

Отчёт по качеству расшифровки структуры белка  
4-оксалоокротонат таутомеразы *Methylobium*  
*petroleiphilum* PM1 (PDB код 4FAZ) методом  
рентгеноструктурного анализа

---

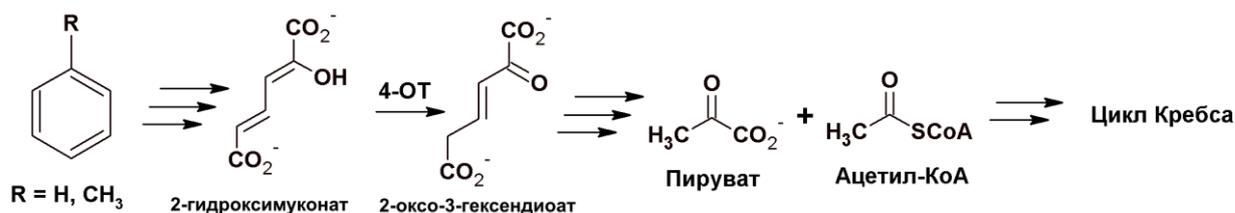
Выполнил:  
студент 4-го курса  
Гафуров Азамат

## Аннотация

В отчёте рассмотрены некоторые индикаторы качества модели структуры 4FAZ , а также проведён анализ индикаторов локального качества структуры для ряда аминокислотных остатков. В ходе работы над отчётом удалось познакомиться с возможностями , которые предоставляют серверы для оценки качества структур белковых молекул .

## Введение

4-оксалокротонат таутомераза (КФ 5.3.2.6., 4-ОТ) – фермент, катализирующий реакцию изомеризации 2-гидрокси-2,4-гександиенодиовой кислоты (2-гидроксимуконата) в 2-оксо-3-гексендионовую кислоту. Этот фермент входит в состав бактериального метаболического пути расщепления ароматических углеводов (Рис. 1)

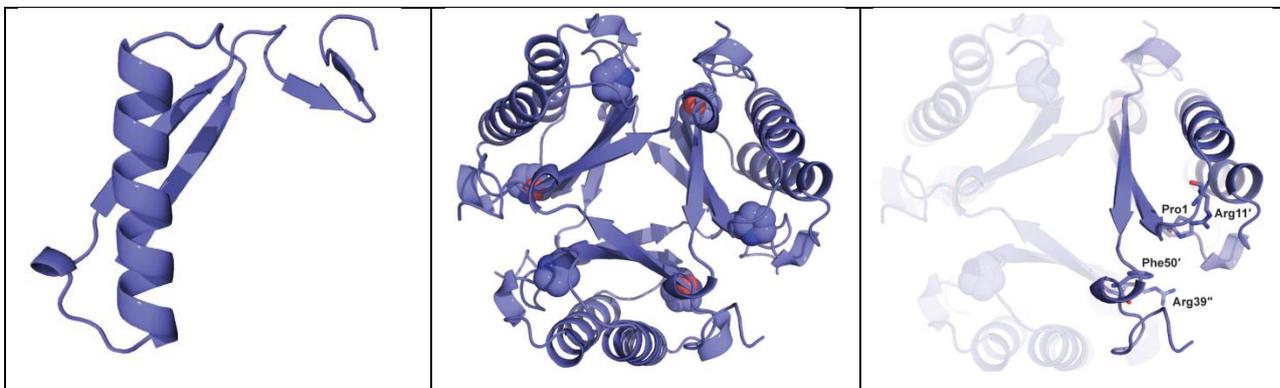


**Рисунок 1** Реакция, катализируемая 4-ОТ, в контексте деградации ароматических углеводов. Картинка взята из [1] и адаптирована.

Изучение 4-ОТ изначально проводили на бактерии *Pseudomonas putida* mt-2 [2], в результате чего выяснилось, что этот фермент объединяет в себе целый ряд исключительных свойств. Во-первых, с длиной последовательности всего 62 аминокислотных остатка, его мономер является самым маленьким среди всех известных ферментов [3]. Во-вторых, в состав его активного центра входит 3 аргинина – полярные аминокислоты, имеющие сильно положительный заряд. В-третьих, главным катализирующим основанием является терминальная аминокислота, а именно N-концевой пролин. И наконец, в-четвертых, 4-ОТ является каталитически совершенным ферментом ( $k_{cat}/K_M = 10^9 \text{ mM}^{-1}\text{s}^{-1}$ ) [4].

Фермент, выделенный из *Pseudomonas putida* mt-2, представляет собой гомогексамер, образующий  $\alpha/\beta$ -баррель. [3]. Мономеры взаимодействуют между собой таким образом, что формируется тример из димеров, между которыми располагаются по

два активных центра (Рис. 2). Помимо всего прочего, 4-ОТ является основателем суперсемейства таутомераз – TSF (Tautomerase Superfamily). Представители суперсемейства TSF имеют  $\beta$ - $\alpha$ - $\beta$  структурный мотив, а также разделяют сходный механизм кето-енольной таутмерии, в ходе которого аминотерминальный пролин играет основную каталитическую роль, выступая в реакции в качестве основания [5].



