

Поиск нового центра связывания артокарпина в структуре NanA из *Streptococcus pneumoniae* с учетом конформационной пластичности белка*

Я.А. Шарапова, Д.А. Суплатов, В.К. Швядас

Факультет биоинженерии и биоинформатики и НИИ ФХБ имени А.Н.Белозерского
МГУ имени М.В.Ломоносова, Москва, Россия

Конформационная пластичность является исключительно важным свойством белка – как для реализации его функции, так и при его эволюции. Тем не менее, в большинстве проводимых исследований при изучении связывания/распознавания лиганда в структуре сайта-мишени, а также при скрининге больших библиотек низкомолекулярных соединений, применяются стратегии, которые можно назвать «статическими» – изучаемые белки упрощенно представляются в виде единственной структуры, полученной, как правило, из банка данных PDB. Последние исследования позволяют говорить о низкой эффективности таких «статических» подходов. Активные и аллостерические центры белков/ферментов следует рассматривать не как статические (жесткие) структуры, а в виде ансамбля конформеров [1].

В этой работе оригинальный комплексный подход, основанный на сочетании методов биоинформатики, молекулярного моделирования и машинного обучения, был применен для поиска нового, ранее не охарактеризованного центра распознавания ингибитора артокарпина в структуре нейраминидазы А (NanA) из патогенной бактерии *Streptococcus pneumoniae*. Нейраминидаза А считается ключевым фактором вирулентности пневмококка и обладает уникальной структурной организацией, а именно пространственной обособленностью лектинового и каталитического доменов, соединенных гибким линкером из 16 аминокислот [2]. Низкомолекулярное вещество артокарпин подавляет ключевую стадию патогенеза – формирование биопленок, однако центр связывания этого ингибитора в структуре NanA и механизм его действия остаются неизвестными [3]. Новизной использованного в работе оригинального подхода к изучению механизмов связывания/распознавания лигандов является учет пластичности структуры белка за счет анализа большого ансамбля конформеров. Поиск нового центра связывания артокарпина в структуре NanA из *S. pneumoniae* включал следующие этапы (Рис. 1): (а) генерацию и предварительный анализ большого ансамбля конформеров фермента выполняли с использованием молекулярного моделирования [2] и биоинформатического анализа [4]; после чего для каждого выбранного конформера (б) аннотировали потенциальные центры распознавания низкомолекулярных лигандов [5] и (в) оценивали возможность связывания артокарпина в этих сайтах; наконец, на основании всей собранной информации (г) выбирали наиболее перспективные ориентации лиганда с использованием методов молекулярного моделирования и машинного обучения. Процедуру завершали выбором одного сайта, комплементарного искомому лиганду. Новый метод является вычислительно сложным и представляет собой последовательное решение различных типов задач компьютерной биологии: разделяемых на подзадачи, решение которых требует обмена данными в процессе исполнения (этапы а и г), а также задачи, которые разделяются на большое количество эквивалентных подзадач, решение которых может быть организовано отдельно по одной подзадаче на одно ядро CPU и без обмена данными в процессе исполнения (этапы б и в). Для решения задачи на суперкомпьютере «Ломоносов-2» оптимальное оборудование выбирали для каждого этапа единого комплексного решения. Этап (а) решали с использованием мощностей раздела «pascal», оснащенного одним процессором Intel Xeon Gold 6126, 96 Гб оперативной памяти и двумя видеокартами Tesla P100 на один узел: GPU+CPU ресурс использовали для молекулярного моделирования в соответствии с ранее описанным протоколом [2], после чего CPU ресурс использовали для сравнительного анализа и выбора репрезентативной выборки конформеров с использованием MPI-совместимой программы parMATT [4]. Использование двух GPU ускорителей одного узла в режиме peer-to-peer по-

* Работа выполнена при финансовой поддержке РФФИ (грант № 19-04-01297).

зволило достигнуть максимальной производительности молекулярного моделирования на этапе (а). Для организации вычислительного процесса этапов (б) и (в) в параллельном режиме на CPU раздела «compute», оснащенного процессором Intel Xeon E5-2697 v3 и 64 Гб оперативной памяти на узел, использовали оригинальный планировщик задач mpiWrapper [6]. Наконец, всю собранную информацию обрабатывали на этапе (г) в параллельном режиме в рамках общей памяти на одном CPU узле раздела «pascal».

