

УДК: 577.322

## Поиск низкомолекулярного лиганда для афинной хроматографической очистки фактора свёртывания крови VIII человека молекулярным докингом *in silico*

Мацкевич В.А.<sup>11</sup>, Орлова Н.А.<sup>1</sup>, Воробьев И.И.<sup>1,2</sup>, Юрьев А.С.<sup>1</sup>,  
Воробьев А.И.<sup>1</sup>

<sup>1</sup>Гематологический научный центр, Российская академия медицинских наук, Москва, 125167, Россия

<sup>2</sup>Институт биоорганической химии им. М.М.Шемякина и Ю.А.Овчинникова, Российская академия наук, Москва, 117997, Россия

**Аннотация.** При отборе группы синтетических лигандов-кандидатов для очистки рекомбинантного фактора свёртывания крови VIII был использован метод докинга *in silico*. Созданные Кесслером и соавторами пептидомиметические лиганды L3 и L4, которые обратимо связывают молекулу фактора VIII, были использованы как образцовые лиганды сравнения. Предположительные сайты связывания этих лигандов были размещены на поверхности доменов C1 и/или C2 молекулы фактора VIII.

В настоящей работе мы смоделировали взаимодействие лигандов L3 и L4 с предположенными теоретически поверхностями связывания, размещёнными на вершинных областях доменов C1 и C2 фактора VIII. Было обнаружено, что оба образцовых лиганда сравнения связываются с картированными поверхностями связывания со значительными композитными индексами докинга. Эти поверхности были использованы для докинга библиотек двух поставщиков коммерчески доступных соединений *in silico*. Конечные композитные индексы докинга наилучших лигандов-кандидатов из обеих библиотек находятся в одном диапазоне, который соответствует образцовым лигандам. Фильтрация исходных библиотек по признаку наличия функциональных групп, за которые лиганды-кандидаты могут быть иммобилизованы на подложку для проверки афинности, специфичности и обратимости связывания, сильно сократило время расчётов при докинге, а добавка (во время подготовки фильтрованной библиотеки к докингу) линкерной группы, за которую лиганд-кандидат мог бы быть иммобилизован на подложку, увеличило химическую релевантность отобранных докингом наилучших лигандов-кандидатов. Группа из малого числа наилучших лигандов-кандидатов будет в дальнейшем проверена на афинность, специфичность и обратимость связывания с фактором VIII.

**Ключевые слова:** докинг, лиганд, фактор VIII, биотехнология.

### ВВЕДЕНИЕ

Докинг *in silico*, или, другими словами, скрининг библиотек химических соединений, является широко распространённым методом поиска новых лекарственных средств [1–3]. В настоящей работе мы представляем пример использования обычных программных инструментов скрининга библиотек химических соединений для поиска

---

<sup>1</sup> matskevichv@gmail.com

кандидатов аффинного лиганда для очистки рекомбинантного фактора свёртывания крови VIII человека.

Фактор свёртывания крови VIII участвует в реакционном каскаде в качестве коактиватора фактора свёртывания IX, увеличивая его каталитическую активность примерно на три порядка. Отсутствие в плазме крови функционально корректного фактора VIII приводит к развитию гемофилии А, потенциально смертельного генетически обусловленного заболевания. Поскольку распространенность гемофилии А составляет около 1 случая на 10000 человек, лекарственные препараты природного или рекомбинантного фактора VIII для ее лечения изготавливаются многими мировыми производителями и представляют интерес для отечественной фармацевтической индустрии.

Свойствами идеального аффинного лиганда являются высокая специфичность, высокая афинность и полная обратимость связывания с целевой молекулой при ощутимом увеличении ионной силы раствора. Идеальная молекула лиганда также должна содержать функциональную группу для иммобилизации лиганда на поверхности хроматографического сорбента. Данная группа не только не должна участвовать во взаимодействии с целевой белковой молекулой, но и должна быть как можно сильнее удалена от поверхности белка для того, чтобы адсорбция и десорбция целевого белка с поверхности аффинного сорбента как можно меньше зависели от длины и химической структуры линкера между поверхностью сорбента и лигандом.

