

Задание 1-3

В данных заданиях необходимо рассмотреть цепь D структуры 1IGY и определить ее домены различными способами. Эта структура соответствует антителу, а именно одной из его тяжелых цепей.

Для выполнения задания воспользуемся заготовкой в колабе. Внесем некоторые изменения в код, чтобы исправить недочеты фильтрации данных, а также допишем функцию для определения *split_value* (см. ссылки в конце отчета) (Рис. 1). Кроме этого, изучим запись о нашей белке в PDB (Рис. 2) и запустим поиск доменов по последовательности в базе Interpro (Рис. 3).

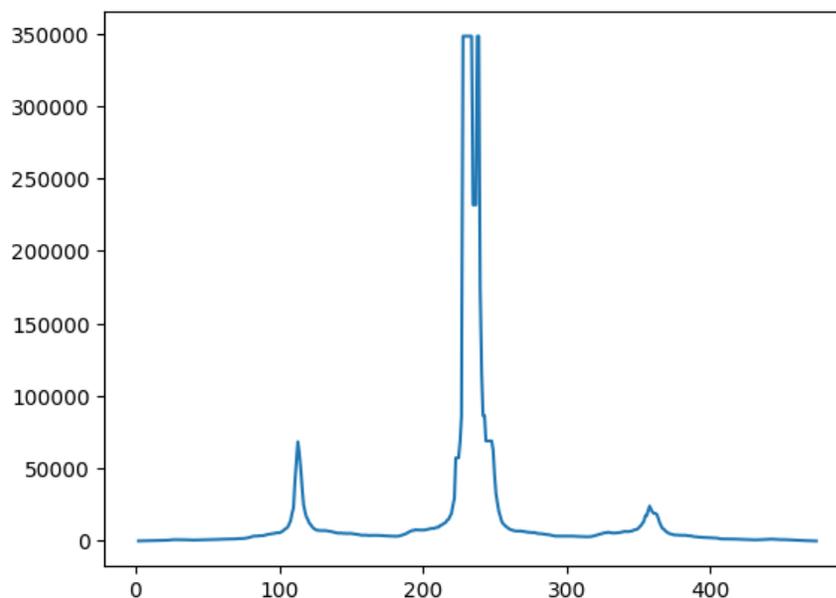


Рис. 1. График *split_value* и разбиение на домены при помощи ДОМАК



Рис. 2. Разбиение на домены в аннотации к структуре в PDB

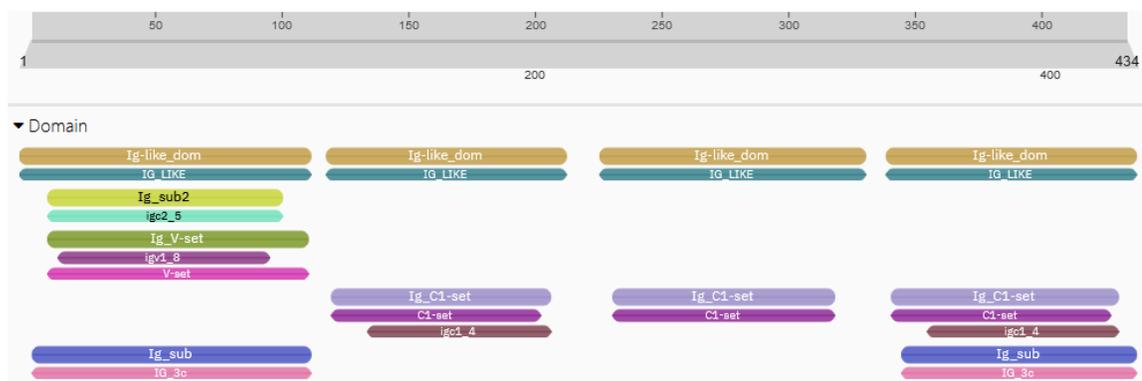


Рис. 3. Разбиение на домены в базе Interpro

В итоге мы можем сравнить качество трех различных способов разбиения на домены. Для этого запишем в таблицу полученные результаты (Таблица 1). Как можно видеть из таблицы, границы доменов, рассчитанные каждым из способов, приблизительно одинаковы. Главное отличие мы видим между границами 2 и 3 доменов. Данное явление можно объяснить тем, что в этом месте расположена длинная петля между двумя «хвостами» цепи антитела (Рис. 4).

Домен	DOMAK	CATH	SCOP	Interpro
1	1-111	1-114	1-114	1-113
2	112-226	115-220	115-216	120-212
3	238-356	221-328	217-332	226-328
4	357-434	329-434	329-434	337-434

Таблица 1. Границы доменов, полученные различными разметками

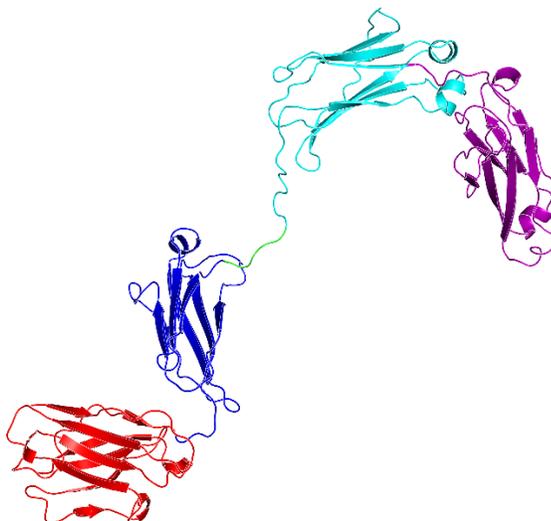


Рис. 4. Визуализация доменов (1 – красный, 2 – синий, 3 – циан, 4 – фиолетовый, петля между 2 и 3 доменами – зеленый)

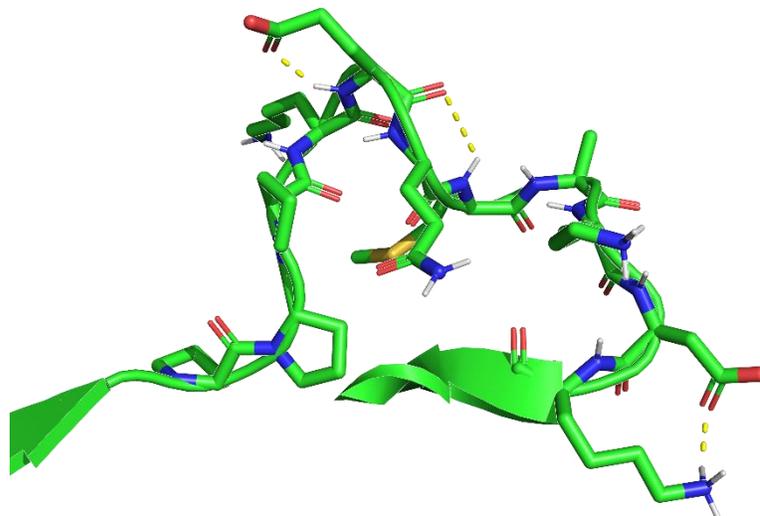


Рис. 6. Структура участка с различиями в разметке

[Скрипт для изменения нумерации аминокислот](#)

[Скрипт для изменения нумерации аминокислот и сохранения только одной цепи](#)

[Файл с измененной нумерацией](#)