

# Практикум 6

PDB ID: 1XGO

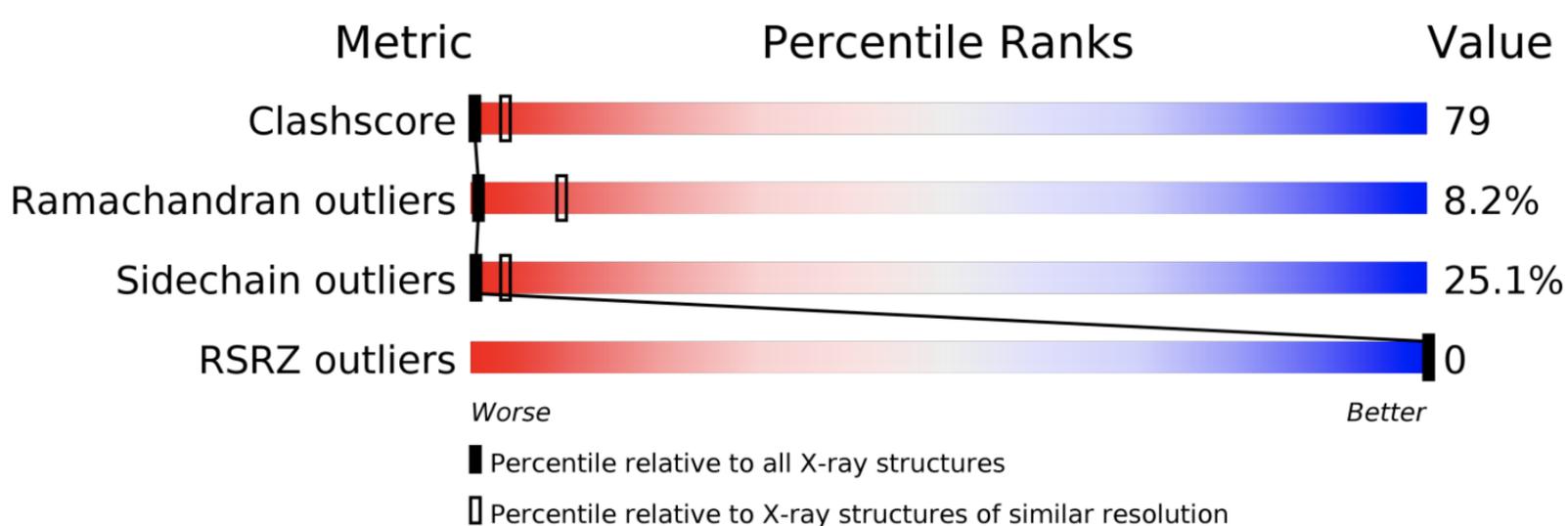
## Задание 1

В практикуме предлагается работать с метионин-аминопептидазой археи *Pyrococcus furiosus*.

Публикация:

Tahir H Tahirov, Hideyuki Oki, et al, **Crystal structure of methionine aminopeptidase from hyperthermophile, *Pyrococcus furiosus***. *Journal of Molecular Biology*, Volume 284, Issue 1, 1998

Разрешение структуры: 3.50 Å (полнота 80.3%). Качество не самое хорошее. Это во многом может объясняться годом получения структуры.



Отчет, сгенерированный автоматической системой проверки wwPDB OneDep

Структура содержит 2312 атома и 295 аминокислот, причем ни для одного атома не было определено альтернативной конформации и в модели нет пропущенных атомов.

R-фактор: 0.182, R-free: 0.274. "Хорошие" значения R-free общепризнанно лежат ниже 0.25, из чего делаем вывод о том, что модель была уточнена не очень правильно. В это время обычный R-фактор принимает допустимое значение. Значит, скорее всего, мы слишком хорошо подогнали данные под эксперимент и отделились от истинной модели.

Clashscore (число перекрываний Ван-дер-Ваальсовых радиусов на 1000 атомов) для этой структуры составляет 79. Это хуже, чем clashscore большинства структур на PDB, более того, хуже, чем большинство структур того же разрешения. Это число перекрываний посчитано после эмпирического добавления водородов с помощью MolProbity и, судя по списку перекрываний, они происходят в подавляющем большинстве случаев из-за водородов.

