

Практикум № 8. Вторичная структура

PDB ID - 1n8z.

1 Работа с разметкой вторичной структуры в ручном режиме

Выбранные области

№	1	2	3
Chain	A	A	B
residues	149-153	121-129	170-177
dssp	OBTTB	HHHHTTT	EEEEOOEEE
stride	CGGGC	HHHHHCC	EEEEEEEE
PSEA	bbccb	aaaaaac	bbbcbbbb
sticks	BBbOO	OOOOOOO	bBBBBBBBB
pic	<p>151</p> <pre> WKVDNALQS :EXXXXXOXX :EOBTTBOOO :ECGGCCCTT)bbbccbbb)BBBbO0000 </pre>	<p>121 131</p> <pre> =PPSDEQLKSGTASVV :XOOHHHXXOXEEEE :O00HHHHTTTEEEEE :CCCHHHHCCEEEE)bccaaaaacbbbbb)b000000000bBBB </pre>	<p>171</p> <pre>)VHTFPAVLQS <EEEXXXEEXX :EEEE000EEEO :EEEEEEETT)bbbcbbbbc)BBBBBBBBbC </pre>

1.1 3_{10} Спираль

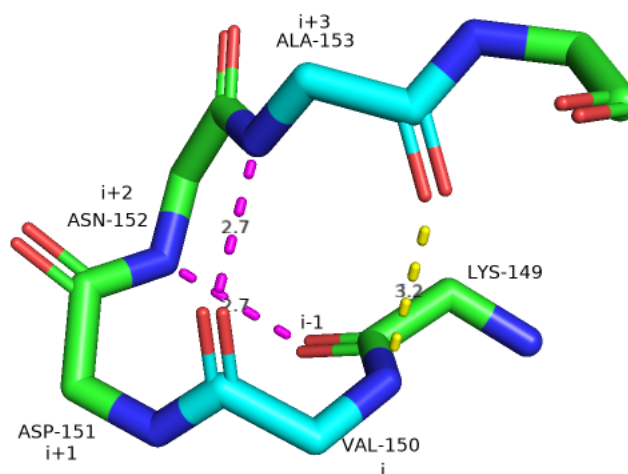


Рис. 1: Остатки 149-153

На рисунке я разметила все водородные связи, которые, по моему мнению, есть в данном участке. Stride размечает остатки 150-152 как 3_{10} спираль. Правда, в этом, безусловно, есть : между кислородом остова LYS-149 и азотом остова ASN-152 ($i-1 \rightarrow i+2$) а также между VAL-150'O и ALA-153'N ($i \rightarrow i+3$) есть водородные связи, что позволяет сказать, что остатки с $i-1$ по $i+3$ образуют спираль (краевые остатки не считаются).

Однако такая разметка никак не отражает водородной связи, которая образуется между кислородом ALA-153 и азотом VAL-150 (помечена желтым на рисунке 1). Две водородные связи этих двух остатков составляют как бы бэга-мостик (остатки покрашены голубым), что и отражено в разметке dssp. Оставшаяся связь в спирали названа T - поворотом, поддержанным водородной связью.

Итого, я бы сказала, что верны разметки Stride и DSSP, но не отражают в полной мере все связи - потому что тут есть одновременно и бэга-мостик, и 3_{10} спираль. Проблема в том, что один из остатков - VAL-150 - и одна из его связей - входит в состав обеих структур, поэтому присваивание однозначной метки тут и в принципе невозможно.

Рассмотрим остальные два алгоритма. PSEA показывает чтото похожее на выдачу DSSP - два остатка (151 и 152) обрамлены бэга структурой. Однако он продолжает бэга лист и далее - остатки 153-156 входят в бэга

лист.



Рис. 2: Остатки 153-156 покрашены оранжевым.

Однако эта область (остатки 154-156) находится на краю белка, и ни с кем не образует на самом деле бэга листа. Этот алгоритм анализирует углы при Ca атомах. Поэтому вполне ожидаемо что иногда остатки, которые могли бы образовывать бэга-лист и имеют подходящие углы, ни в какие взаимодействия не вступают.

Разметка sticks, включила остаток 151 в протяженный бэга лист. Посмотрим, почему.

