

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

имени М.В.ЛОМОНОСОВА

---

***ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ***

**Отчёт по качеству расшифровки структуры с PDB-кодом 1E0,  
полученной методом рентгеноструктурного анализа**

Работа студентки 4 курса

Макаровой Надежды

Москва

2017 г.

## СОДЕРЖАНИЕ

Сокращения.....	2
Аннотация .....	3
Введение.....	3
Результаты и обсуждение .....	5
Общая информация о модели .....	5
Индикаторы качества модели .....	5
R-factor и R-free .....	5
Пространственный R-фактор.....	6
Z-score.....	7
Карта Рамачандрана.....	8
Статистика сервиса MolProbity.....	9
Поиск и анализ маргинальных остатков в структуре.....	11
PDB-REDO.....	21
Выводы .....	23
Заключение.....	23
Список литературы.....	24

## СОКРАЩЕНИЯ

**EDS**, Electron Density Server

**PDB**, Protein Data Bank

**RSR**, пространственный R-фактор (Real Space R-factor)

## АННОТАЦИЯ

В данной работе представлен отчет о качестве структуры с PDB-кодом 1IE0, расшифрованной методом рентгеноструктурного анализа. Был произведен поиск, систематизация и интерпретация информации по оценке качества модели 1IE0. Особое внимание уделяется локальным индикаторам качества структуры и анализу маргинальных остатков. По результатам работы сделан вывод о качестве модели 1IE0.

## ВВЕДЕНИЕ

Для анализа был выбран белок S-рибозилгомоцистеин лиаза (далее LuxS) из организма грам-положительной почвенной бактерии сенной палочки *Bacillus subtilis*. Этот фермент (EC 4.4.1.21) участвует в интересном явлении, которое называется чувство кворума (quorum sensing). Это способность некоторых бактерий общаться и координировать своё поведение за счёт секреции молекулярных сигналов. Главное назначение этого процесса - координировать определённое поведение или действия между бактериями того же вида или подвида в зависимости от плотности их популяции. Quorum sensing происходит при регуляции таких процессов, как агрегация клеток, споруляция, биолюминесценция и др. Большинство видов бактерии обладают способностью к quorum sensing, Существует много разных способов «общения» между бактериями, однако самый распространенный осуществляется с помощью LuxS.

Чтобы понять какую роль играет белок LuxS в этом процессе, посмотрим на реакцию, которую он катализирует (Рисунок 1). Он расщепляет S-рибозил-L-гомоцистеин (далее SRH) с образование гомоцистеина и 4,5-дигидрокси-2,3-пентандиона (далее DPD), который является предшественником ключевой сигнальной молекулы AI-2 (autoinducer), с помощью которого бактерии могут «оценить» плотность популяции и отреагировать путем изменения экспрессии генов[1].