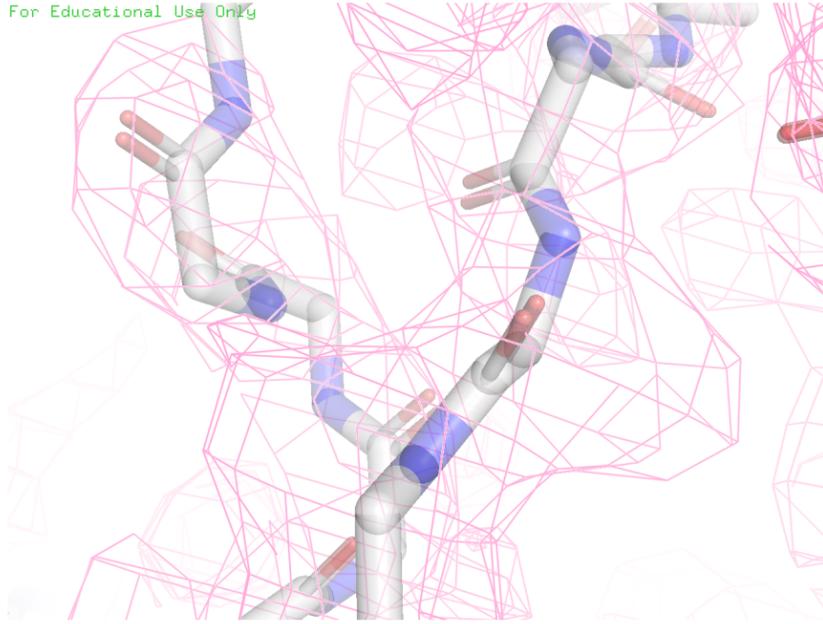
Было найдено 24 маргинала по картам Рамачандрана из 293 (99%) проверенных остатков что составляет 8% от общего числа остатков. Это высокое количество сравнительно с множеством структур на PDB, в том числе с подмножеством структур того же разрешения.

Был найден 61 маргинал по ротамерам боковых радикалов из 243 проверенных остатков что составляет 25% от общего числа остатков. Из отчета PDB я посчитала, что самые популярные остатки из них это лейцин (11) и изолейцин (10).

Согласно отчету, маргиналов по RSRZ в модели нет. Это означает что все остатки в модели хорошо соотносятся в восстановленной из эксперимента электронной плотностью. Действительно, согласно отчету PDB остатков без подтверждающей их электронной плотности в модели нет.

Однако, картина ЭП в приближении выглядит так: (близко к центру белка, не маргинальные остатки). Из-за низкого разрешения реальное положение атомов невозможно восстановить из электронной плотности.

For Educational Use Only



Подрезка = 1.5, carve = 2

## Задание 2

Рассмотрим подробнее маргинальные остатки. Из 24 маргиналов по картам Рамачандрана и 61 маргинала по ротамерам боковых радикалов 4 остатка пересекаются: Thr71, Gln218, Val276 и Lys287.

Проанализируем структуру с помощью сторонних сервисов по валидации. Начнем с сервиса [MolProbity](#).

Проанализируем геометрию без учета всех межатомных контактов (для этого потребуется только исходный pdb) и с их учетом (преварительно добавив водороды с помощью того же сервиса). В процессе добавления H сервис также проводит проверку на инверсии His/Asn/Gln и выдает список позиций, требующих разворота. В случае с рассматриваемой структурой это His86, Asn175 и Asn242. Двое последних были отмечены в отчете PDB как маргиналы по ротамерам боковых радикалов.

Protein Geometry	Poor rotamers	66	27.16%	Goal: <0.3%
	Favored rotamers	133	54.73%	Goal: >98%
	Ramachandran outliers	24	8.19%	Goal: <0.05%
	Ramachandran favored	204	69.62%	Goal: >98%
	Rama distribution Z-score	-6.45 ± 0.33		Goal: abs(Z score) < 2
	Cβ deviations >0.25Å	1	0.37%	Goal: 0
	Bad bonds:	0 / 2353	0.00%	Goal: 0%
	Bad angles:	11 / 3178	0.35%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	1 / 14	7.14%	Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	18	6.2%	Goal: <1.0%
	CA Geometry outliers	6	2.06%	Goal: <0.5%
Additional validations	Chiral volume outliers	0/365		

Выдача без учета всех межатомных контактов.

All-Atom Contacts	Clashscore, all atoms:	78.83		16 <sup>th</sup> percentile* (N=37, 3Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	66	27.16%	Goal: <0.3%
	Favored rotamers	133	54.73%	Goal: >98%
	Ramachandran outliers	24	8.19%	Goal: <0.05%
	Ramachandran favored	204	69.62%	Goal: >98%
	Rama distribution Z-score	-6.45 ± 0.33		Goal: abs(Z score) < 2
	MolProbity score <sup>^</sup>	4.30		12 <sup>th</sup> percentile* (N=342, 3.50Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.73%	Goal: 0
	Bad bonds:	0 / 2353	0.00%	Goal: 0%
Bad angles:	11 / 3178	0.35%	Goal: <0.1%	
Peptide Omegas	Cis Prolines:	1 / 14	7.14%	Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	18	6.2%	Goal: <1.0%
	CA Geometry outliers	6	2.06%	Goal: <0.5%
Additional validations	Chiral volume outliers	0/365		
	Waters with clashes	0/0	0.00%	See UnDowser table for details

In the two column results, the left column gives the raw count, right column gives the percentage.

Выдача с учетом всех межатомных контактов.

По суммарным таблицам можно сделать вывод о том, что добавление H и разворота вышеупомянутых маргинальных остатков не улучшило показателей модели (однако сервис ниже дополнительно сообщает о том, что после этих процедур clashscore для всех атомов улучшился на 0.64). Число маргиналов по картам Рамачандрана совпадает с указанным в отчете PDB, однако число маргиналов по ротамерам боковых радикалов, найденных MolProbity, на 5 больше.