**Рисунок 2** Слева направо: пространственная структура мономера 4-ОТ; пространственная структура гексамера 4-ОТ, каталитический пролин изображен в шаровой модели; активный центр на стыке димеров, боковые цепи аминокислот, входящие в состав активного центра, представлены в каркасной модели [6].

В ходе поиска в базах данных новых представителей суперсемейства TSF в бактериях штамма *Chloroflexus aurantiacus* J-10-fl был открыт фермент, катализирующий каноническую 4-ОТ реакцию. Однако он оказался гетерогексамером, в котором чередуются  $\alpha$  и  $\beta$  субъединицы [7].

Подобный поиск был проведен в бактериях штамма *Methylibium petroleiphilum* strain PM1. Впервые их выделили из биофильтров очистных сооружений сточных вод нефтеперерабатывающих заводов. Они оказались способными эффективно метаболизировать некоторые компоненты бензина, в том числе ароматические соединения. Также впоследствии было обнаружено, что у них есть два альтернативных варианта пути мета-расщепления ароматических углеводородов. Чтобы разобраться в природе этих путей гены таутомеразы I и таутомеразы II, соответствующие альтернативных метаболическим путям, были заклонированы. Далее были экспрессированы белки и изучены их структурные и кинетические параметры [8].

В итоге выяснилось, что по структуре оба исследуемых фермента однозначно относятся к семейству 4-ОТ. Однако по кинетическим параметрам таутомераза II отличается от 4-ОТ более значительно, чем таутомераза I. В ходе работы также было показано, что таутомераза I может работать с каноническими субстратами 4-ОТ, в то

время как таутомераза II катализирует изомеризацию 2-гидроксимуконатов, замещенных в 3-м или 4-м положениях, но только не те, которые замещены в 5-м положении.

## Результаты и обсуждение

### Общая информация о модели и индикаторы её качества в целом и по отдельным аминокислотным остаткам

Асимметрическая единица 4FAZ представлена тремя идентичными цепями (А, В и С). Длина одной цепи составляет 62 аминокислотных остатка. Биологическая единица является собой гомогексамер. 4FAZ содержит лиганд – сульфат-анион ( $\text{SO}_4^{2-}$ ).

Модель структуры 4FAZ была получена в 2012 году [9]. Общая информация о модели 4FAZ систематизирована в таблице 1.

Таблица 1. Общая информация о модели 4FAZ		
Разрешение	1.57 Å	
Диапазон разрешения	47.22 Å – 1.57 Å	
Число измеренных рефлексов	163467	
Число использованных рефлексов	24295	
Полнота данных	98.5%	
Кристаллографическая группа	C 2 2 2 <sub>1</sub>	
Параметры ячейки	<b>Длина (Å)</b>	<b>Угол (°)</b>
	a = 60.17	$\alpha = 90.00$
	b = 76.17	$\beta = 90.00$
	c = 69.88	$\gamma = 90.00$

Для решения фазовой проблемы использовался метод молекулярного замещения (MR – Molecular Replacement) с использованием структуры 4-OT, полученной ранее (1BJP, [10]).

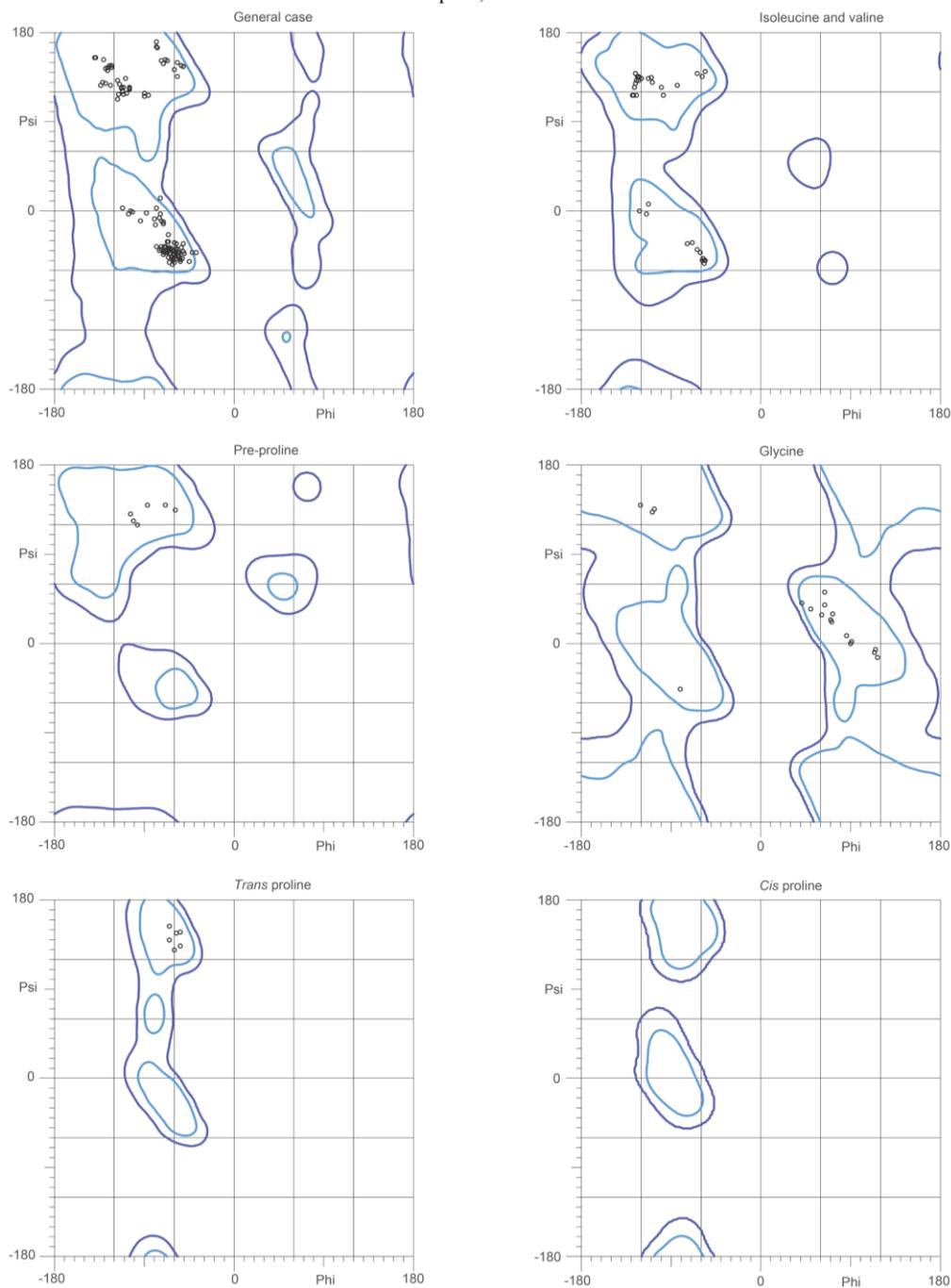
Для данной структуры R-фактор равняется 0.229. Это значительно превышает значение 0.157, что свидетельствует о плохом соответствии кристаллографической модели набору экспериментальных структурных факторов<sup>1</sup>.  $R_{\text{free}}$  равняется 0.275, что также является довольно слабым показателем<sup>2</sup>. Плохие значения R-фактора, и  $R_{\text{free}}$  вероятно связаны с тем, что для оптимизации была подобрана не совсем подходящая

