

Отчет по расшифровке структуры изомеразы L-рамнозы ЗГУ

Выполнил:

Студент 4 курса ФББ

Сигорских Андрей

Москва, 2017г.

Содержание

Список сокращений.....	3
I. Введение.....	3
I.1. О белке.....	3
I.2. О кристалле.....	6
II. Основные показатели качества модели.....	8
III. Маргинальные остатки.....	15
A: GLU224.....	16
A: LEU252.....	17
C: VAL317.....	18
D: PRO225.....	19
D: ASN344.....	20
IV. PDB redo.....	21
Выводы.....	23
Список источников.....	24

Список сокращений

- L-RhI – изомераза L-рамнозы (L-rhamnose isomerase);
- *P. stutzeri* - *Pseudomonas stutzeri*;
- EDTA - Этилендиаминтетрауксусная кислота;
- R_{free} – свободный R-фактор

I. Введение

I.1. О белке

Изомераза L-рамнозы – широко распространенный у многих микроорганизмов фермент, принимающий участие в пути метаболизма рамнозы. Катализирует обратимую изомеризацию L-рамнозы в L-рамнулозу. Гомотетрамер, длина одной цепи 438 аминокислотных остатков, вес ~48 КДа.

Рассмотренный в данной работе белок из *P. stutzeri* интересен тем, что помимо этой реакции также способен катализировать обратимую изомеризацию D-аллозы в D-псикозу (см. рис.1). Данные сахара в природе встречаются в малых количествах, и белок *P. stutzeri* L-RhI используется в промышленности для производства редких сахаров.

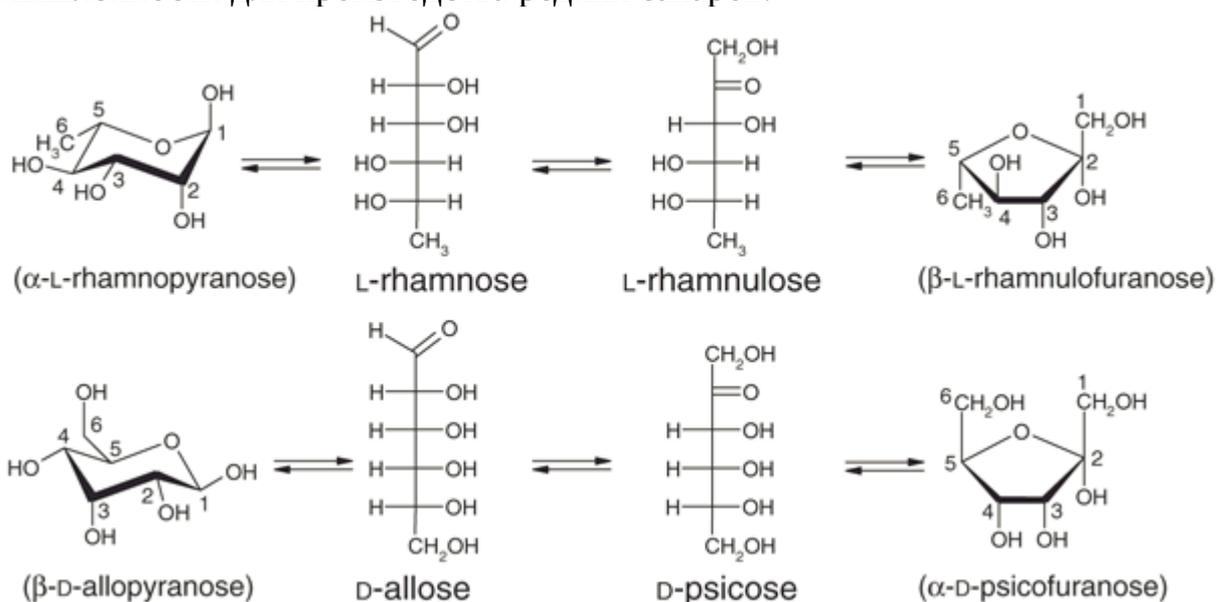


Рисунок 1. Реакции, катализируемые *P. stutzeri* L-RhI

Осуществляемая L-RhI реакция альдозно-кетозной изомеризации, скорее всего, протекает по опосредованному ионами металла механизму “гидридного сдвига” (см. рис. 2), для изучения которого и была получена структура ЗИТУ [1].

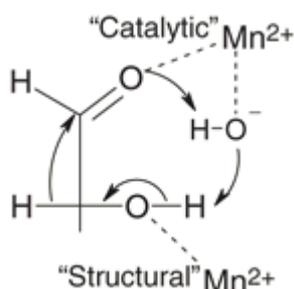


Рисунок 2. Предположительный механизм "гидридного сдвига"

L-RhI способна связывать ионы марганца, меди, кобальта и цинка, а её ферментативная активность сильно зависит от типа иона: соотношение активностей для белков, связавших вышеописанные ионы, составляет 100: 35: 19: 10, соответственно [2]. Исходя из этих данных, исследование проводилось на ферменте, связавшем ионы марганца.

В отличие от остальных ионов, фермент может связать ионы марганца в двух альтернативных формах в каждом из четырех активных центров: в так называемой “AD-форме” первый (“структурный”) ион Mn_1 координирован шестью связями с четырьмя остатками Glu₂₁₉(OE), Asp₂₅₄(OD), His₂₈₁(ND), Asp₃₂₇(OD) и двумя молекулами воды W_1 и W_2 , а второй (“каталитический”) Mn_2 координирован связями с двумя остатками His₂₅₇(NE), Asp₂₈₉(OD1) и четырьмя молекулами воды (W_2 , W_3 , W_4 и W_5). Дистанция между ионами составляет 4.2 Å. В так называемой “BC-форме” структурный ион координирован также, а каталитический – четырьмя остатками His₂₅₇(NE), Asp₂₈₉(OD1), Asp₂₈₉(OD2), Asp₂₉₁(OD) и двумя молекулами воды (W_6 , W_7). В этом случае дистанция между ионами составляет 5.2 Å (см. рис. 3)[1]. Активный сайт фермента связывает субстрат в AD-форме.

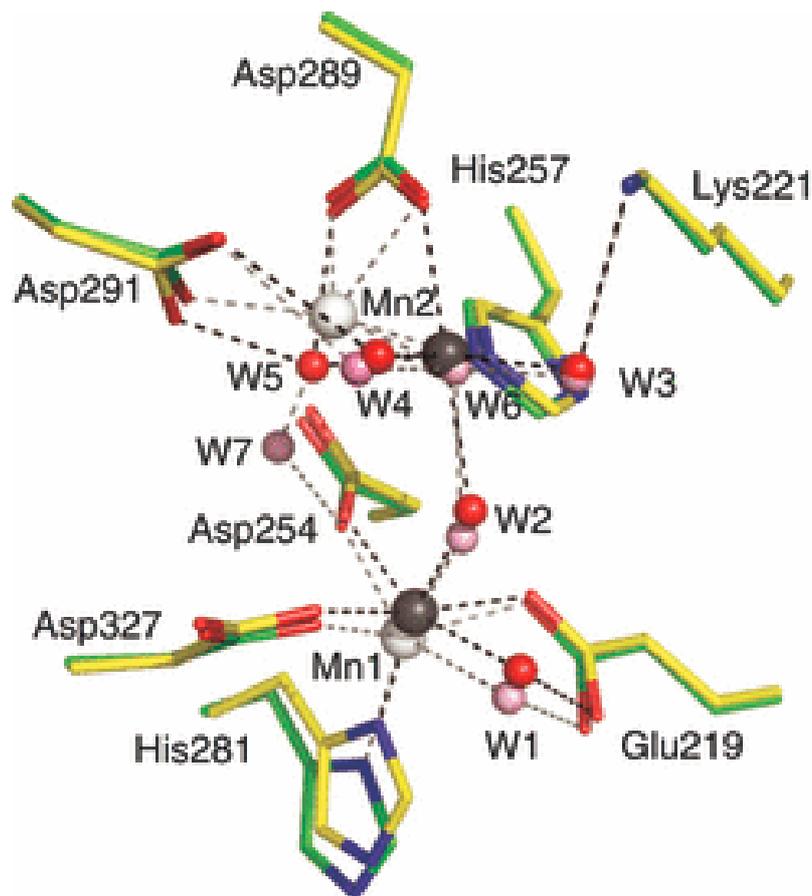


Рисунок 3. AD и BC-формы связывания ионов марганца в активном сайте фермента. AD-форма: жёлтые углероды остова, черные ионы марганца, красные молекулы воды, черные пунктирные линии координационных связей. BC-форма: зелёные углероды остова, белые ионы марганца, розовые молекулы воды, серые пунктирные линии координационных связей.

