

Практикум 8

Вторичная структура

Задание 1. Работа с разметкой вторичной структуры в ручном режиме

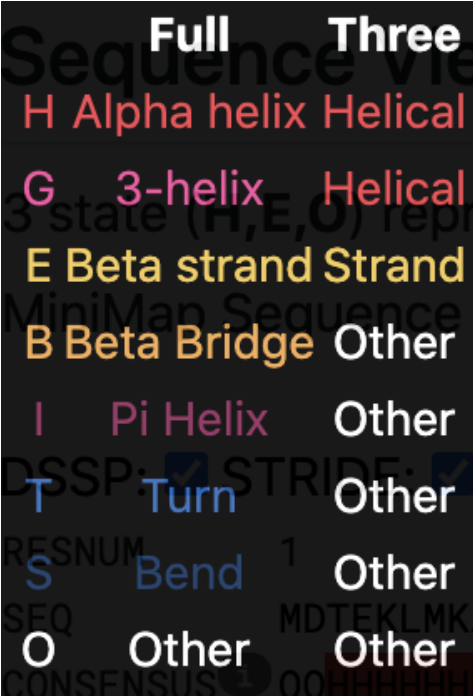
Мой PDB ID: 1XGO (из практикума 6 “Валидация”). Модель белка состоит из одной цепи.

При помощи инструмента 2Struc генерируем аннотацию вторичной структуры.

```
RESNUM      1      11      21      31      41      51      61      71      81      91     101     111     121     131     141
SEQ         MDTEKLMKAGEIAKKVREKAIKLARPGMLLLELAESIEKMIMELGGKPAPFVNLSEINEIAAHYTRYKGDITVLKEGDYLIKIDVGVHIDGFIADTAVTVRVGMEEDELMEEAAKEALNAAISVARAGVEIKELGKAIENEIRKRG
CONSENSUS  ooXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
DSSP        ooXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
STRIDE      cXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

```
141      151      161      171      181      191      201      211      221      231      241      251      261      271      281      291
FKPIVNLSGHKIERYKLHAGISIPNIYRPHDNYVLKEGDVFAIEPFATIGAGQVIEVPPTLIYMYRDPVVRVAQARFLLAKIKREYGTLPFAYRWLQNDMPGQLKLALKTLEKAGAIYGYVPLKEIRNGIVAQFHTIIVEKDSVIVTTE
OXXXXXXXXXOXXXXXXXXXXXXXXXOXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
OEEooooeeovtvtoooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtooooooTtoooooo
ccsttcEEEEETEEEEEEECSCCTTTEETTTEEEESCCECCTTEEEESCSCCCEEECCCCCXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

Выдача 2Struc. Картинка разнесена на две строки для удобства.



Расшифровка обозначений в выдаче

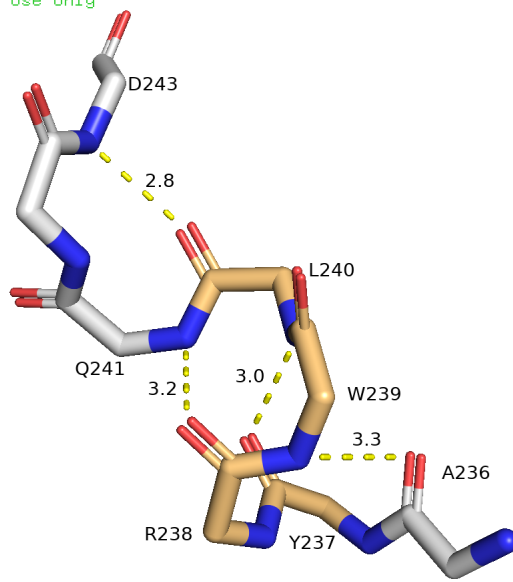
Участок 1

Остатки 235-240



Для остатка 237 (L) DSSP детектировал излом (изменение направления более чем на 70 градусов) и три последовательных поворота 238,239,240 - участка, не составляющих спираль достаточной длины. Алгоритм STRIDE же обозначил этот участок (237-240) как 3₁₀-спираль.