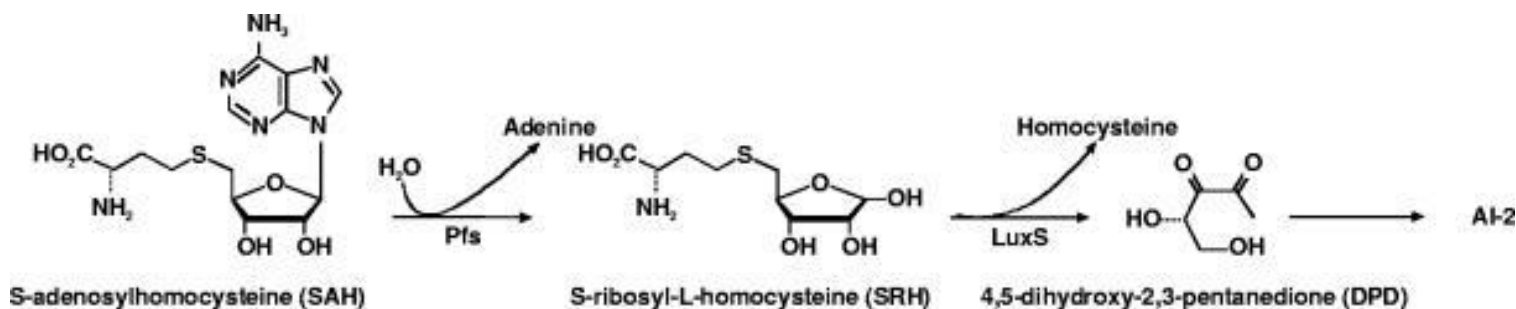
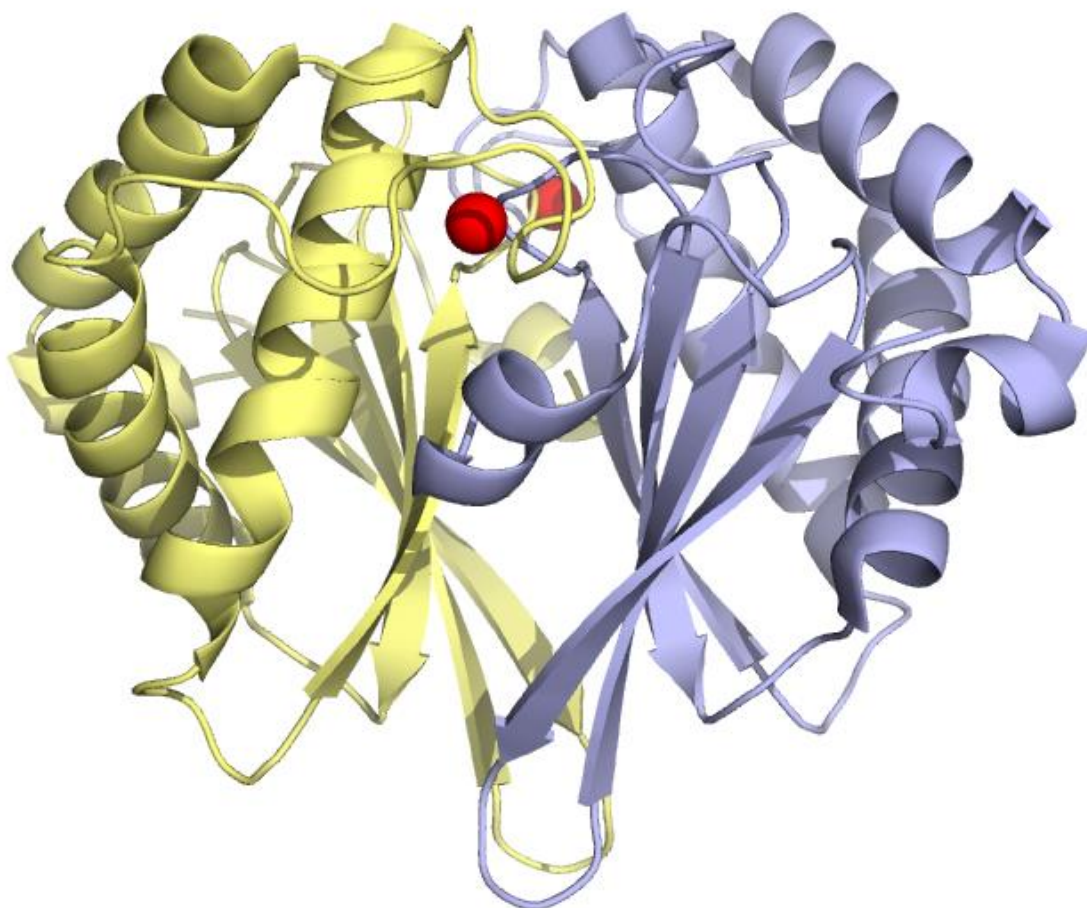


Рисунок 1. Реакция, осуществляемая лиазой LuxS. Взято из [2].

LuxS – металлсодержащий фермент и представляет собой гомодимер со структурой, уникальной для белков данного небольшого семейства. Два идентичных активных сайта образуются в месте контакта двух субъединиц, последовательность, которых содержит инвариантный мотив НХХЕН, где два иона  $Fe^{2+}$  тетраэдрически скоординированы. [1] Структура димера изображена на Рисунке 2.



*Рисунок 2. Структура димера LuxS. Субъединицы выделены светло-желтым и голубым цветом. Ион металла, который находится в активном центре, выделен красным. Модель PDB - 1IE0.*

Для оценки качества была взята модель LuxS с идентификатором PDB – 1IE0. Работа была сделана довольно давно в 2001 году [3]. Авторы работы - Hilgers, M.T., Ludwig, M.L. Тогда было мало известно о механизме действия LuxS. Поэтому целью эксперимента можно считать получение дополнительной информации о работе фермента и использование ее для дальнейшей разработки модификаторов LuxS, предотвращающих патогенез бактерий.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Общая информация о модели

Модель 1IE0 представляет собой мономер, состоящий из одной цепи. Длина – 156 а.о. Биологическая единица однако представляет собой гомодимер. 84-ый остаток цистеина в данной структуре модифицирован в цистеинилсульфоновую кислоту (происходит из-за окисления железа и инактивирует фермент). Также присутствует ион  $Zn^{2+}$ . Тут интересно заметить, что сейчас доказано, что в активном центре находится все-таки ион  $Fe^{2+}$ . Однако фермент может связывать ионы цинка, железа и кобальта.

Статья [3] о расшифровке опубликована в 2001 году в журнале (2001) Proc.Natl.Acad.Sci.USA 98: 11169-11174. Авторы статьи - Mark T. Hilgers и Martha L. Ludwig.