По результатам рентгеноструктурного исследования был предложен следующий механизм катализа L-RhI: до связывания субстрата активный центр осциллирует между AD и BC-формами (Рис 4, А, D). При смене от BC к AD, молекула воды W_4 (отмечена красным эллипсом) активируется, теряет протон и превращается в гидроксид-ион, координированный $Mn_2[3]$. Кетоза с фуранозным кольцом связывается AD-формой (Рис 4, В), при этом остаток Asp_{327} участвует в раскрытии кольца, а ионы Mn_1 и Mn_2 координируют атомы O_1 , O_2 и O_3 . После раскрытия каталитическая молекула воды W_4 участвует в переносе протона с O_1 на O_2 , производя альдозу с O_2 -гидроксилом (Рис 4, С). Альдоза с пиранозным кольцом, скорее всего, связывается с BC-формой (Рис 4, Е), при этом O_1 и O_3 координируются Mn_1 . Каталитическая молекула воды W_4 раскрывает кольцо, после чего Mn_2 переходит в AD-форму, активирует её, что приводит к атаке протона с O_2 на O_1 , производя кетозу (Рис 4, F).

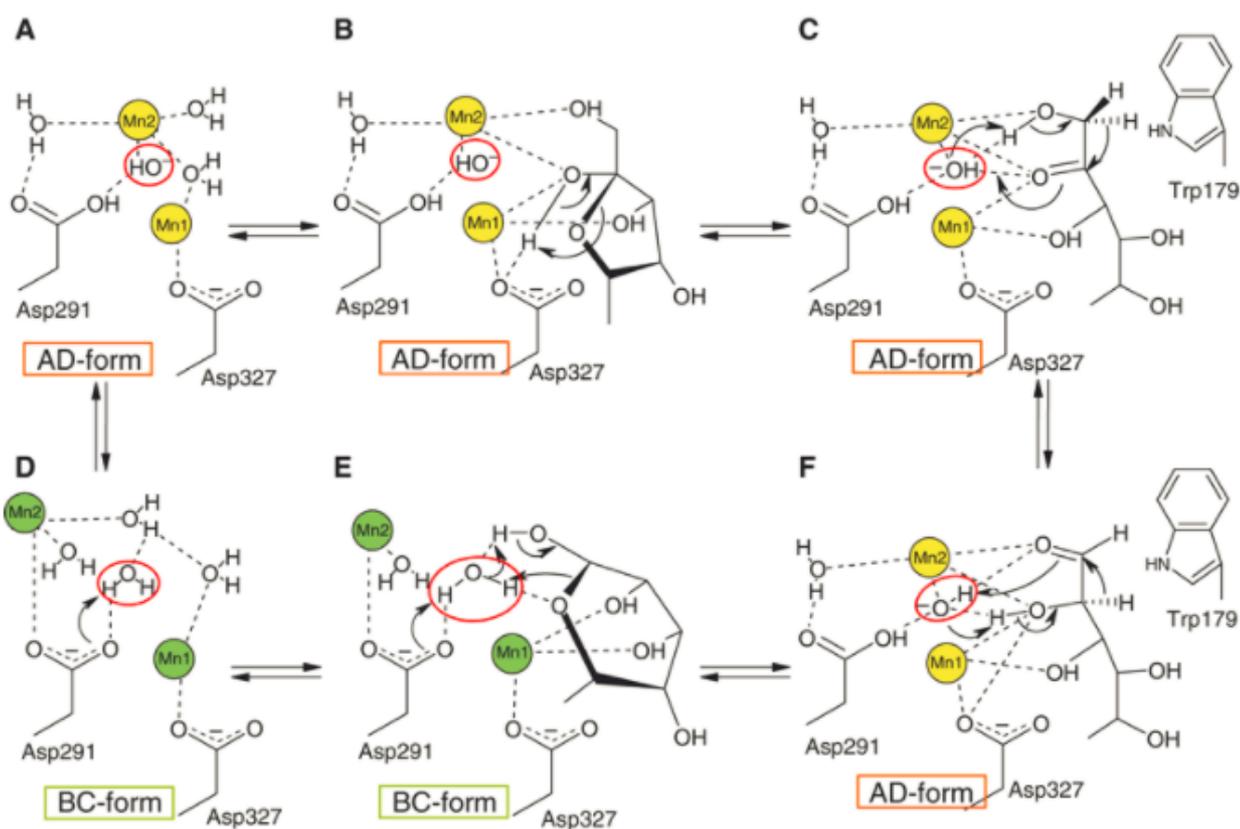


Рисунок 4. Механизм катализа *P.stutzeri* L-Rhl

1.2. О кристалле

Структура 3ITY – одна из структур, полученных в ходе работы [1] по определению механизма катализа *P.stutzeri* L-Rhl, вышедшей в 2010 году. Авторы - Yoshida H, Yamaji M, Ishii T, Izumori K, Kamitori S.

Данная структура не содержит ионов металла. Очищенные фракции белка диализовали против буфера (5 mM Tris/HCl и 5 mM EDTA, pH 8.0) для удаления ионов металлов. После этого смесь диализовали против 5 mM Hepes, pH 8.0 для удаления EDTA, а белок был сконцентрирован до 1–2 mg·mL⁻¹. Кристаллы белка были выращены методом диффузии паров (vapor diffusion) с участием смеси 7–8% полиэтиленгликоль 20 000 и 50 mM Mes buffer (pH 6.3).

В ходе РСА, кристаллы белка были охлаждены до 100 К в жидком азоте. Данные получены с помощью прибора PHOTON FACTORY BEAMLINE BL-6A (Цукуба, Япония).

Фазовая проблема решалась методом молекулярного замещения с использованием структуры белка дикого типа 2HCV. В ходе эксперимента

было измерено 140561 рефлексов с разрешением от 46.94 до 1.84 Å; полнота данных составляет 97.0%, то есть разрешению структуры 1.84 Å можно доверять. Из всех измеренных рефлексов 10% (14044) были удалены из датасета для оптимизации, и были использованы для подсчета R_{free} . Значения R-фактора и R_{free} составляют 0.158 и 0.190 ± 0.002 , соответственно; значения не очень сильно отличаются, а значение R_{free} не превышает порог для “очень хороших” структур (0.2), что говорит о высоком качестве структуры.

Асимметрическая ячейка включает в себя тетрамер белка, являющийся биологической единицей. Некристаллографических симметрий не содержит. Кристалл имеет симметрию $P2_1$, параметры элементарной ячейки приведены в таблице 1:

Таблица 1. Параметры элементарной ячейки

Длина вектора (Å)	a: 74.691	b: 104.785	c: 104.785
Угол (°)	α : 90	β : 107.98	γ : 90

