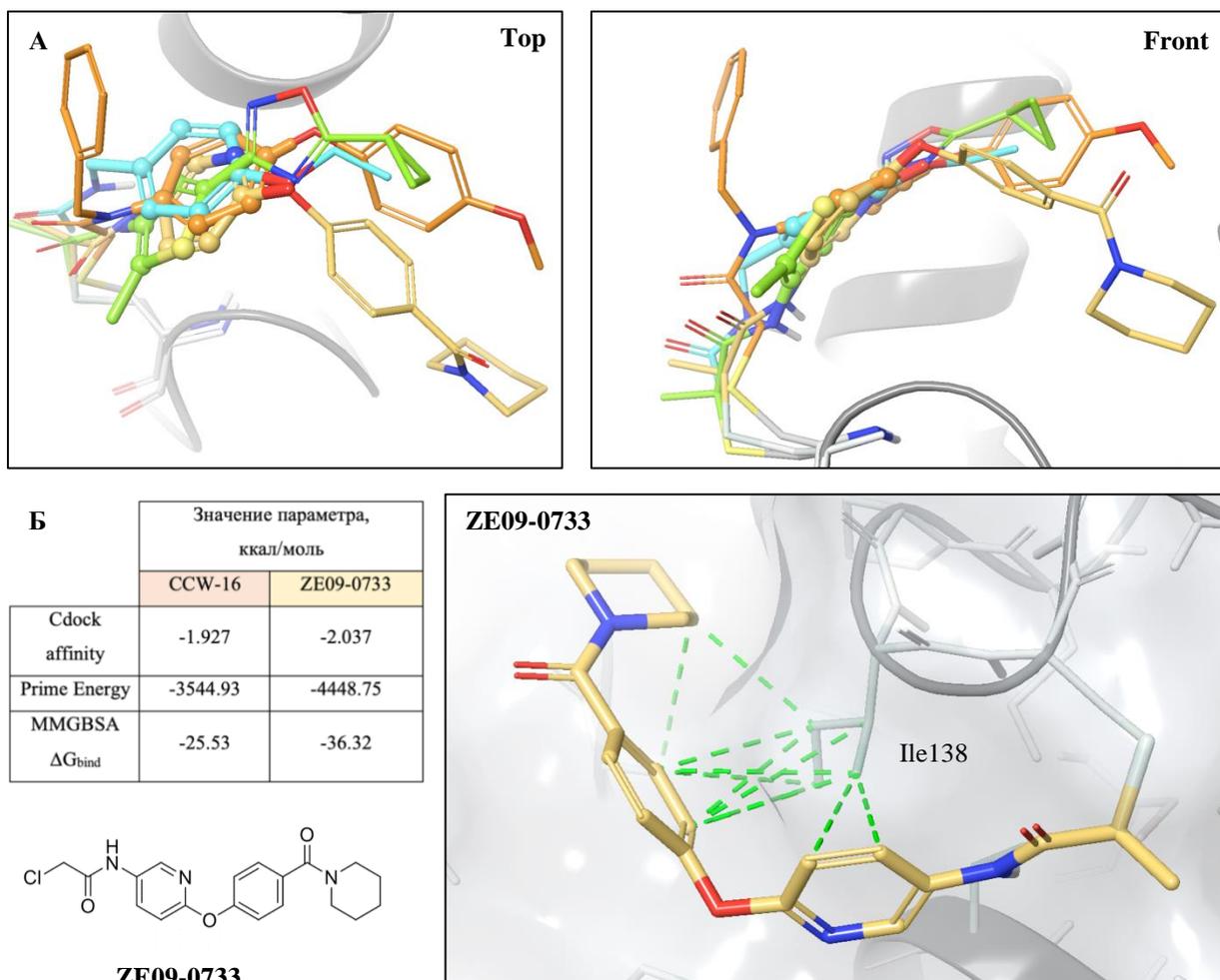


Рисунок 25 – Координация атома цинка в лигазе RNF4

Весомый вклад в связывание, безусловно вносят локальные гидрофобные взаимодействия. Нативный лиганд реализует большую их часть за счет расположенного вблизи остатка Pe138 дифенилэфирного фрагмента, как описывалось на рисунке 21. Среди структур библиотеки можно выделить соединения, демонстрирующие высокую степень гомологии укладки гидрофобных групп в данной области, например: пиримидин ZE09-0733, тиофен ZE09-1329, фенил T001-1882 (рисунок 26). Причем ZE09-0733, имея подобный CCW-16 мотив двух ароматических систем, соединенных эфирной связью как нельзя лучше воспроизводит позу нативного лиганда, демонстрируя даже большую степень покрытия гидрофобной бороздки на поверхности RNF4, чем можно объяснить умеренно высокие значения Cdock и Prime energy, а также самое низкое среди соединений библиотеки значение свободной энергии связывания MM GBSA ΔG_{bind} (рисунок 26).