Метод решения фазовой проблемы - множественное изоморфное замещение (MIR).

Всего измерено 21965 рефлексов.

Разрешение структуры -  $1.60\text{\AA}$ , максимальное разрешение для использованных рефлексов —  $21.69\text{\AA}$ . В этот диапазон попадают 92% всех рефлексов.

Тип симметрии ячейки - P 6<sub>5</sub> 2 2.

R-factor = 0.174, R-free = 0.204.

Некристаллографических симметрий асимметрической единицы не наблюдается.

### Индикаторы качества модели

#### R-factor и R-free

Стандартный фактор достоверности (R-фактор) – это величина, которая характеризует соответствие модели ее экспериментальным значениям (модулям структурных факторов). Это мера различия оптимизированной модели и экспериментальных данных. Чем меньше R-фактор, тем достовернее модель. R-factor < 25% свидетельствует о хорошем соответствии.

R-free вычисляется так же, как и R-фактор, только по специально оставленной части рефлексов (в нашем случае 2202 рефлекса (10.03%)), по которым не проводилась оптимизация. Условно приемлемый диапазон: <25%. Дополнительно учитывается разница между R-factor и R-free, чтобы оценить вероятность переоптимизации.

$(R\_free - R\_factor) < 10\%$

В данном случае R-factor = 0.174 (<25%), R-free = 0.204 (<25%), R-factor-R-free = 0.03 (<10%)

Следовательно, можно сделать вывод, что соответствие между моделью и экспериментом хорошее, а также можем утверждать, что переподгонки не было.

### Пространственный R-фактор

Пространственный R-фактор (Real Space R-factor, RSR) характеризует соответствие «экспериментальной» электронной плотности и электронной плотности, построенной по модели. Хорошими считаются значения от 10% до 20%. RSR используется не для оценки качества всей структуры, как R-free и R-фактор, а для оценки локального качества. То есть, для отдельных аминокислотных остатков. На сервере EDS [4] можно получить диаграмму со значениями RSR для каждого остатка. Пространственный R-фактор для всех остатков структуры 1IE0 показан на Рисунке 3.

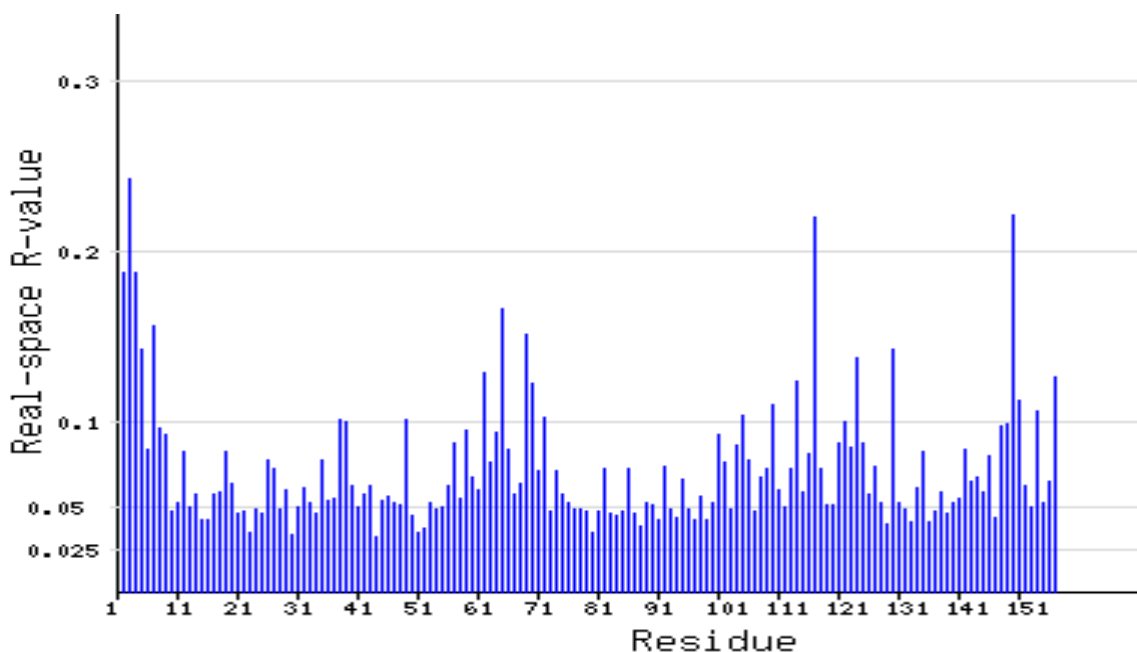


Рисунок 3. Пространственный R-фактор для всех аминокислотных остатков структуры 1IE0.

