

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Имени М.В. ЛОМОНОСОВА

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Отчёт по качеству расшифровки структуры с PDB-кодом 1E5K,
Полученной методом рентгеноструктурного анализа**

Работа студента 4 курса
Сафронова Григория

**Москва
2017 г.**

АННОТАЦИЯ

В этой работе представлен анализ качества расшифровки структуры 1E5K, рассмотрены основные индикаторы качества модели, а также несколько маргинальных остатков.

ВВЕДЕНИЕ

Для анализа был выбран белок *Molybdopterin-synthase adenyltransferase* из организма *E.coli*. Этот фермент (EC 2.7.7.80) катализирует реакцию:



Этот фермент аденирует С-конец небольшой субъединицы молибдоптеринсинтазы[1].

РЕЗУЛЬТАТЫ

1. Общая информация о модели

Структура 1E5K была создана в 2000 году и описана в статье “Crystal Structure of the Molybdenum Cofactor Biosynthesis Protein MobA from *Escherichia coli* at Near-Atomic Resolution” авторов Stevenson, C.E.M.; Sargent, F.; Buchanan, G.; Palmer, T.; Lawson, D.M.

В таблице 1 представлена краткая характеристика структуры белка[2].

Таблица 1. Основные параметры структуры 1E5K.

Параметр	Значение
Разрешение структуры	1,35 Å
Количество измеренных рефлексов	38232
Полнота данных	97,6%
Диапазон разрешений	36.66 - 1.35 Å
Тип симметрии	P 2 ₁ 2 ₁ 2
Параметры кристаллографической ячейки	a=76.57 Å, b=41.76 Å, c=54.53 Å, вектора ортогональны
Метод решения фазовой проблемы	Многоволновое аномальное рассеяние

2. Значения индикаторов качества модели в целом

На рисунке 1 показаны значения индикаторов относительно всех структур PCA. Из всех значений хорошими можно назвать только Clashscore и Sidechain outliers.