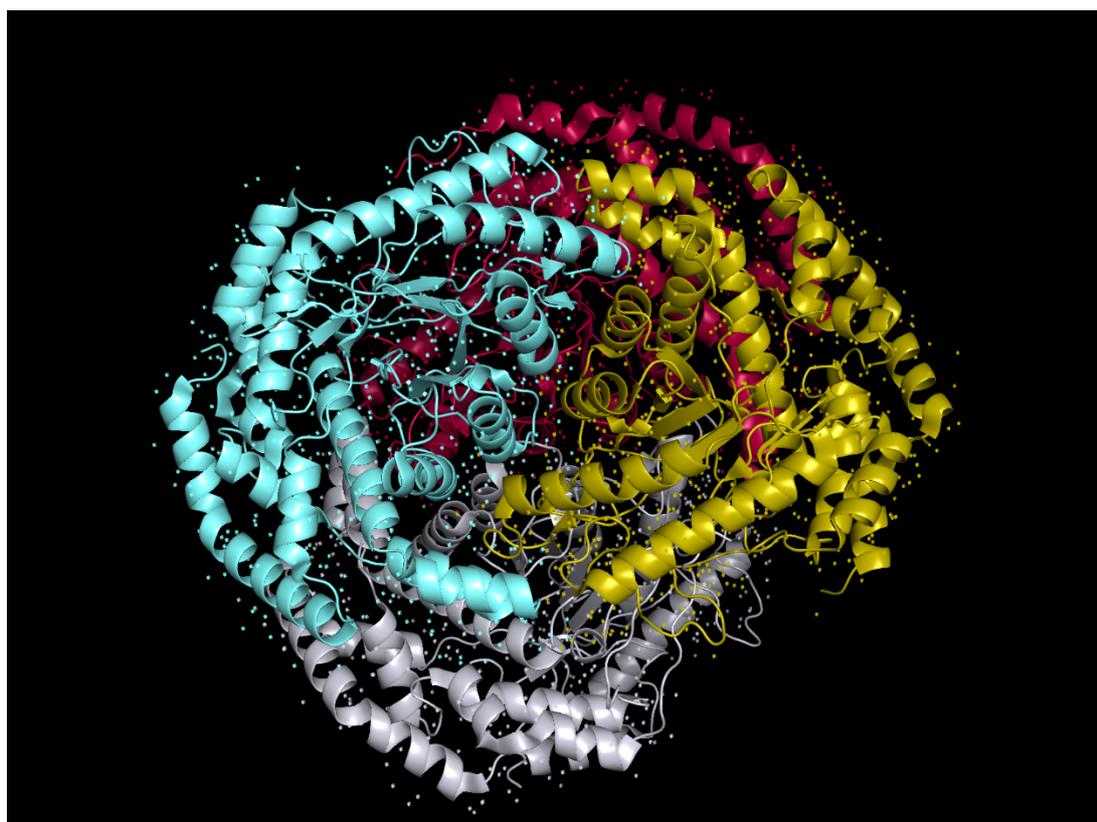


Рисунок 5. Структура Zity, раскраска по цепям

II. Основные показатели качества модели.

Исходя из приведенного на странице структуры в PDB[4] графика Percentile plot (рис. 6), структура Zity лучше большинства остальных структур, а по ряду показателей – и структур с таким же разрешением, однако провалы по параметрам Clashscore и Ramachandran outliers дополнительного исследования вопроса. Для этого обратимся к сервису MolProbity[5].

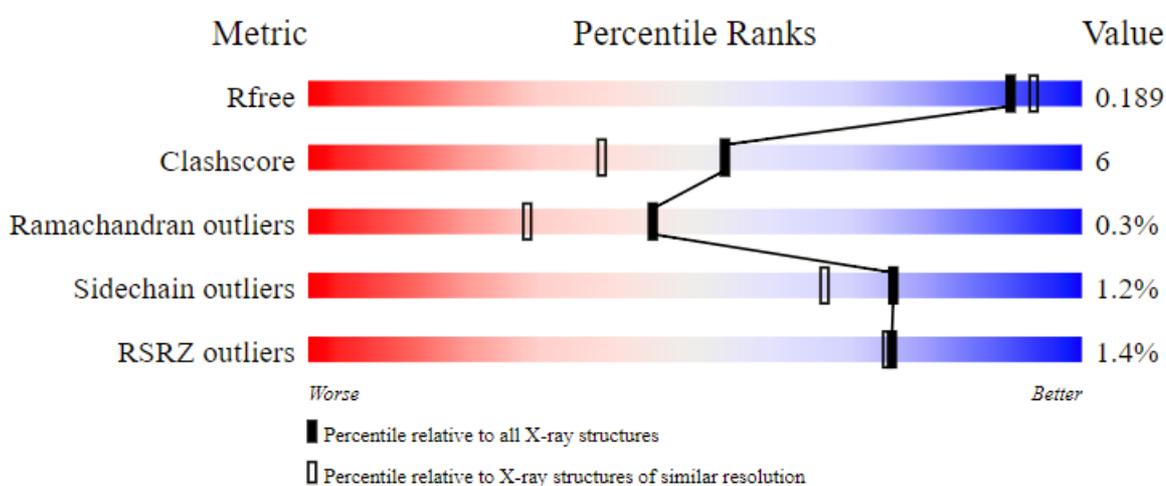


Рисунок 6. Параметры качества структуры Zity по сравнению с другими структурами из PDB

После загрузки структуры на сервер MolProbity была проведена процедура добавления водородов с одновременным исправлением инверсий остатков Asn/Gln/His. Список остатков, подвергшихся инверсиям, приведен ниже:

- A 15 ASN
- A 160 HIS
- A 189 GLN
- A 258 HIS
- A 324 HIS
- A 344 ASN
- B 8 GLN

- B 15 ASN
- B 160 HIS
- B 189 GLN
- B 258 HIS
- B 324 HIS
- B 344 ASN
- C 8 GLN
- C 15 ASN
- C 160 HIS
- C 189 GLN
- C 258 HIS
- C 324 HIS
- C 344 ASN
- D 15 ASN
- D 160 HIS
- D 189 GLN
- D 258 HIS
- D 324 HIS
- D 344 ASN

После “исправления” структуры сервис дал подробную информацию (см. таблицу 2). В целом, структура оценена очень хорошо, хотя порог >98% по количеству “очень хороших” ротамеров и углов на карте Рамачандрана кажется мне несколько завышенным. По остальным параметрам структура является одной из лучших среди всех структур с подобным разрешением:

Таблица 2. Характеристика сервиса MolProbity структуры Zity с исправленными инверсиями остатков

All-Atom Contacts	Clashscore, all atoms:	4.27		98 th percentile* (N=809, 1.84Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	16	1.21%	Goal: <0.3%
	Favored rotamers	1275	96.44%	Goal: >98%
	Ramachandran outliers	5	0.30%	Goal: <0.05%
	Ramachandran favored	1637	97.38%	Goal: >98%
	MolProbity score ^a	1.39		97 th percentile* (N=11316, 1.84Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 13360	0.00%	Goal: 0%
	Bad angles:	2 / 18113	0.01%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	8 / 72	11.11%	Expected: ≤1 per chain, or ≤5%

Также, сервис позволяет визуализировать карту Рамачандрана с положением остатков на ней (рис. 7). Сервис нашел всего 5 аутлаеров, причем 4 из них – это одна и та же аминокислота 224 GLU, находящаяся перед пролином, на разных цепях:

- A 224 GLU (76.8, 116.2)
- B 224 GLU (76.5, 117.1)
- C 224 GLU (78.1, 118.0)
- C 317 VAL (-32.1, 98.6)
- D 224 GLU (77.9, 118.2)