Выбранный в данной работе для проведения докинга белок является важным фармацевтическим препаратом. Известно, что при получении фармацевтических препаратов фактора VIII с помощью иммуноаффинных сорбентов нередко случаи возникновения аллергических реакций на остаточный иммуноглобулин мыши [4]. Наилучшей альтернативой традиционным иммунсорбентам являются аффинные сорбенты с иммобилизованными пептидомиметиками [5] или небиологическими малыми молекулами [6]. Такие сорбенты в наименьшей степени подвержены деградации ферментами и не загрязняют продукт иммуногенными биополимерами.

## ОБОРУДОВАНИЕ И МАТЕРИАЛЫ

Расчеты проводились при помощи трех четырёхъядерных рабочих станций с архитектурой Intel x86, работающих под управлением ОС Linux (сборка Fedora Core 12 x86 32 bit). В процессе подготовки файлов библиотек химических соединений и трёхмерной структуры белка к докингу использовались внутренние утилиты оболочки для докинга Maestro [7]. Для фильтрации виртуальных библиотек химических соединений и конвертации записей в формат SMILES использовалась утилита OpenBabel [8]. Модификация записей молекул в формате SMILES производилась при помощи мощного и удобного текстового редактора Notepad++ [9]. Модифицированные записи, содержащие данные о линкерных группах, были подготовлены к докингу при помощи утилиты LigPrep [10].

В работе была использована пространственная структура фактора VIII PDB код 2R7E [11], которая была модифицирована следующим образом: домены, не несущие (по литературным данным [12]) на своей поверхности сайтов связывания референсных лигандов, были удалены ручным редактированием координатного файла. Далее координатный файл был подготовлен к докингу при помощи утилиты ProteinPrep из пакета Maestro.

Для фильтрации результатов докинга и отождествления отобранных соединений с фактически синтезированными была использована онлайн-система поиска и заказа у поставщиков по двумерным химическим формулам eMolecules [13].

## МЕТОДЫ И РЕЗУЛЬТАТЫ

Компьютерное моделирование было проведено в две последовательные стадии: докинга виртуальной библиотеки химических соединений производства компании Asinex (РФ) против пространственных рецепторных решёток доменов С1 и С2 фактора VIII, и последующего докинга виртуальной библиотеки производства компании ChemBridge (РФ) против тех же пространственных рецепторных решёток доменов С1 и С2 фактора VIII, проведенного с учетом анализа результатов первого докинга. Общая схема проведения докинга приведена на рис. 1.