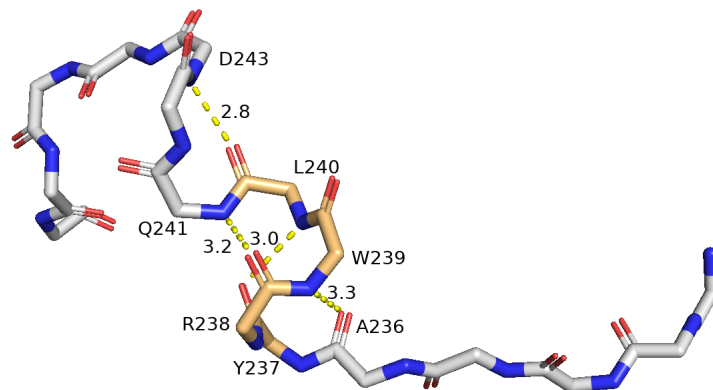
For Educational Use Only



Остов окрестности рассматриваемого участка. Светло-оранжевый - сам рассматриваемый участок (237-240).

Определенная структура очень похожа на 3_10 спираль, определенную алгоритмом STRIDE. Косвенными доказательствами тому может служить то, что она довольно не протяженная, что характерно для 3_10 спиралей, при этом она находится в месте изгиба:

For Educational Use Only



Расширенное окружение рассматриваемого участка.

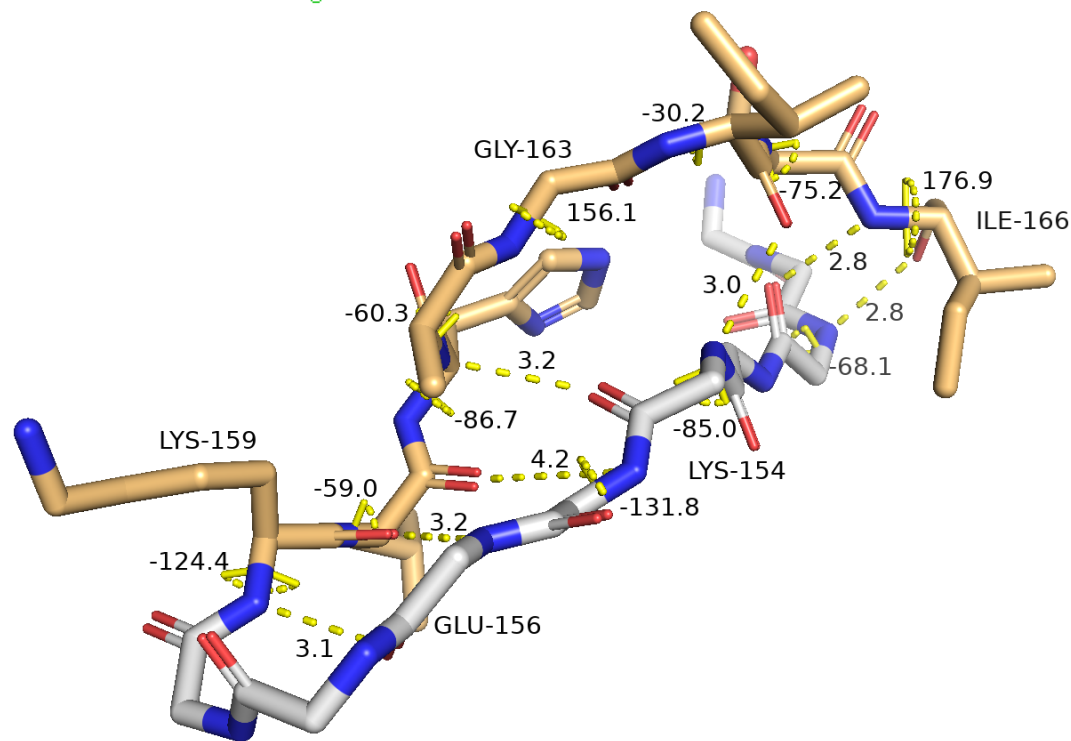
Участок 2

Остатки 159-166

```
161
KLHAGISI
XXXXXXXXX
BTTSOOO
EEEEEEEE
```

Алгоритм STRIDE находит на этом участке бета-лист. Алгоритм DSSP находит для 159Lys бета-мостик, однако затем распознает два поворота (участки, не составляющие спирали достаточной длины) и излом (изменение направления более чем на 70 градусов).

Визуализируем. Попытаемся найти участок последовательности, могущий образовывать бета-лист с данным.