MolProbity Ramachandran analysis

3ityFH.pdb, model 1

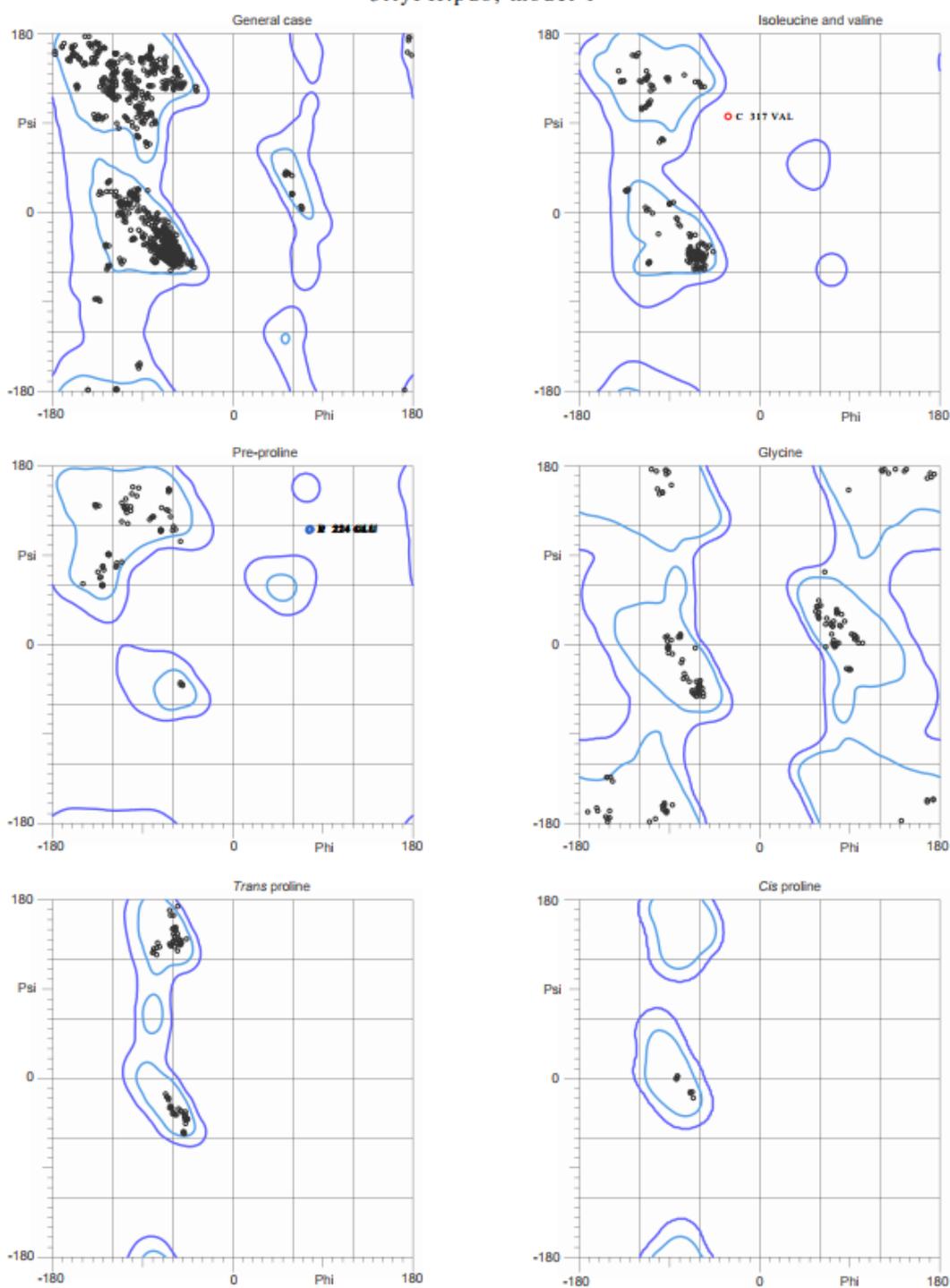


Рисунок 7. Положение аминокислотных остатков 3ity на картах Рамачандрана.

PDB validation report содержит интересную информацию об остатках в виде графика (рис.8): верхняя линия показывает качество покрытия остатков белка электронной плотностью, для остатков представлено количество параметров, по которым остаток является аутлайером – 0 – зеленый цвет, 1 – желтый, 2 – оранжевый, 3 и более – красный. Остатки, положение которых плохо описывается электронной плотностью помечены красными точками над ними. Остатки, не вошедшие в модель, показаны серым. Наиболее интересной в данном случае является цепь C, содержащая 3% плохо покрытых электронной плотностью остатков (и единственная имеющая “оранжевый” остаток в своём составе). Остальные цепи содержат такой же процент “желтых” остатков и < 1% плохо покрытых плотностью.

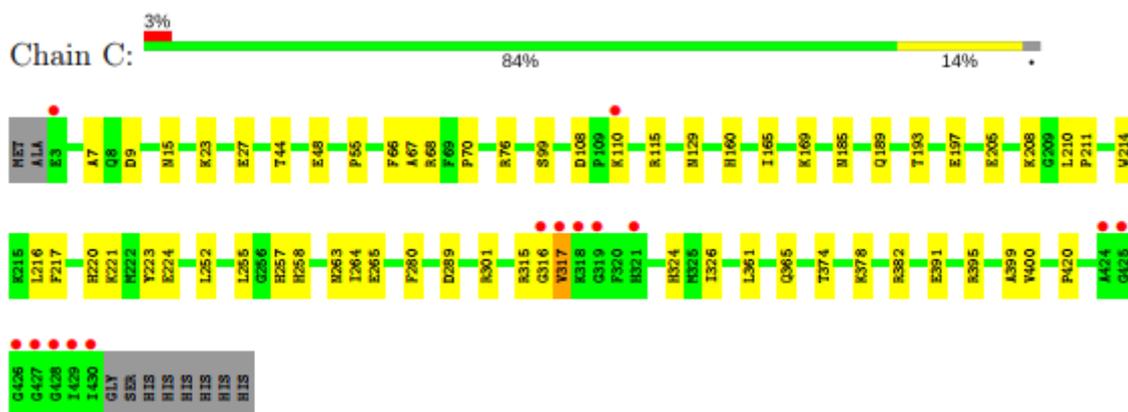


Рисунок 8. Визуализация проблемных мест в структуре 3ity по данным pdb validation report

За дополнительной информацией также можно обратиться к серверу EDS [6]. Например, для всех цепей можно посмотреть графики температурного фактора (рис.9), Real-space R-value (рис. 10), Z-score (рис.11, а, б).

График температурного фактора свидетельствует об относительно малой подвижности всех остатков, среднее значение не превышает 16 Å.

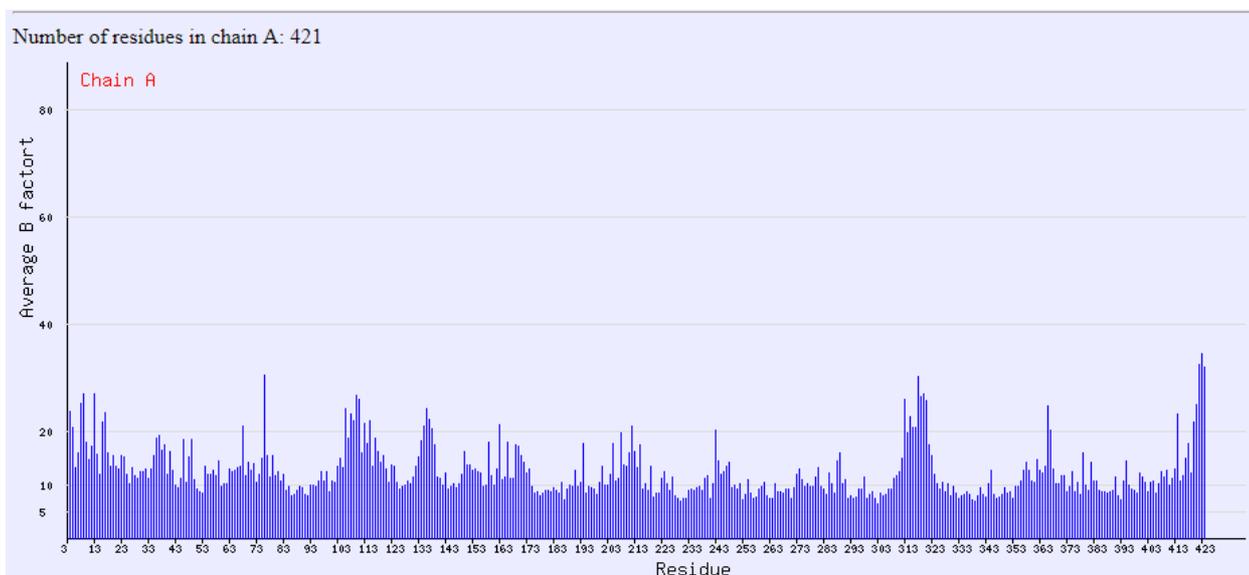


Рисунок 9. График температурного фактора для цепи A структуры Zity

Среднее значение параметра Real-space R-value равно 0.089, и большая часть остатков довольно мало отклоняются от этого значения, хотя во всех цепях встречаются редкие аутлаеры со значениями 0.2 и выше. Особенно выражено это в уже упомянутой ранее цепи C, в которой ряд маргинальных остатков имеет значение параметра выше 0.4:

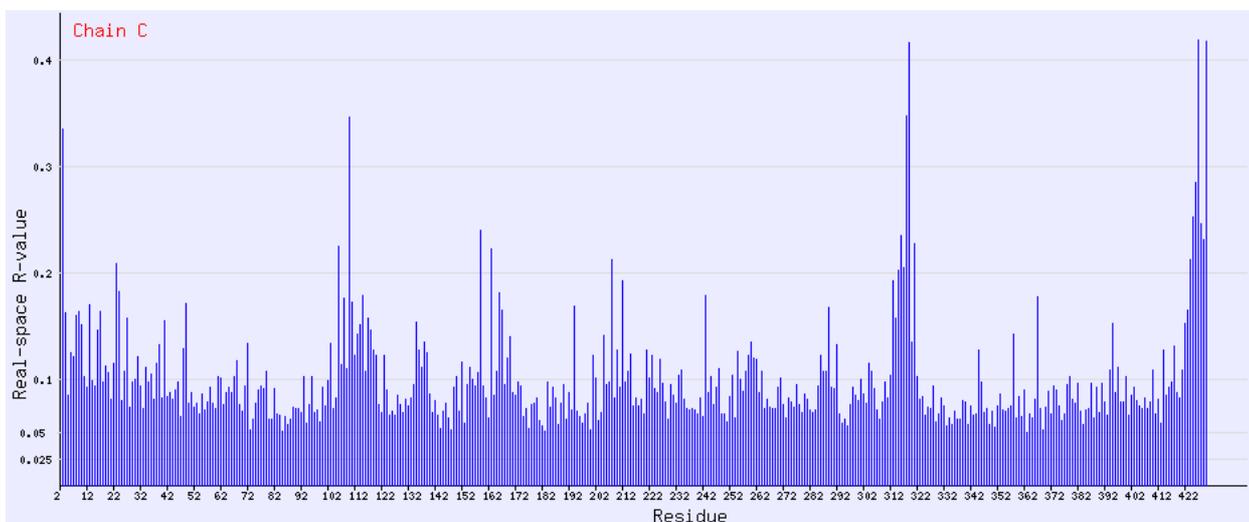


Рисунок 10. График Real space R-value для цепи C структуры Zity

Эти же аутлаеры в цепи C хорошо прослеживаются и на графике Z-score; стоит отметить, что по этому параметру последние остатки являются аутлаерами с значениями $Z \geq 3$ во всех цепях, а цепь C особенно выделяется своим “хвостом” потому, что у неё на нём находится больше остатков. Подробная информация о значениях Z-score по цепям приведена в таблице 3:

