

ОТЗЫВ

**официального оппонента на диссертацию на соискание ученой
степени кандидата биологических наук
Богдановой Елизаветы Александровны
на тему: «Предсказание аффинности в белок-белковых комплексах на
основе межатомных расстояний с использованием трёхмерной
свёрточной нейронной сети»
по специальности 1.5.8. – «Математическая биология, биоинформатика»**

Изучаемые Богдановой Е.А. взаимодействия между белковыми молекулами лежат в основе большинства клеточных процессов, таких как реакции матричного синтеза, сигнальные каскады и многие другие пути, запускающие программы управления геномными процессами. Использование белок-белковых взаимодействий в качестве мишени для разработки новых терапевтических средств является актуальным направлением в фармакологии. В связи с этим, предсказание энергии связывания белок-белковых взаимодействий важно для таких областей, как биоинформатика, молекулярная биология, фармакология и др. Автор предлагает новый подход анализа структур белок-белковых молекул и предсказания их аффинности, который может ускорить ранние этапы разработки пептидных и белковых лекарственных средств. Следовательно, очевидна практическая значимость работы Елизаветы Александровны.

Диссертационная работа Богдановой Е.А. построена по классической схеме. Она включает следующие разделы: Введение, Обзор литературы, Материалы и методы, Результаты и их обсуждение, Заключение, Выводы, Список использованной литературы и Приложение. Также в работе представлен список используемых сокращений, что облегчает процесс изучения диссертации. Работа изложена на 141 странице, содержит 44 иллюстрации, 10 таблиц, 2 приложения и цитирует 187 литературных источников.

Обзор литературы знакомит читателя с существующими подходами по предсказанию энергии связывания в белок-белковых комплексах, а также с их областью применения и ограничениями. Также в данной главе описаны основные особенности и разновидности белок-белковых взаимодействий, учитывая как связывание между белковыми молекулами, так и белков с пептидами.

Глава “Материалы и методы” содержит описание различных биоинформатических подходов, используемых в данной работе для обработки пространственных структур белковых комплексов. В главе также представлены технологии и инструменты, использованные для разработки и обучения трехмерной сверточной нейронной сети, предсказывающей аффинность связывания белковых молекул, содержит описание использованных методик.

В главе «Результаты и их обсуждение» автор приводит результаты, полученные в процессе исследования. В первом подразделе описывается предложенная автором методика автоматической локализации интерфейса связывания и представления химических свойств взаимодействующих молекул. Ее ключевой особенностью является то, что она учитывает важные для белок-белковых взаимодействий контакты (водородные связи, гидрофобные контакты и др.), сохраняя при этом информацию о пространственном расположении атомов белок-белкового комплекса.

Далее Елизавета Александровна представляет и описывает новый разработанный предсказательный алгоритм, оценивающий энергию связывания между белковыми молекулами в комплексе. Приводится описание архитектуры нейронной сети, состоящей из нескольких сверточных слоев, что позволяет анализировать сложные нелинейные зависимости между атомами. Обсуждаются этапы обучения модели, включая использование методов аугментации для расширения обучающей выборки и предотвращения переобучения.

Разработанный предсказательный алгоритм был апробирован автором на нескольких независимых разнородных тестовых наборах и показал лучшую точность предсказания среди существующих альтернативных методов. Также стоит отметить проведенный автором анализ важности каналов, что позволяет оценить, какие взаимодействия определяются предсказательным алгоритмом как наиболее значимые для предсказания аффинности связывания.

Во второй половине главы автором излагаются возможности практического применения разработанной модели. Так, рассматривается применение алгоритма для оценки влияния точечных мутаций на аффинность связывания в комплексах ACE2-RBD спайкового белка коронавируса. Показана высокая точность предсказания, превосходящая остальные опробованные алгоритмы, что закладывает основу для такой области использования алгоритма, как оценка вклада точечных мутаций в энергию связывания. В конце главы обсуждается возможность использования алгоритма для анализа стабильности белок-белковых комплексов, образованных различными вариантами гистонов, что может иметь важное значение для понимания их роли в регуляции генов и поддержании структуры хроматина.