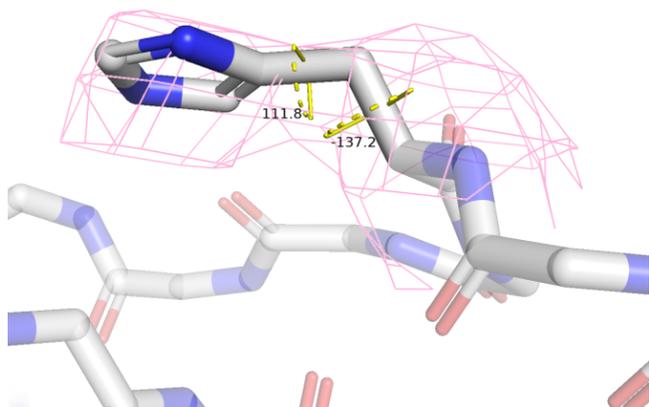
Маргинальных остатков настолько много, что рассмотреть все из них не представляется возможным в рамках данного практикума. В дальнейшем будем рассматривать только те маргиналы, которые считаются маргиналами и в отчете PDB, и в отчете MolPrody. В первую очередь проверим, есть ли функционально важные остатки среди маргинальных. Список функционально важных остатков возьмем с сайта [Uniprot](#). Было найдено два остатка,

которые были посчитаны выбросами и в отчете PDB, и в MolPrody. Причем это именно те остатки, которые отвечают за связывание с субстратом.

1) 161His - маргинал по ротамерам боковых радикалов

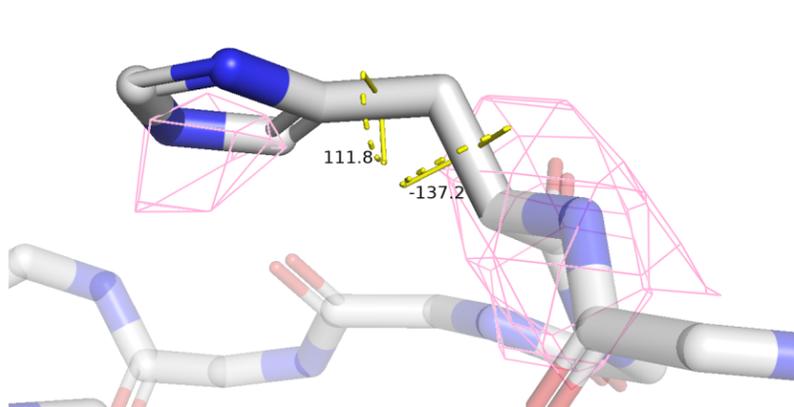
На уровне подрезки 2 видно, что боковая цепь гистидина слабо покрыта электронной плотностью, и наличие нестандартных углов  $\chi$  объясняется ошибкой расшифровки. По идее кольцо должно быть расположено зеркально и азот ND1, который находится со стороны азота остова, должен находиться со стороны кислорода остова.