<sup>1</sup> Существует эмпирическое правило, согласно которому хорошее значение R-фактора не должно сильно превышать [разрешение / 10].

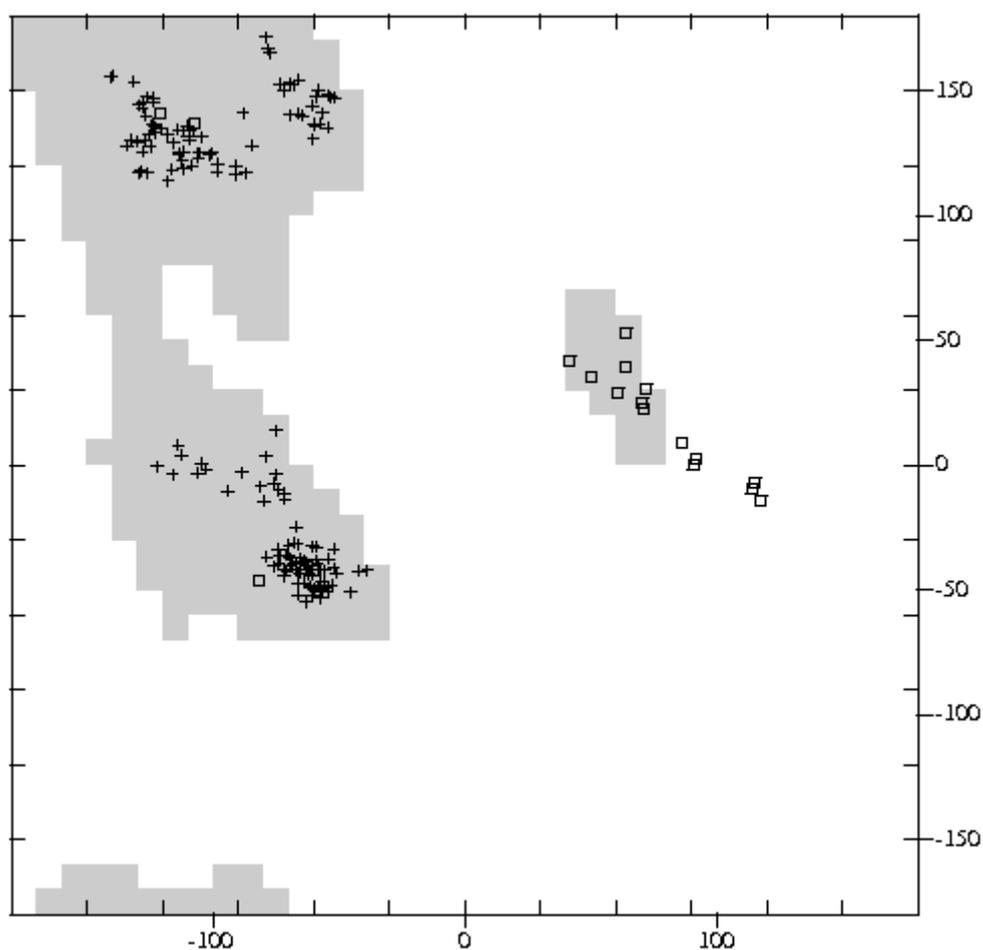
<sup>2</sup> Для структур с разрешением 2.0 Å и лучше эмпирическое правило гласит, что  $R_{\text{free}}$  не должен превышать значения [разрешение / 10] более, чем на 0.05.