В результате проведенных исследований описан новый, ранее не изучавшийся сайт связывания ингибитора артокарпина в структуре NanA из пневмококка, а также динамический характер его функционирования. Совместное использование различных вычислительных архитектур суперкомпьютера «Ломоносов-2» для исполнения различных типов задач компьютерной биологии позволило выйти на качественно новый уровень понимания конформационной пластичности и особенностей структурной организации функционально важных участков связывания в NanA из *S. pneumoniae*. Полученные результаты представляют фундаментальный интерес в контексте изучения структурно-функциональных взаимосвязей нейраминидаз/сиалидаз из патогенных организмов, а также имеют практическую значимость для поиска новых путей их регуляции и разработки эффективных антибиотиков.

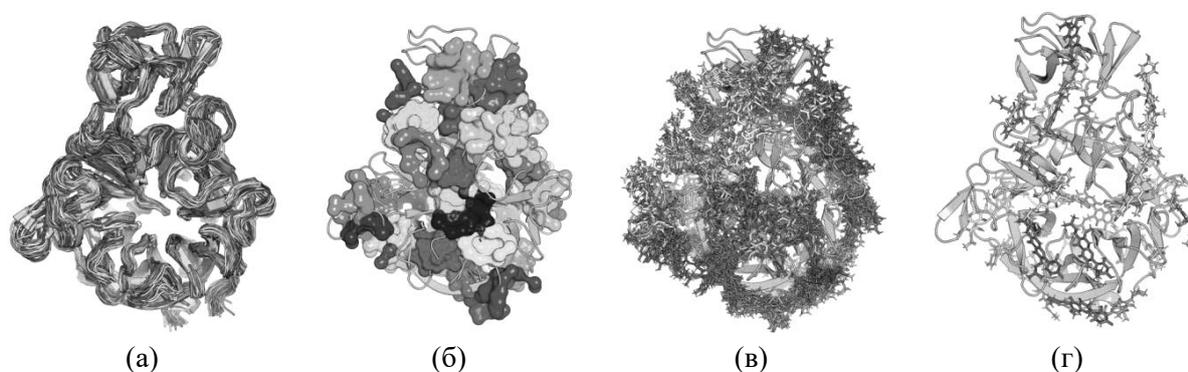


Рис. 1. Этапы решения задачи поиска нового центра связывания артокарпина в структуре NanA из *S. pneumoniae*.

Литература

1. Суплатов Д.А., Шведас В.К. Изучение функциональных и аллостерических сайтов в суперсемействах белков // *Acta Naturae*. 2015. Т. 7, №. 4. С. 39–52.
2. Sharapova Y., Suplatov D., Švedas V. Neuraminidase A from *Streptococcus pneumoniae* has a modular organization of catalytic and lectin domains separated by a flexible linker // *The FEBS journal*. 2018. Vol. 285, No. 13. P. 2428–2445.
3. Walther E., Richter M., Xu Z. et al. Antipneumococcal activity of neuraminidase inhibiting artocarpin // *International Journal of Medical Microbiology*. 2015. Vol. 305, No. 3. P. 289–297.
4. Shegay M.V. Suplatov D.A., Popova N.N., Švedas V.K., Voevodin V.I.V. parMATT: parallel multiple alignment of protein 3D-structures with translations and twists for distributed-memory systems // *Bioinformatics*. 2019. DOI:10.1093/bioinformatics/btz224.
5. Suplatov D., Kirilin E., Arbatsky M., Takhaviev V., Švedas V. pocketZebra: a web-server for automated selection and classification of subfamily-specific binding sites by bioinformatic analysis of diverse protein families // *Nucleic acids research*. 2014. Vol. 42, No. W1. P. W344–W349.
6. Suplatov D., Popova N., Zhumatiy S., Voevodin V.I., Švedas V. Parallel workflow manager for non-parallel bioinformatic applications to solve large-scale biological problems on a supercomputer // *Journal of bioinformatics and computational biology*. 2016. Vol. 14, No. 2. P. 1641008.