

Разметка вторичной структуры

Знакомство с укладками

Для выполнения задания было получено 10 записей PDB, содержащих структуры, имеющие по последовательностям не более 40% попарного сходства. Было сделано выравнивание их всех по структуре (3ptwA01) (Рис. 1).

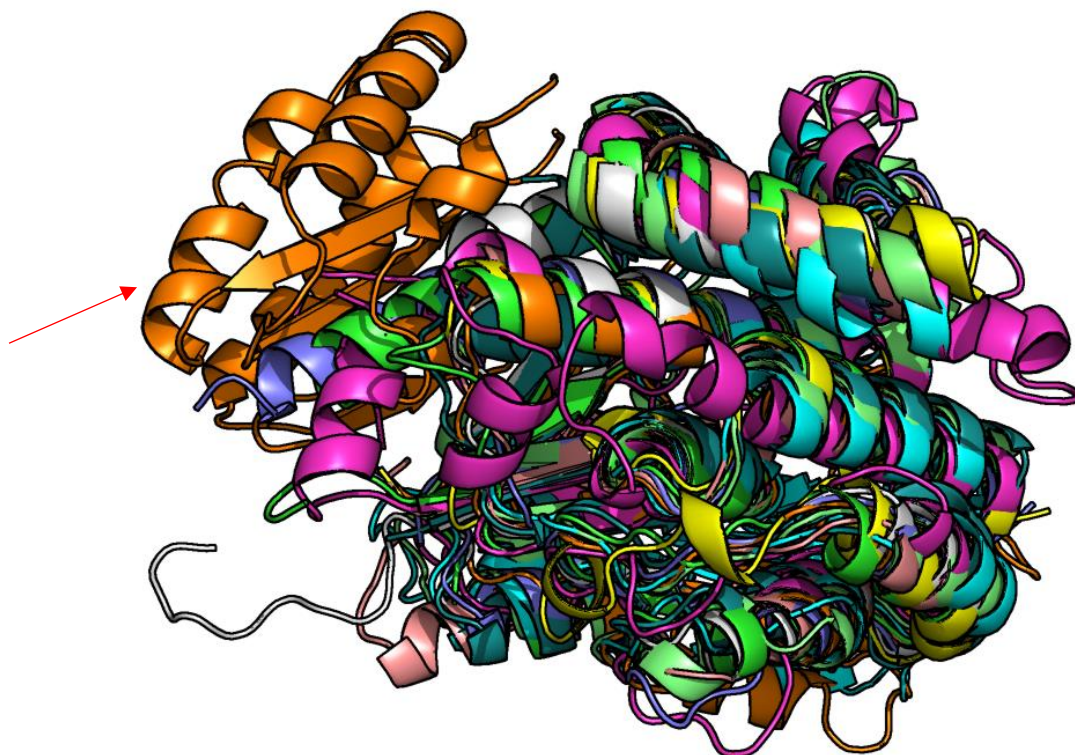


Рисунок 1. Выравнивание всех структур по 3ptwA01 (синий).

В целом можно заметить сходство, однако есть и различия: петли, например, все идут по-разному, а у структуры 4amhA00 есть целая дополнительная часть из 3 альфа-спиралей и бета-листа (оранжевый, обозначена стрелкой), которой нет в других структурах. Также есть различия в укладке альфа-спиралей.

Работа с разметкой вторичной структуры в ручном режиме

Для сравнения был выбран участок 4qbuA02 и 5czcB01 (Рис. 2). В структуре 5czcB01 он представлен бета-листом (242 – 254 ак), тогда как во второй структуре это просто петля

(323-335ак). Этот участок пространственно находится в одном месте, но отличается по последовательности.