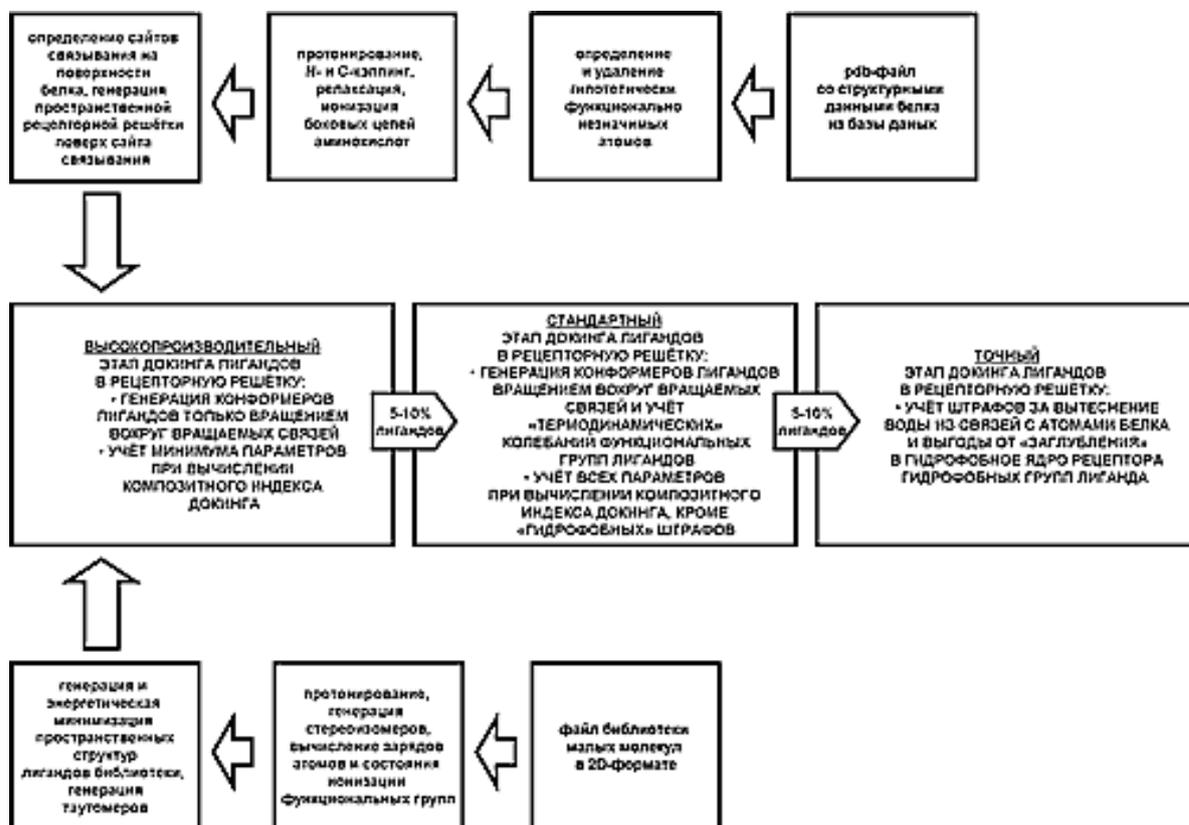
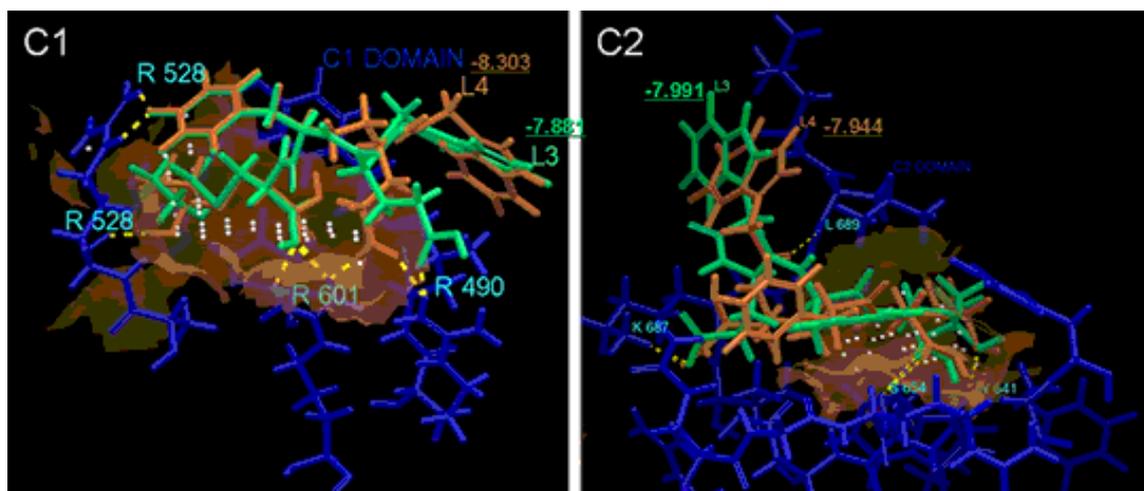


Рис. 1. Общая схема докинга.

Первоначально известные лиганды фактора VIII L3 и L4 были использованы для проверки релевантности выбранных гипотетических сайтов связывания на поверхности доменов С1 и С2. Выбор сайтов связывания был проведен независимо двумя путями: во-первых, анализом литературных данных и поверхностей доменов фактора VIII, участвующих в межмолекулярных взаимодействиях, и во-вторых, использованием утилиты SiteMap (которая картирует сайты связывания лигандов, анализируя распределение заряда и гидрофобных областей по поверхности белковой молекулы) [14]. Оба пути привели к поверхностям доменов С1 и С2, образующим область контакта молекулы белка с клеточной мембраной [11,15], и несущим множество ионизированных боковых цепей аминокислот вместе с редкими для нахождения на поверхности белка гидрофобными боковыми цепями аминокислот. Над поверхностями потенциальных сайтов связывания были расположены границы рецепторных решёток кубической формы со стороной 34 ангстрема.