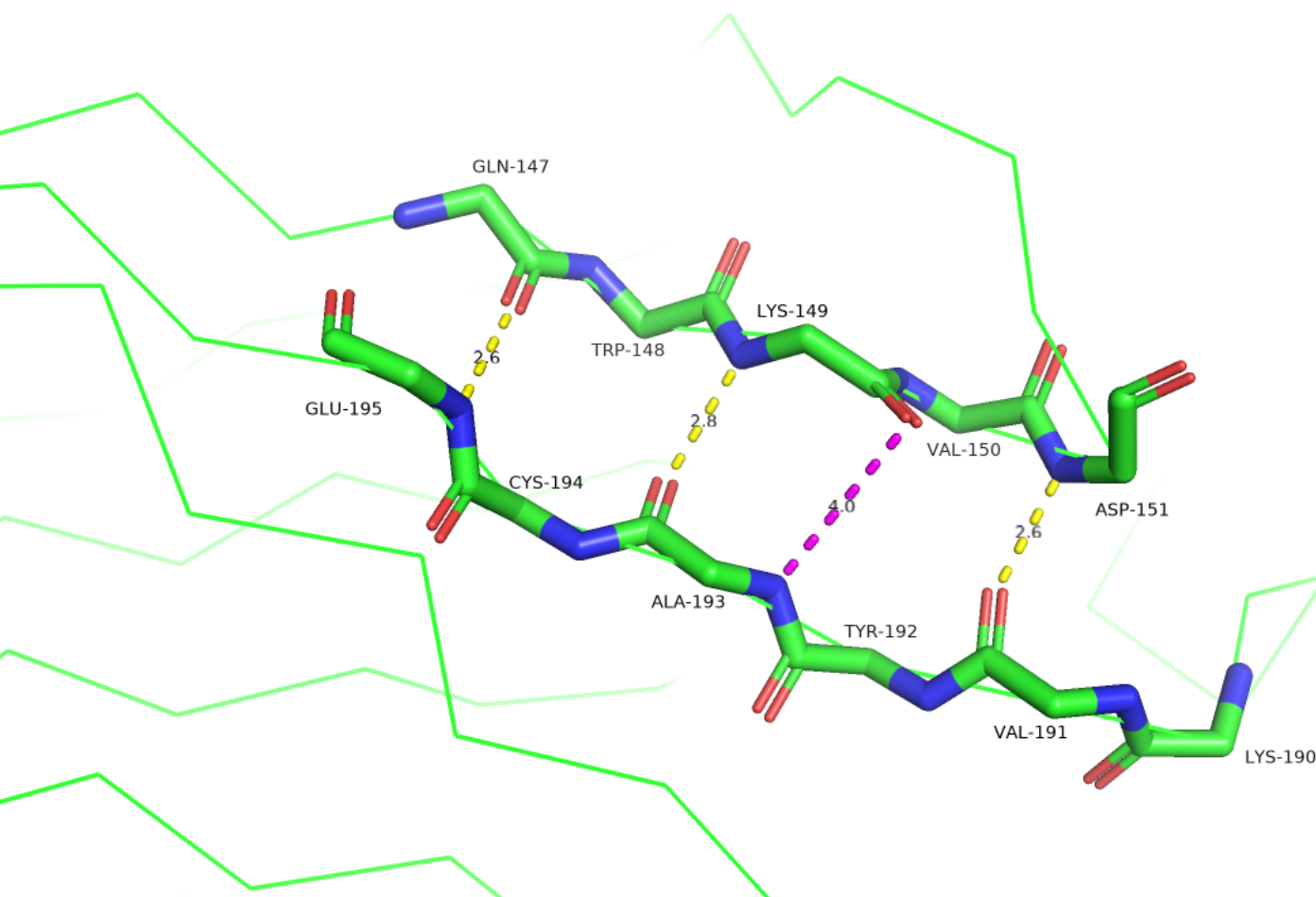


Рис. 3: Бэ́та-ли́ст по версии sticks

Видно, что такая разметка позволяет учесть водородную связь между ASP-151 и VAL-191. Однако точно нет водородной связи, которая показана ярко-розовым между LYS-149 и ALA-193. Так что такая разметка тоже не совсем верна.

В общем, в этом случае все неправы. Думаю, Dssp в данном случае наиболее полно отразил все водородные связи, stride выявил необычную структуру, sticks предсказал неверную водородную связь на краю бэ́та-листа, а PSEA не смог понять, что имеет дело со структурой на краю белка. Предсказание DSSP думаю тут самое логичное.

1.2 α -спираль или поворот?