For Educational Use Only



Подрезка = 1, carve = 2

For Educational Use Only



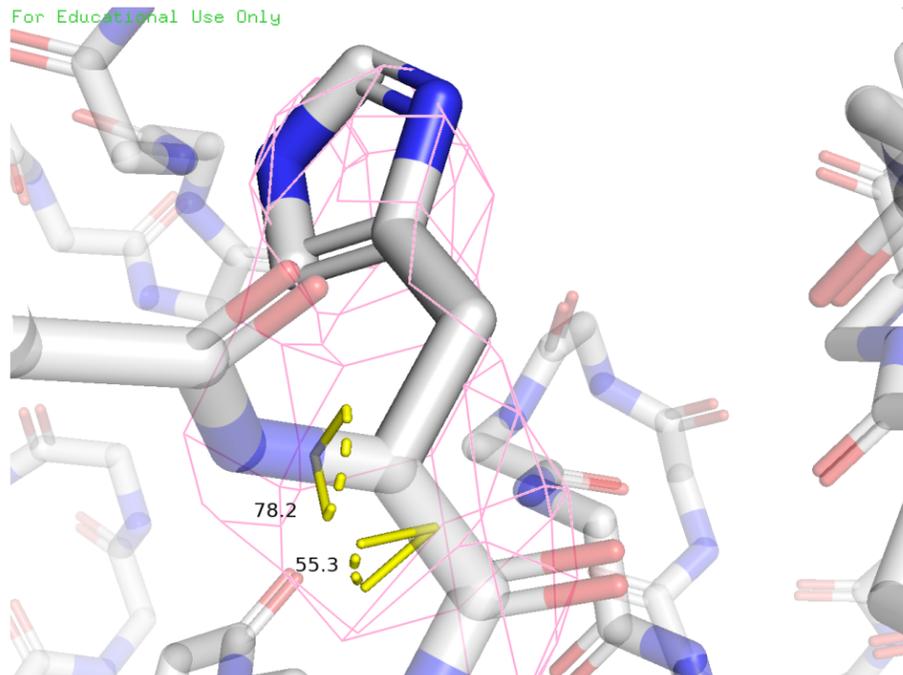
Подрезка = 2, carve = 2

2) 62His - маргинал по картам Рамачандрана

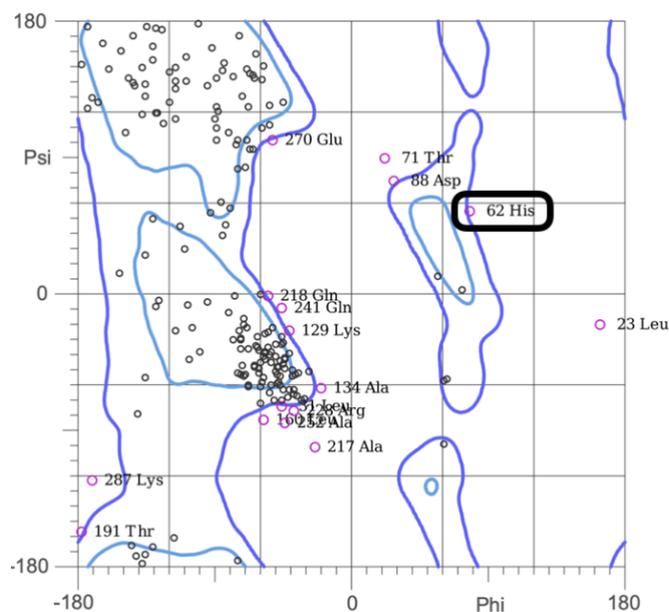
phi: 78.2, psi: 55.3

Остаток находится внутри “облака” электронной плотности, однако, возможно из-за низкого разрешения, совершенно не четко ею описывается.

For Educational Use Only



Подрезка = 2, carve = 1.5

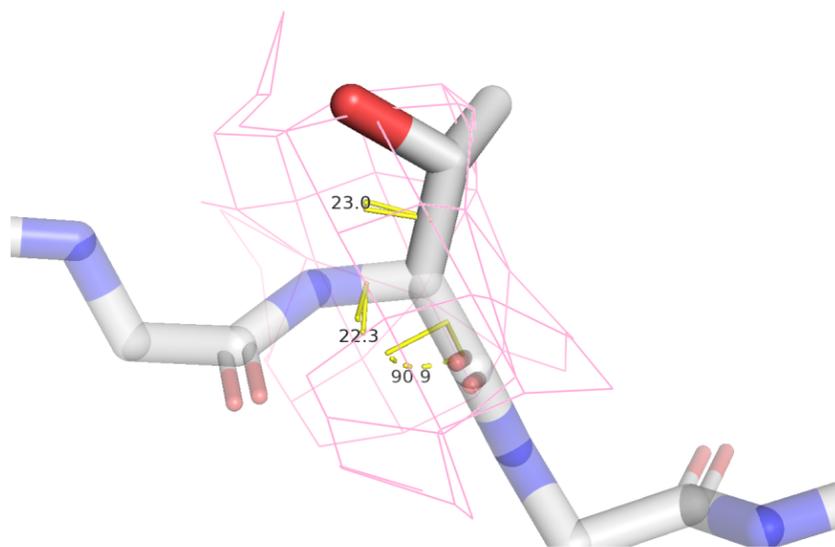


Карта Рамачандрана из отчета MolPrody. Рамкой выделен рассматриваемый остаток.

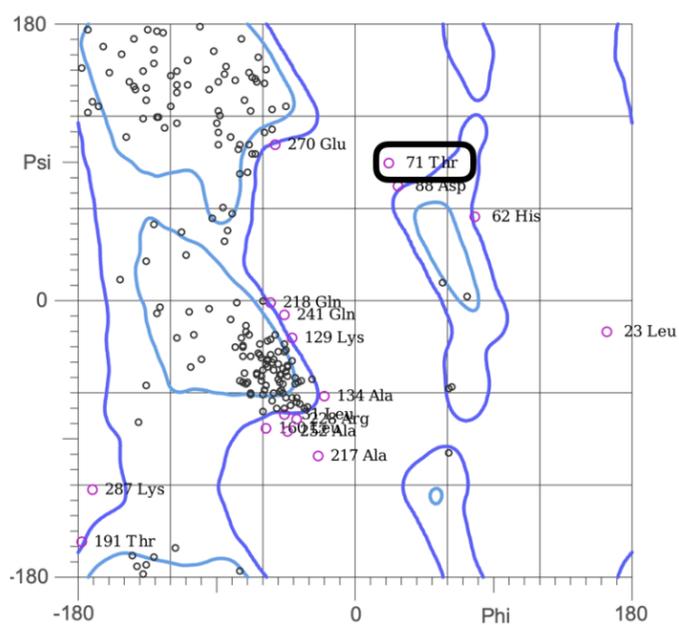
\*) 71Thr - маргинал по ротамерам боковых радикалов и по картам Рамачандрана

Из-за низкого разрешения вновь электронная плотность помогает восстановить лишь примерное пространственное положение отдельного остатка, но никак не составляющих ее атомов.