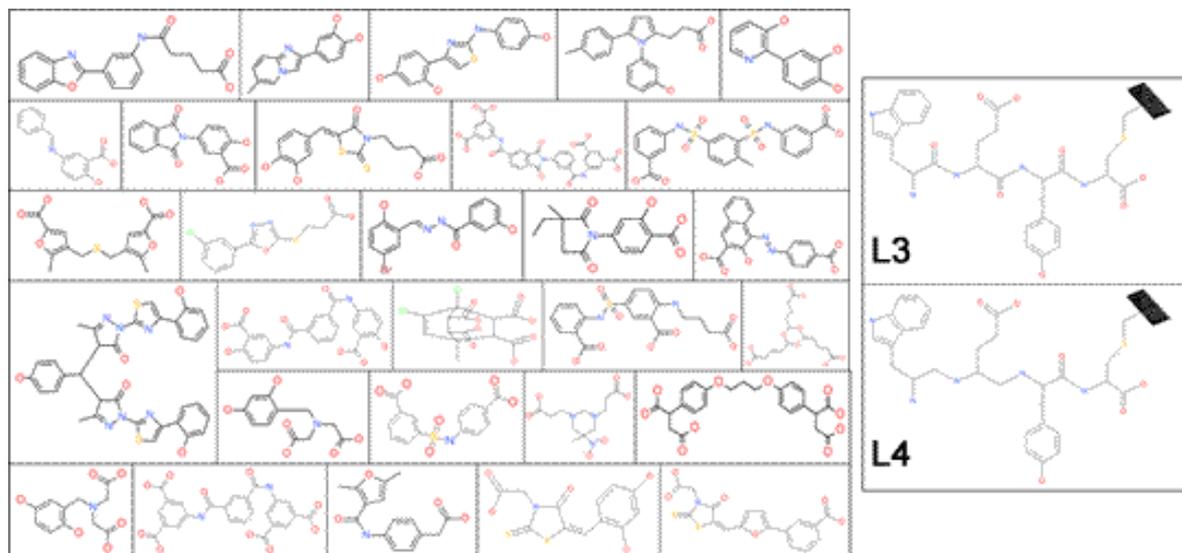
Докинг лигандов L3 и L4 в рецепторные решётки предполагаемых сайтов связывания доменов С1 и С2 дал величины композитного индекса Glidescore в диапазоне между  $-7$  и  $-9$ , что, по предыдущему опыту докинга авторов, является хорошим свидетельством высокоафинного связывания с данными сайтами. Наиболее выгодные энергетически позы лигандов L3 и L4 показаны на рис. 2.



**Рис. 2.** Лиганды L3 и L4 на картированных сайтах связывания доменов C1 и C2. L3 обозначен салатным цветом, L4 – оранжевым. Соответствующим лиганду цветом также показан композитный индекс показанной позы в рецепторной решётке соответствующего сайта (C1 – слева, C2 – справа).

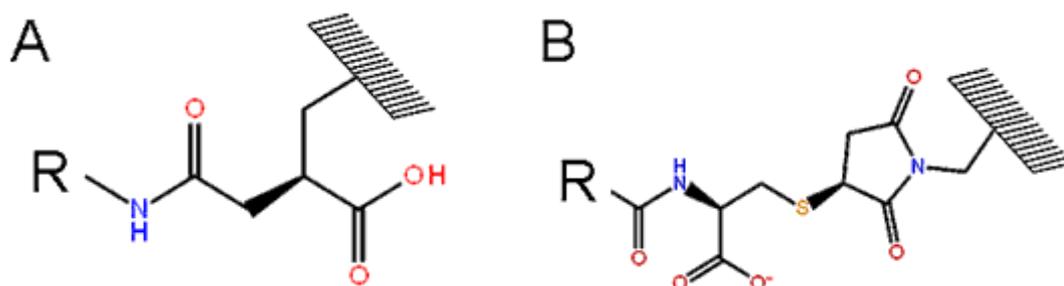
После получения композитных индексов докинга Glidescore в программе Glide [16] для лигандов L3 и L4 стало возможным количественно оценивать результаты докинга виртуальных библиотек химических соединений.

