

# Практикум 8

## Задание 1

В этом практикуме я использовала структуру из практикума 6 - 1DMN. Для нее с помощью 2Struc была генерирована аннотация вторичной структуры.

Были выбраны 2 генерации: с помощью DSSP и STRIDE.

В целом, аннотации этих программ похожи, но все же есть некоторые различия в принадлежности таким структурам как, например, альфа-спирали и бета-листы. Давайте их рассмотрим ниже.

1) Смотрим на 108-D, точнее на аспарагиновую кислоту. По выдаче DSSP она не принадлежит бета-листу, а по STRIDE - принадлежит. Посмотрев на структуру подробнее вручную, я увидела, что во-первых, начиная с PHE-107 и до GLN-114 атомы вывернуто (рисунк 2), что нарушило расстояния между ними, соответственно и длины водородных связей, которые должны быть в б-листе. И если 109-111 остатки считаются поворотом по выдаче, но 107,108,112,113 входят в б-лист по DSSP. Поэтому в данном случае я считаю аннотацию STRIDE более верной, так как она говорит, что эти остатки принадлежат чему-то другому, а не бета-листу.

| RESNUM    | 75   | 85 | 95   | 105          | 115 | 125          |              |    |              |          |    |          |     |          |     |    |     |   |
|-----------|--|----|------|--------------|-----|--------------|--------------|----|--------------|----------|----|----------|-----|----------|-----|----|-----|---|
| SEQ       | TGPVRASHNGCGAMPFRVEMVWNGQPCALDVIDVMRFDEHGRIQTMQAYWSEVNL SV |    |      |              |     |              |              |    |              |          |    |          |     |          |     |    |     |   |
| CONSENSUS | i 0X00   | EE | XXXX | EEEEEEEEEEEE | TT  | EEEEEEEEEEEE | X            | TT | XX           | EEEEEEEE | 00 | GG       | GEE | 0        |     |    |     |   |
| DSSP      | i 0  | S  | 00   | EE           | 0   | SS           | EEEEEEEEEEEE | TT | EEEEEEEEEEEE | 0        | TT | S        | 0   | EEEEEEEE | 00  | GG | GEE | 0 |
| STRIDE    | i  | C  | CCC  | E            | ET  | TTTT         | EEEEEEEEEEEE | TT | EEEEEEEEEEEE | TT       | TT | EEEEEEEE | CC  | GG       | GEE | C  |     |   |

Рис.1 Участок выдачи 2Struc.

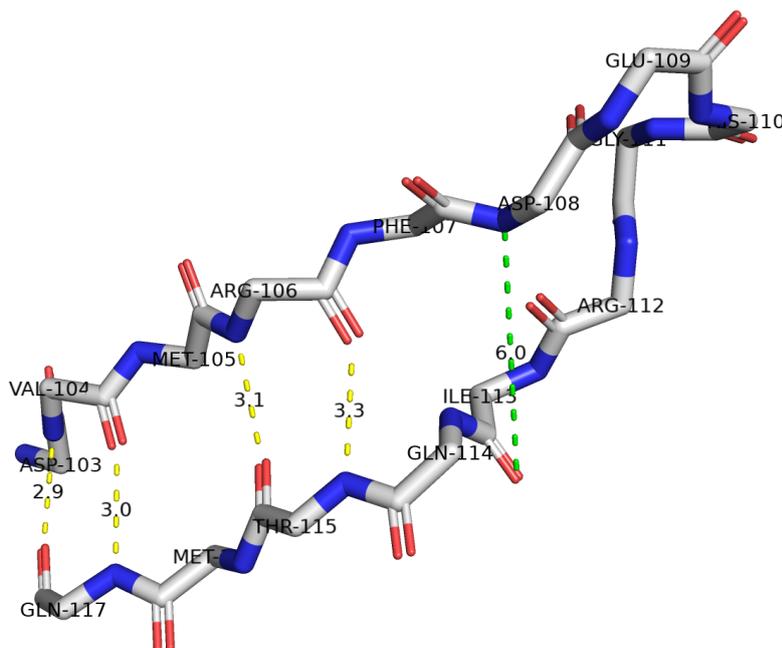


Рис. 2. Бета-лист, предпологаемо содержащий ASP-108 и ARG-112. (желтым- «хорошие» водородные связи, зеленым- нарушенная поворотом кислорода в другую сторону)

2) Аналогично для 112-R. Из-за того, что ILE-115 развернуло, то это мешает, и аргинин тоже не может образовать связи с нужным остатком (рисунки 2).

