

**Использование 3D-мотивов для компьютерной инженерии дисульфидных связей в белках с учетом структурной подвижности их основной цепи**

**Научный руководитель – Суплатов Дмитрий Андреевич**

***Тимонина Дарья Сергеевна***

*Аспирант*

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

*E-mail: darik93@yandex.ru*

Дисульфидные связи между тиольными группами двух остатков цистеина являются общепризнанными факторами стабильности белка, а также могут играть существенную роль в его функции и регуляции. Для предсказания конкретных позиций в белке, замена которых на цистеины приведет к полезному изменению свойств за счет образования S-S связи, разработаны компьютерные подходы, основанные на анализе его статической кристаллографической структуры. Накопленные за последнее время результаты экспериментальных исследований, в которых введение дисульфидного мостика в структуру белка за счет точечных мутаций приводило к значительному конформационному сдвигу атомов основной цепи, позволяют говорить об ограниченности таких «статических» подходов.

Для компьютерной инженерии дисульфидных связей в белках с учетом конформационной подвижности основной цепи предложен новый подход, основанный на использовании 3D-мотивов — паттернов локальной структуры, описывающих два ковалентно связанных цистеина. Были проанализированы все известные конфигурации дисульфидных связей в кристаллографических структурах с высоким разрешением, с использованием метода машинного обучения DBSCAN проведена их кластеризация по обычным и двугранным углам, описывающим конкретную геометрию S-S мостика. Полученная в результате кластеризации библиотека содержала 273 «типовых» 3D-мотивов, описывающих 72% известной вариабельности дисульфидных связей. Для выработки статистической модели, описывающей допустимую структурную подвижность атомов основной цепи пары несвязанных аминокислотных остатков, мутация которых на цистеины способна привести к образованию дисульфидной связи, были проанализированы PDB структуры суперсемейств по классификации CATH и собран набор из 123 уникальных дисульфидных связей и эквивалентных им несвязанных аминокислотных остатков в структурах гомологичных белков. В соответствии с этой моделью пара аминокислотных остатков в структуре изучаемого белка принимается в качестве перспективного сайта для инженерии дисульфидной связи, если соответствующие атомы основной цепи позиций-кандидатов совмещаются хотя бы с одним из 273 3D-мотивов с  $RMSD \leq 0.70 \text{ \AA}$ , что соответствует  $P(x > X) = 0.05$  нормального распределения с параметрами  $\mu = 0.39 \text{ \AA}$  и  $\sigma = 0.19 \text{ \AA}$ , либо отвергается в ином случае. Оценки чувствительности и специфичности нового подхода при тестировании составили 99.89% и 77.29%, соответственно. Новый метод применен на этапе структурной фильтрации в новом веб-сервере Yosshi (<https://biokinet.belozersky.msu.ru/yosshi>) для инженерии дисульфидных связей в белках на основании биоинформатического анализа их гомологов в суперсемействе.

Работа выполнена при поддержке РФФИ (грант 18-29-13060) с использованием оборудования ЦКП сверхвысокопроизводительными вычислительными ресурсами МГУ.