В докинге были использованы библиотеки коммерчески доступных химических соединений Asinex (примерно 240 тыс. химических соединений, библиотека Asinex Gold) и ChemBridge (примерно 1 млн. химических соединений, получена из онлайн-репозитория ZINC [17]). Докинг библиотек химических соединений производился по схеме последовательного обогащения более энергетически-выгодными лигандами-комплексообразователями, связывающимися с рецепторными сайтами. Первый докинг библиотеки Asinex начался с библиотеки из примерно 500 тыс. подготовленных к докингу химических соединений на входе в первую высокопроизводительную стадию, продолжился в стандартном режиме с 20 тыс. отобранных на первой стадии соединений, и закончился в точном режиме с 2 тыс. отобранных на второй стадии соединений. Каждая стадия докинга против одной рецепторной решётки требовала примерно 12 дней чистого машинного времени. Три стадии докинга последовательного обогащения по величине индекса Glidescore для двух рецепторных решёток и с учётом нескольких сбоев и, как следствие, повторов расчетов, длились около 12 недель чистого машинного времени на одной четырёхъядерной рабочей станции. В результате по значению композитного индекса Glidescore, генерируемого для каждого соединения в ходе докинга, среди химических соединений были отобраны 40 наиболее энергетически выгодных молекул-кандидатов. Среди найденных 40 молекул-кандидатов 24 имели в составе карбоксильную группу, причём 16 из этих 24 соединений имели 2 и более карбоксильных групп. Величина Glidescore для данных 40 кандидатов соответствовала диапазону Glidescore конформеров L3 и L4 в использованных для докинга рецепторных сетках. Соединения с наилучшими (наиболее отрицательными) значениями Glidescore показаны на рис.3. Для сравнения химического строения найденных докингом лигандов на рис.3 также приведена химическая структура лигандов L3 и L4.



**Рис. 3.** Кандидатные лиганды, отобранные из библиотеки химических соединений Asinex Gold. L3 (справа сверху) и L4 (справа внизу) – структурные формулы соответствующих лигандов.

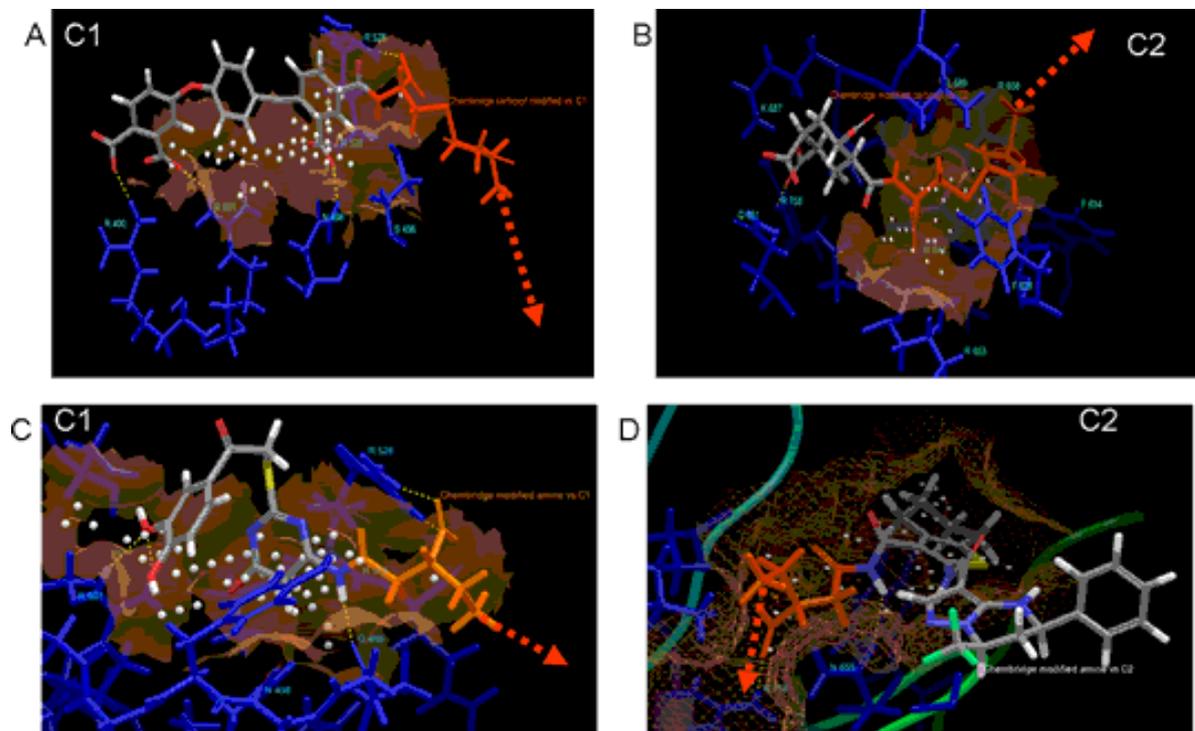
Поскольку проведение докинга по исходной схеме для большей по размеру виртуальной библиотеки химических соединений ChemBridge потребовало бы около 60 недель вычислений одной рабочей станции, было предположено, что перед началом докинга библиотека может быть профильтрована для удаления молекул, непригодных для иммобилизации на подложке. Также для проведения второго этапа докинга были использованы 3 рабочие станции. После проведения пробной фильтрации исходной библиотеки было обнаружено около 120 тыс. молекул, содержащих первичную аминогруппу или карбоксильную группу. Данные молекулы были пакетно модифицированы путем включения в их состав конечного участка соответствующей линкерной группы, тем самым была повышена химическая релевантность моделирования связывания белка с иммобилизованными малыми молекулами. Варианты модификаторов первичных аминогрупп и карбоксильных групп показаны на рис. 4.



