

**Реализация метода предсказания сайтов связывания цинка на основе пространственной структуры белка****Научный руководитель – Головин Андрей Викторович****Буянова Мишель Фёдоровна***Студент (специалист)*

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

*E-mail: emkeller@ya.ru*

Ионы цинка, находящиеся в составе комплекса с белком, имеют большое значение в физико-химических процессах различной природы. Показано, что  $Zn(II)$ , проявляя структурную функцию, участвует в фолдинге белковых доменов и сборке макромолекулярных комплексов. Примером тому служит индуцируемая ионом цинка агрегация бета-амилоида, приводящая к развитию болезни Альцгеймера [1]. Находясь в активном центре, цинк способен выступать каталитически активным кофактором ферментов всех шести классов. Функции, выполняемые катионом  $Zn^{2+}$ , являются критичными для работы белков, в состав которых он входит. Таким образом, точное предсказание цинк-связывающих сайтов может помочь как в изучении белков с неизвестной функцией, так и в дизайне новых, обладающих желаемыми свойствами.

Однако экспериментальное обнаружение цинка затруднено. Электронное строение  $Zn(II)$  исключает использование ЭПР-спектроскопии, а его комплексы не имеют поглощения в УФ и микроволновой областях. Использование же метода ЯМР возможно только для небольших соединений ввиду особенностей изотопа  $^{67}Zn$ . Более того, экспериментальные методы не позволяют избежать потери ионов при очистке и кристаллизации. Описанные ограничения приводят к необходимости разработки метода предсказания сайтов *in silico*.

Существующие технические подходы основываются как на последовательности белка, так и на его пространственной структуре. Значительная часть авторов отдает предпочтение приемам машинного обучения, таким как случайные леса, метод опорных векторов, байесовский классификатор [2]. В то же время воспользоваться разработанными алгоритмами затруднительно, так как они не находятся в свободном доступе или сервисы, их предоставляющие, не функционируют.

В рамках данной работы проведен анализ координационных окружений цинка в белках на основе структур из банков данных PDB и PDB\_REDO. Построены распределения наблюдаемых расстояний и углов в зависимости от координаторов; проведена оценка влияния различных аминокислот и их атомов на прочность координационной сферы. Разработан алгоритм, который, базируясь на предпочтительной геометрии сайтов связывания, с использованием элементов теории графов определяет предполагаемых координаторов и положение цинка в структуре и демонстрирует результаты не ниже уровня заявленных в литературе. Для удобства использования реализация доступна как в качестве CLI-инструмента, так и в виде веб-сервиса.

В дальнейшем планируется расширить список доступных катионов, создав таким образом единый инструмент для анализа координации металлов в белках.

**Источники и литература**

- 1) S.A. Kozin, Y.V. Mezentsev, A.A. Kulikova, M.I. Indeykina, A.V. Golovin, A.S. Ivanov, P.O. Tsvetkov, S.A. Makarov. Zinc-induced dimerization of the amyloid-b metal-binding domain 1–16 is mediated by residues 11–14 // Mol. Biosyst. 7 (2011):1053–55

- 2) Zheng C, Wang M, Takemoto K, Akutsu T, Zhang Z, et al. An Integrative Computational Framework Based on a Two-Step Random Forest Algorithm Improves Prediction of Zinc-Binding Sites in Proteins // PLoS ONE 7(11)

### Иллюстрации

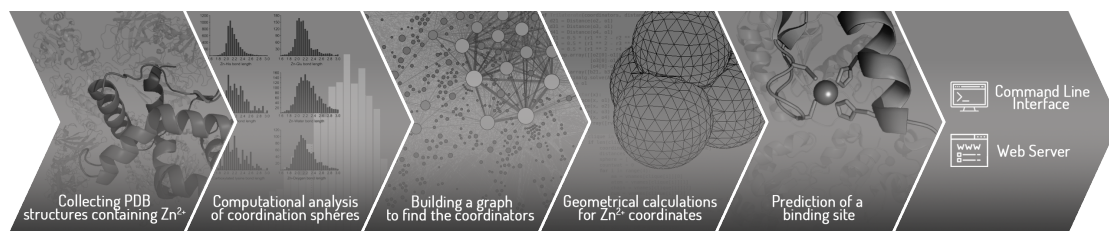


Рис. 1. Краткая схема проделанной работы