По Рисунку 3 можно определить аминокислотные остатки, для которых пространственный фактор больше порогового значения 20%. Это Ser3, Glu117 и Glu150. Также близкое к 20% значение пространственного фактора имеют Pro2, Val4. Среднее значение RSR по всем аминокислотным остаткам для структуры 1IE0 указанное на сервере EDS составляет 7.2% со стандартным отклонением 3.7%. Следовательно, можем сказать, электронная плотность атомов модели хорошо соответствует «экспериментальной» электронной плотности.

## Z-score

Для более точного определения остатков, плохо вписанных в электронную плотность используется показатель Z-score RSR. Это оценка RSR относительно других структур. Для вычисления Z остатка его RSR сравнивается со средним RSR для того же типа остатков по выборке из PDB с аналогичным разрешением. Если RSR плохой, а RSR-Z – хороший, то значит координаты атомов расшифрованы плохо, но не хуже, чем в других подобных структурах. Высокие положительные значения  $Z > 2$  свидетельствуют о том, что остаток плохо вписан в электронную плотность по сравнению с другими структурами с тем же разрешением.

Z-score для всех остатков структуры 1IE0 показан на Рисунке 4.

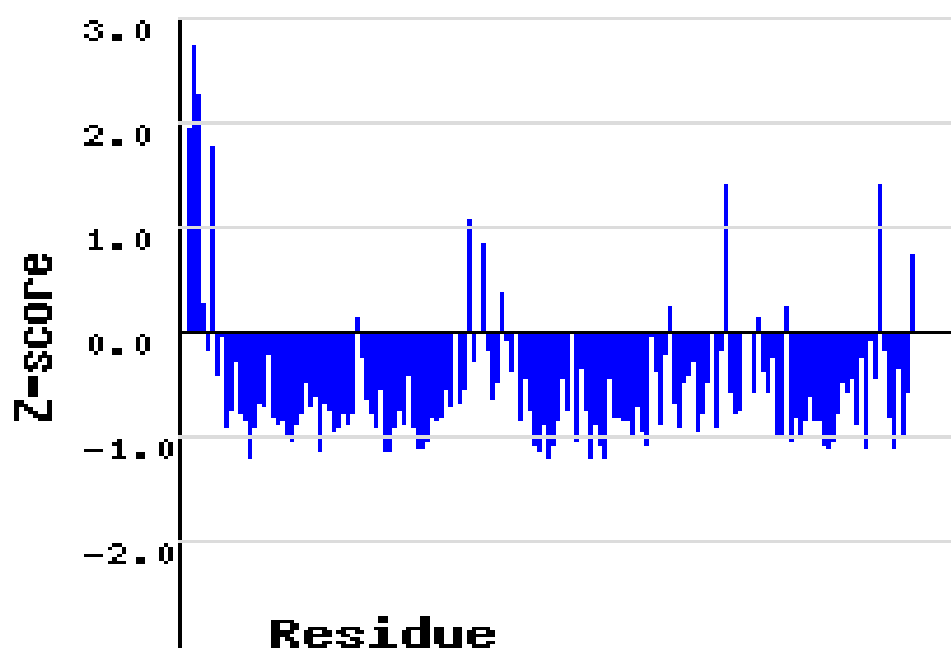


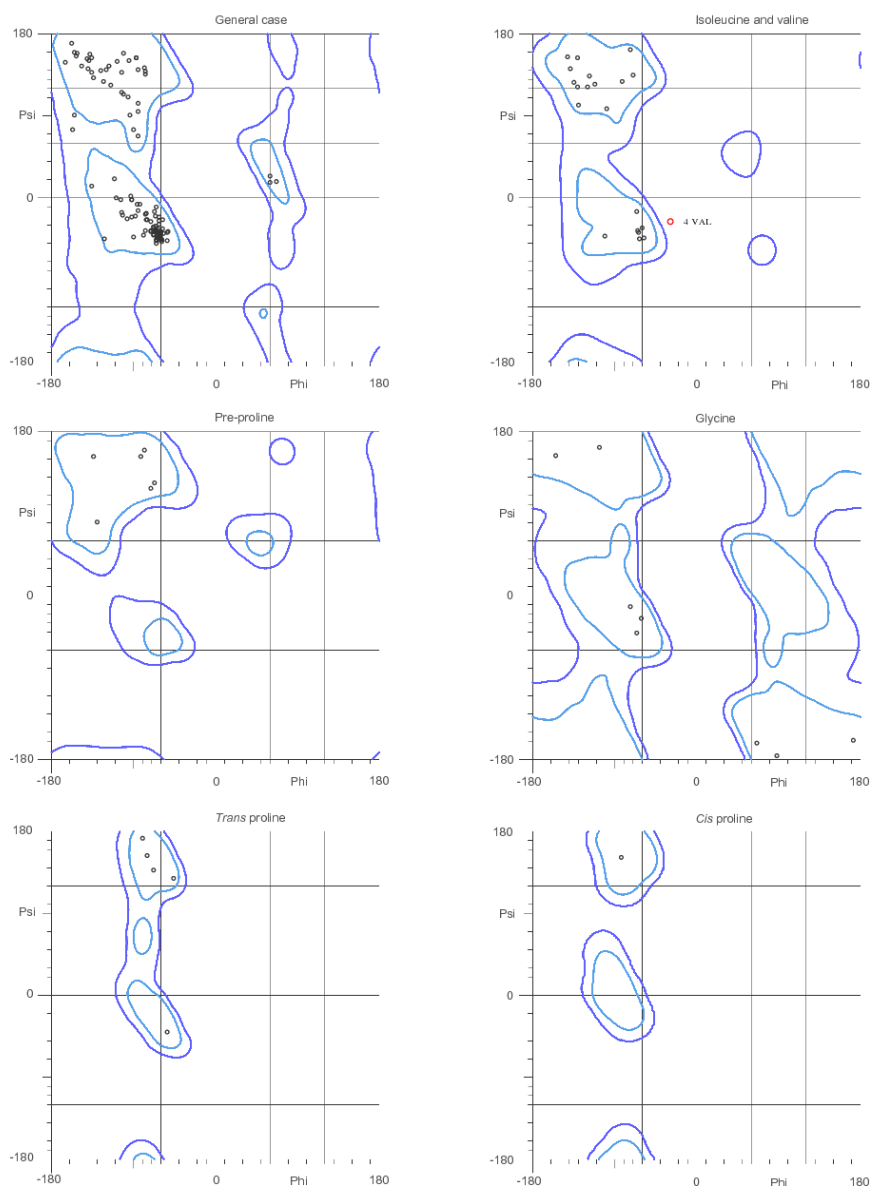
Рисунок 4. Z-score для всех аминокислотных остатков структуры 1IE0.

Как можно видеть на Рисунке 4. На N-конце белка есть несколько остатков, у которых Z-score больше и близкий к пороговому значению 2. Это Pro2, Ser3, Val4, Phe7.

## Карта Рамачандрана

Карты торсионных углов (карты Рамачандрана) служат хорошей оценкой моделей, независимой от оптимизации. Они помогают найти маргинальные остатки по конформации остова молекулы. Карты Рамачандрана для модели 1IE0 были построены с помощью сервиса MolProbity. Они представлены на Рисунке 5.

98% аминокислотных остатков находятся в предпочитаемой области. В модели 1IE0 есть один остаток, находящийся в запрещенной области. То есть он обладает конформацией, сильно отклоняющейся от принятой. Это Val4.