For Educational Use Only



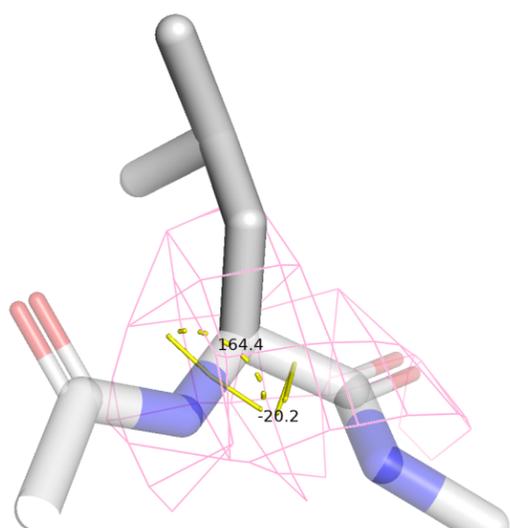
Подрезка = 1.5, carve = 2



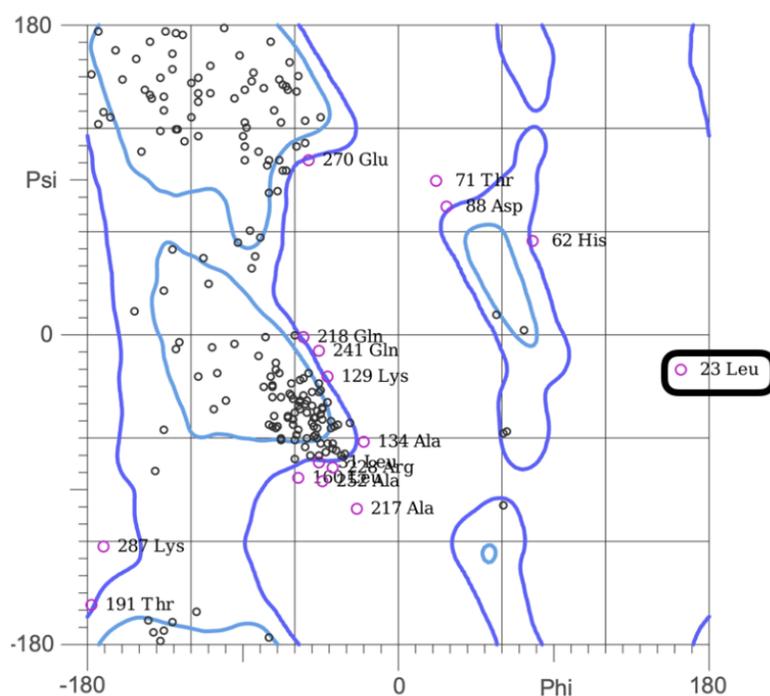
Карта Рамачандрана из отчета MolPrody. Рамкой выделен рассматриваемый остаток.

\*) 23Leu - остаток с экстремальным положением на карте Рамачандрана

For Educational Use Only



Подрезка = 2, carve = 1.5

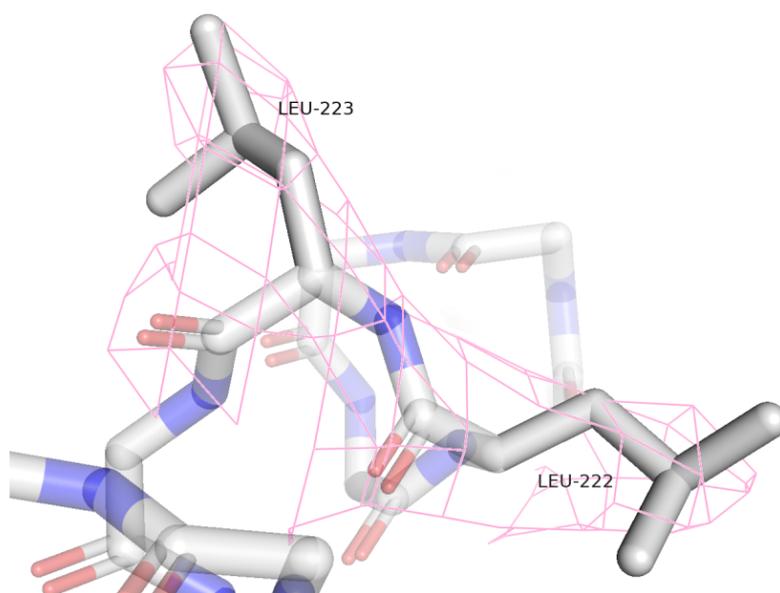


Карта Рамачандрана из отчета MolPrody. Рамкой выделен рассматриваемый остаток.

\*) 223Leu и 222Leu - идущие подряд остатки лейцина, маргиналы по ротамерам боковых радикалов

Я хотела рассмотреть эти остатки потому что, как сказано выше, большинство маргиналов по ротамерам боковых радикалов - лейцины. В данном случае это даже лейцины, идущие подряд.

For Educational Use Only



Ионов в структуре нет.

## Задание 3

Электронная плотность хоть и обволакивает остатки, но разрешение очень низкое и восстановить положение отдельных атомов из ЭП невозможно. Это приводит к неадекватному построению модели с низкими показателями качества и крайне высоким процентов маргинальных остатков. Маргинальными являются и функционально важные остатки, а именно два гистидина, отвечающих за соединение субстрата. Делать биологические выводы на основании данной структуры можно лишь только очень поверхностные, касающиеся общей “формы” белка, но для наскольконибудь более глубокого анализа она не пригодна.