А – наложение структур **ZE09-0733** (желтый), **CWW-16** (оранжевый), **ZE09-1329** (зеленый) и **T001-1882** (голубой) в связывающих конформациях на поверхности RNF4 (вид сверху и сбоку); Б – укладка **ZE09-0733**, его структура и значения параметров скоринга в сравнении с нативным лигандом

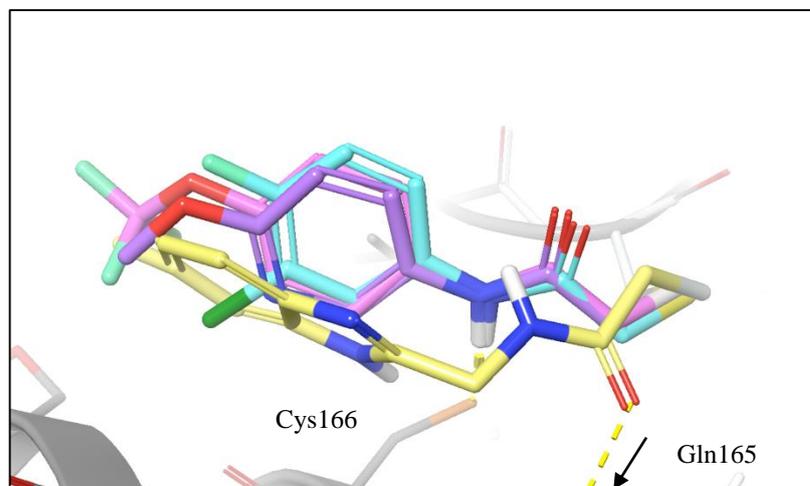
Рисунок 26 – Структуры, демонстрирующие локальные гидрофобные взаимодействия; зеленый пунктир – гидрофобные контакты

С учетом того, что глобальной целью данной работы являлся поиск соединений, которые могут стать потенциальными новыми лигандами RNF4 для конструирования PROTAC, рекрутирующих данную лигазу, нам было важно следовать следующим критериям при отборе хитов.

Во-первых, соединения должны иметь схожий с CCW-16 режим ориентации на поверхности белка. Во-вторых, иметь удобный мотив для прикрепления линкера; такой мотив может изначально быть в структуре

лигандов (пиперазиновые фрагменты, метокси- или сложноэфирные группы) или же легко добавлен – главное, чтобы векторы выхода линкеров примерно совпадали с вектором нативного. В-третьих, среди таких соединений особое внимание следовало уделить тем, кто демонстрирует лучший по отношению к CCW-16 скоринг.

В составе библиотеки было довольно много небольших фрагментных лигандов – бензимидазол T001-1322, замещенные арилы 8011-6560, R152-1954 и некоторые другие. Несмотря на то, что режимы их ориентации хорошо совпадают с режимом CCW-16, а наличие гидрофобных групп обеспечивает хорошие контакты в соответствующей области – данные структуры определенно будут требовать дальнейшей доработки и модификации, в ходе которых их мотивы связывания могут измениться в худшую сторону. Некоторые из таких лигандов представлены на рисунке ниже и были отброшены нами при отборе хитов (рисунок 27).



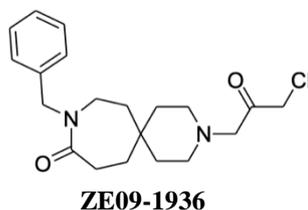
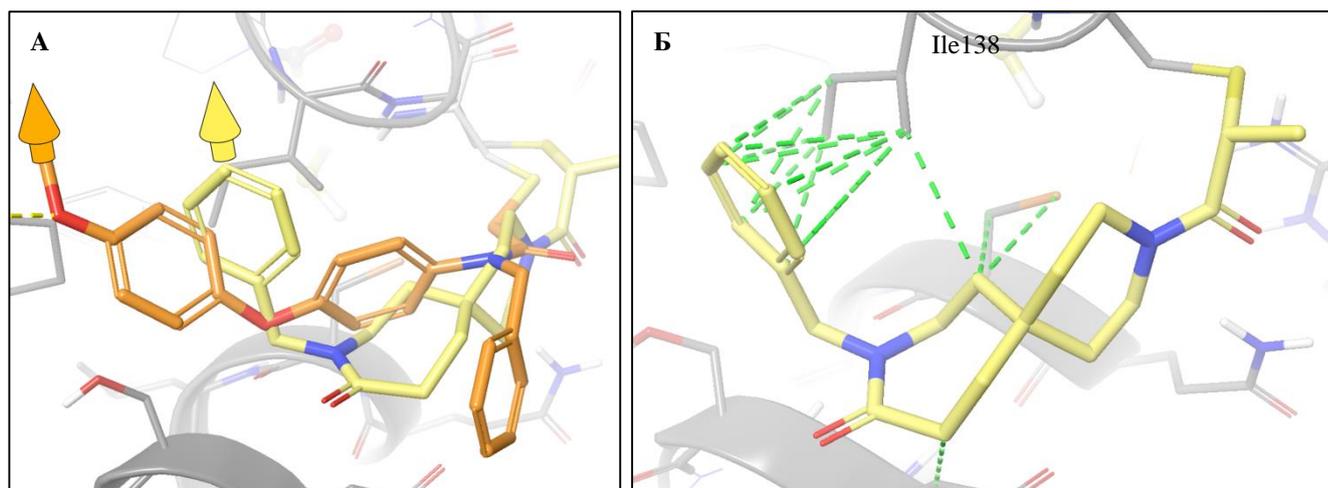
T001-1322 (желтый), **8011-6560** (розовый), **R152-1954** (голубой), **T001-0952** (фиолетовый)

Рисунок 27 – Связывание фрагментных лигандов

Первое отобранное нами соединение – ZE09-1936; 3,9-диазаспиро[5.6]додекан-10-оновый фрагмент простирается вдоль той же ложбинки (Pro178 и Ser170, Ile138), а бензильная группа, помимо того что

обеспечивает удобный предполагаемый вектор выхода линкера, образует множество гидрофобных контактов с Ile138.