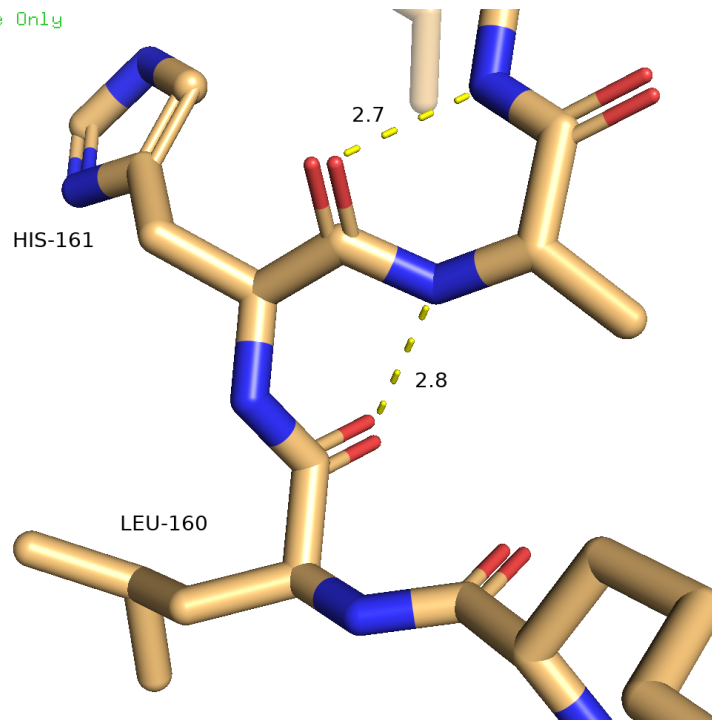
Светло-оранжевый - рассматриваемый участок, серый - остов участка, предположительно образующий бета-лист с рассматриваемым

Азот остова 159Lys действительно, по-видимости, формирует водородную связь с кислородом остова 156Glu с образованием структуры типа бета-поворота. Поэтому классификация остатка 159Lys и как бета-мостика, и как компонента бета-слоя, имеет место быть.

Для последующих остатков рассматриваемого участка видно, что несмотря на наличие остова на достаточно близком расстоянии (подписанные расстояния близки к пороговым, однако вполне могут свидетельствовать о наличии водородной связи), антипараллельный бета-лист вряд ли формируется, потому что

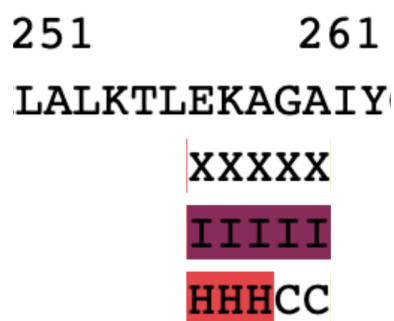
- есть много пропусков связей (т.н. лестницы часто не формируется)
- торсионные углы не похожи на те, что мы ожидаем наблюдать в бета-листе

Вариант, предложенный алгоритмом DSSP более реалистичен. Действительно, очень похоже, что 160Leu и 161His формируют повороты - водородные связи с азотом остова остатка $i+2$.