## Задание 4

Сравним модель из PDB с переделанной моделью с сайта [PDB Redo](#).

	PDB	PDB Redo
Разрешение	3,50 Å	3,51 Å
R-фактор	0,2062	0,2188
R-free	0,2474	0,2570

R-free, подсчитанный авторами, PDB Redo посчитало смещенным. Было развернуто 4 ротамера и положение 191 остатка было изменено для того чтобы они лучше соответствовали электронной плотности.

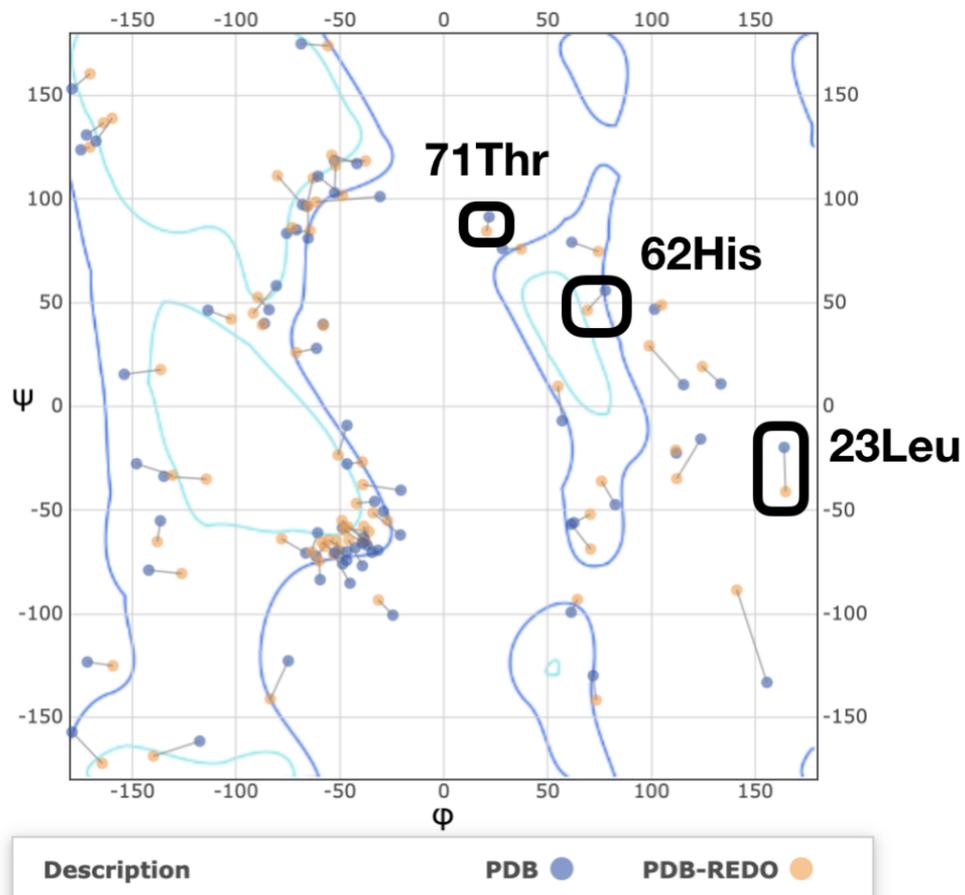
В целом, показатели качества незначительно ухудшились вопреки ожиданиям. Однако, возможно, это во многом следствие изменения положения 191 остатков. Структура была опубликована в 1998 году, когда в распоряжении ученых было меньше вычислительных мощностей и биоинформатический аппарат не был так развит. Так что, возможно, ухудшение показателей качества в одну десятую компенсируются более высоким реальным качеством определенной структуры.

Проверим это. Скачаем pdb “Re-refined and rebuilt structure”.

All-Atom Contacts	Clashscore, all atoms:	14.83	97 <sup>th</sup> percentile* (N=37, 3Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	57	23.46% Goal: <0.3%
	Favored rotamers	139	57.20% Goal: >98%
	Ramachandran outliers	12	4.10% Goal: <0.05%
	Ramachandran favored	219	74.74% Goal: >98%
	Rama distribution Z-score	-5.95 ± 0.36	Goal: abs(Z score) < 2
	MolProbity score <sup>^</sup>	3.51	66 <sup>th</sup> percentile* (N=342, 3.25Å - 3.76Å)
	Cβ deviations >0.25Å	0	0.00% Goal: 0
	Bad bonds:	0 / 2353	0.00% Goal: 0%
	Bad angles:	1 / 3178	0.03% Goal: <0.1%
Peptide Omegas	Cis Prolines:	1 / 14	7.14% Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	18	6.2% Goal: <1.0%
	CA Geometry outliers	3	1.03% Goal: <0.5%
Additional validations	Chiral volume outliers	0/365	
	Waters with clashes	0/0	0.00% See UnDowser table for details

Выдача MolProbity с учетом всех межатомных контактов для структуры полученной PDB Redo.

