

Практикум 7. Структурные домены

Определение структурных доменов

Была исследована структура с PDB ID **6XLW**, относящаяся к человеческому фактору сплайсинга мРНК **U2AF2**, который также участвует в 3'-процессинге. Структура была получена с использованием метода рентгеноструктурного анализа. Белок был закристаллизован в комплексе с олигонуклеотидом **AdML**.

Зависимость значения **Split value** от номера остатка представлена на **рис. 2**. Согласно алгоритму **DOMAK**, в структуре предположительно выделяются два основных домена: **домен А** (остатки 3–106) и **домен В** (остатки 110–192). Границы доменов визуализированы на **рис. 1**.

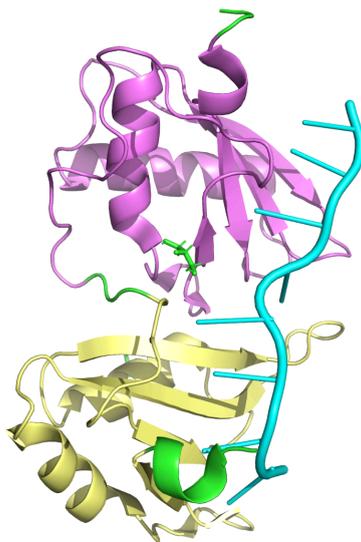


Рисунок 1. Структура U2AF2. Домен А выделен розовым цветом, домен В - желтым.

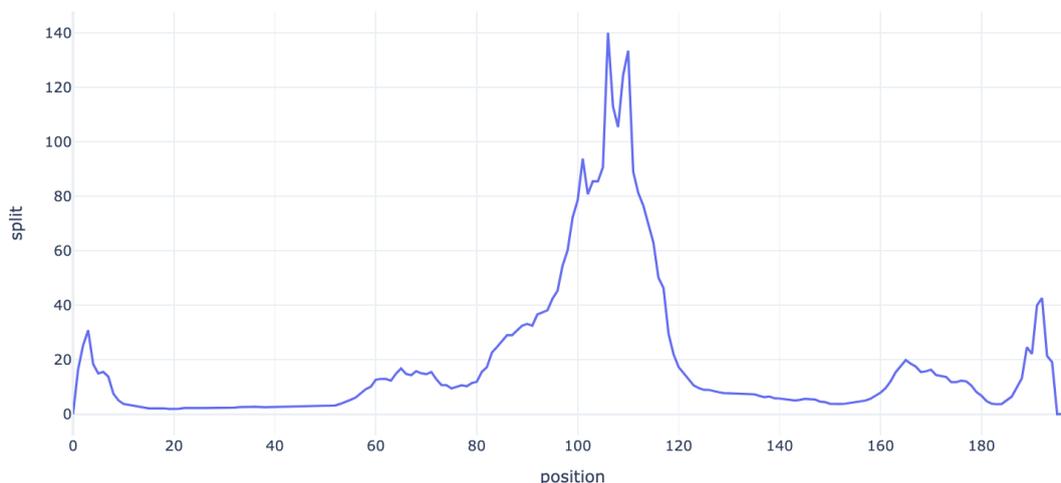


Рисунок 2. График зависимости Split Value от позиции.

Сравнение с базами данных SCOP2 и CATH

Результаты определения доменов с помощью DOMAK сравнили с базами данных SCOP2 и CATH (табл. 1). В CATH информация о доменах данного белка отсутствует. Согласно SCOP2, домен А соответствует [РНК-связывающему домену RBD](#).

Таблица 1. Сравнение полученных данных с данными SCOP2 и CATH.

Источник данных	Домен А	Домен В
ДОМАК	3-106	110-192
SCOP2	1-85	-
CATH	-	-

Визуальное сравнение (рис. 3) показывает, что протяженный фрагмент в 21 остаток на конце домена А (по ДОМАК) не входит в домен по версии SCOP2, в то время как начальный фрагмент (2 аминокислотных остатка) отнесен в SCOP2 к домену, но он является N-концом белка, поэтому в целом скорее не несет важной функциональной роли.

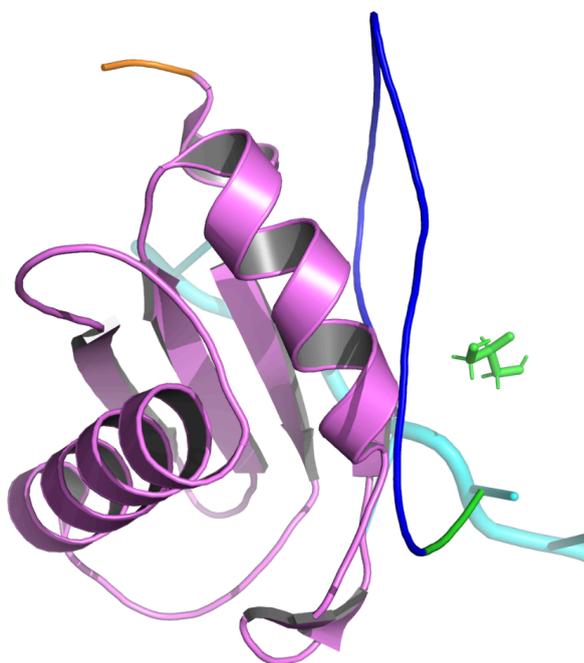


Рисунок 3. Сравнение интерпретации результатов работы ДОМАК с данными SCOP2 для домена А; оранжевым отмечен фрагмент, вошедший в домен по версии SCOP2, но не отмеченный в интерпретации ДОМАК; синим отмечен фрагмент, отнесенный в ДОМАК к домену А, но не относящийся к нему по версии SCOP2.