98.0% (150/153) of all residues were in favored (98%) regions.

99.3% (152/153) of all residues were in allowed (>99.8%) regions.

There were 1 outliers (phi, psi):

4 VAL (-29.9, -26.7)

*Рисунок 5. Карты Рамачандрана для структуры 1IE0. Выбывающий остаток Val4 обозначен красным кругом.*



## Статистика сервиса MolProbity

Далее был проведен анализ структуры с помощью сервиса MolProbity [5]. Название модели было подано на вход программе с указанием добавить к модели атомы водорода. Сводная статистика анализа структуры приведена в Таблице 1.

Таблица 1. Сводная статистика анализа структуры 193l с помощью сервиса MolProbity (добавлены атомы водорода). Красным отмечены отрицательные параметры модели, желтым - средние, зеленым – хорошие.

All-Atom Contacts	Clashscore, all atoms:	6.51	88 <sup>th</sup> percentile* (N=718, 1.60Å ± 0.25Å)	
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	3	2.22%	Goal: <0.3%
	Favored rotamers	128	94.81%	Goal: >98%
	Ramachandran outliers	1	0.66%	Goal: <0.05%
	Ramachandran favored	148	98.01%	Goal: >98%
	MolProbity score <sup>^</sup>	1.62	81 <sup>st</sup> percentile* (N=7200, 1.60Å ± 0.25Å)	
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	3 / 1269	0.24%	Goal: 0%
	Bad angles:	4 / 1716	0.23%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	1 / 6	16.67%	Expected: ≤1 per chain, or ≤5%

### Clashscore

В первой графе (Clashscore, all atoms) представлено число недопустимых наложений атомов на 1000. В данном случае – это 6,51, то есть т.е. лишь 0.651% атомов структуры имеют недопустимые наложения (т.е. их Ван-дер-Ваальсовы радиусы перекрываются больше, чем на 0.4 Å). В таблице 2 представлена информация об атомах, которые имеют недопустимые наложения. Перцентиль 88 процентов (данной структуры по отношению к структурам примерно такого же разрешения) говорит о том, у 88% структур Clashscore хуже данного, а хуже – значит больше. Если учесть, что лучшее значение перцентили – это 100%, то можно заключить, значение Clashscore для данной структуры довольно хороший.