ZE09-1936, однако, обладает практически самым высоким значением скоринга (CovDock affinity = -1.993 ккал/моль; ММ GBSA ΔG_{bind} = -20.32 ккал/моль). Типичное для многих лигандов взаимодействие с Cys166 не реализуется (рисунок 28).



А – Наложение нативного лиганда (оранжевый) и **ZE09-1936** (желтый); Б – образование гидрофобных взаимодействий в области Ile138

Рисунок 28 – Связывание ZE09-1936 на поверхности RNF4

Ранее уже обсуждалось связывание ранжируемого выше всех по значению ММ GBSA ΔG_{bind} лиганда ZE09-0733. Невыгодная с точки зрения конъюгации линкера геометрия пиперидинамида, на наш взгляд, может стать точкой дальнейшей оптимизации этой структуры (рисунок 29).

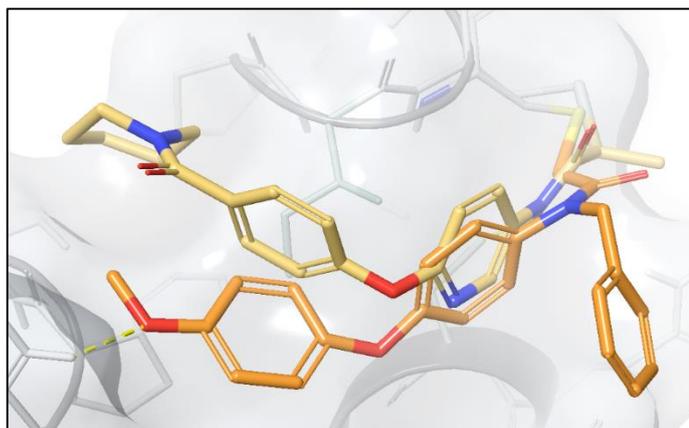
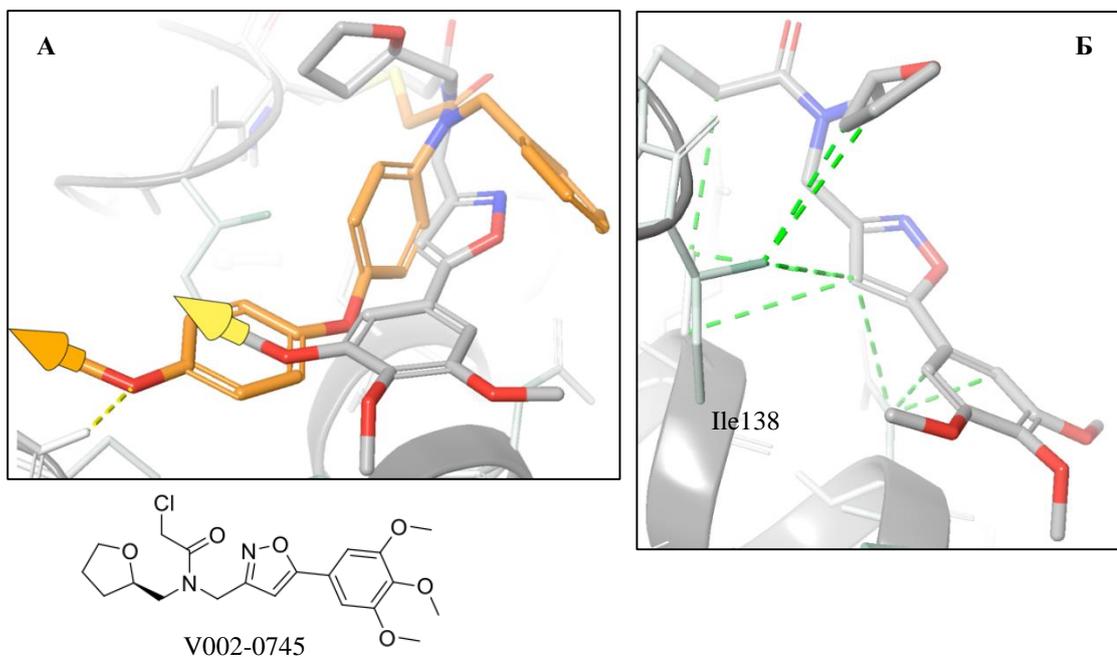


Рисунок 29 – Наложение структур CCW-16 (оранжевый) и ZE09-0733 (желтый)

Хорошо себя показали лиганды, содержащие метоксизамещенные ароматические кольца, например – V002-0745. Изоксазольный фрагмент укладывается вдоль Pe138, образования связей с Cys166 не происходит, а направление метила одной из метоксигрупп совпадает с CCW-16 (рисунок 30). Данный лиганд превосходит нативный по всем показателям скоринга.

Следующие две структуры также были отобраны нами как потенциальные рекрутеры RNF4 – ZE09-1329 и Y020-9951. Тиофен-оксадиазол-содержащий ZE09-1329 ориентирует свой циклопропил в ту же полость, что и нативный лиганд. Вероятно, такой вектор выхода будет удачным.