Таблица 3. Значения параметра Z-score для цепей структуры Zity

Chain	Nres	<Z>	Sigma	%<-1	%<0	%>1	%>2
A	421	- 0.57	0.76	12.11	89.55	0.95	0.24
B	421	- 0.53	0.80	14.01	88.84	2.61	0.95
C	428	- 0.26	0.92	5.61	77.57	5.61	3.27
D	419	- 0.54	0.84	15.51	86.40	1.67	0.95
All	1689	- 0.47	0.83	11.78	85.55	2.72	1.36

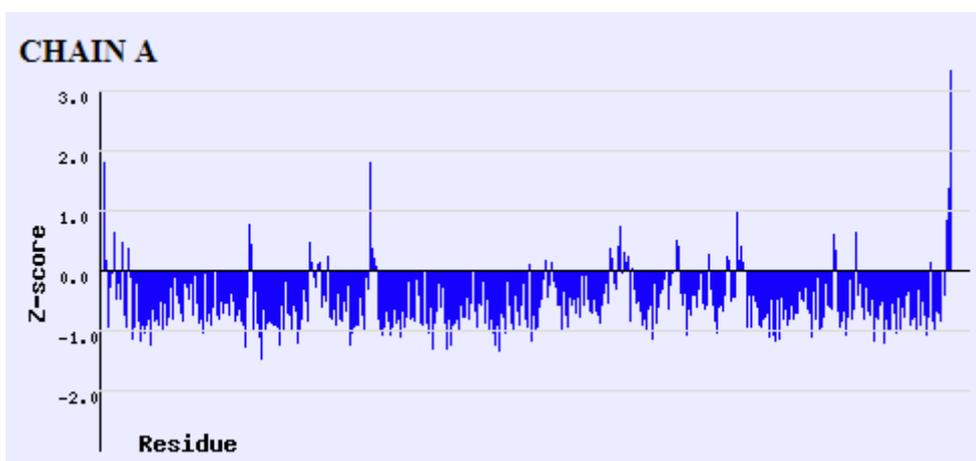


Рисунок 11а. График параметра Z-score для цепи А структуры 3ity

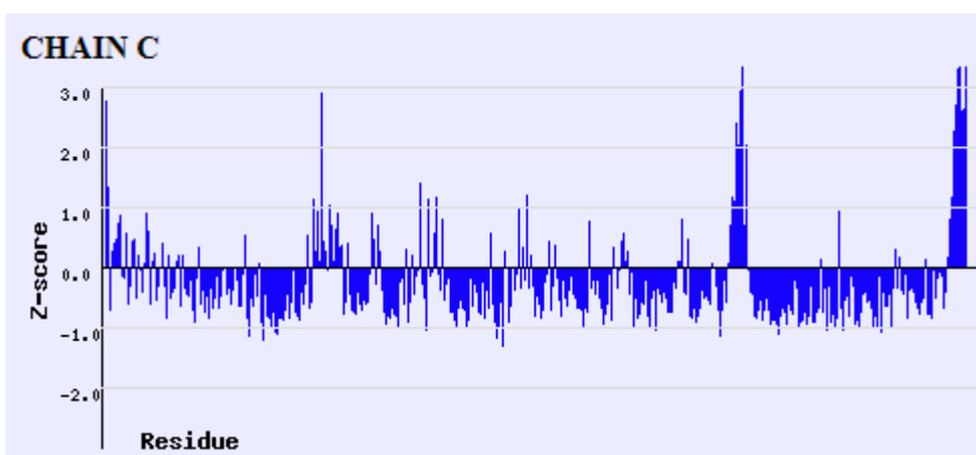


Рисунок 11б. График параметра Z-score для цепи С структуры 3ity

Согласно имеющейся на сервере карте Рамачандрана, белок имеет 34 маргинальных остатка, однако этой карте меньше доверия, чем выдаче MolProbity, потому что последняя учитывает куда больше информации (например, строит отдельные карты для изолейцинов и валинов, или пред-

пролиновых остатков). Все остатки, являющиеся маргинальными по карте Рамачандрана по мнению MolProbity, являются таковыми и по карте EDS.

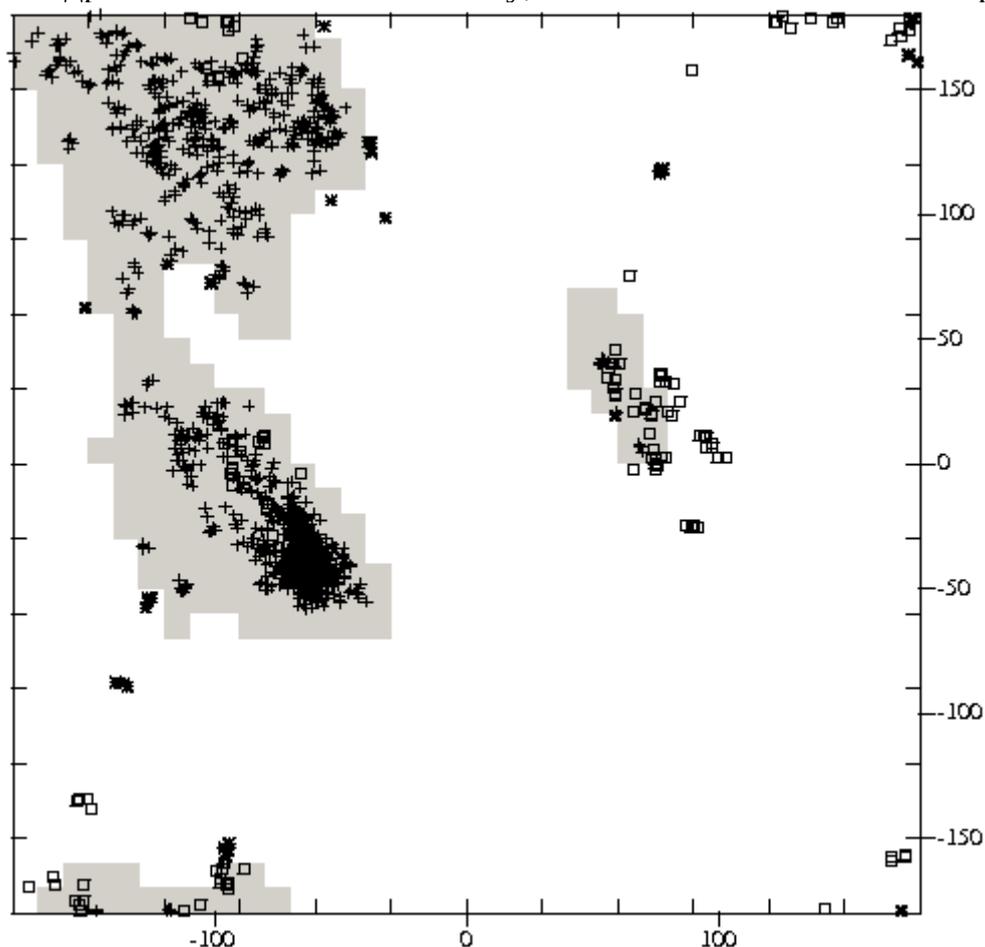


Рисунок 12. Карта Рамачандрана структуры 3ity, полученная с сервера EDS

III. Маргинальные остатки.

Приведем примеры нескольких маргинальных остатков в структуре (по версии MolProbity, по структуре с добавленными водородами и исправленными инверсиями боковых цепей) в таблице 4:

Таблица 3. Примеры маргинальных остатков в структуре 3ity

Остаток	Причина маргинальности
A: SER99	Ротамер: $\chi_1 = 257.2$
A: GLU224	Карта Рамачандрана: пре-пролиновый остаток, 76.8, 116.2
A: LEU252	Clash HG атома с A:PHE217 HB3 на 0.61 Å
A: GLU366	Clash HG2 атома с кислородом воды A:HOH 575 на 0.57 Å

B: LYS110	Ротамер: $\chi = 309.2, 200.2, 305.4, 76$
C: ASN191	Плохие углы в связи N-CA-C
C: VAL317	Плохо покрыт электронной плотностью; по карте Рамачандрана: (изолейцин либо валин) - $-32.1, 98.5$
C: THR374	Ротамер: $\chi = 217.1$
D: PRO225	Цис-пролин
D: ASN344	Инверсия боковой цепи (флип)

Рассмотрим подробнее пять отмеченных зелёным остатков:

A: GLU224

Согласно [7], предпочитаемые области на карте Рамачандрана пре-пролиновых остатков имеют три отличия от областей для остатков, находящихся перед другими аминокислотами. Во-первых, исключается большая горизонтальная полоса по $-40^\circ < \psi < 50^\circ$, отделяющая α_L и α регионы. Во-вторых, α_L – регион поднят гораздо выше. Наконец, имеется выступающий участок плотности предпочитаемых конформаций под β -областью, получивший название ζ -региона (см. рис. 13).

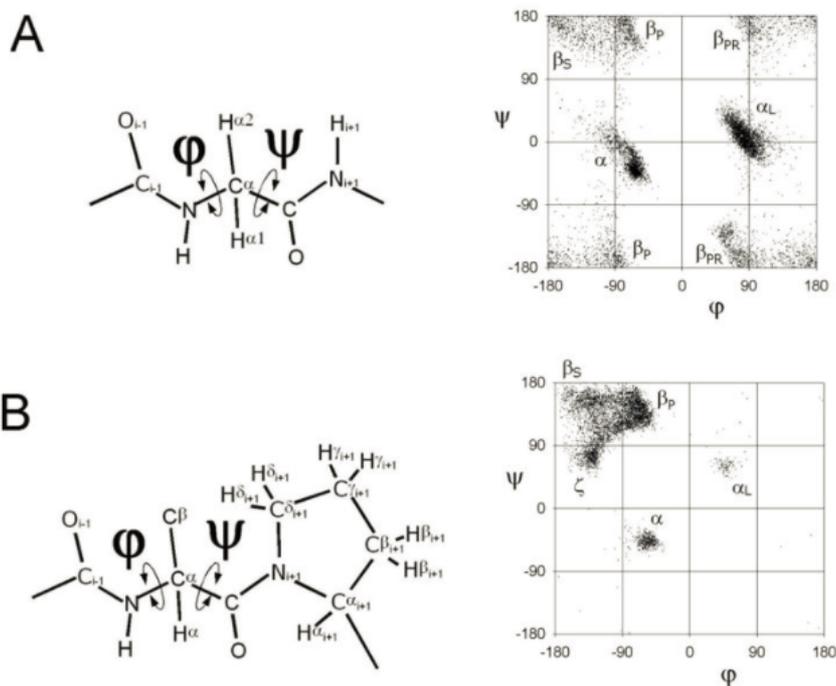


Рисунок 13. Сравнение карт Рамачандрана для глицина (A) и пред-пролиновых остатков (B) [7].

Примечателен тот факт, что 224-й глутамат является маргиналом по данному параметру во всех четырех цепях белка с одинаковыми значениями торсионных углов (см. рис. 7, “pre-proline”). В результате исследования остатков и их окружения выяснилось следующее:

- Электронная плотность на уровне подрезки 2.5 покрывает углероды боковой цепи не очень хорошо, но позволяет определить положение кислородов карбоксильной группы во всех четырех случаях;
- Во всех случаях эти кислороды находятся в положении, благоприятном для образования сразу двух водородных связей с NE и NH2 азотами 65-го аргинина соседней цепи белка: глутамат цепи А с аргинином цепи В и наоборот, и тоже самое для цепей С и D (см. рис. 14).

Исходя из вышеперечисленного, можно сделать вывод, что данные остатки являются не ошибками расшифровки и не артефактами кристаллизации, а участвуют в стабилизации димеров в составе тетрамера и находятся в модели абсолютно правомерно.

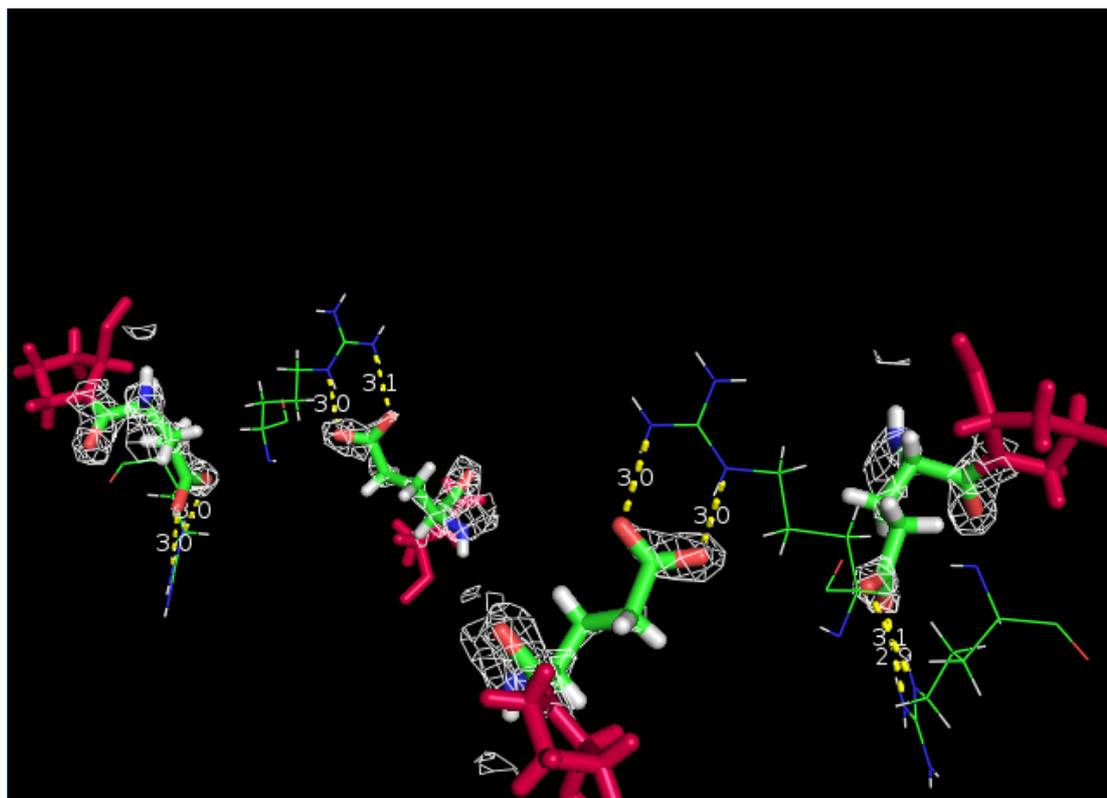


Рисунок 14. Остатки 224GLU, их электронная плотность на уровне подрезки 2.5 и их предполагаемые водородные связи с остатками 65ARG. Пронины 225PRO, которым предшествуют данные остатки, окрашены малиновым.

А: LEU252

Ван-дер-Ваальсовы сферы атомов действительно визуалью пересекаются (за радиус сферы было взято значение 1.2 Å), а попытка подсчитать

дистанцию командой `dist` дала ответ 0.0, то есть clash в модели действительно присутствует. Электронная плотность на уровне подрезки 3.0 покрывает почти все углероды боковой цепи остатка PHE217 и позволяет достаточно точно определить их координаты, а на цепи LEU252 полностью отсутствует, следовательно, её положение определялось авторами, и в их вычисления могла закрасться ошибка (так как водороды были добавлены сервисом MolProbity и в изначальной структуре отсутствовали) (см. рис. 15).

