

Сравнение программ для молекулярного докинга с точки зрения точности предсказания предреакционных комплексов сериновых и цистеиновых протеаз с их пептидными субстратами

Научный руководитель – Злобин Александр Сергеевич

Мурзин Владислав Андреевич

Студент (специалист)

Московский государственный университет имени М.В.Ломоносова, Факультет
биоинженерии и биоинформатики, Москва, Россия

E-mail: 89086373434@yandex.ru

Сериновые и треониновые протеазы - это классы протеаз, использующие боковые цепи соответствующих аминокислот в качестве нуклеофилов, атакующих карбонильный атом углерода в пептидной связи. Для успешного совершения нуклеофильной атаки требуется правильное расположение гидролизуемого пептида в субстрат-связывающей бороздке фермента. Нуклеофильный атом кислорода или серы должен быть приближен к атакуемому карбонильному атому углерода, а карбонильный атом кислорода должен быть стабилизирован двумя водородными связями с остатками так называемого оксианионного сайта.

Молекулярный докинг - это метод, позволяющий *in silico* получить пространственную структуру молекулярного комплекса, имея на вход структуры исходных компонент. Однако молекулярный докинг далеко не всегда даёт реалистичные предсказания, и нужен способ оценки алгоритмов докинга с точки зрения точности результатов.

Как правило, подобные задачи решаются при помощи сравнения полученных *in silico* результатов с экспериментально полученными структурами комплексов. В случае предреакционных комплексов каталитически активных протеаз с их пептидными субстратами экспериментальные данные получить невозможно, поэтому точность алгоритмов докинга нужно оценивать по косвенным признакам качества предсказания.

AlphaFold - недавно появившаяся программа, предсказывающая пространственную структуру белка исходя из его аминокислотной последовательности. Благодаря опции моделирования в ней мультисубъединичных белков появляется возможность по-новому предсказывать геометрию предреакционного комплекса, представляя протеазу и пептид как две цепи одного белка.

В данной работе нами было проведено сравнение точности работы 7 различных бесплатных (для академических пользователей) программ молекулярного докинга с точки зрения предсказания структур предреакционных комплексов 7 протеаз с различной организацией активного центра с разрезаемыми пептидами. Также было проведено сравнение предсказаний программ для молекулярного докинга с предсказаниями, полученными при помощи AlphaFold для изучения возможности использования последнего для моделирования предреакционных комплексов протеаз.

Использованные протеазы: трипсин *Bos taurus*, субтилизин BPN' *Bacillus amyloliquefaciens*, седолизин *Pseudomonas* sp, ассемблин Human betaherpesvirus 5, папаин *Carica papaya*, протеаза Tobacco etch virus, каспаза-7 *Homo sapiens*.

Использованные программы: AlphaFold, Attract, CABSdock, FroDock, HadDock, HawkDock, HPerDock, ZDock.

Параметры сравнения: расстояние от нуклеофильного атома кислорода/серы до карбонильного атома углерода ($\leq 4\text{\AA}$), угол для соответствующей атаки (O-C-Nu, 60-120 [U+1D52]), длины двух потенциальных оксианионных водородных связей ($\leq 3,5\text{\AA}$).

Для каждой программы и для каждой протеазы (с соответствующим пептидом) был получен набор из 100 результатов (10 повторностей с 10 лучшими предсказаниями для тех программ, которые выдают иной результат при повторном запуске; 100 лучших предсказаний для остальных). Для запуска использовались веб-сервисы (когда это возможно); параметры по умолчанию; в случае, если программа принимает на вход также информацию о сближенных остатках - контактирующими считались остаток с N-конца от атакующей пептидной связи в пептидном субстрате, нуклеофильный остаток протеазы и два остатка протеазы, участвующие в формировании оксианионного сайта. Для выдачи каждой программы был написан скрипт на языке Python, который при помощи библиотеки Prody рассчитывает количество предсказаний, подходящих по всем вышеупомянутым параметрам.

По результатам работы ни одна из семи выбранных программ для молекулярного докинга не выдала значительного числа предсказаний, полностью удовлетворяющих условиям выше. Чуть лучше остальных справилась программа HPerDock. Сравнение по тем же критериям предсказаний AlphaFold показало, что последний в целом лучше справляется с поставленной задачей.

Источники и литература

- 1) R Ménard, A C Storer, Oxyanion hole interactions in serine and cysteine proteases, 1992, 373(7):393-400. doi: 10.1515/bchm3.1992.373.2.393
- 2) Piyush Agrawal, Harinder Singh, Hemant Kumar Srivastava, Sandeep Singh, Gaurav Kishore, Gajendra P S Raghava, Benchmarking of different molecular docking methods for protein-peptide docking, 2019, 19(Suppl 13):426. doi: 10.1186/s12859-018-2449-y
- 3) Gaoqi Weng, Junbo Gao, Zhe Wang, Ercheng Wang, Xueping Hu, Xiaojun Yao, Dongsheng Cao, Tingjun Hou, Comprehensive Evaluation of Fourteen Docking Programs on Protein-Peptide Complexes, 2020, 16(6):3959-3969. doi: 10.1021/acs.jctc.9b01208