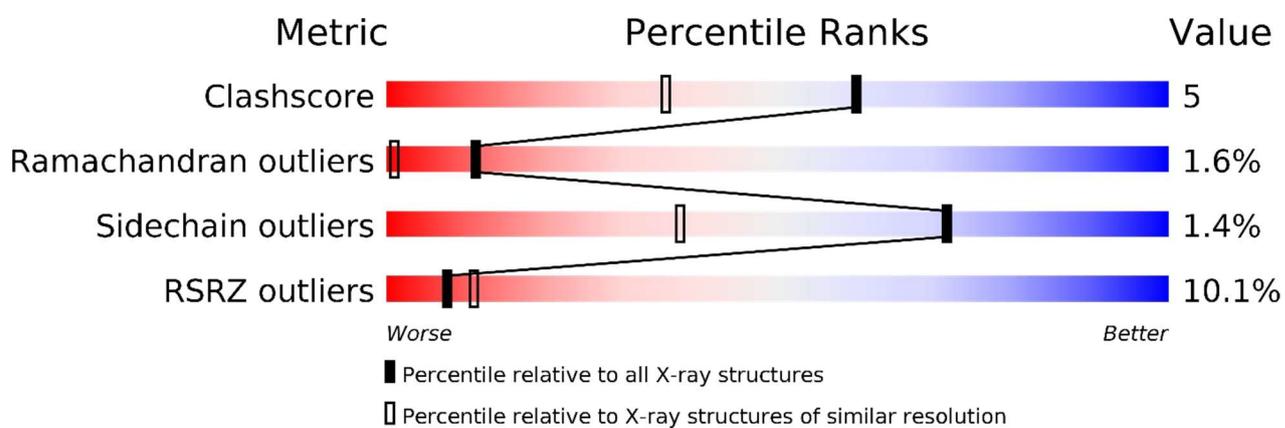
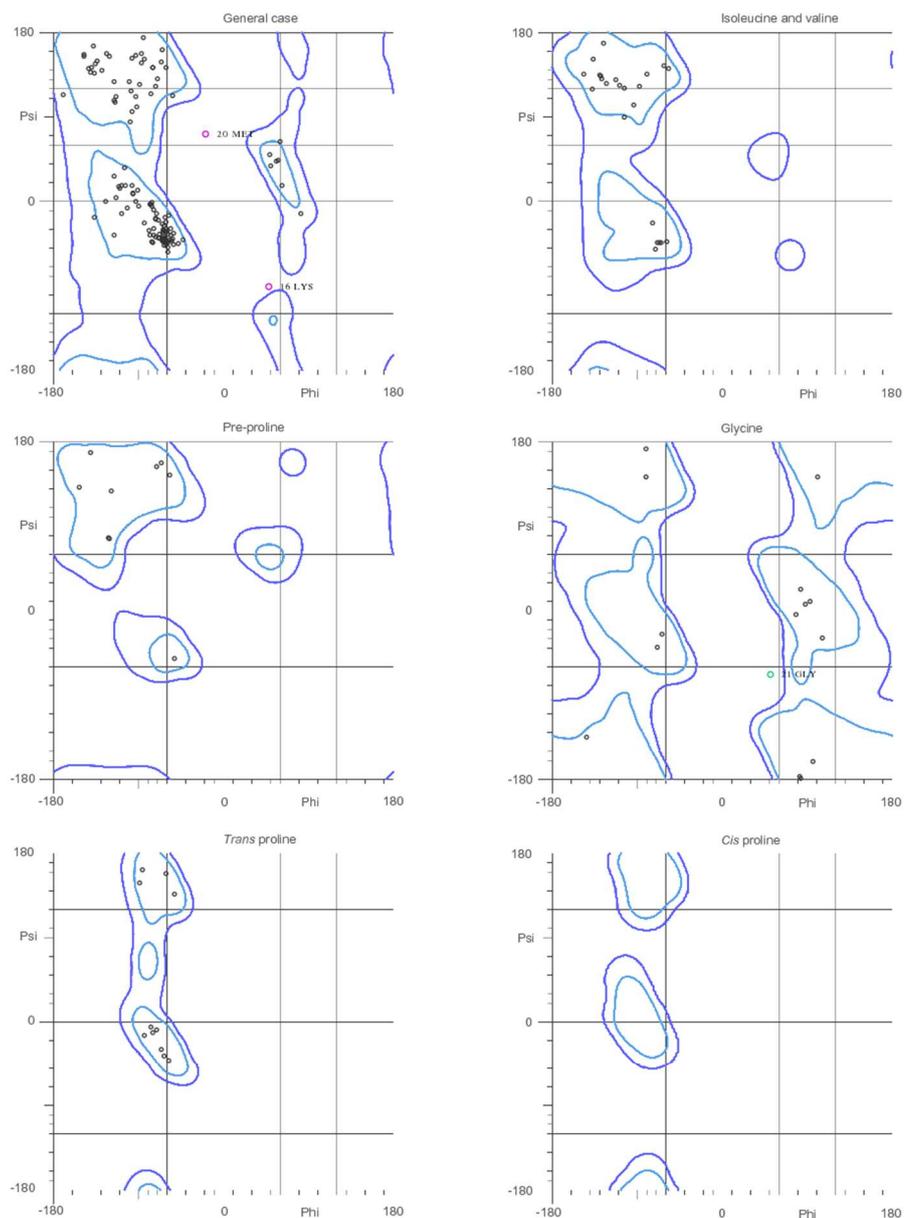


Рис. 1 Значение различных индикаторов качества модели.

Значения R-фактора и R-free составляет 0,184 и 0,213, соответственно [3]. Так как значения R-free <25%, а R-free – R <10%, можно сделать вывод, что у модели высокое качество.