структура (1BJP). Для структур с разрешением 2.0 Å и лучше, что меньше 0.25, значит значение хорошее.  $R_{free}$  – R-фактор = 0.046, что говорит об отсутствии переоптимизации.

Для построения карты Рамачандрана были использованы серверы MolProbity и EDS. На обеих картах Рамачандрана, показывающих положение пар торсионных углов  $\phi/\psi$  полипептидной цепи, все остатки лежат в допустимой области (Рис. 3,4).



**Рисунок 3** Карта Рамачандрана для структуры белка 4FAZ. Множество пар торсионных углов  $\phi/\psi$  для всех типов аминокислотных остатков разбито для удобства на 6 групп. Карта построена сервисом MolProbity.



**Рисунок 4** Карта Рамачандрана для структуры белка 4FAZ, построенная с помощью сервиса EDS.

Пространственный R-фактор (RSR) – показатель того, насколько построенная модель соответствует « экспериментальной» электронной плотности. Среднее значение R-фактора (RSR-фактора) равно 0.117 (со стандартным отклонением 0.058), что говорит о неплохом соответствии атомов модели электронной плотности, определенной в эксперименте.

**Таблица 2. Величины некоторых индикаторов качества структуры в целом для модели 4FAZ**

R-фактор	0.229
R <sub>free</sub>	0.275
R <sub>free</sub> – R	0.046
RSR-фактор (среднее значение и стандартное отклонение)	0.117 (0.058)
Число маргиналов по карте Рамачандрана	0
Число остатков в предпочитаемой области карты Рамачандрана	178/180
Валентные углы, существенно отклоняющиеся от теории	1/2005
Ковалентные связи, существенно отклоняющиеся от теории	4/1448
Необычные .ротамеры	1

На рисунке 5 представлены графики значений RSR-фактора для остатков в цепях А, В и С белка. Максимальные значения наблюдаются для остатков Arg62, Gly61 и Glu48 в цепи А, Asn37 в цепи В, Arg62 и Glu36 в цепи С (Таблица 3).

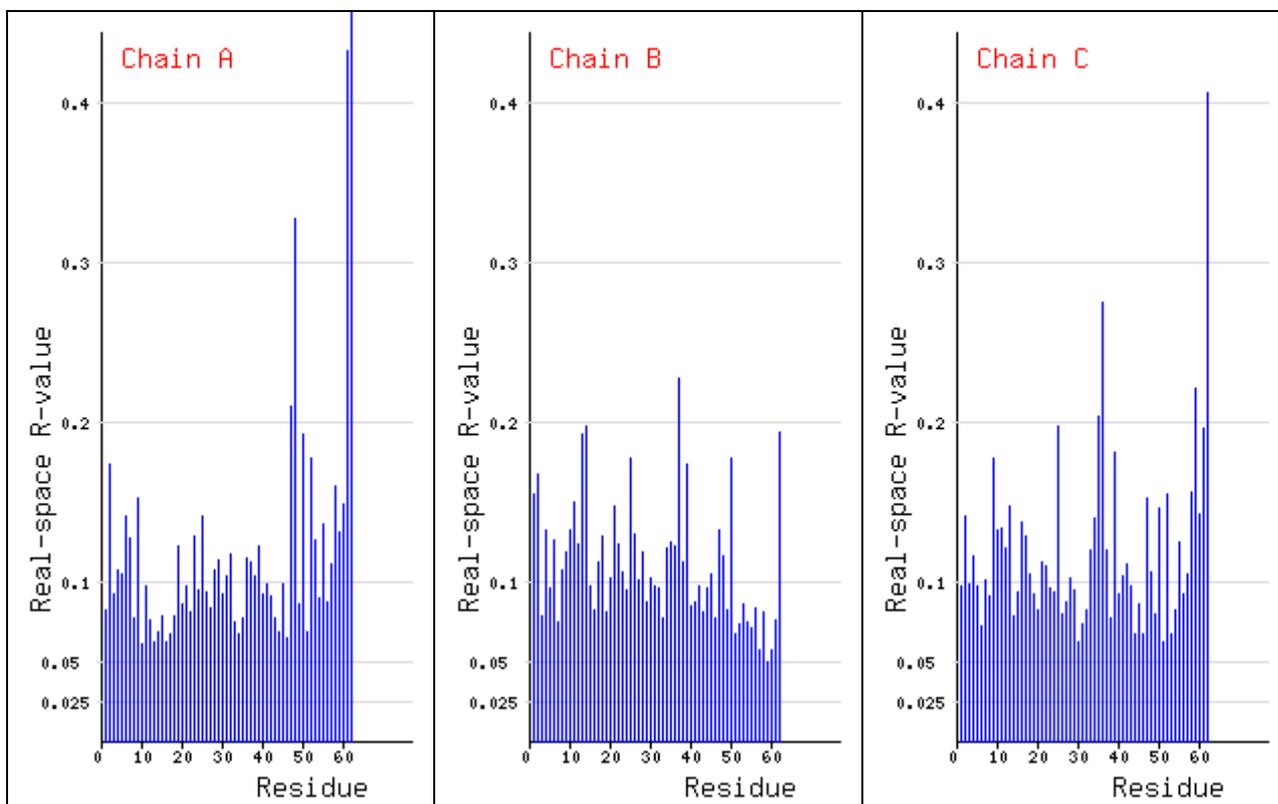


Рисунок 5 Значения RSR-фактора для аминокислотных остатков А, В и С цепей в модели 4FAZ.