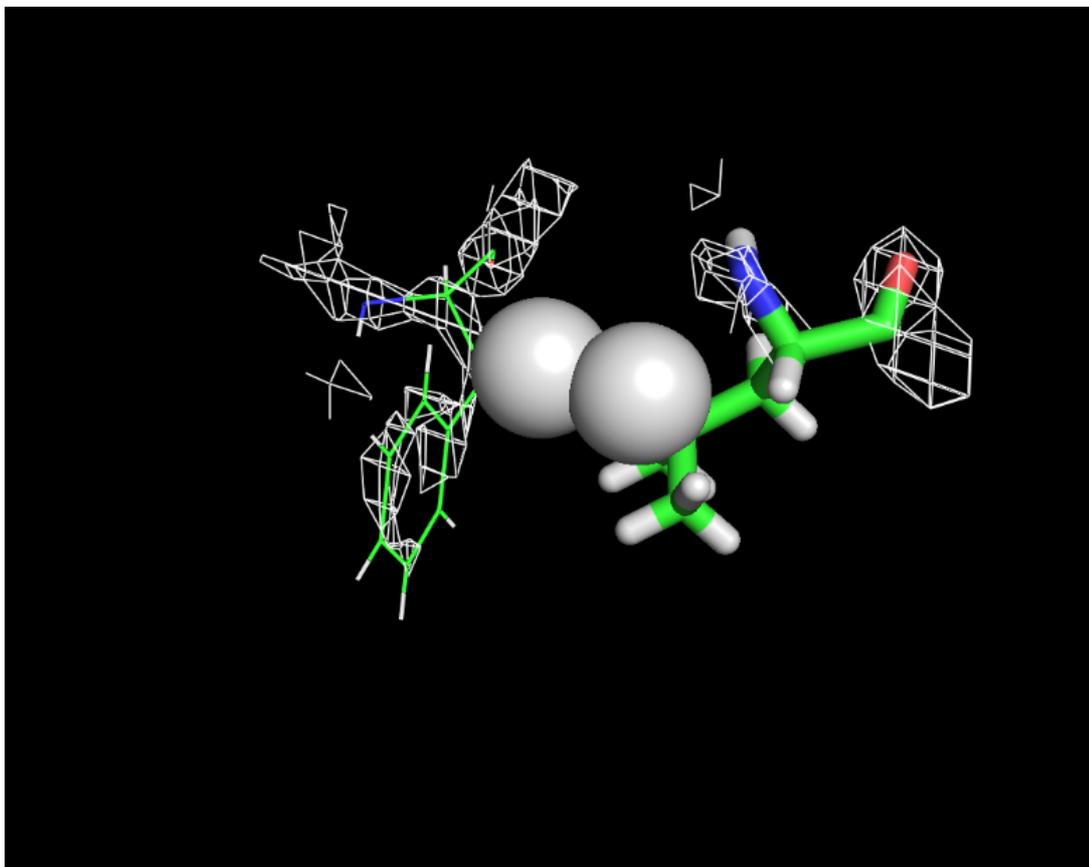


Рисунок 15. Ван-дер-Ваальсовы радиусы интересующих водородов и электронная плотность на уровне подрезки 3 для остатков LEU252 и PHE217.

C: VAL317

В цепи C остатки 316-320 не покрыты электронной плотностью уже на уровне подрезки 1.5 (см. рис. 16). Это значит, что их координаты были вписаны в модель исходя из соображений авторов. На рис. 10 и 11б как раз для этих остатков наблюдаются аномально высокие значения Real-space R-value и Z-score соответственно, что дополнительно говорит об их маргинальности.

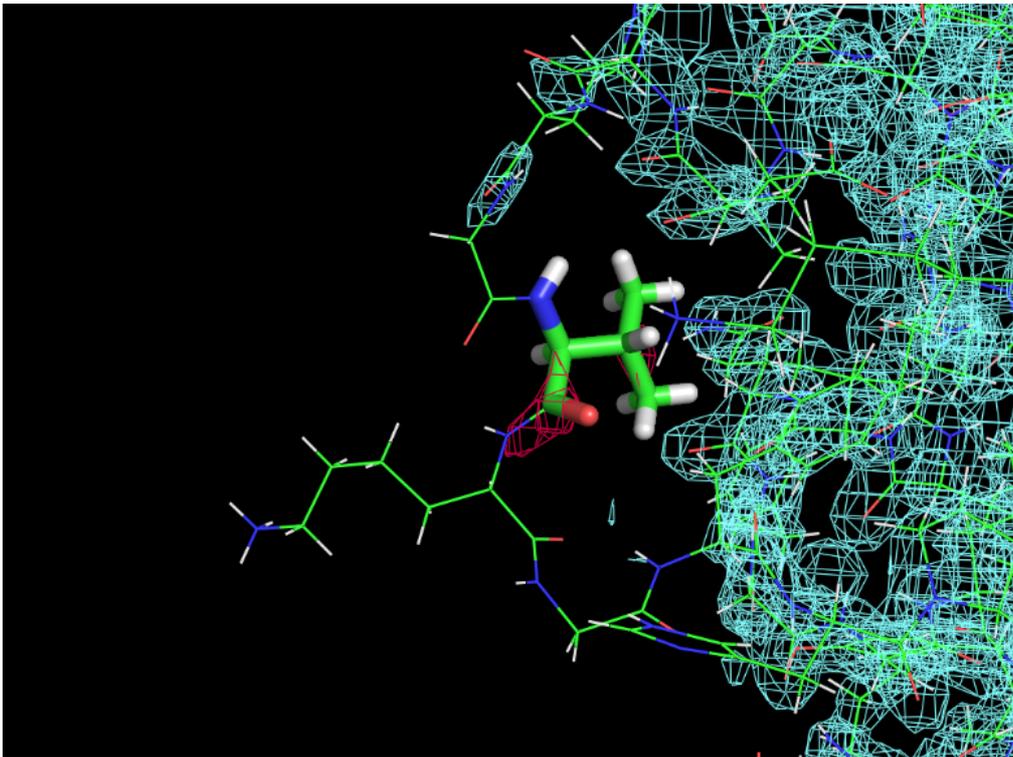


Рисунок 16. Покрытие электронной плотностью цепи C (аквамариновый) и отдельно – остатков 316-320 (малиновый) на уровне подрезки 1.5. Остаток VAL317 показан визуализацией sticks.

D: PRO225

Данный остаток довольно слабо покрыт электронной плотностью на уровне подрезки 2.5, однако соседние с ним остатки покрыты плотностью довольно хорошо (см. рис. 17), значения Z-score и Real-space R-value для него не выбиваются из нормы (-0.83 и 0.061 соответственно), и он не является аутлаером по карте Рамачандрана для цис-пролинов, так что скорее всего этот остаток расшифрован правомерно.

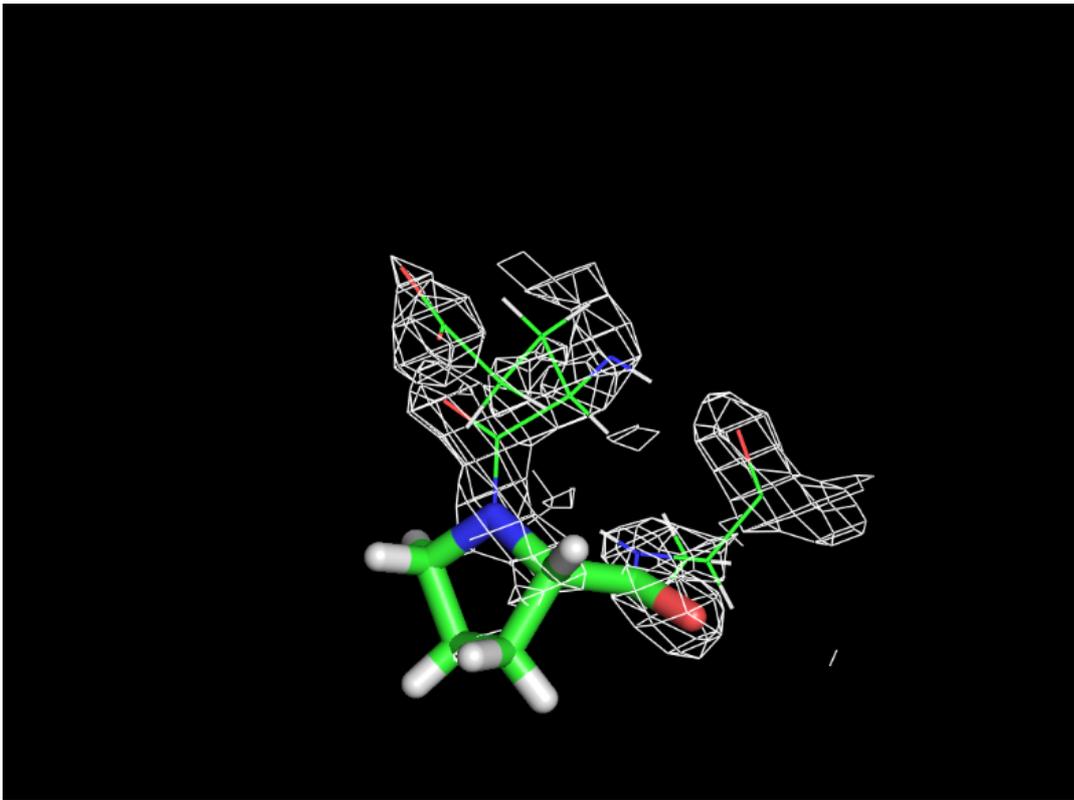


Рисунок 17. Покрытие электронной плотностью остатка PRO225 и соседних с ним на уровне подрезки 2.5.