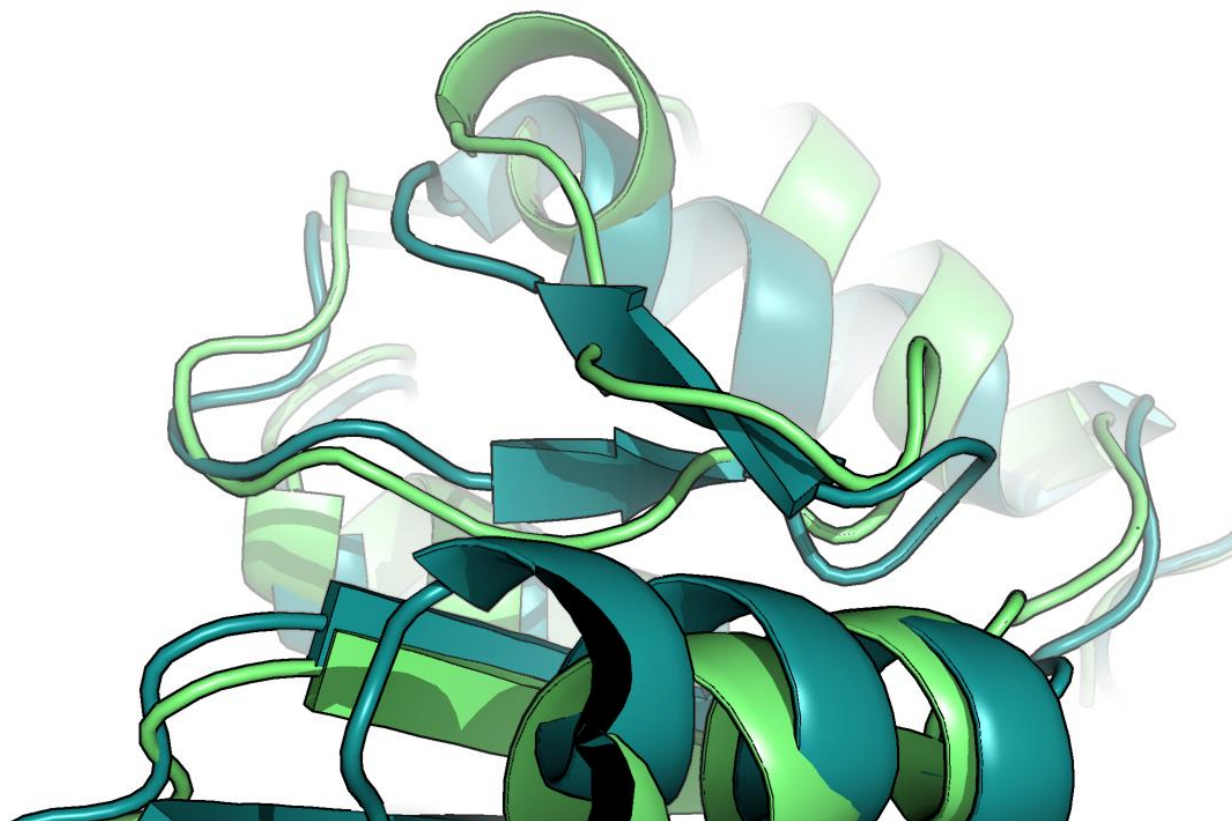


Рисунок 2. Сравнение участка 4qbuA02(зеленый) и 5czcB01(deepsteel).

При рассмотрении водородных связей в этом участке (Рис. 3) можно заметить, что в структуре 4qbuA02 невозможно образование еще одной водородной связи, необходимой для образования бета-листа. Эта водородная связь присутствует в структуре 5czcB01. Кроме того, последовательность аминокислотных остатков также отличается, поэтому расположение этого участка в пространстве не одинаково, что также препятствует образованию бета-листа. Таким образом, я бы сказала, что алгоритм Rm01 сработал верно (интересное замечание, что бета-тяж присутствует у 4 структур из 10).