**Рис. 4.** Структурные формулы дистальных участков линкерных групп, модифицирующих первичные аминогруппы (А) и карбоксильные группы (В) в составе исследуемых молекул. Для иммобилизации лигандов предполагалось использовать N-гидроксисукцинимидную группу в случае первичных аминогрупп и цистеиновое производное малеинового ангидрида в случае карбоксильных групп. «R» обозначен остаток молекулы после присоединения к ней соответствующего иммобилизационного линкера к реакционно-способной группе.

После проведённой в программе OpenBabel автоматической фильтрации карбоксил-содержащих и амино-содержащих лигандов из 2.4 млн. соединений библиотеки ChemBridge осталось в сумме около 100 тыс. соединений. Файл с этими соединениями был преобразован программой OpenBabel в формат SMILES (линеаризованный формат химических формул, представляет из себя текстовую строку, допускающую над собой текстовые операции поиска-замены) для последующей модификации соединений

атомами линкерных групп при помощи Notepad++ [9]. После пакетной модификации этих соединений по карбокси- и аминогруппам данная библиотека фильтрованных и модифицированных лигандов прошла заново процедуру подготовки к докингу в программе Schrodinger и достигла размеров примерно 210 тыс. соединений. Был проведён докинг этой библиотеки в решётки докинга вершин доменов C1 и C2 на двенадцати вычислительных ядрах. Трёхстадийный докинг этой библиотеки в решётки C1 и C2 в общей сложности (три стадии последовательного обогащения для каждой из решёток с модифицированной и подготовленной библиотекой) занял около 2 недель чистого машинного времени. При анализе полученных таблиц наилучших кандидатов с третьей стадии докинга библиотеки модифицированных молекул было установлено, что поляризованные атомы линкеров в большей части соединений участвуют в образовании стабилизирующих водородных связей с атомами боковых цепей фактора свёртывания VIII (см. рис. 5), тем самым была повышена химическая релевантность моделирования *in silico* и уменьшена до приемлемой продолжительность докинга. Величины Glidescore для кандидатных лигандов также оказались в области значений этого индекса как для конформеров L3 и L4, так и для выбранных кандидатов докинга библиотеки Asinex.



**Рис. 5.** Предполагаемые пространственные структуры комплексов кандидатных лигандов с поверхностью фактора VIII. Атомы линкерных групп обозначены оранжевым цветом. Область связывания на поверхности белка обозначена коричневым, предполагаемое направление хода алкильной цепи линкера – красной стрелкой с прерывистой линией. Названия аминокислот, атомы боковых цепей которых участвуют в образовании стабилизирующих водородных связей с лигандом, обозначены голубым шрифтом. Панель А - комплекс карбоксил-содержащего лиганда с сайтом связывания C1. Панель «В» - комплекс карбоксил-содержащего лиганда с сайтом C2. Панели «С» и «D» - комплексы аминосодержащих лигандов с сайтами C1 и C2, соответственно.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Докинг *in silico* лигандов в рецепторные сайты молекул-мишеней сейчас всё более и более преобладает над методом химического скрининга в первичной фильтрации перспективных соединений при поиске новых лекарственных средств. Кроме высокой сложности и ресурсоёмкости химического скрининга, другим ограничением