Пиперизонный скаффолд Y020-9951 простирается далеко в область растворителя, однако связь центрального амида с Cys166 обеспечивает верное направление метоксизамещенного фенила. Данные лиганды также обладают хорошими значениями скоринга (рисунок 31).



А – наложение **V002-0745** (серый) и **CCW-16** (оранжевый); Б – Образование гидрофобных контактов **V002-0745**

Рисунок 30 – Связывание V002-0745 на поверхности RNF4

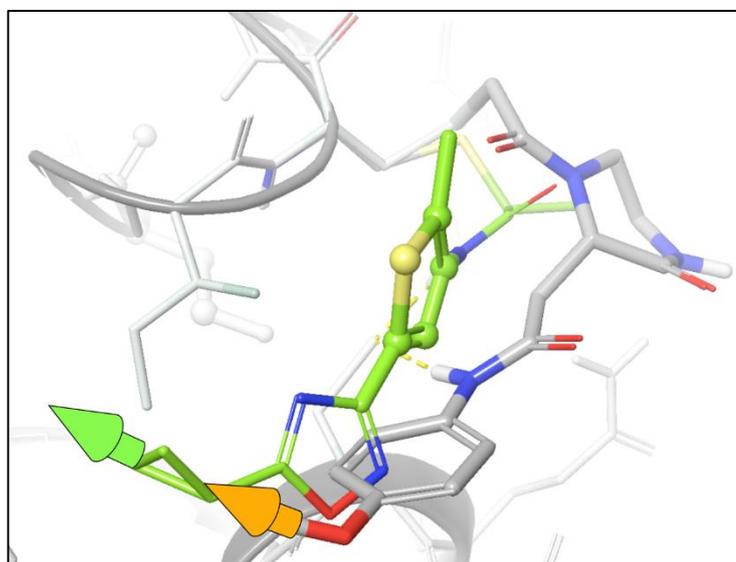
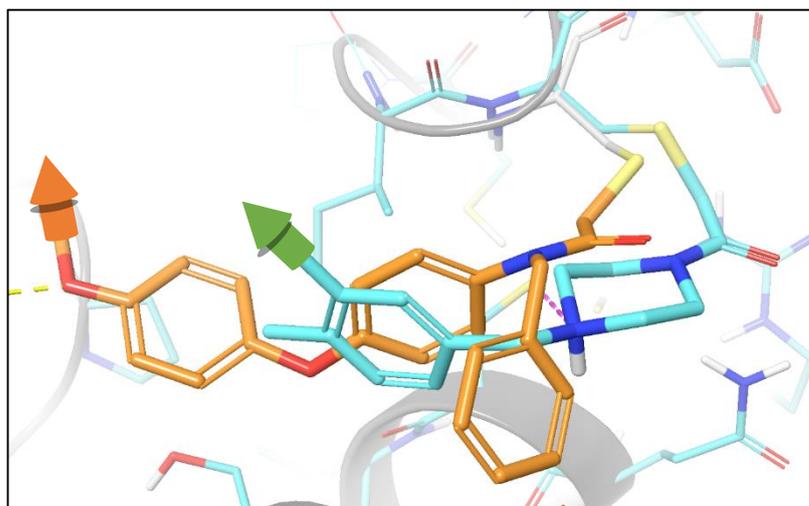


Рисунок 31 – Связывание **ZE09-1329** (зеленый) и **Y020-9951** (серый) на поверхности RNF4

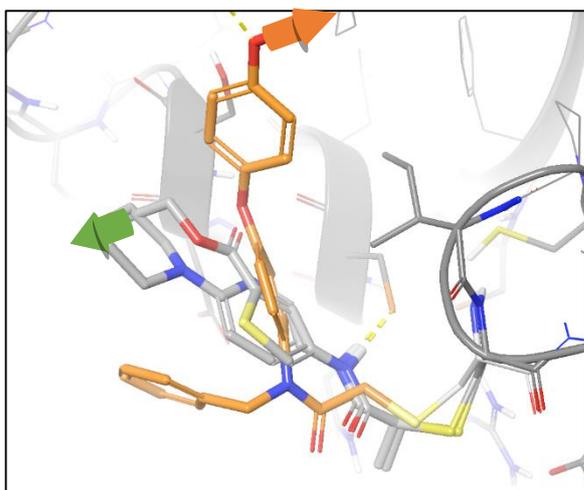
Обсуждаемый ранее пиперазин T001-1615, образующий с Cys166 солевой мостик, показывает практически полное наложение замещенного бензольного кольца с центральным кольцом дифенилэфирного фрагмента нативного лиганда. 3-е положение кольца T001-1615 потенциально может обеспечивать ориентацию в том же направлении, что и метокси-группа CCW-16. Кроме того, ароматический фрагмент пиперазин-содержащего соединения образует множество гидрофобных контактов (рисунок 32).



	MMGBSA ΔG_{bind} , ккал/моль	Cdock affinity, ккал/моль	Prime Energy, ккал/моль
CCW-16	-25.53	-1.927	-3544.93
T001-1615	-20.32	-2.252	-4429.13

Рисунок 32 – Наложение T001-1615 (голубой) и нативного лиганда (оранжевый) в связывающих конформациях

Напоследок отметим еще два соединения – ZE09-1182 и ZE09-1286, ранжируемых скоринговой функцией примерно одинаково. Оба они демонстрируют наложение гидрофобных ядер (1182– пиридин, 1286 - тиофен) подобно CCW-16, однако ориентация концевых групп отличается (1182– морфолин, 1286 - этилацетильная) (рисунок 33).



