

***IN SILICO* АНАЛИЗ БЕЛКОВ-АНТИМИШЕНЕЙ: МЕТОДОЛОГИЯ МОЛЕКУЛЯРНОГО ДОКИНГА ДЛЯ ОЦЕНКИ СВЯЗИ С ОСТРОЙ ТОКСИЧНОСТЬЮ**

Моргунов И. А.¹, Никитин И. Д.¹

Научный руководитель – д.х.н, чл.-корр. РАН Федоров М. В.¹

¹ИППИ им. А.А. Харкевича РАН

morgunov@iitp.ru

Исследование было проведено в рамках государственного задания Министерства науки и высшего образования Российской Федерации для ИППИ РАН (FFNU-2025-0040).

Введение

Разработка интерпретируемых моделей в вычислительной токсикологии требует интеграции биологических дескрипторов, таких как профили взаимодействия с белками-антимишенями. Эти белки, ассоциированные с нежелательными эффектами, играют ключевую роль в понимании механизмов токсичности. Однако методология их анализа, включая молекулярный докинг, сталкивается с вызовами: выбор структур, подготовка мишеней и учет видовых различий. В настоящей работе мы фокусируемся на методологических аспектах докинга для панели из 44 белков, связанных с токсичностью, с целью оценки их значимости в контексте острой токсичности (LD50, мыши, внутривенно). Это позволяет глубже понять, как обработка данных о мишенях влияет на качество предсказаний.

Основная часть

Для анализа использовали датасет из 12 654 молекул с экспериментальными значениями LD50 [1]. Подготовка лигандов включала конвертацию SMILES в 3D-структуры с помощью Open Babel v3.1.0: генерацию геометрий с оптимизацией по MMFF94s, протонирование при pH 7.4 и расчет зарядов по методу Gasteiger. Особое внимание уделялось удалению молекул с неопределенными хиральными центрами или E/Z-конфигурациями, что снизило шум в данных.

Панель из 44 белков-антимишеней была выбрана на основе публикации четырех ведущих фармацевтических компаний [2]. Структуры получены из PDB, за исключением рецептора вазопрессина V1A (модель по гомологии X5D2B0). Подготовка мишеней в PyMOL v3.1.3 включала удаление растворителей, артефактов и внешних лигандов для создания апо-форм. Определение ортостерических сайтов и координат грида (центр и размер) проводилось с учетом предыдущих исследований (см. Таблицу S1 в [оригинальной работе]). Конвертация в PDBQT с помощью Meeke учитывала альтернативные конформации аминокислот (выбор наиболее вероятных). Для учета видовых различий (человек vs. мышь) провели парное выравнивание последовательностей с помощью BLAST v2.7.0: идентичность составила 81–99%, подтвердив применимость человеческих структур.

Молекулярный докинг выполняли с Vina-GPU 2.0 для ускорения расчетов. Положительные скоринговые значения заменяли на ноль. Анализ результатов показал вариабельность аффинитетов: для hERG/KCNH2 и AVPR1A наблюдались более низкие средние скоринги, указывающие на сильное связывание с токсичными лигандами, в то время как для KCNQ1 и EDNRA распределения скорингов были шире, требуя дополнительной фильтрации. Применение Vutina-кластеризации (Tanimoto 0.65, ECFP4) выявило 9665 кластеров, что подчеркнуло структурное разнообразие и помогло идентифицировать кластеры с аномальными аффинитетами.

Выводы

Методология докинга, включая тщательную подготовку мишеней и учет видовых особенностей, позволила получить надежные аффинитеты для 44 антимишеней. Результаты по отдельным белкам (например, более сильная ассоциация для hERG и AVPR1A) демонстрируют потенциал для механистического профилирования. Датасет с докинг-скорами доступен на <https://github.com/chemagents/ld50-antitargets>, что способствует дальнейшим исследованиям в области объяснимой токсикологии.

Литература

1. Jain S. и др. Large-scale modeling of multispecies acute toxicity end points using consensus of multitask deep learning methods // *Journal of Chemical Information and Modeling*. – 2021. – Т. 61. – №. 2. – С. 653-663. – DOI: 10.1021/acs.jcim.0c01164.
2. Bowes J. и др. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling // *Nature Reviews Drug Discovery*. – 2012. – Т. 11. – №. 12. – С. 909-922. – DOI: 10.1038/nrd3845.