применимости этого метода является невозможность очистки продуктов от предшественников и побочных продуктов синтеза. Вследствие этого, комбинаторный синтез иммобилизованных молекул обычно проводят в две стадии, используя единственное стартовое химическое соединение. Это ограничивает размер иммобилизованной библиотеки несколькими сотнями или тысячами соединений. Получаемое химическое разнообразие такой библиотеки несравнимо с достигаемым разнообразием обычных библиотек соединений, имеющих размер  $10^5$ – $10^6$  соединений и основанных на множестве стартовых соединений. В то же самое время, совершенствование вычислительных технологий и средств многократно снизили время и стоимость выполнения фильтрации библиотек миллионного размера *in silico* на уже достаточном уровне химической релевантности расчетов.

При проведении докинга лигандов из стандартной библиотеки химических соединений следует принимать во внимание возможность иммобилизации обнаруженных кандидатных лигандов на подложке или поверхности пористого сорбента. Мы предположили, что при наличии у химических соединений первичной аминогруппы или карбоксильной группы их иммобилизация может быть проведена широко распространенными и хорошо изученными методами, при этом могут быть использованы коммерчески доступные микропланшеты с активированной поверхностью, предназначенные для ковалентной иммобилизации пептидов и белков. Использование таких микропланшетов вместо индивидуальных порций активированных пористых сорбентов позволит значительно увеличить скорость экспериментального скрининга кандидатов.

Поскольку химическая модификация линкерными группами ионизируемых амино- или карбокси- групп в составе скринируемых соединений может значительно изменить их свойства, к формулам соединений в составе отфильтрованных суб-библиотек были добавлены дистальные атомы линкерных групп.

В результате проведенной модификации число атомов в полученных «молекулах» было увеличено в 1,5–2 раза по сравнению с исходными молекулами, что пропорционально увеличило продолжительность скрининга. Одновременно с этим комплексы белка с лигандами, в которых реакционноспособная группа лиганда располагалась бы недоступным для модификации образом, были полностью исключены из рассмотрения. Это сократило количество соединений на входе в первый этап докинга в 10 раз, пропорционально укоротив время первого вычислительного этапа. Интересно отметить, что для всех отобранных кандидатных лигандов было обнаружено, что точка прикрепления алкильной цепочки линкера к его дистальной группе предположительно ориентирована в сторону раствора и не образует непосредственного контакта с поверхностью белка. Предположительно, данные соединения не будут иметь стерических затруднений при взаимодействии с поверхностью белка.

Обычная практика проведения высокопроизводительного скрининга библиотек химических соединений направлена на отбор молекул, плотно связывающихся с ферментами в области их активного сайта или блокирующих белок-белковые взаимодействия. Как правило, первичный скрининг библиотек может обнаружить молекулы, обладающие константой диссоциации  $10^{-5}$  –  $10^{-7}$  М. Для получения молекул, связывающихся с белком более плотно, т.е. в диапазоне величин констант диссоциации  $10^{-9}$  –  $10^{-11}$  М, проводят синтез и скрининг библиотек производных молекул или скрининг производных молекул *in silico* и последующий синтез небольшого набора отобранных молекул. В случае поиска обратимо связывающихся с белком лигандов данный этап исследований не нужен, поскольку для проведения практической ассоциации и диссоциации белка с аффинным сорбентом область констант диссоциации  $10^{-5}$  –  $10^{-7}$  М является оптимальной. Таким образом, основной целью дальнейшего отбора лигандов будет являться не поиск лиганда с высокой активностью,

а экспериментальный поиск лигандов с наилучшей кинетикой адсорбции белка и полнотой десорбции в условиях, не повреждающих целевой белок.

В настоящий момент большинство рекомбинантных белков для медицинского применения получают в культивируемых клетках животных. Очистка таких белков до фармакопейных требований невозможна без использования стадии аффинной хроматографии. Для проведения этой стадии традиционно используют моноклональные антитела, но сорбенты на их основе обладают рядом существенных недостатков – ограниченным сроком службы, относительно медленной скоростью адсорбции целевого белка. Кроме того, при элюции целевого белка происходит и частичная утечка иммобилизованных антител. В ряде случаев загрязнение продукта даже небольшими количествами иммуноглобулинов мыши вызывает аллергические реакции у пациентов [4].