	MMGBSA ΔG_{bind} , ккал/моль	Cdock affinity, ккал/моль	Prime Energy, ккал/моль
CCW-16	-25.53	-1.927	-3544.93
ZE09-1182	-24.02	-1.996	-4443.45
ZE09-1286	-23.33	-2.058	-4483.33

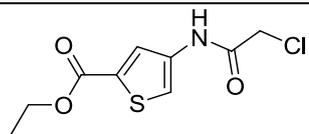
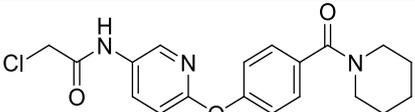
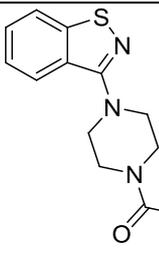
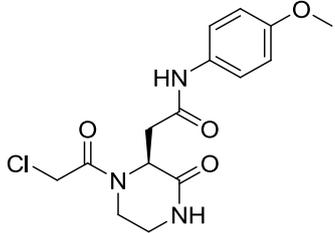
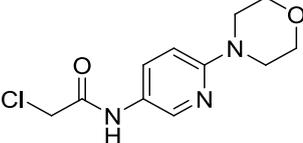
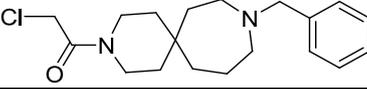
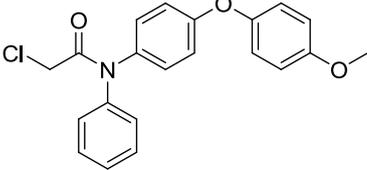
Рисунок 33 – Связывание ZE09 –1182 и 1286 и CCW-16 на поверхности RNF4

Отобранные лиганды представлены в таблице 3 и отсортированы по значению CovDock. Отрадно, что многие из них демонстрируют свободную энергию связывания MM GBSA ниже, чем у CCW-16.

Таблица 3 – Соединения-хиты

Структура	Код соединения	MM GBSA ΔG_{bind} , ккал/моль	Prime Energy, ккал/моль	CovDock affinity, ккал/моль
1	2	3	4	5
	ZE09-1329	-26.46	-4455.49	-2.319
	T001-1615	-20.67	-4452.56	-2.252
	V002-0745	-27.84	-4425.49	-2.103

Продолжение таблицы 3

1	2	3	4	5
	ZE09-1286	-23.33	-4483.33	-2.058
	ZE09-0733	-36.32	-4448.75	-2.037
	ZE09-1379	-16.93	-4393.04	-2.026
	Y020-9951	-25.91	-4481.56	-2.014
	ZE09-1182	-24.02	-4443.45	-1.996
	ZE09-1936	-20.32	-4429.13	-1.993
	CCW-16	-25.53	-3544.93	-1.927

Таким образом, были отобраны соединения, демонстрирующие схожие с нативным лигандом режим связывания (особенно в области гидрофобных взаимодействий с Pe138), а также выходной вектор точки конъюгации линкера.

3 Экспериментальная часть

3.1 Компьютерное моделирование

Все расчеты в рамках выполненной выпускной квалификационной работы проводили с применением программного пакета Schrodinger Suite 2023-1 на рабочей станции Mac Pro 2013, снабженной процессором Intel Xeon E5 (6 ядер, тактовая частота 3.5 ГГц) под управлением ОС Ubuntu 22.04.

3.1.1 Подготовка структур белка

Кристаллическая структура RNF 4 (PDB ID: 4PPE) была загружена из Protein Data Bank [57]. Подготовка осуществлялась с помощью модуля ProteinPreparationWizard [62]. Добавлены отсутствующие атомы водорода и боковые цепи, уточнены порядки связей, состояние протонирования уточнено с использованием алгоритма Epik [65] при диапазоне pH = 7 ± 2 . Атом Zn, участвующий в координации Cys136 и Cys139, молекулы воды и посторонние гетероатомы удалены. Структура RNF4 была подвергнута дополнительной минимизации с использованием силового поля OPLS4 [66].

3.1.2 Подготовка виртуальных библиотек

Для проведения виртуального скрининга из доступных каталогов ChemDiv [56] была загружена библиотека Covalent Fragments Smart Library [67] в формате sdf.

Структуры лигандов были подготовлены с помощью модуля LigPrep; хиральность определялась так, как было указано в 2D представлении, состояние протонирования определено с использованием алгоритма Epik [65] при диапазоне pH = 7 ± 2 . Дополнительно проводился расчет физико-химических свойств с помощью модуля QikProp [60]. Из библиотеки были удалены все лиганды, содержащие мотивы ковалентного связывания, отличные от хлорацетильного.

3.1.3 Ковалентный докинг

После проведения подготовки и фильтрации все структуры библиотеки, а также известный ковалентный рекрутер CWW-16 [53] были подвергнуты процедуре ковалентной стыковки с помощью модуля CovalentDocking [63]. В качестве реактивного аминокислотного остатка был выбран Cys 139; область стыковки была определена как куб с ребром 20 Å, центр которого совпадал с положением Cys. Тип реакции был выбран как «Нуклеофильное замещение» – Cl, Br, F или I установлены в качестве реактивных атомов лиганда.