Таблица 2. Информация об атомах с аномальным clashscore. свидетельствующем о перекрывании атомных радиусов.

```
Bad Clashes >= 0.4 Angstrom:
A 301 GOL C1 A 301 GOL O1 :1.455
A 2 PRO HA A1188 HOH O :0.996
A 49 LYS HE2 A1174 HOH O :0.749
A 5 GLU OE2 A1169 HOH O :0.659
A 74 PHE CD2 A 96 PRO HB3 :0.651
A 113 VAL O A 138 LYS HE2 :0.609
A 301 GOL C1 A 301 GOL HO1 :0.605
A 4 VAL HG12 A1140 HOH O :0.593
A 59 LEU O A 63 THR HG23 :0.557
A 2 PRO HB3 A 89 TYR CE1 :0.553
A 65 ARG HG2 A1036 HOH O :0.529
A 2 PRO CA A1188 HOH O :0.503
A 3 SER O A 5 GLU N :0.490
A 40 PHE O A 301 GOL O3 :0.475
A 142 ARG HD2 A1168 HOH O :0.460
A 84 OCS OD1 A1081 HOH O :0.444
clashscore = 6.51
clashscore (B factor cutoff = 40) = 2.210433
```

## Ротамеры

Ротамеры - это боковые цепи в типичных для данного типа остатка конформациях. Параметр «Poor rotamers» - это число остатков с маргинальными по отклонению от ротамеров боковыми цепями. Этот параметр позволяет найти маргинальные остатки по боковым цепям. Кроме того, он служит хорошей оценкой качества модели, потому что, как и карта Рамачандрана, он не зависит от оптимизации.

Структура 1Е0 содержит 3 (2,22%) остатка (Val4, Asp37, Ser66), конформация боковых цепей которых существенно отличается от ротамеров (хорошее значение < 1% остатков). Полученный процент маргинальных остатков по ротамерам можно считать приемлемым.

## Другие параметры

Менее информативные параметры указывают на то, что в структуре 1Е0 число C\_beta с неприемлемым отклонением от ожидаемого положения равно 0. Плохим значением является число ковалентных связей, существенно отклоняющихся от теории, равное 3 из 1269 (0,24%) (в идеале оно должно быть равным 0). Число валентных углов, существенно отклоняющихся от теории, равно 4 из 1716 (0,23), что тоже не очень хорошо, так как в идеале эта величина должна быть менее 0,10 %

В протоколе проверки качества WHAT\_CHECK представлены результаты проверки модели и по другим критериям, не описанным в данной работе. Однако проверку по большинству из них модель прошла успешно, в принципе, как и практически по всем рассмотренным выше параметрам. Подозрительные участки структуры представлены ниже.

## Аминокислотные остатки с подозрительным окружением

Остатки, окружение которых вызывает подозрение, обнаруженные с помощью протокола WHAT\_CHECK: Gln46, Asn44, Lys26, Met82 (значение параметра менее, чем -5). Возможно, что какие-то из этих остатков образуют симметричные контакты, образуют контакты с ионами или лигандами, находятся в активном центре или же их координаты определены неверно.

## Повороты пептидной цепи (flip)

Инверсия боковой цепи детектируется с помощью поиска водородных связей остатков. Если при повороте боковой группы остатка на 180° образование водородных связей энергетически более выгодно, значит, вероятно, электронная плотность была интерпретирована неправильно и имеет место инверсия

При анализе структуры сервисом MolProbity20 и изучении протокола WHAT CHECK было найден 1 остаток (His132) с предполагаемой инверсией боковой цепи.

Вода без водородных связей

Была обнаружена молекула воды (номер 1161), не образующая водородные связи. Расстояние между молекулой воды НОН 1161 и ближайшим остатком из цепи А составляет 4.5Å, а это расстояние слишком велико для водородной связи. На расстоянии около 4Å от этой молекулы находится кольцо фенилаланина соседней молекулы.

### Поиск и анализ маргинальных остатков в структуре

При поиске маргинальных остатков в структуре 1Е0 была проанализирована выдача сервиса MolProbity, выдача сервера EDS, выдача сервера WHAT CHECK, а также полный отчет о валидации структуры 1Е0, представленный на сайте PDB [6]. Найденные маргинальные остатки приведены в Таблице 3, с указанием свидетельства о маргинальности. Ни один из остатков, которые считаются необходимыми для функции белка, не был отмечен как маргинальный.