### 3) Альфа спираль и остаток 49-G

DSSP относит его к спирали, STRIDE не относит. Я может быть тоже отнесла бы остаток глицина к альфа-спирали, так как длина водородной связи соответствует допустимым значениям, а также на каждый виток приходится 3,6 аминокислотного остатка, шаг винта (т.е. минимальное расстояние между двумя эквивалентными точками) между GLY-49 и ILE-53 составляет 5.5Å. Но, углы  $\phi$  и  $\psi$  сильно не соответствуют значениям торсионных углов альфа спирали (примерно  $-60^\circ$ ,  $-45^\circ$ ) (рисунки 4), поэтому я склоняюсь к тому, что этот остаток все же не входит в спираль, и соглашаюсь с STRIDE.

| RESNUM    | 22   | 32        | 42          | 52       | 65             |
|-----------|--|-----------|-------------|----------|----------------|
| SEQ       | ARYIELVDVGDIEAIVQMFADDATVEDPFGQPPIHGREQIAAFFRQGLKVRACLTP |           |             |          |                |
| CONSENSUS | i HHHHHHHHXO   | HHHHHHXXE | EEEEEEEXXTT | XOOEEX   | HHHHHHHHHHHXO  |
| DSSP      | i HHHHHHHHTO   | HHHHHHTTE | EEEEEEESS   | TTTSOOEE | HHHHHHHHHHHOS  |
| STRIDE    | i HHHHHHHHHC   | SNHHHHHHC | EEEEEEETT   | TTTTCC   | EECHHHHHHHHHHC |

Рис. 3. Участок выдачи 2Struc.

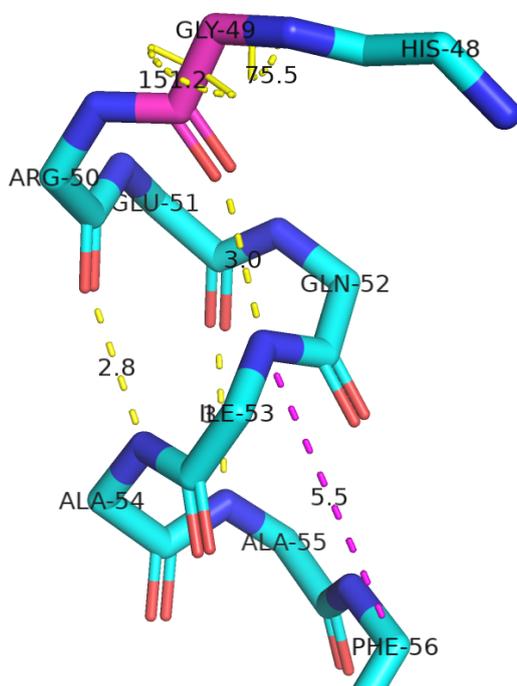


Рис. 4. Участок альфа-спирали. Розовым-рассматриваемый остаток GLY-49. Желтый пунктир- водородные связи, розовый- расстояние между витками. Также желтым- торсионные углы.

## Задание 2

Мне был выдан набор из 30 белков и для каждого из них я сгенерировала разметку вторичной структуры с помощью DSSP.

С помощью python были рассчитаны склонности каждого типа аминокислоты образовывать тот или иной тип вторичной структуры (amino acid secondary structure propensity) по следующей формуле:

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

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

Таб. 1. Склонности к образованию типов структур (бета-лист ('E') , альфа-спираль ('H'), петля ('C'))

|   | Н        | Е        | С        |
|---|----------|----------|----------|
| A | 1.442494 | 0.843170 | 0.792581 |
| C | 0.754509 | 1.699654 | 0.838963 |
| D | 0.733869 | 0.584790 | 1.354388 |
| E | 1.372226 | 0.608190 | 0.942646 |
| F | 0.945210 | 1.291737 | 0.902953 |
| G | 0.499007 | 0.747414 | 1.428702 |
| H | 0.708907 | 1.061802 | 1.155094 |
| I | 1.020506 | 1.876615 | 0.592034 |
| K | 1.191471 | 0.530270 | 1.091364 |
| L | 1.284826 | 1.223073 | 0.720456 |
| M | 1.331886 | 0.898704 | 0.837068 |
| N | 0.882196 | 0.520480 | 1.290151 |
| P | 0.313577 | 0.284235 | 1.753990 |
| Q | 1.300346 | 0.693335 | 0.949447 |
| R | 1.292873 | 0.954485 | 0.836447 |
| S | 0.776163 | 0.780720 | 1.239504 |
| T | 0.840398 | 0.930110 | 1.131805 |
| V | 0.847792 | 1.966134 | 0.660237 |
| W | 1.526877 | 1.064618 | 0.639744 |
| Y | 0.947764 | 1.546094 | 0.786712 |

Рассмотрим несколько случаев на примерах.

В спиральных чаще всех встречается аланин, в бета-листах лидер по встречаемости валин. Это можно объяснить тем, что наличие неполярных боковых групп, видимо, характерно для бета-структур, в которых в целом вообще есть гидрофобные взаимодействия. Петли чаще всего образует глицин. Глицин действительно может быть склонен к образованию петель, ведь у него есть водородный радикал, а его можно назвать достаточно «гибким».