В целом, глава разносторонне демонстрирует, как разработанный алгоритм может быть использован для глубокого анализа белок-белковых взаимодействий, а также для предсказания аффинности связывания и оценки влияния мутаций, что открывает новые возможности для исследований в области молекулярной биологии и биоинформатики.

«Заключение» и «Выводы» в полной мере обоснованы представленными данными, полученными в ходе выполнения работы. Данные достоверны и получены автором впервые.

Диссертация и автореферат структурированы и оформлены согласно требованиям нормативных документов. Содержание автореферата диссертации в полной мере отражает основные аспекты диссертационной работы. Цель, задачи, положения, выносимые на защиту и выводы, приведенные в автореферате, соответствуют таковым в диссертации.

Работа прошла необходимую апробацию. Результаты работы опубликованы в 4 статьях рецензируемых научных журналов, рекомендованных ВАК и индексируемых в базах РИНЦ, Scopus и Web of Science. Материалы работы успешно докладывались на профильных научных конференциях. Достоверность и научная значимость полученных автором результатов не вызывают сомнений, а выводы полностью соответствуют полученным результатам. Автором был предложен и разработан новый подход к анализу белок-белковых взаимодействий и предсказанию их аффинности, основанный на сверточной нейронной сети. Новизна предложенной методологии также не вызывает сомнений.

Разработанная модель глубокого обучения ProBan, опубликованная в журнале WoS Q1, была протестирована пока на комплексах ACE2-RBD и гистоновых комплексах и были получены значимые научные результаты. Применение разработанной модели к другим белковым комплексам будет логическим предложением данного научного исследования. Детальное изучение предсказаний модели, позволит улучшить ее качество, а также использовать модель для открытия новых, неизвестных ранее белковых взаимодействий. Универсальность подходов, взятых в основу модели, увеличивает ее практическую значимость.

По тексту представленной диссертации можно сделать некоторые замечания:

- 1) В главе, посвященной обзору литературы, местами используются длинные, сложные многосоставные предложения с нарушенным согласованием, что затрудняет их понимание.
- 2) На Рис. 29-32 числа на подписи к одной из горизонтальных осей сливаются, что осложняет их чтение, и, как следствие, восприятие масштаба изображенных молекул.
- 3) В тексте неоднократно употребляется термин «производительность» относительно предсказательных алгоритмов, однако, не хватает пояснения, что именно он обозначает в используемом контексте.

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования. Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В.Ломоносова к работам подобного рода. Содержание диссертации соответствует специальности 1.5.8. – «Математическая биология,

биоинформатика» (по биологическим наукам), а также критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В.Ломоносова, а также оформлена согласно требованиям Положения о совете по защите диссертаций на соискание ученой степени кандидата наук Московского государственного университета имени М.В.Ломоносова.

Таким образом, соискатель Богданова Елизавета Александровна заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. – «Математическая биология, биоинформатика».

Официальный оппонент:

кандидат физико-математических наук,
заведующий международной лабораторией биоинформатики Института искусственного интеллекта и цифровых наук Факультета компьютерных наук Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Высшая школа экономики»

ПОПЦОВА Мария Сергеевна

Подпись

19.02.25

Дата подписания

Контактные данные:

тел.: 7(962) 909 5153, e-mail: mpoptsova@hse.ru

Специальность, по которой официальным оппонентом защищена диссертация:

03.01.02 – Биофизика

Адрес места работы:

109028, г. Москва, Покровский бульвар, д. 11,

Национальный исследовательский университет «Высшая школа экономики», Факультет компьютерных наук, Институт искусственного интеллекта и цифровых наук, Международная лаборатория биоинформатики

Тел.: +7 (495) 531-00-00 27335; e-mail: mpoptsova@hse.ru

Подпись заверяю