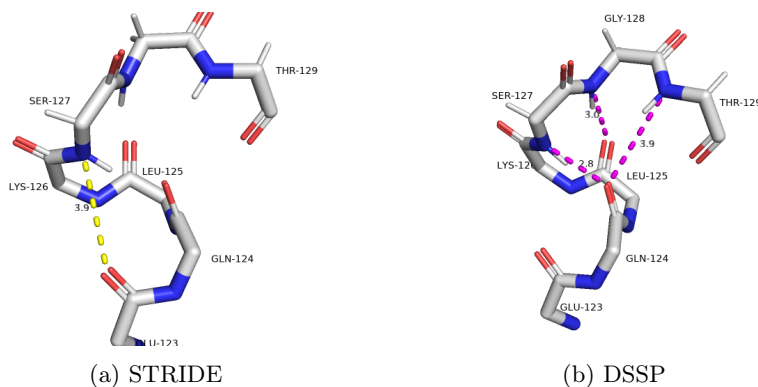


Рис. 4: Конец альфа спирали - остатки 124-129 цепи A

Данную область Stride размечает как альфа спираль, тянущуюся со 122 по 126 остаток, то есть последний остаток с водородной связью - SER-127. Эта водородная связь отображена на рисунке 2a желтым цветом. Однако даже глазами видно, что серин развернут так, что водород от него смотрит совсем на другой кислород - на кислород GLN-124. Для следующих остатков stride ничего не предсказывает.

Я предполагаю, что связи такие, как отображены на рисунке 2b розовым цветом. То есть SER-127 не относится к предыдущей альфа-спирали, а заключает водородную связь с GLN-124. Также, как мне кажется, есть водородная связь между GLY-128 LUE-125. То есть я думаю, что эту область вполне можно назвать поворотом с поддержкой водородными связями, как это размечает DSSP. Либо даже эти остатки (124-128) снова можно назвать 3_{10} спиралью.

PSEA предсказывает длинную альфа-спираль до 128 остатка. То есть водородные связи в таком случае должны были бы идти так:

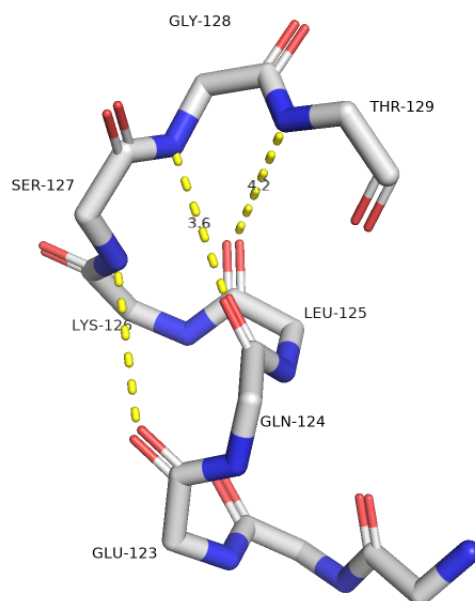


Рис. 5: Водородные связи по предсказанию PSEA

С одной стороны, можно было бы сказать, что расстояния (4.2, 3.6, 3.9) все еще удовлетворяют длинам водородных связей. Однако как видно на рисунке 4а, водороды, скорее всего повернуты не так, как нужно для образования этих трех связей. Интересно, что по углам Са атомов этого нельзя предсказать - все таки спираль оказывается немного сверх закрученной на своем конце.

sticks вообще не предсказывает тут альфа-спирали, а зря, все-таки остатки 121-126 ее образуют.

Итого, в данном случае, разметка DSSP кажется более разумной.