В подборе примеров маргинальных остатков был использован рисунок из PDBReport [6], который наглядно демонстрирует отклонения (геометрические) и несоответствие электронной плотности (см. Рисунок 7). Остатки, выделенные красным, имеют отклонения по трем критериям, желтые по одному, зеленый цвет означает отсутствие маргинальности по всем критериям. Красные точки означают, что соответствующие остатки имеют RSR Z-score > 2.

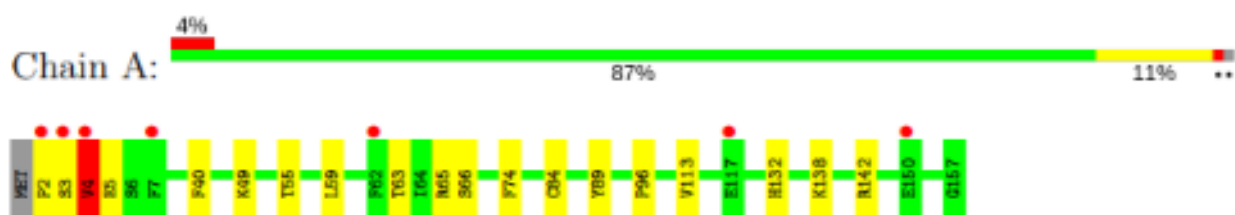


Рисунок 6. Качество отдельных остатков в структуре.

Таблица 3. Некоторые маргинальные остатки в структуре IIE0. Салатовым фоном выделены остатки для дальнейшего рассмотрения

№	Остаток	Свидетельства о маргинальности
1	Pro2	1. RSR $\cong$ 0.2 2. Z-score $\cong$ 0.2
2	Ser3	1. RSR > 0.2 2. Z-score >0.2
3	Val4	1. RSR $\cong$ 0.2 2. Z-score >0.2 3. Находится в запрещенной зоне карты Рамачандрана 4. Имеет боковую цепь в запрещенной конформации,
4	Phe7	1. Z-score $\cong$ 0.2
5	His132	1. Подозрение на необходимость инверсии боковой цепи
6	HOH1161	1. Молекула воды, которая не образует водородных связей
7	Asp37	1. Имеет боковую цепь в запрещенной конформации,
8	Asn44	1. Неблагоприятное окружение
9	Gln46	1. Неблагоприятное окружение
10	Lys26	1. Неблагоприятное окружение
11	Met82	1. Неблагоприятное окружение
12	Ser66	1. Имеет боковую цепь в запрещенной конформации,
13	Glu117	1. RSR > 0.2
14	Glu150	1. RSR > 0.2

Были проанализированы подробно пять (Ser3, Val4, Lys26, Met82, His132) остатков из приведённых в Таблице 3. Были сделаны выводы о том, с чем связана та или иная аномалия в каждом конкретном случае.

### Ser3

Таблица 4. Характеристики остатка Ser3.

Показатель	Значение
RSR	0.242
Z-score	2.739
Температурный фактор B	67.6
Положение	На поверхности белка

Серин 3 расположен в самом начале полипептидной цепи в начале небольшой альфа-спирали. У него есть один полярный контакт с молекулой воды (см. Рис. 8а). Остаток был

признан маргинальным потому, что плохо описывалась экспериментальная электронная плотность. Обратимся за информацией о температурном факторе остатка к серверу EDS. У этого остатка самый большой температурный фактор (67,6). Это означает, что остаток более чем в среднем подвержен тепловым колебаниям и электронная плотность сравнительно более «размазана». Некоторые проблемы с электронной плотностью наблюдаются и при увеличении уровня срезки (Рисунок 7).