Карта Рамачандрана:



95.2% (177/186) of all residues were in favored (98%) regions.
98.4% (183/186) of all residues were in allowed (>99.8%) regions.

There were 3 outliers (phi, psi):

16 LYS (48.1, -91.4)
20 MET (-19.2, 72.0)
21 GLY (51.4, -68.9)

Рис. 2 Карта Рамачандрана для 1e5k

На рисунке 2 можно заметить 2 маргинальных остатка в запрещенной области и еще 1 в неблагоприятной области.

Таблица 2. Критерии качества модели, выданные сервисом MolProbity. Зеленым цветом выделены значения, удовлетворяющие условиям хорошей модели, желтым – значения, отклоняющиеся от условий хорошей модели и красным – значения, не удовлетворяющие условиям хорошей модели.

All-Atom Contacts	Clashscore, all atoms:	5.29		88 th percentile* (N=466, 1.35Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	1	0.69%	Goal: <0.3%
	Favored rotamers	139	96.53%	Goal: >98%
	Ramachandran outliers	3	1.61%	Goal: <0.05%
	Ramachandran favored	177	95.16%	Goal: >98%
	MolProbity score [^]	1.62		70 th percentile* (N=3057, 1.35Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	6 / 1480	0.41%	Goal: 0%
	Bad angles:	21 / 2021	1.04%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 11	0.00%	Expected: ≤1 per chain, or ≤5%

3. Анализ маргинальных остатков

Для поиска маргинальных остатков использовался график из PDBReport[2], который демонстрирует геометрические отклонения и несоответствия электронной плотности (рис. 3)

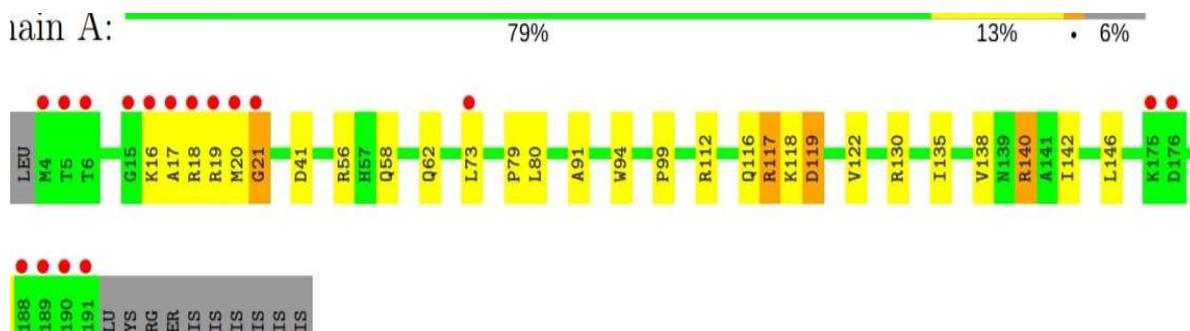


Рис. 3 Качество отдельных остатков в структуре. Зеленые – не маргинальные, желтые – обладают маргинальностью по 1 критерию, оранжевые – маргинальность по 2 критериям, красные точки – соответствующие остатки имеют RSR Z-score > 2.

Таблица 3. Некоторые маргинальные остатки в структуре 1E5K.

№	Остаток	Критерий маргинальности
1	GLY21	RSRZ, Карта Рамачандрана
2	MET20	RSRZ, Карта Рамачандрана
3	ALA17	RSRZ
4	THR5	RSRZ
5	LYS16	RSRZ, Карта Рамачандрана
6	ASP119	Ротамер
7	ARG140	Выброс по углу

4. Анализ пяти маргинальных остатков.

На рисунке 4 изображен остаток Gly21, маргинальный по RSRZ и карте Рамачадрана. Электронная плотность не кладется на него полностью. Скорее всего данный остаток является маргинальным из-за особенностей структуры.