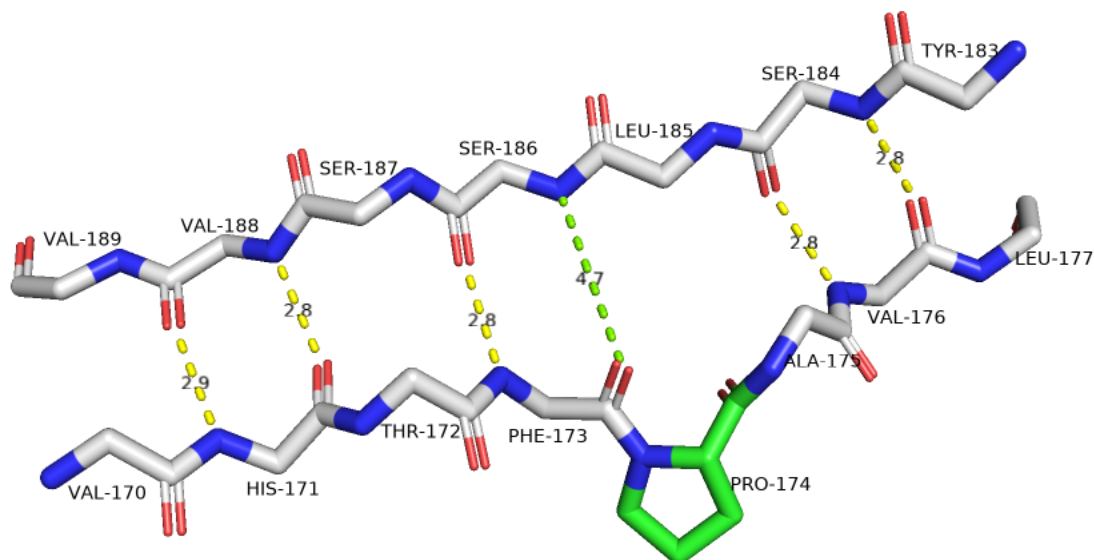
1.3 β слой

Рис. 6: Край бэ́та-листа - цепь В

Тут разметка STRIDE говорит, что остатки 170-177 целиком входят в бэ́та-лист. Разметка DSSP пропускает остатки 173-175. И правильно делает - потому что 174ый остаток - это пролин! Из-за того, что на этом месте стоит пролин, PHE-173 оказывается вывернут и вторая водородная связь с SER-186 (нужная для образования формального бэ́та-мостика) не может образоваться (показана зеленым на рисунке).

Разметка Stride никак не отображает этой особенности, в то время как dssp - да! Хотя в dssp тоже не отражен тот факт, что фенилаланин входит в бэ́та-слой, но как бы только одной связью.

Разметка rpea тут кажется ближе всего - она показывает протяженный бэ́та-слой, в который посередине не входит 174 пролин. Конечно, тут не отражено то, что вторая связь фенилаланина не образует водородной связи, но зато видно особенность - некоторое выпетливание из одного остатка посреди протяженной структуры. Видимо по Ca атомам это можно хорошо предсказать.

sticks, также, как и stride предсказывает всюду бэта лист.

Итого, это опять спорный случай, где однозначно присвоить метку нельзя. Однако DSSP и PSEA чуть лучше отразили то, что происходит что-то необычное, в то время как stride и sticks тяготеют к тому, чтобы как можно больше остатков засунуть в привычную структуру.

2 Работа с разметкой вторичной структуры в автоматическом режиме

	aa	Hpro	Epro	Spro
0	Q	1.11	0.80	0.98
1	Y	1.06	1.57	0.77
2	A	1.40	0.82	0.75
3	P	0.46	0.47	1.59
4	S	0.76	0.58	1.32
5	I	1.11	1.81	0.65
6	K	1.13	0.73	0.99
7	F	1.00	1.39	0.87
8	D	0.74	0.63	1.32
9	T	0.82	1.18	1.08
10	L	1.32	0.85	0.80
11	G	0.52	0.59	1.51
12	M	1.42	1.12	0.63
13	V	0.89	2.10	0.73
14	E	1.21	0.71	0.93
15	C	0.70	1.34	1.12
16	N	0.76	0.62	1.31
17	W	0.98	1.29	0.92
18	R	1.19	1.14	0.81
19	H	1.02	1.29	0.89

Рис. 7: Propensity of aminoacids. Hpro - альфа-спирали, Epro - бэта-листы, Spro - петли.

В этом задании была посчитана склонность аминокислоты образовывать ту или иную структуру, по формуле:

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

Где

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

n_{ik} - количество остатков i в датасете, образующих тип вторичной структуры j

n_i - общее количество остатков i в датасете

N_k - общее количество остатков, образующих тип вторичной структуры j во всем датасете

N - это общее количество остатков в датасете (описание формул я привожу для себя)

В моей выборке оказалось, что к образованию альфа-спиралей тяготеют аланин, глутамин и лейцин. Я думаю тут нельзя никакого особенного вывода сделать - в альфа спиральных бывает мало пролинов - это видно в Pr_0 для пролина, которое равно 0.46, а так никаких ограничений на аминокислоты в альфа-спиральных быть не должно - спираль образуется только остовом, альфа спирали бывают и глубоко внутри и снаружи белка.

В бета-листах у меня получились тяготеющие остатки - валин, изолейцин, фенилаланин. Их объединяет то, что они гидрофобны. Думаю это означает что в мою выборку попали структуры с бета-листами внутри - то есть бета-листы редко бывают экспонированы наружу, и скорее образуют структурированное гидрофобное ядро. Я правда не уверена, что это может быть верно вообще для всех структур.

Петли в моей выборке чаще образуются пролином (наверное потому, что в альфа-спираль и бета-лист ему залезть тяжело), а также серин, глутамат и аспарат - то есть заряженные или способные образовывать водородные связи своими боковыми цепями остатки. Думаю, что это логично - наверное, петли чаще оказываются экспонированы в водное окружение (если белок растворимый), так что чаще в петлях не будет чего-нибудь гидрофобного.