Стыковка проводилась в режиме точности «Pose Prediction» (самый точный режим) с дополнительной оценкой свободной энергии связывания методом MM GBSA. Остальные настройки модуля стыковки были установлены по умолчанию и не изменялись.

Заключение

Разработка рекрутеров RNF4 является перспективным направлением медицинской химии. Убиквитилирование регулирует важные для развития рака сигнальные пути и во многих случаях нацеливает белки на деградацию. Биологически RNF4 усиливает фенотип опухоли и необходим для пролиферации и выживания агрессивных раковых клеток, а ее экспрессия потенцирует онкогенез в культуре.

По результатам проделанной работы можно сделать следующие выводы:

- проведен ковалентный докинг виртуальной библиотеки соединений (542 соединения), содержащих хлорацетамидные фрагменты в качестве ковалентно-реактивных групп. Для сравнения результатов стыковки осуществлено моделирование связывания известного лиганда RNF4 – нам удалось воспроизвести опубликованный режим связывания на поверхности лигазы;

- в качестве потенциальных рекрутеров отобраны 9 соединений, обеспечивающие схожую с нативным лигандом ориентацию. Они показывают скоринг лучше, чем у нативный лиганд. Значения CovDock affinity (диапазон 1.993-2.319 ккал/моль) и Prime Energy (4393.04-4483.33 ккал/моль), а у нативного CovDock affinity = -1.927 ккал/моль и Prime Energy = -3544.93 ккал/моль. Значение свободной энергии связывания MM GBSA ΔG_{bind} по сравнению с нативным выше у соединений ZE09-1329, V002-0745, ZE09-0733, Y020-9951;

- отобранные соединения потенциально содержат удобные точки и векторы конъюгации линкеров. Полученные данные имеют большое значение для дальнейшей модификации структуры найденных соединений, с целью создания новых PROTAC.

Список используемой литературы

1. Kramer L. T. Expanding the landscape of E3 ligases for targeted protein degradation / L. T. Kramer, X. Zhang // *Curr. Chem. Biol.* – 2022. – Vol. 2. – P. 100-120.
2. Kumalo H. M. Theory and Applications of Covalent Docking in Drug Discovery: Merits and Pitfalls / H. M. Kumalo, S. Bhakat, E. S. Soliman // *Molecules.* – 2015. – Vol. 20, № 2. – P. 1984–2000.
3. Burslem G.M. Proteolysis-targeting chimeras as therapeutics and tools for biological discovery / G.M. Burslem, C.M. Crews // *Cell.* – 2020. – Vol. 181, № 1. – P. 102-114.
4. Zhong Y. Emerging targeted protein degradation tools for innovative drug discovery: From classical PROTACs to the novel and beyond / Y. Zhong, F. Chi, H. Wu et al. // *Eur. J. Med. Chem.* – 2022. – Vol. 231. – P. 114-142.
5. Hu J. Discovery of ERD-308 as a highly potent proteolysis targeting chimera (PROTAC) degrader of estrogen receptor (ER) / J. Hu, B. Hu, M. Wang et al. // *J. Med. Chem.* – 2019. – Vol. 62, № 3. – P. 1420-1442.
6. Sun Y. Degradation of Bruton's tyrosine kinase mutants by PROTACs for potential treatment of ibrutinib-resistant non-Hodgkin lymphomas / Y. Sun, N. Ding, Y. Song et al. // *Leukemia.* – 2019. – Vol. 33, № 8. – P. 2105-2110.
7. Wispelaere M. Small molecule degraders of the hepatitis C virus protease reduce susceptibility to resistance mutations / M. Wispelaere, G. Du, K.A. Donovan et al. // *Nat. Commun.* – 2019. – Vol. 10, № 1. – P. 3468.
8. LaPlante G. Targeting the ubiquitin-proteasome system for cancer therapeutics by small-molecule inhibitors / G. LaPlante, W. Zhang // *Cancers.* – 2021. – Vol. 13, № 12. – P. 1-43.
9. Wang C. PROTACs technology for targeting non-oncoproteins: advances and perspectives / C. Wang, Y. Zhang, D. Xing et. al. // *Bioorg. Chem.* – 2021. – Vol. 114, № 10. – P. 105-109.

