

ТОЧНОСТЬ АНКЕРНЫХ МОТИВОВ КАК КРИТЕРИЙ КАЧЕСТВА ДОКИНГА ПРИ РАЗРАБОТКЕ ЭВОЛЮЦИОННО УСТОЙЧИВЫХ ИНГИБИТОРОВ ВИРУСНОГО ВХОДА

Фофонов М.А.¹, Сахабеев Р.Г.¹

Научный руководитель – канд. биол. наук, доцент Сидоров И. П.¹

¹Санкт-Петербургский государственный технологический институт (технический университет)
spicmaff@gmail.com

Введение

Эволюция вирусов, использующих белки входа для проникновения в клетку хозяина, протекает в широком спектре режимов: от антигенного дрейфа до скачкообразных событий через рекомбинацию и реассортацию. Терапевтические агенты, нацеленные на вирусный вход, неизбежно становятся частью этого эволюционного ландшафта. Эпистаз меняет стоимость мутаций в разных генетических фонах, что осложняет прогноз устойчивости и объясняет повторяющиеся случаи утраты эффективности препаратов при появлении новых вариантов. Средняя длительность клинической разработки противовирусных препаратов составляет 77.2 месяца, причем для HIV-препаратов она увеличилась с 41.7 до 91.7 месяцев за три десятилетия. Стоимость клинической фазы в среднем составляет 117.4 млн долларов США, а каждый день задержки обходится примерно в 800 тыс. долларов в упущенных продажах [1].

Рецептор-миметические ловушки – перспективная стратегия блокирования вирусного входа через имитацию консервативного рецепторного интерфейса. Близость к нативному рецепторному интерфейсу критична для сохранения широты нейтрализации вариантов. Функциональные ограничения рецептор-связывающих сайтов создают фитнес-стоимость эскапе-мутаций: перекрытие с консервативными эпитопами замедляет ускользание вируса. Конъюгация пептидов на микрочастицы усиливает ингибирующую активность на несколько порядков за счет одновременного связывания нескольких белков на поверхности вириона [3]. Однако корректность таких разработок зависит от точности структурных моделей комплекса вирусный белок–рецептор.

Основная часть

Мы провели мотив-центричную валидацию докинга HIV-1 gp120–CD4. Для дегликозилированных моделей ядра gp120 (HXB2) и доменов CD4 D1–D2 сгенерирован ансамбль rigid-body комплексов в ClusPro, pyDock и HDOCK. Модели сопоставлены с экспериментальным референсом 1GC1 и независимым предсказанием AlphaFold 3. Глобальная оценка выполнена по DockQ, CAPRI, площади интерфейса (Δ SASA) и PRODIGY. Функциональная корректность формализована через ключевой мотив CD4-binding site: расстояния C α –C α и контактный отпечаток Phe43/Arg59 (CD4) с Asp368/Trp427/Val430 (gp120).

Для референсной структуры 1GC1 мотив характеризуется: площадью интерфейса 957 Å², ΔG – 13.4 ккал/моль, расстояниями Phe43–Asp368/Trp427/Val430 – 5.903/7.016/7.493 Å, Arg59–Asp368 – 11.550 Å. AlphaFold 3 воспроизводит геометрию с отклонениями 0.2 – 0.3 Å. Лучшая модель докинга достигает оценки medium DockQ \approx 0.70, имеет крупнейшую площадь интерфейса \approx 1201 Å² и наиболее благоприятную оценку PRODIGY. Однако локальные метрики демонстрируют выраженное искажение Phe43-кластера: расстояния отклоняются от референсных более чем на 25%, при нулевом перекрытии ключевых контактов с 1GC1. Краткая MD-релаксация (\sim 200 пс) не приводит к сближению с референсной геометрией, указывая на устойчивость альтернативного

локального минимума. Таким образом, глобальные критерии недостаточны для функциональной валидации: необходима мотив-центричная проверка геометрии и hotspot-контактов с верификацией по экспериментальному и ML-референсу.

Это обосновывает разработку пайплайна с мотив-центричной валидацией на всех этапах генерации и отбора пептидов. Для каждой вирусной системы идентифицируются ключевые остатки на рецепторном интерфейсе, определяющие связывание: структурная близость к вирусному белку, эволюционная консервативность, функциональная значимость замен для связывания и критерии очищающего отбора. Библиотека пептидов генерируется тремя способами: линейные фрагменты рецептора, разрывные мотивы, соединенные через гибкие линкеры, и *de novo* спроектированные белковые каркасы.

Эволюционная устойчивость оценивается через построение взвешенного графа мутаций вирусного белка: вершины соответствуют одиночным мутантам, а ребра связывают варианты, различающиеся ровно на одну аминокислоту, с весами, отражающими фитнес-стоимость каждого перехода [2]. Эта модель фитнес-ландшафта позволяет формализовать мутационные траектории *escape* как задачу поиска кратчайшего пути. Динамическая модель описывает конкуренцию дикого типа и *escape*-мутантов в организме хозяина с учетом фармакокинетики обоих форматов и окна селекции мутантов – диапазона концентраций, в котором дикий тип подавлен, но устойчивые варианты еще способны к размножению [3]. Центральная выходная метрика – время до *escape*, определяющая устойчивость терапии к эволюционному прорыву. Валидация проводится ретроспективно: предсказанные позиции сопоставляются с данными глубокого мутационного сканирования, а ранжирование *escape*-барьеров – с экспериментально наблюдавшимися *escape*-мутациями антител.

Выводы

Показано, что глобальные метрики докинга не гарантируют функциональную корректность интерфейса вирусный белок–рецептор: формально удовлетворительные оценки DockQ и PRODIGY могут соответствовать ненативной конфигурации ключевых связывающих мотивов. Это обосновывает необходимость мотив-центричных критериев валидации с верификацией по экспериментальным и высокоточным ML-референсам.

Литература

1. Sertkaya A., Beleche T., Jessup A., et al. Costs of Drug Development and Research and Development Intensity in the US, 2000-2018 // JAMA network open. 2024. Vol. 7, no. 8. <https://doi.org/10.1001/jamanetworkopen.2024.15445>
2. Hsu C., Nisonoff H., Fannjiang C., et al. Learning protein fitness models from evolutionary and assay-labeled data // Nature biotechnology. 2022. Vol. 40, no. 7. <https://doi.org/10.1038/s41587-021-01146-5>.
3. Drlica K., Zhao X. Mutant selection window hypothesis updated // Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2007. Vol. 44, no. 5. <https://doi.org/10.1086/511642>.