D: ASN344

В неоптимизированной с помощью MolProbity структуре для данного остатка электронная плотность довольно точно покрывает координаты азота и кислорода в боковой цепи, однако не позволяет сказать, какой из них где, следовательно, нельзя однозначно сказать, произошла ли инверсия при его расшифровке или нет (см. рис. 18).

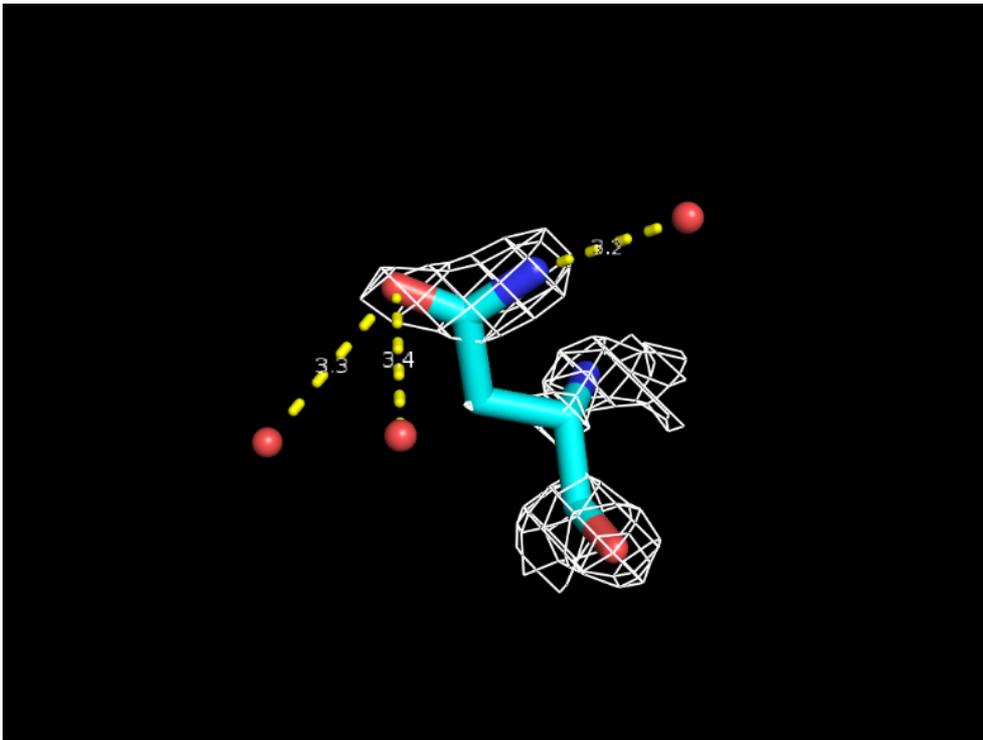


Рисунок 18. Остаток ASN344 цепи D, его электронная плотность на уровне подрезки 2.5 и водородные связи с ближайшими молекулами воды.

IV. PDB redo.

База PDB redo[8] содержит исправленные структуры из PDB. В том числе там есть и структура Zity. Согласно странице, посвященной структуре, оптимизация улучшила большинство показателей кроме связанных с водородными связями (см. рис. 19):

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
<i>R</i>	0.1524	0.1290
<i>R-free</i>	0.1798	0.1532
<i>Bond length RMS Z-score</i>	0.240	0.552
<i>Bond angle RMS Z-score</i>	0.571	0.737
Model quality (raw scores percentiles)		
<i>Ramachandran plot appearance</i>	61	85
<i>Rotamer normality</i>	54	91
<i>Coarse packing</i>	N/A	N/A
<i>Fine packing</i>	40	48
<i>Bump severity</i>	86	96
<i>Hydrogen bond satisfaction</i>	82	73
<i>WHAT_CHECK</i>	Report	Report

Рисунок 19. Отчет PDB redo об оптимизации структуры Zity.

Оптимизированную структуру можно скачать на той же странице. Структура была загружена на сервер MolProbity и подвергнута всем вышеупомянутым процедурам кроме инверсий цепей. По результатам проверки (см. Таблицу 4) видно, что, хотя у модели и улучшились clash score и общий MolProbity score и стало на одного меньше аутлаеров по картам Рамачандрана, состояние других показателей слегка ухудшилось, в том числе появились ранее отсутствовавшие “плохие связи”.

Таблица 4. Характеристика сервиса MolProbity исправленной PDB redo структуры Zity

All-Atom Contacts	Clashscore, all atoms:	0.93		100 th percentile* (N=809, 1.84Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	14	1.06%	Goal: <0.3%
	Favored rotamers	1276	96.52%	Goal: >98%
	Ramachandran outliers	4	0.24%	Goal: <0.05%
	Ramachandran favored	1647	97.98%	Goal: >98%
	MolProbity score [^]	0.80		100 th percentile* (N=11316, 1.84Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.13%	Goal: 0
	Bad bonds:	2 / 13360	0.01%	Goal: 0%
	Bad angles:	4 / 18113	0.02%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	8 / 72	11.11%	Expected: ≤1 per chain, or ≤5%

Выводы

1. Структура Zity расшифрована достаточно хорошо для заявленного разрешения 1.84 Å. Явные ошибки и/или грубые попытки “подогнать” качество структуры по каким-либо параметрам обнаружены не были.
2. В структуре присутствует несколько аутлаерных остатков по различным параметрам, в большинстве случаев их маргинальность объясняется отсутствием достаточной для их расшифровки электронной плотности и попытками авторов разместить боковые цепи остатков вручную. Этим же можно объяснить и пониженное качество участка цепи C, в которой авторам пришлось разместить в одном месте сразу несколько остатков. 224-й остаток глутамата, аутлаерный по карте Рамачандрана пре-пролиновых остатков во всех четырех цепях, скорее всего маргинальным не является и служит для стабилизации димеров в составе тетрамера за счет образования водородных связей с 65-м аргинином.
3. Оптимизация в сервере PDB redo улучшила модель по нескольким показателям, но ухудшила по ряду других; как раз эту оптимизацию можно назвать “подгонкой” качества (под параметр clashscore). Пользоваться оптимизированной структурой из PDB redo не рекомендую.

СПИСОК ИСТОЧНИКОВ

- [1] Yoshida H, Yamaji M, Ishii T, Izumori K, Kamitori S. Catalytic reaction mechanism of *Pseudomonas stutzeri* L-rhamnose isomerase deduced from X-ray structures, FEBS J. 277 1045-57 (2010)
- [2] Leang K, Takada G, Fukai Y, Morimoto Y, Granstrom TB & Izumori K (2004) Novel reactions of l-rhamnose isomerase from *Pseudomonas stutzeri* and its relation with d-xylose isomerase via substrate specificity. Biochim Biophys Acta 1674, 68–77.
- [3] Fenn TD, Ringe D & Petsko GA (2004) Xylose isomerase in substrate and inhibitor Michaelis states: atomic resolution studies of a metal-mediated hydride shift. Biochemistry 43, 6464–6474.
- [4] <https://www.ebi.ac.uk/pdbe/entry/pdb/3ity>
- [5] <http://molprobit.biochem.duke.edu/>
- [6] <http://eds.bmc.uu.se/eds/>
- [7] Bosco K Ho and Robert Brasseur, The Ramachandran plots of glycine and pre-proline, BMC Structural Biology (2005) 5:14
- [8] http://www.cmbi.ru.nl/pdb_redo/