Показатель clashscore уменьшился более чем в 5 раз. Маргиналов по картам Рамачандрана стало ровно в 2 раза меньше. Маргиналов по ротамерам боковых радикалов стало на 9 меньше (13% от изначального количества). Также показатели связей и углов в модели пришли в норму. Хотя показатели качества все еще не лучшие, однако видно, что портал PDB Redo смог улучшить модель.



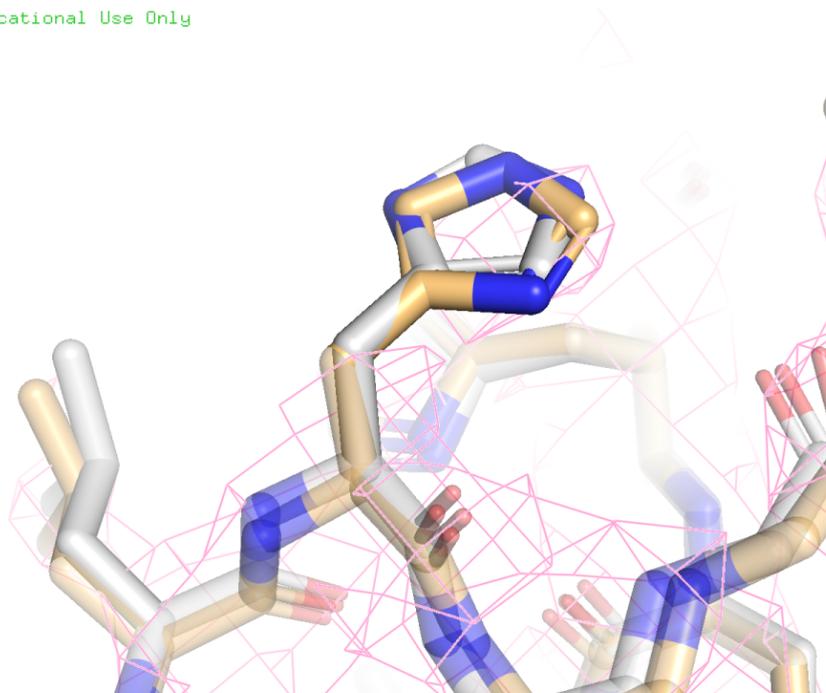
Изменение положения остатков на карте Рамачандрана, выделены рассматриваемые в практикуме маргинальные остатки

Далее серым будут изображены углеродные атомы исходной структуры, светло-оранжевым - структуры, полученной PDB Redo. Подрезка = 1.5, carve = 1.5.

1) 161His - маргинал по ротамерам боковых радикалов

Теперь кольцо гистидина действительно находится зеркально по отношению к прежнему положению и соответствует нормальной конформации этого остатка. В отчете MolProbity новое положение подписано как разрешенное.