Z-score для RSR-фактора позволяет сравнить электронную плотность со средним значением функции электронной плотности для остатков того же типа в структурах с разрешением 1.40-1.60 Å. Значения этого параметра для Gly61 и Arg62 в цепи А равны 7.19 и 6.42, соответственно, 3.43 для Trp50 цепи В и 5.16 для Arg62 цепи С (Рис. 6). Высокие значения Z-score ( $Z > 2$ ) свидетельствуют о том, что остаток плохо вписан в электронную плотность по сравнению с другими структурами с близким разрешением.

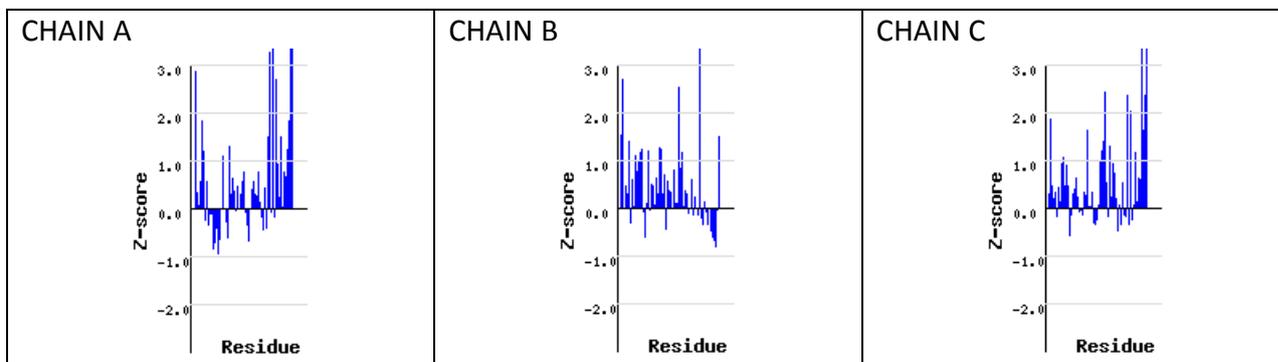


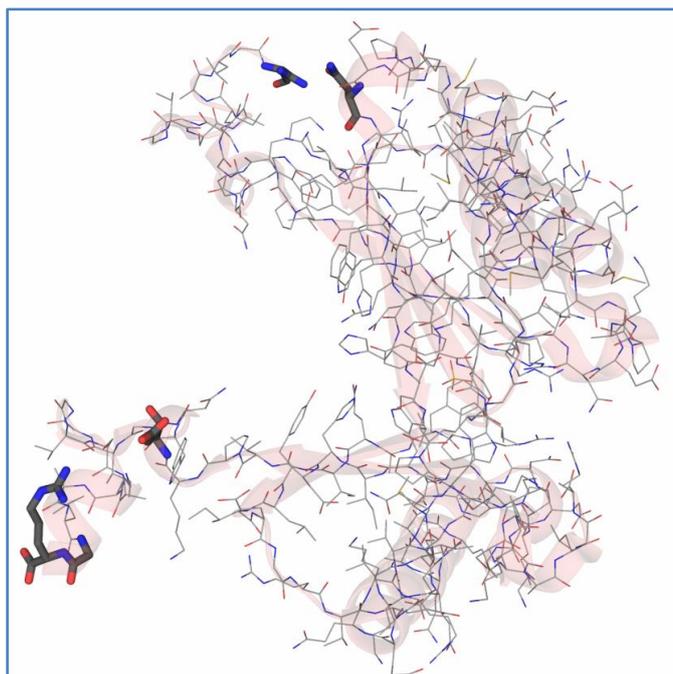
Рисунок 6 Значения Z-score для аминокислотных остатков А, В и С цепей в модели 4FAZ. У 8.06% остатков Z-score > 2.

В таблице 3 приведено 12 примеров маргинальных остатков.