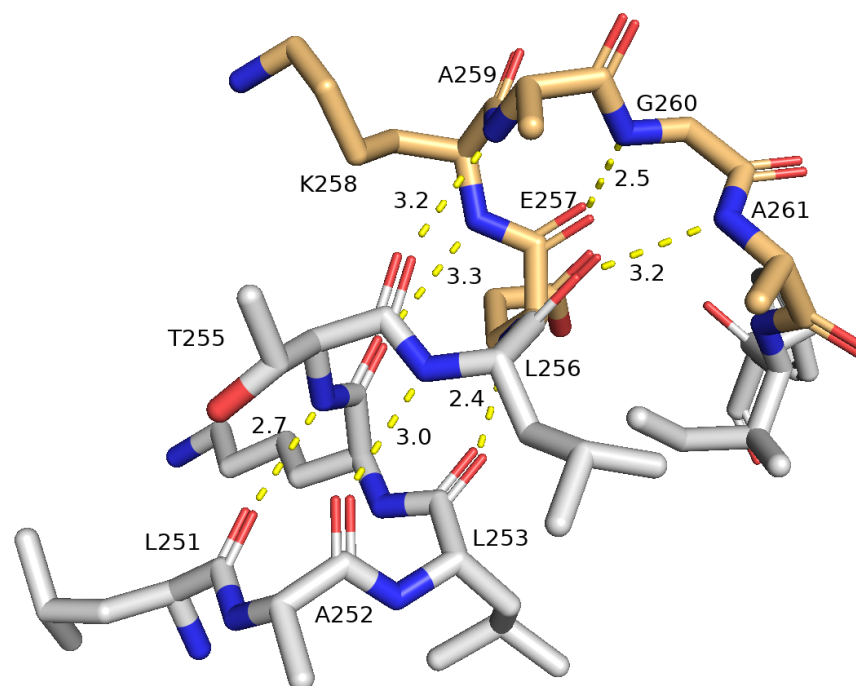
Участок 3

Остатки 257-261



Рассматриваемому участку предшествует протяженный участок альфа-спирали, однозначно определенный обоими алгоритмами и содержащийся в консенсусной аннотации. После рассматриваемого участка находится последовательность из остатков, включенных в бета-слой (так же однозначно как и предыдущая альфа-спираль). То есть, мы имеем дело с переходным участком.

For Educational Use Only



Светло-оранжевый - рассматриваемый участок

Наблюдения подтверждают, что предшествующий участок - альфа-спираль, атомы остова связаны водородными связями с атомами остова $i+4$ остатка. Рассматриваемый участок комплексный. Разберем по остаткам.

- 257E связан водородной связью с 260G ($i + 3$)
- 258K связан водородной связью с 254K ($i + 4$)
- 259A связан водородной связью с T255 ($i + 4$)
- 260G связан водородной связью с 257E ($i + 3$)
- 261A связан водородной связью с 256L ($i + 5$)

Сложно говорить о какой-либо четкой закономерности, дающей основание включить рассматриваемый участок в состав альфа или ρ -спирали. Скорее, из-за связи 257E-260G я бы присвоила этим остаткам 3-поворот, а 258K и 259A сочла бы продолжением предшествующей этому участку альфа-спирали.

STRIDE мог охарактеризовать этот участок как альфа-спираль целиком из-за предшествующего протяженного участка альфа-спирали, которую он продолжил. DSSP мог охарактеризовать этот участок как ρ -спираль из-за наличия $i+5$ связи 261A-256L. Почему он решил присвоить это обозначение всему участку я не знаю.

Задание 2

Имитируем то, как рассчитываются склонности аминокислотных остатков образовывать те или иные типы вторичной структуры, на выборке из 30 моделей белков. Считаем, что склонность аминокислотного остатка i образовывать вторичную структуру j является интегральным свойством этого остатка и не зависит от того, в каком белке он находится.

Генерируем разметки вторичных структур для данных 30 моделей с помощью dssp (v3.0.0). Получившийся файл переводим в более удобный для анализа формат и расцениваем все, что не является альфа-спиралью или бета-листом, как петлю.

Получившиеся разметки сливаем в один датасет из предположения выше.

Получившаяся таблица значений склонности аминокислотного остатка образовывать определенную структуру выглядит так:

	Бета-лист	Альфа-спираль	Петля
S	0.77	0.83	1.21
C	1.76	0.40	0.92
Y	1.49	0.92	0.78
F	1.32	1.10	0.78
H	1.07	0.74	1.10
P	0.51	0.46	1.54
E	0.61	1.52	0.92
G	0.63	0.34	1.55
K	0.80	1.09	1.06
A	0.74	1.43	0.91
I	1.59	1.12	0.62
W	1.23	1.06	0.85
N	0.66	0.57	1.41
L	1.11	1.27	0.80
R	0.86	1.36	0.88
V	2.00	0.81	0.58
Q	0.80	1.39	0.89
D	0.61	0.90	1.26
T	1.12	0.82	1.03
M	1.25	1.40	0.66

В таблице приведены значения $P_{ik} = (n_{ik}/n_i) / (N_k/N)$

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 где P_{ik} это propensity аминокислотного остатка i образовывать тип вторичной структуры j

Наиболее склонными образовывать альфа-спираль оказались остатки глутамата, аланина, метионина, глутамина и аргинина. Наиболее склонными образовывать бета-лист оказались остатки валина, цистеина, тирозина, изолейцина и фенилаланина. Все эти аминокислоты не заряжены и кроме тирозина неполярны. Для заряженных аминокислот D E K R H показатели propensity для бета-листов сравнительно низкие.

В петлях наиболее часто оказываются предпочтительны пролин и глицин, которые к тому же довольно не предпочтительны в бета-листах и альфа-спиралях. Конформация пролина вызывает нарушение хода альфа-спирали и бета-листов. Глицин же может вызывать дестабилизацию указанных выше вторичных структур из-за повышенной гибкости.