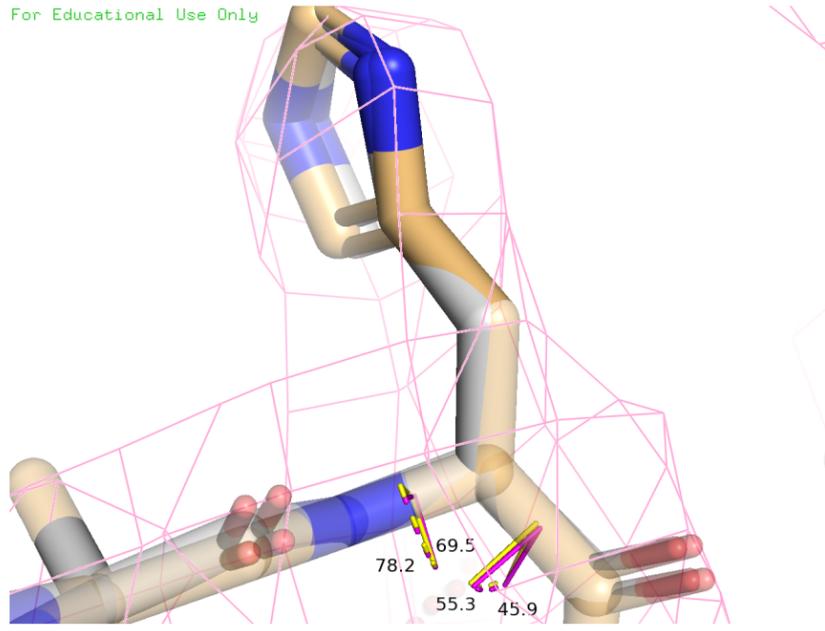
For Educational Use Only



2) 62His - маргинал по картам Рамачандрана

Положение остатка 62His тоже приобрело статус разрешенного вместо маргинального

For Educational Use Only

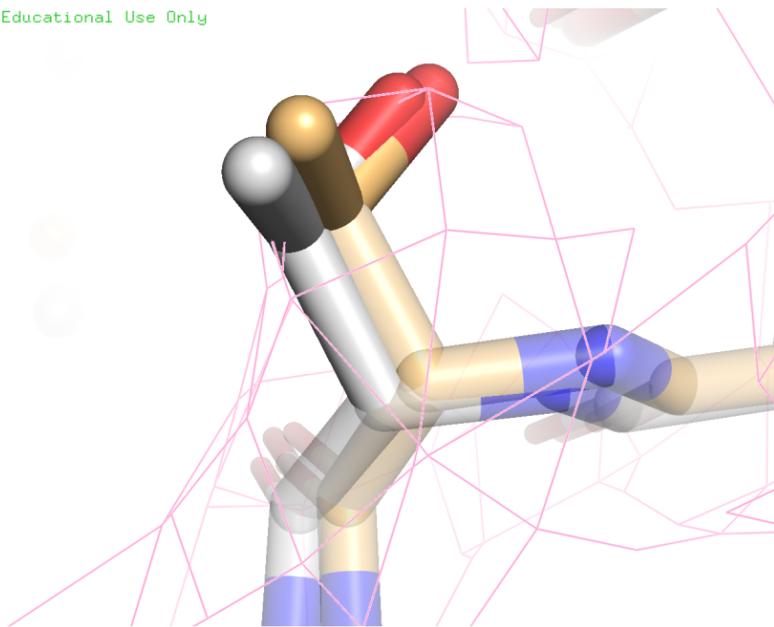


Желтый и подписи справа - измерения для исходной структуры, magenta и подписи слева - для PDB Redo

\*) 71Thr - маргинал по ротамерам боковых радикалов и по картам Рамачандрана

Этот остаток так и остался маргинальным

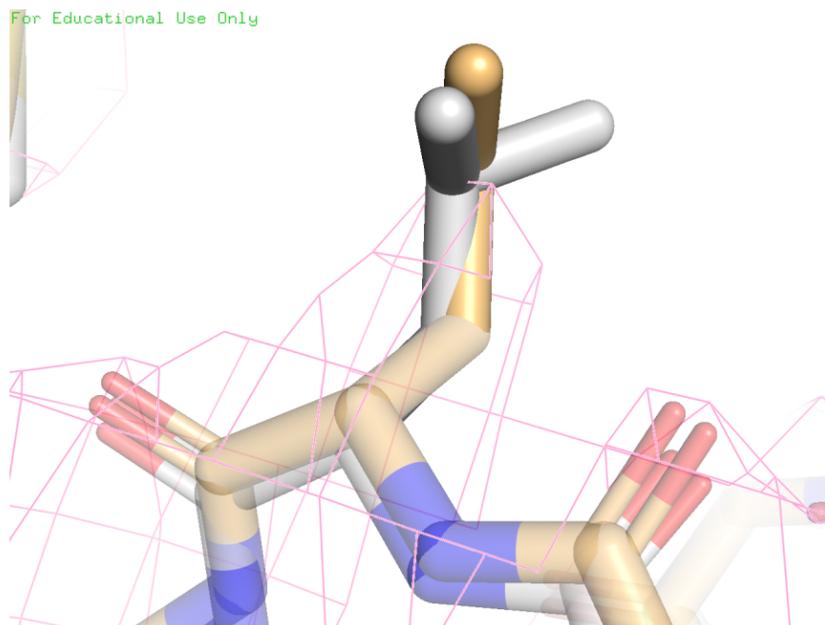
For Educational Use Only



\*) 23Leu - остаток с экстремальным положением на карте Рамачандрана

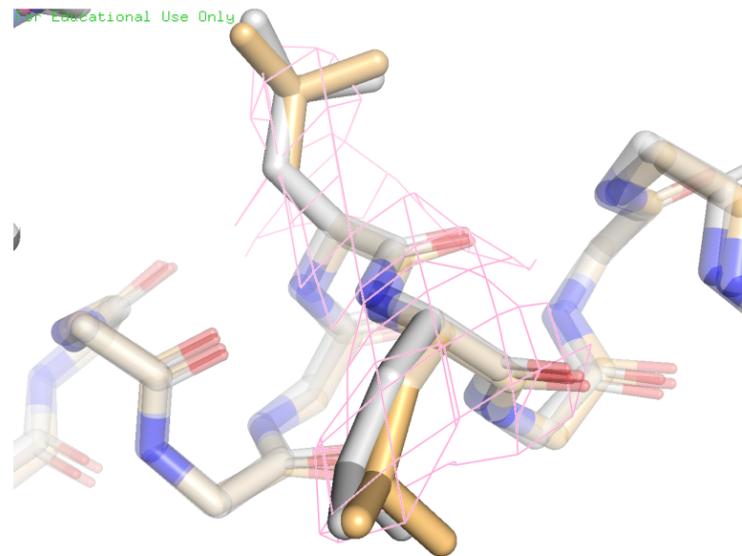
Этот остаток так и остался маргинальным

For Educational Use Only



\*) 223Leu и 222Leu - идущие подряд остатки лейцина, маргиналы по ротамерам боковых радикалов

Эти остатки так и остались маргинальными



С главной задачей PDB Redo справилось: оно хоть немного, но улучшило показатели качества структуры, более того, функционально важные остатки 62His и 161His перестали быть маргинальными. Однако, качество структуры все еще плохое и, на мой взгляд, плохо подходит для изучения данного белка.