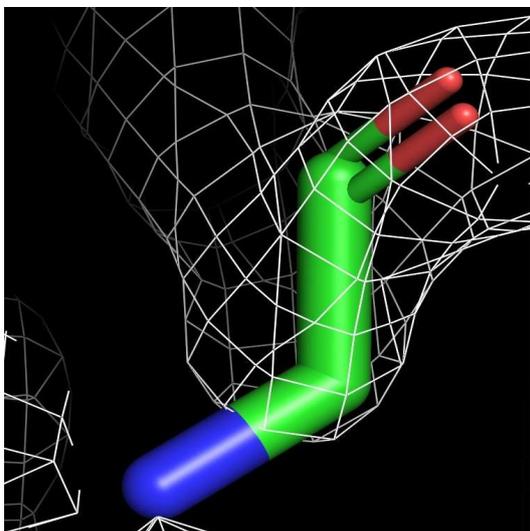


Рис. 4 Остаток Gly21 и электронная плотность. Уровень подрезки 1σ .

На рисунке 5 изображен маргинальный остаток, который почти-что не попадает на электронную плотность. Скорее всего из-за ошибки расшифровки данных.

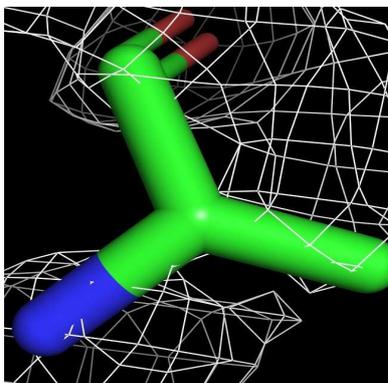


Рис. 5 Остаток Ala17 и электронная плотность. Уровень подрезки 1σ .

Остаток Thr 5 (рис. 6) отнесен к маргинальным скорее всего из-за ошибок расшифровки данных.

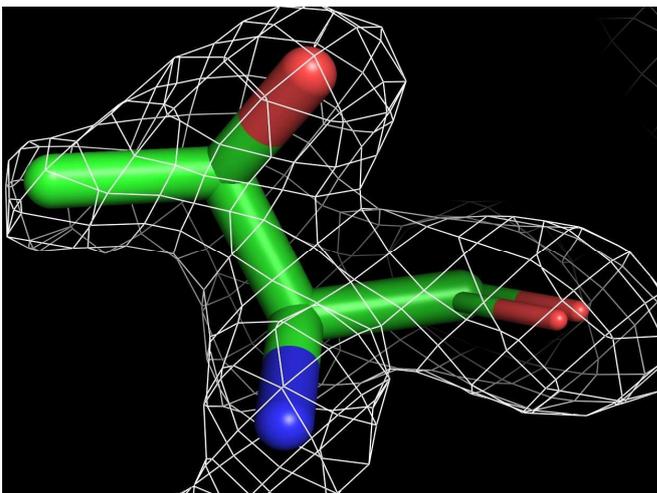


Рис. 6 Остаток Thr 5 и электронная плотность. Уровень подрезки 1σ .

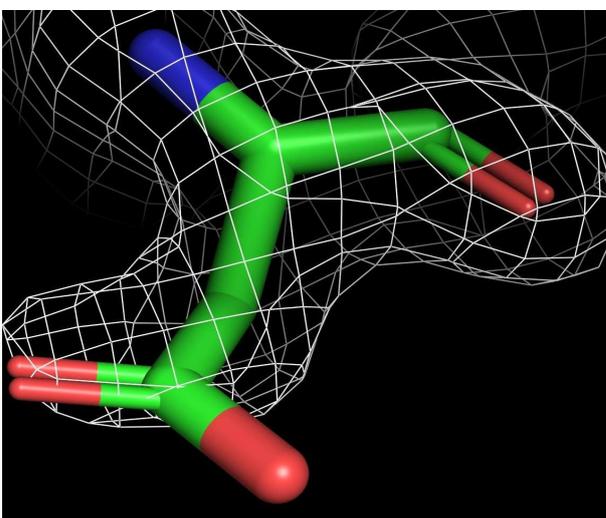


Рис. 7 Остаток Asp 119 и электронная плотность. Уровень подрезки 1σ .