Аффинные сорбенты на основе низкомолекулярных лигандов непептидной природы наилучшим образом подходят для очистки белков фармацевтического применения. Основным ограничением для их широкого применения является высокая трудоемкость экспериментального скрининга кандидатных лигандов. Приведенный в настоящей работе алгоритм проведения скрининга кандидатных лигандов *in silico* может быть полезен для поиска лигандов к различным рекомбинантным белкам медицинского применения.

Работа выполнена при поддержке Правительства Москвы, государственный контракт № 8/3-125Н-09 от 10 июня 2009 г.

#### СПИСОК ЛИТЕРАТУРЫ

1. Koborova O.N., Filimonov D.A., Zakharov A.V., Lagunin A.A., Ivanov S.M., Kel A., Poroikov V.V. In silico method for identification of promising anticancer drug targets. *SAR and QSAR Environmental Research*. 2009. V. 20. P. 755–766.
2. Lagunin A., Filimonov D.A., Poroikov V.V. Multi-targeted natural products evaluation based on biological activity prediction with PASS. *Current Pharmaceutical Design*. 2010. V. 16. P. 1703–1717.
3. Zoete V., Grosdidier A., Michielin O.J. Docking, virtual high throughput screening and *in silico* fragment-based drug design. *Cellular and Molecular Medicine*. 2009. V. 13. P. 238–248.
4. Larson P., Zhang C., Gorina E., Burmeister Getz E. IgG formation to mammalian proteins in hemophilia A patients following treatment with a new recombinant human factor VIII. *Journal of Thrombosis and Haemostasis*. 2004. V. 2. P. 1011–1012.
5. Perugini M., Varelias A., Sadlon T., D'Andrea R.J. Hematopoietic growth factor mimetics: From concept to clinic. *Cytokine & Growth Factor Reviews*. 2009. V. 20. P. 87–94.
6. Morrill P., Gupta G., Sproule K., Winzor D., Christensen J., Mollerup I., Lowe C. Rational combinatorial chemistry-based selection, synthesis and evaluation of an affinity adsorbent for recombinant human clotting factor VII. *Journal of Chromatography B*. 2002. V. 774. P. 1–15.
7. Schrödinger Maestro, version 9.0. URL: <https://www.schrodinger.com/productsguide/> (дата обращения 29.12.2010).
8. Guha R., Howard M.T., Hutchison G.R., Murray-Rust P., Rzepa H., Steinbeck C., Wegner J. K., Willighagen E. The Blue Obelisk Interoperability in Chemical Informatics. *Journal of Chememical Information and Modeling*. 2006. V. 46. P. 991–998.
9. Notepad++. URL: <http://notepad-plus-plus.org/home> (дата обращения: 22.12.2010).
10. Schrödinger LigPrep, version 2.3. URL: <https://www.schrodinger.com/productsguide/> (дата обращения 29.12.2010).

11. Shen B.W., Spiegel P.C., Chang C.H., Huh J.W., Lee J.S., Kim J., Kim Y.H., Stoddard B.L. The tertiary structure and domain organization of coagulation factor VIII. *Blood*. 2008. V. 111. P. 1240–1247.
12. Knor S., Khrenov A., Laufer B., Benhida A., Grailly S.C., Schwaab R., Oldenburg J., Beaufort N., Magdolen V., Saint-Remy J.M., Saenko E.L., Hauser C.A., Kessler H. Efficient factor VIII affinity purification using a small synthetic ligand. *Journal of Thrombosis and Haemostasis*. 2008. V. 6. P. 470–477.
13. eMolecules. URL: <http://www.emolecules.com> (дата обращения: 22.12.2010).
14. Schrödinger SiteMap, version 2.3. URL: <https://www.schrodinger.com/productsguide/> (дата обращения 29.12.2010).
15. Ngo J.C., Huang M., Roth D.A., Furie B.C., Furie B. Crystal structure of human factor VIII: implications for the formation of the factor IXa-factor VIIIa complex. *Structure*. 2008. V. 16. P. 597–606.
16. Schrödinger Glide, version 5.5. URL: <https://www.schrodinger.com/productsguide/> (дата обращения 29.12.2010).
17. Irwin J.J., Shoichet B.K. ZINC – a free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*. 2005. V. 45. P. 177–182.

Материал поступил в редакцию 24.12.2010, опубликован 28.01.2011.