Согласно литературным данным [1], в факторе U2AF2 содержится два основных РНК-связывающих домена, первый начиная с аланина 148 по глицин 237, второй - с аланина 258 по аланин 342. Так как в эксперименте была получена частичная структура, то в нашем случае эти аннотированные домены находятся с 7 по 96 остатки и с 116 по 200.

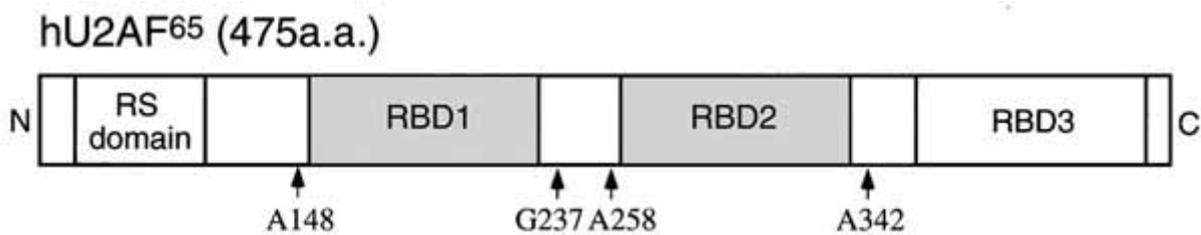


Рисунок 4. Схематическое изображение доменной структуры hU2AF65. RS-домен обозначает богатый аргинином/серином домен, RBD — РНК-связывающий домен. Стрелками показаны N- и C-концы полипептидов, использованных в данном исследовании.

Сравнение с INTERPRO

[INTERPRO](#) в предложенной последовательности выделил два домена, соответствующие доменам А и В (см. табл. 2). Визуализация аналогичного сравнения на рис. 5.

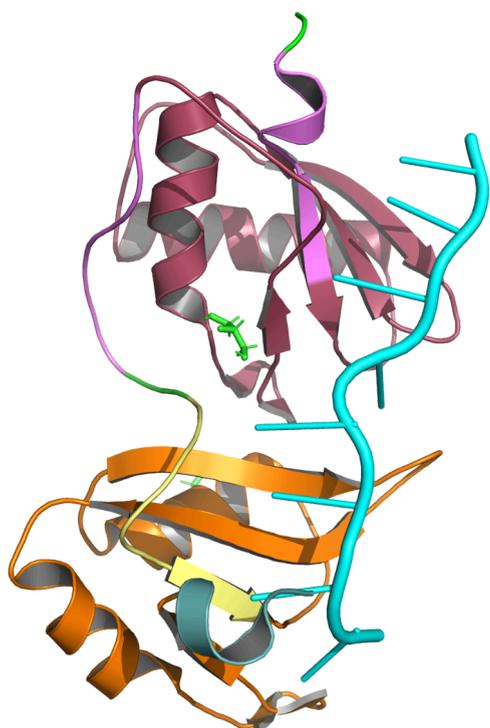


Рисунок 5. Сравнение интерпретации результатов работы ДОМАК с результатом поиска в INTERPRO;

- фиолетовым отмечен фрагмент, вошедший в домен А по версии интерпретации ДОМАК, но не найденный INTERPRO;
- желтым отмечен фрагмент, отнесенный в ДОМАК к домену В, но не относящийся к нему по версии INTERPRO;
- голубым отмечен фрагмент, входящий в домен В по версии INTERPRO, но не отнесенный к нему в интерпретации ДОМАК;
- пересекающиеся части: домен А выделен малиновым, домен В — оранжевым;

Анализ домена А в INTERPRO полностью попал в границы результатов ДОМАК. INTERPRO также исключил из домена А часть протяженного фрагмента на конце белка, как и ДОМАК. Однако, INTERPRO не включил в домен А не только этот фрагмент, но и прилегающие спираль и бета-лист, что представляется нелогичным, учитывая их регулярную структуру. В случае домена В, INTERPRO не включил часть протяженного участка (что согласуется с ДОМАК), а также последовавший за ним короткий бета-лист, но в отличие от ДОМАК, отнес к домену В весь С-концевой фрагмент, содержащий небольшую альфа-спираль.

Таблица 2. Сравнение полученных данных с данными INTERPRO.

Источник данных	Домен А	Домен В
ДОМАК	3-106	110-192
INTERPRO	12-94	122-200

Если опираться на литературные данные, то все программы не справились успешно разметить доменную структуру, однако отличия как правильно по краям доменов, что можно простить 😊

[ссылка на colab](#)

[ссылки на сессии pytoml](#)

- [1] T. Ito, Y. Muto, M. R. Green, and S. Yokoyama, "Solution structures of the first and second RNA-binding domains of human U2 small nuclear ribonucleoprotein particle auxiliary factor (U2AF65)," *EMBO J.*, vol. 18, no. 16, pp. 4523–4534, Aug. 1999, doi: 10.1093/emboj/18.16.4523.