10. Pettersson M. PROteolysis TArgeting Chimeras (PROTACs) — Past, present and future / M. Pettersson, C. M. Crews // *Drug Discov. Today Technol.* – 2019. – Vol. 31. – P. 15-27.
11. Bondeson D.P. Catalytic in vivo protein knockdown by small-molecule PROTACs / D.P. Bondeson, A. Mares, I.E. Smith et al. // *Nat. Chem. Biol.* – 2015. – Vol. 11, № 8. – P. 611-617.
12. Bond M. J. Proteolysis targeting chimeras (PROTACs) come of age: entering the third decade of targeted protein degradation / M. J. Bond, C. M. Crews // *RSC Chem. Biol.* – 2021. – Vol. 2. – P. 725-742.
13. Samarasinghe K.T.G. Targeted protein degradation: a promise for undruggable proteins / K.T.G. Samarasinghe, C.M. Crews. // *Cell Chem. Biol.* – 2021. – Vol. 28, № 7. – P. 934-951.
14. Hetz C. Mechanisms, regulation and functions of the unfolded protein response / C. Hetz, K. Zhang, R.J. Kaufman. // *Nat. Rev. Mol.* – 2020. – Vol. 21. – P. 421–438.
15. Sontag E.M. Mechanisms and functions of spatial protein quality control / E.M. Sontag, R.S. Samant, J. Frydman // *Annu. Rev. Biochem.* – 2017. – Vol. 86, № 1. – P. 1-26.
16. Yerbury J. J. Walking the tightrope: proteostasis and neurodegenerative disease / J. J. Yerbury, L. Ooi, A. Dillin et al. // *J. Neurochem.* – 2016. – Vol. 137, № 4. – P. 489-505.
17. Pohl C. Cellular quality control by the ubiquitin-proteasome system and autophagy / C. Pohl, I. Dikic // *Science.* – 2019. – Vol. 366. – P. 818–822.
18. Lecker S. H. Protein degradation by the ubiquitin–proteasome pathway in normal and disease states / S. H. Lecker, A. L. Goldberg, W. E. Mitch // *J. Am. Soc. Nephrol.* – 2006. – Vol. 17, № 7. – P. 1807-1819.
19. Hu H. Ubiquitin signaling in immune responses / H. Hu, S.C. Sun // *Cell Res.* – 2016. – Vol. 26. – P. 457–483.
20. Deshaies R.J. RING domain E3 ubiquitin ligases / R.J. Deshaies, C. A. P. Joazeiro // *Annu. Rev. Biochem.* – 2009. – Vol. 78. – P. 399-434.

21. Wang Y. HECT E3 ubiquitin ligases – emerging insights into their biological roles and disease relevance / Y. Wang, E. I. Kane, A. Zhou et al. // *J. Cell Sci.* – 2020. – Vol. 133, № 7. – P. 1-13.

22. Chamberlain P.P. Development of targeted protein degradation therapeutics / P.P. Chamberlain, L.G. Hamann // *Nat. Chem. Biol.* – 2019. – Vol. 15. – P. 937–944.

23. Fryrear K. A. The Sumo-targeted ubiquitin ligase RNF4 regulates the localization and function of the HTLV-1 oncoprotein Tax/ K. A. Fryrear, X. Guo, O. Kerscher et al. // *J. Blood.* – 2012. – Vol. 119, № 5. – P. 1173-1181.

24. Luo K. CDK-mediated RNF4 phosphorylation regulates homologous recombination in S-phase / K. Luo , M. Deng , Y. Li et al. // *Nucleic Acids Res.* – 2015. – Vol. 43, № 11. – P. 5465-5475.

25. Fernandez A. R. SUMO chain-induced dimerization activates RNF4 / A. R. Fernandez, A. Plechanovova, N. Hattersley et al. // *Mol. Cell.* – 2014. – Vol. 53, № 6. – P. 880-892.

26. Bassermann F. The ubiquitin proteasome system - implications for cell cycle control and the targeted treatment of cancer / F. Bassermann, R. Eichner, M. Pagano // *Biochim. Biophys. Acta.* – 2014. – Vol. 1843, № 1. – P. 150-162.

27. Holland J.D. Wnt signaling in stem and cancer stem cells / J.D. Holland, A. Klaus A. N. Garratt et al. // *Curr. Opin. Cell Biol.* – 2013. – Vol. 25, № 2. – P. 254-264.

28. Clevers H. Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control / H. Clevers, K.M. Loh, R. Nusse // *Science.* – 2014. – Vol. 346. – P. 1-8.

29. Thomas J. J. RNF4-dependent oncogene activation by protein stabilization / J. J. Thomas, M. Abed, J. Heuberger et al. // *Cell Rep.* – 2016. – Vol. 16, № 12. – P. 3388-3400.

30. Leach A. Molecular Modelling: Principles and Applications / A. Leach. // Pearson Prentice Hall. – 2009. – Vol. 2.

31. Gleeson M. P. In-silico ADME models: A general assessment of their utility in drug discovery applications / M. P. Gleeson, A. Hersey, S. Hannongbua // *Curr. Top. Med. Chem.* – 2011. – Vol. 11, № 4. – P. 358-381.
32. Wang A. Open-source browser-based tools for structure-based computer-aided drug discovery / A. Wang, J. D. Durrant // *Molecules.* – 2022. – Vol. 27, № 14. – P. 1-14.
33. Lengauer T. Computational methods for biomolecular docking / T. Lengauer, M. Rarey // *Curr. Opin. Struct. Biol.* – 1996. – Vol. 6, № 3. – P. 402-406.
34. Garrett M. M. Molecular Docking / M. M. Garrett, M. L. Wilby // *Met. in Mol. Biol.* – 2008. – Vol. 443. – P. 365-382.
35. Diller D. J. High throughput docking for library design and library prioritization / D. J. Diller, K. M. Merz // *Proteins.* – 2001. – Vol. 43, № 2. – P. 113-124.
36. Jain A. N. Scoring functions for protein-ligand docking / A. N. Jain // *Curr. Protein Pept. Sci.* – 2006. – Vol. 7, № 5. – P. 407-420.
37. Guedes I. A. Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges / I. A. Guedes, F. S. S. Pereira, L. E. Dardenne // *Front. Pharmacol.* – 2018. – Vol. 9. – P. 1089.
38. Genheden S. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities / S. Genheden, U. Ryde // *Expert. Opin. Drug. Discov.* – 2015. – Vol. 10, № 5. – P. 449-461.
39. Meng E. C. Automated docking with grid-based energy evaluation / E. C. Meng, B. K. Shoichet, I. D. Kuntz // *J. Comput. Chem.* – 1992. – Vol. 13. – P. 505–524.
40. Weiner S. J. An all atom force field for simulations of proteins and nucleic acids / S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case // *J. Comput. Chem.* – 1986. – Vol. 7. – P. 230–252.