<b>Таблица 3. Примеры маргинальных аминокислотных остатков</b>	
<b>Маргинальный остаток</b>	<b>Критерий выбора</b>
Arg62 A	RSR = 0.479; Z-score(RSR) = 6.42
Gly61 A	RSR = 0.433; Z-score(RSR) = 7.19
Glu48 A	RSR = 0.328; Z-score(RSR) = 3.26
Asn37 B	RSR = 0.227; Z-score(RSR) = 2.53
Arg62 C	RSR = 0.406; Z-score(RSR) = 5.16
Glu36 C	RSR = 0.275
Trp50 B	Z-score(RSR) = 3.43
Lys68 C	Плохие $\chi$ углы: 227.2,205.5,212.6,64.4
His43 A	SD длины связи (CD2--NE2) = 5.1 $\sigma$
Trp41 B	SD длины связи (CD2--CE2) = 4.8 $\sigma$
His43 C	SD длины связи (ND1--CE1) = 5.4 $\sigma$
Glu9 B	SD угла (CA-CB-CG) = 4.3 $\sigma$

### Детальный анализ 5 маргинальных аминокислотных остатков

Далее был проведен анализ 5 аминокислотных остатков, вызвавших наибольшее подозрение. Такими оказались остатки, маргинальные по параметрам RSR и Z-score(RSR) (Таблица 3). Пространственное расположение выбранных маргинальных остатков показано на рисунке 7.



**Рисунок 7** Пространственное расположение выбранных маргинальных остатков в структуре 4FAZ

Остатки Arg62C (Рис. 8), Gly61A и Arg62A (Рис. 9), Asn37B (Рис. 10), Glu48A (Рис. 11), в модели 4FAZ и изображения электронной плотности при разных уровнях подрезки.

При уровне  $1.0\sigma$  электронная плотность вокруг атомов боковой группы аминокислотных остатков Arg62C, Gly61A и Arg62A не различима. Для Arg62C видна на уровне  $0.5\sigma$ , а для Gly61A и Arg62A только при  $0.25\sigma$ ; при этом положение атомов боковой группы Arg62C плохо соответствует электронной плотности.

Для Asn37B и Glu48A, имеющие менее выделяющиеся значения RSR и Z-score(RSR), картина соответствия электронной плотности боковым группам более приемлема, но, тем не менее, недостатки довольно заметны.

### Arg62C

- RSR = 0.406;
- Z-score (RSR) = 5.16.

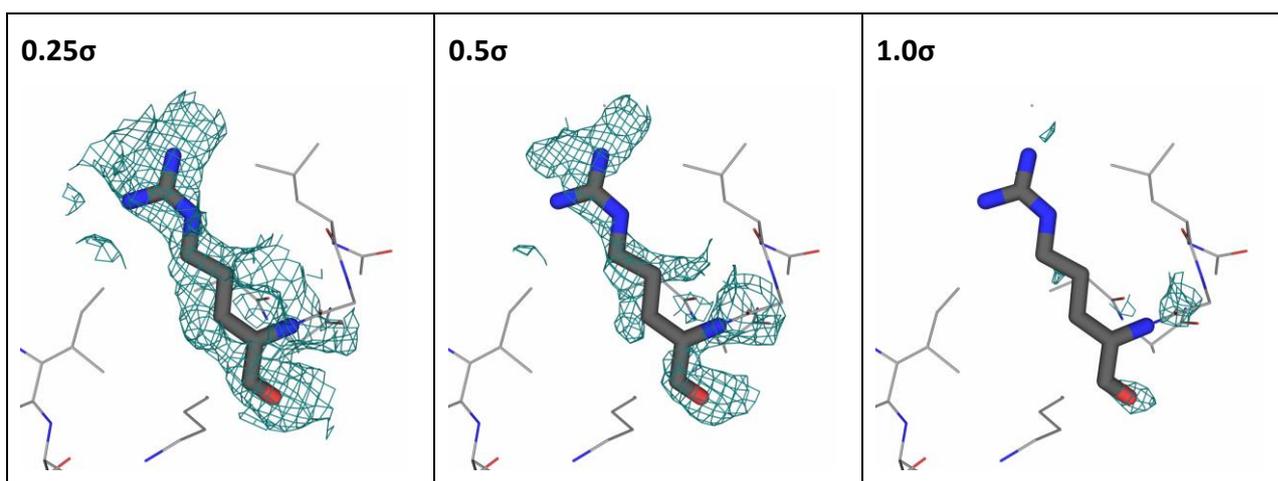


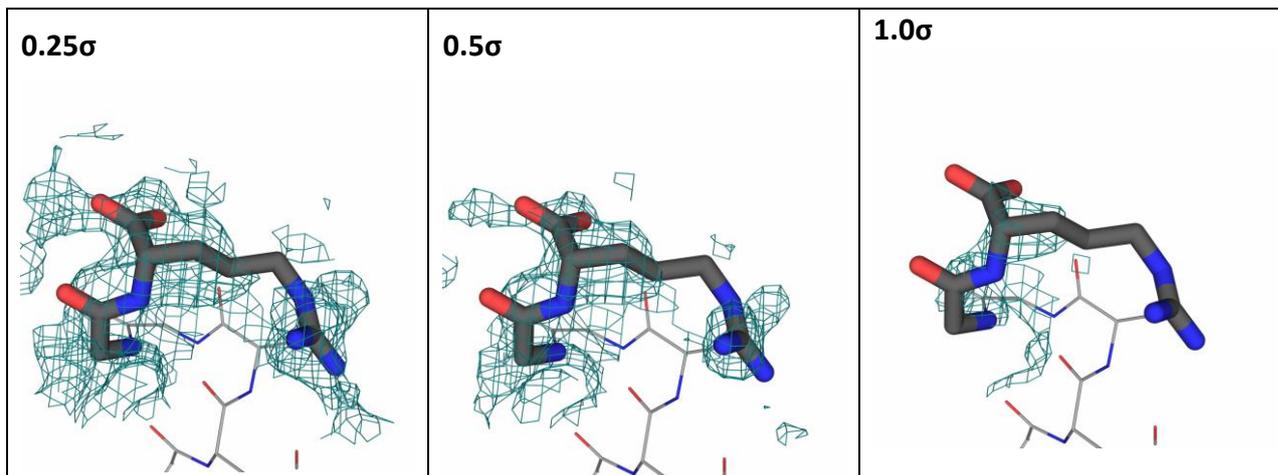
Рисунок 8 Остаток Arg62 цепи C в модели 4FAZ и изображение электронной плотности при разных уровнях подрезки.