На рисунке 8 изображен маргинальный остаток с пустыми карманами в электронной плотности. Остаток отнесен к маргинальным из-за ошибок расшифровки данных

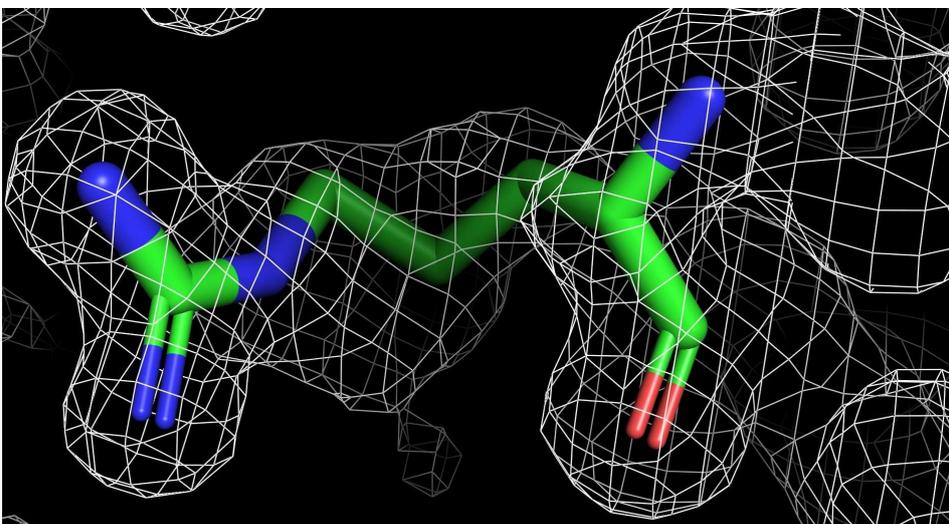


Рис. 8 Остаток Arg 140 и электронная плотность. Уровень подрезки 1σ .

5. Сравнение моделей из PDB и PDB-REDO

В базе PDB_redo [4] проведена оптимизация структуры 1e5k, что не привело к сильному улучшению качества вписывания остатков в энергетическую плотность, хотя значения R-фактор и R-free стали заметно меньше.

Таблица 4. Оптимизация 1e5k.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
R	0.1709	0.1501
R-free 	0.2324	0.1762
Bond length RMS Z-score	0.557	0.704
Bond angle RMS Z-score	1.295	0.887
Model quality (raw scores percentiles)		
Ramachandran plot appearance	43	43
Rotamer normality	97	94
Coarse packing	N/A	N/A
Fine packing	89	93
Bump severity	72	75
Hydrogen bond satisfaction	46	40
WHAT_CHECK	Report	Report

ВЫВОДЫ

Качество структуры трудно назвать хорошим, хотя оно обладает высоким разрешением и малым количеством несоответствий по карте Рамачандрана. Для биоинформатических исследований стоит использовать улучшенную структуру из PDB-REDO.

СПИСОК ЛИТЕРАТУРЫ

1. Stevenson, C.E.M.; Sargent, F.; Buchanan, G.; Palmer, T.; Lawson, D.M. (2000) CRYSTAL STRUCTURE OF THE MOLYBDENUM COFACTOR BIOSYNTHESIS PROTEIN MOBA (PROTEIN FA) FROM ESCHERICHIA COLI AT NEAR ATOMIC RESOLUTION
2. <https://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1e5k>
3. <https://www.ebi.ac.uk/pdbe/entry/pdb/1e5k>
4. <https://pdb-redo.eu/db/1e5k>