41. Huang S. Y. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions / S. Y. Huang, S. Z. Grinter, X. Zou // *Phys. Chem. Chem. Phys.* – 2010. – Vol. 12, № 40. – P. 12899-12908.

42. Friesner R. A. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes / R. A. Friesner, R. B. Murphy, M. P. Repasky et al. // *J. Med. Chem.* – 2006. – Vol. 49, № 21. – P. 6177-6196.

43. Muegge I. PMF scoring revisited / I. Muegge // *J. Med. Chem.* – 2006. – Vol. 49, № 20. – P. 5895–5902.

44. Charifson P. S. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins / P. S. Charifson, J. J. Corkery, M. A. Murcko et al. // *J. Med. Chem.* – 1999. – Vol. 42. – P. 5100–5109.

45. Wang R. How does consensus scoring work for virtual library screening? An idealized computer experiment / R. Wang, S. Wang // *J. Chem. Inf. Comput. Sci.* – 2001. – Vol. 41. – P. 1422–1426.

46. Mah R. Drug discovery considerations in the development of covalent inhibitors / R. Mah, J.R. Thomas, C.M. Shafer // *Bioorg. Med. Chem. Lett.* – 2014. – Vol. 24, № 1. – P. 33-39.

47. Ray S. New electrophiles and strategies for mechanism-based and targeted covalent inhibitor design / S. Ray, A. S. Murkin // *Biochem.* – 2019. – Vol. 58, № 52. – P. 5234–5244.

48. Baillie T. A. Targeted covalent inhibitors for drug design / T. A. Baillie // *Angew Chem. Int. Ed. Engl.* – 2016. – Vol. 55, № 43. – P. 13408–13421.

49. Ghosh A. K. Covalent inhibition in drug discovery / A. K. Ghosh, I. Samanta, A. Mondal, W. R. Liu // *Chem. Med. Chem.* – 2019. – Vol. 14, № 9. – P. 889–906.

50. Scarpino A. WIDOCK: a reactive docking protocol for virtual screening of covalent inhibitors / A. Scarpino, L. Petri, D. Knez et al. // *J. Comput. Aided Mol. Des.* – 2021. – Vol. 35, № 2. – P. 223–244.

51. Lai A. C. Induced protein degradation: An emerging drug discovery paradigm / A. C. Lai, C. M. Crews // *Nat. Rev. Drug. Discov.* – 2017. – Vol. 16. – P. 101-114.

52. Wang S. Advanced activity-based protein profiling application strategies for drug development / S. Wang, Y. Tian, M. Wang e. al. // *Front. Pharmacol.* – 2018. – Vol. 9. – P. 353.

53. Ward C. C. Covalent ligand screening uncovers a RNF4 E3 ligase recruiter for targeted protein degradation applications / C. C. Ward, J. I. Kleinman, S. M. Brittain et al. // *ACS Chem. Biol.* – 2019. – Vol. 14, № 11. – P. 2430–2440.

54. Grocock L.M. RNF4 interacts with both SUMO and nucleosomes to promote the DNA damage response / L.M. Grocock, M. Nie, J. Prudden et al. // *EMBO Rep.* – 2014. – Vol. 15. – P. 601-608.

55. Boike L. Advances in covalent drug discovery / L. Boike, N. J. Henning, D. K. Nomura // *Nat. Rev. Drug Discov.* – 2022. – Vol. 21. – P. 881–898.

56. ChemDiv. URL: <https://www.chemdiv.com/>

57. RCSB Protein Data Bank (PDB). URL: <https://www.rcsb.org/>

58. Prime MM-GBSA. URL: <https://www.schrodinger.com/kb/1875>

59. LigPrep 3.4 User Manual, Glide version 6.4, Prime version 3.7, Schrödinger, LLC, New York, NY, 2015.

60. QikProp 4.4 User Manual, Glide version 6.4, Prime version 3.7, Schrödinger, LLC, New York, NY, 2015.

61. Lipinski C. A. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings / C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney // *Adv. Drug Deliv. Rev.* – 2001. – V. 46, № 1-3 – P. 3-26.

62. Protein Preparation Guide Schrödinger Suite 2012 Update 2, Glide version 6.4, Prime version 3.7, Schrödinger, LLC, New York, NY, 2012.

63. Zhu K. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring / K. Zhu, K. W. Borrelli, J. R. Greenwood et al. // *J. Chem. Inf. Model.* – 2014. – Vol. 54, № 7. – P. 1932-1940.

64. CovalentDocking. URL: <https://www.schrodinger.com/science-articles/covdock>

65. Shelley J. C. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules / J. C. Shelley, A. Cholleti, L. L. Frye et al. // J. Comput . Aided. Mol. Des. – 2007. – Vol. 21, № 12. – P. 681-691.

66. Lu C. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space / C. Lu, C. Wu, D. Ghoreishi et al. // J. Chem. Theory. Comput. – 2021. – Vol. 17, № 7. – P. 4291-4300.

67. <https://www.chemdiv.com/catalog/focused-and-targeted-libraries/covalent-fragments-smart-library/>