### Arg62A

- RSR = 0.479;
- Z-score (RSR) = 6.42.

### Gly61A

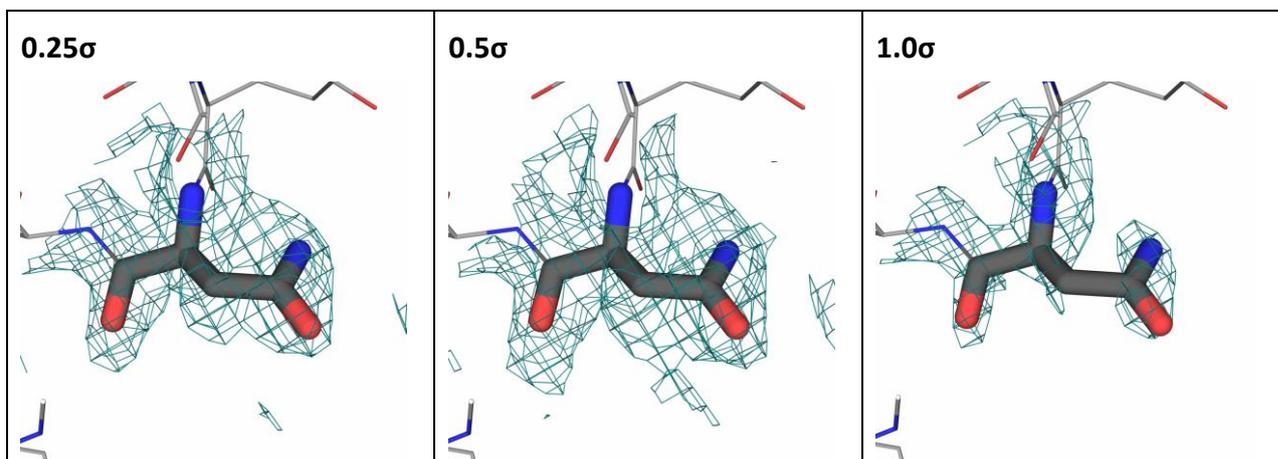
- RSR = 0.433;
- Z-score (RSR) = 7.19.



**Рисунок 9** Остатки Gly61A и Arg62A цепи A и изображение электронной плотности при разных уровнях подрезки.

### Asn37B

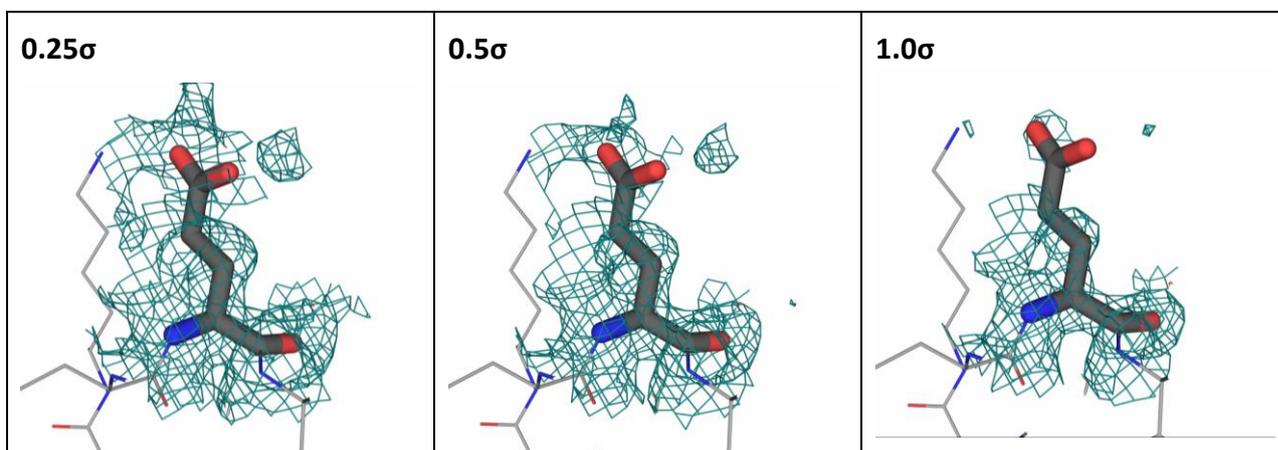
- RSR = 0.227;
- Z-score (RSR) = 2.53.