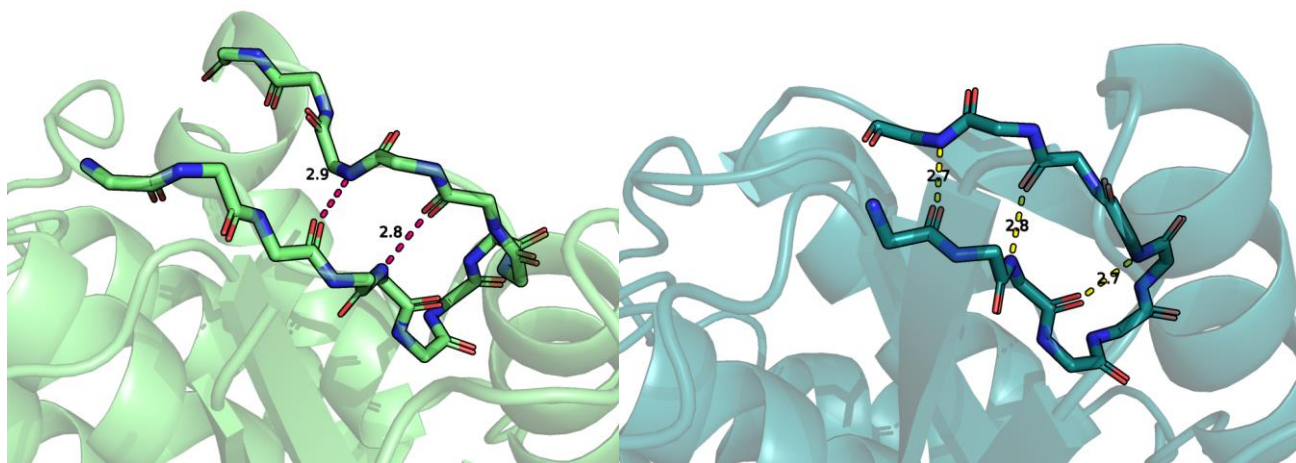


Рисунок 3. Водородные связи участка 4qbuA02(зеленый) и 5czcB01(deersteel).

Работа с разметкой вторичной структуры в автоматическом режиме

Для каждого из 10 pdb была сгенерирована разметка вторичной структуры с помощью dssp и рассчитана склонности каждого типа аминокислоты образовывать тот или иной тип вторичной структуры (amino acid secondary structure propensity):

$$P_{ik} = (n_{ik}/n_i) / (N_k/N)$$

- Где P_{ik} это propensity аминокислотного остатка i образовывать тип вторичной структуры j
- n_{ik} это количество остатков i в датасете, образующих тип вторичной структуры j
- n_i это общее количество остатков i в датасете
- N_k это общее количество остатков, образующих тип вторичной структуры j во всем датасете
- N это общее количество остатков в датасете

Согласно полученным результатам, чаще образуют альфа-спирали, чем бета-листы и петли такие остатки, как L, R и E. Чаще образуют бета-листы - V, I, C и F. Чаще образуют петли - P, S, H, D и N. (Рис. 4).

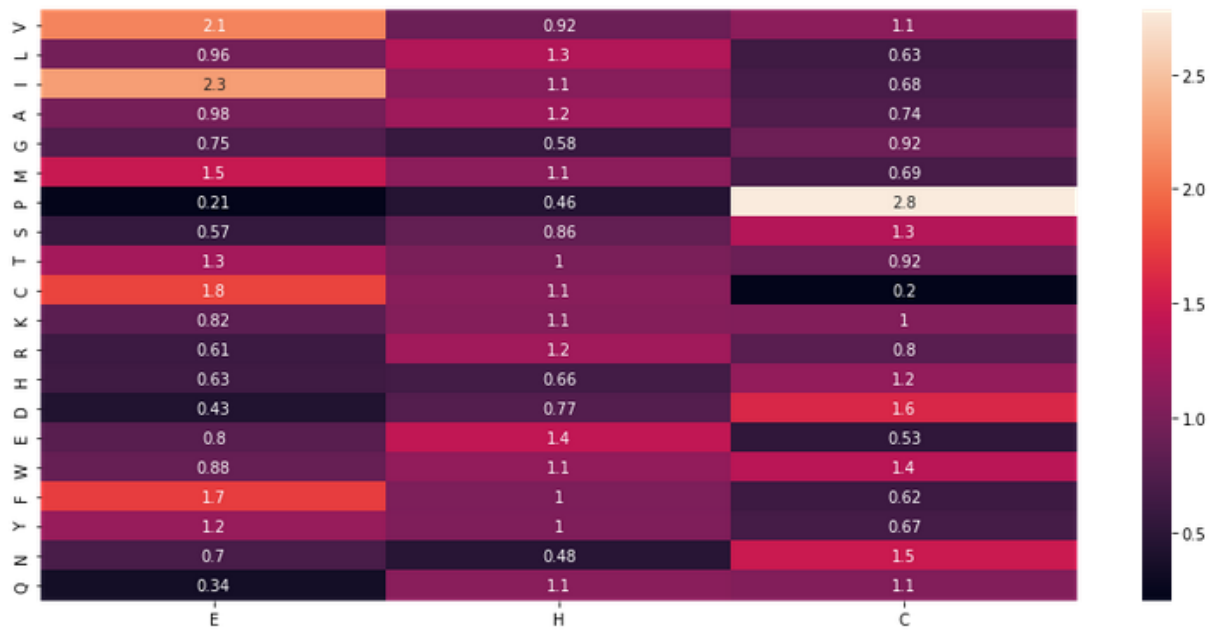


Рисунок 4. Heatmap. Склонность образовывать типы вторичной структуры для различных аминокислот (E — лист, H — спираль, C — петли).