**Рисунок 10** Остаток Asn37 цепи B в модели 4FAZ и изображение электронной плотности при разных уровнях подрезки.

### Glu48A

- RSR = 0.328;
- Z-score (RSR) = 3.26.-



**Рисунок 10** Остаток Glu48 цепи A в модели 4FAZ и изображение электронной плотности при разных уровнях подрезки.

## Сравнение модели из PDB с моделью из PDB\_REDO

Было проведено сравнение модели 4FAZ из PDB с оптимизированной моделью из PDB\_REDO. Значения некоторых параметров для оценки качества двух моделей приведены в таблице 5.

Таблица 5. Параметры модели 4FAZ из PDB и из PDB_REDO		
Параметр	PDB	PDB_REDO
R-фактор	0.229	0.204
$R_{\text{free}}$	0.275	0.233
$R_{\text{free}} - R$	0.046	0.029
Число маргиналов по карте Рамачандрана	0	0
Валентные углы, существенно отклоняющиеся от теории	1/2005	4/2005
Ковалентные связи, существенно отклоняющиеся от теории	4/1448	6/1448
Необычные ротамеры	1	3

Значения R-фактора и  $R_{\text{free}}$  улучшились, и разница их также уменьшилась. Значит, новая модель лучше подогнана под экспериментальные данные. Но при этом увеличилось число маргинальных аминокислотных остатков по валентным углам и длине ковалентных связей, а также увеличилось число необычных ротамеров.

4FAZ представляет собой модель высокого разрешения. Однако качество её разрешения можно оценить как «среднее», о чём свидетельствуют индикаторы оценки качества этой структуры. Это вероятно связано с неудачным выбором модели для оптимизации. Из наблюдаемых в 4FAZ аномалий можно отметить высокие показатели RSR-фактора и Z-score (RSR). В модели структуры удалось также обнаружить остатки, аномальные по длинам и углам ковалентных связей и являющие собой необычные ротамеры. Стоит отметить, что сервер PDB\_REDO значительно улучшил такие показатели модели, как R-фактор и  $R_{\text{free}}$ , но при этом появились маргинальные остатки по другим параметрам.

## Список литературы

- [1] J. P. Huddleston, E. a. Burks, and C. P. Whitman, "Identification and characterization of new family members in the tautomerase superfamily: Analysis and implications," *Arch. Biochem. Biophys.*, vol. 564, pp. 189–196, 2014.

- [2] C. P. Whitman, "The 4-oxalocrotonate tautomerase family of enzymes: How nature makes new enzymes using a beta-alpha-beta structural motif," *Arch. Biochem. Biophys.*, vol. 402, pp. 1–13, 2002.
- [3] L. Chen, G. Kenyon, F. Curtint, S. Harayamat, M. Bembenekt, H. Gholamhossein, and C. Whitman, "4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer," *J. Chem. Inf. Model.*, no. The xylH gene encoding 4-oxalocrotonate tautomerase (4-OT) has been located on a subclone of the *Pseudomonas putida* mt-2 TOL plasmid pWW0 and inserted into an *Escherichia coli* expression vector. Several of the genes of the metafission pathway encoded by p, pp. 17716–21, 1992.
- [4] C. P. Whitman, B. a. Aird, W. R. Gillespie, and N. J. Stolowich, "Chemical and enzymic ketonization of 2-hydroxymuconate, a conjugated enol," *J. Am. Chem. Soc.*, vol. 113, no. c, pp. 3154–3162, 1991.
- [5] J. J. Almrud, A. D. Kern, S. C. Wang, R. M. Czerwinski, W. H. Johnson, A. G. Murzin, M. L. Hackert, and C. P. Whitman, "The Crystal Structure of YdcE, a 4-Oxalocrotonate Tautomerase Homologue from *Escherichia coli*, Confirms the Structural Basis for Oligomer Diversity †, ‡," *Biochemistry*, vol. 41, pp. 12010–12024, 2002.
- [6] C. P. Whitman, "Functional ( Mis ) Assignment in the Tomaymycin Biosynthetic Pathway," pp. 111–128, 2013.
- [7] E. A. Burks, C. D. Fleming, A. D. Mesecar, C. P. Whitman, and S. D. Pegan, "Kinetic and Structural Characterization of a Heterohexamer 4- Oxalocrotonate Tautomerase from *Chloroflexus aurantiacus* J-10-fl: Implications for Functional and Structural Diversity in the Tautomerase Superfamily," *Biochemistry*, pp. 5016–5027, 2010.
- [8] C. R. Terrell, E. A. Burks, C. P. Whitman, and D. W. Hoffman, "Structural and kinetic characterization of two 4-oxalocrotonate tautomerase in *Methylibium petroleiphilum* strain PM1.," *Arch Biochem Biophys*, 2013.
- [9] C. Terrell, D. Hoffman, and C. Whitman, "Kinetic and structural characterization of the 4-oxalocrotonate tautomerase isozymes from *Methylibium petroleiphilum*." *PDB article*, 2012.
- [10] A. Taylor, R. Czerwinski, W. Johnson Junior, C. Whitman, and M. Hackert, "CRYSTAL STRUCTURE OF 4-OXALOCROTONATE TAUTOMERASE INACTIVATED BY 2-OXO-3-PENTYNOATE AT 2.4 ANGSTROMS RESOLUTION." *PDB article*, 1998.