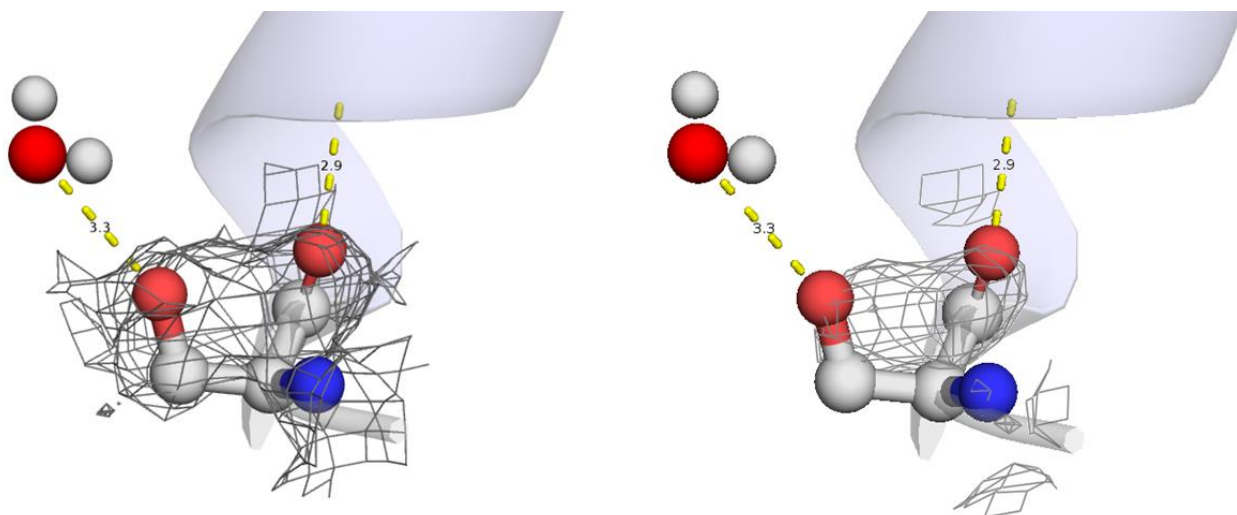


Рисунок 7. Визуализация функции электронной плотности в области Ser3 на уровне подрезки 0.5  $\sigma$  (слева) и 1  $\sigma$  (справа).

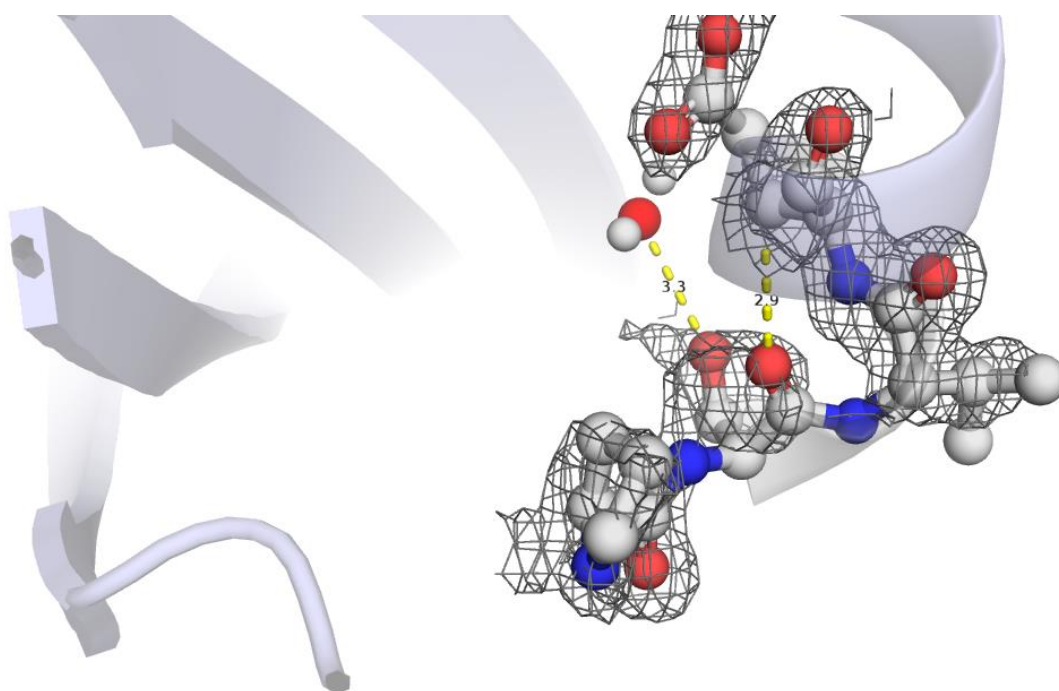


Рисунок 8. Визуализация функции электронной плотности в области первых шести аминокислотных остатка на уровне подрезки  $\sigma$  0.7.

Если посмотреть на соседние остатки (Рисунок 8), то можно также наблюдать, что электронная плотность вписана недостаточно точно. Также для этих остатков значения температурного фактора завышены, что свидетельствует о подвижность целого начального участка белка.

**Вывод:** Остаток Ser3 был неточно восстановлен по электронной плотности. Вероятно, это связано с большой подвижностью N-концевого участка полипептидной цепи, вследствие чего электронная плотность была слишком расплывчата и сигнал был неотличим от шума.

#### Val4

*Таблица 5. Характеристики остатка Val4.*

Показатель	Значение
RSR	0,187
Z-score	2,25
Clashscore	0,593
Температурный фактор B	49,49
Outlier по конформации остова	Да (запрещенная область)
Outlier по конформации боковой цепи	Да (запрещенная область)
Положение	На поверхности белка

Рекордсменом по количеству пройденных критериев маргинальности стал именно данный остаток – валин 4. Это можно объяснить, как минимум, двумя обстоятельствами. Во-первых, он находится в подвижной N-концевой части белка. Во-вторых, у него нет атомов, с повышенной электронной плотностью (например, как присутствие атомов кислорода у Ser3 в предыдущем случае).

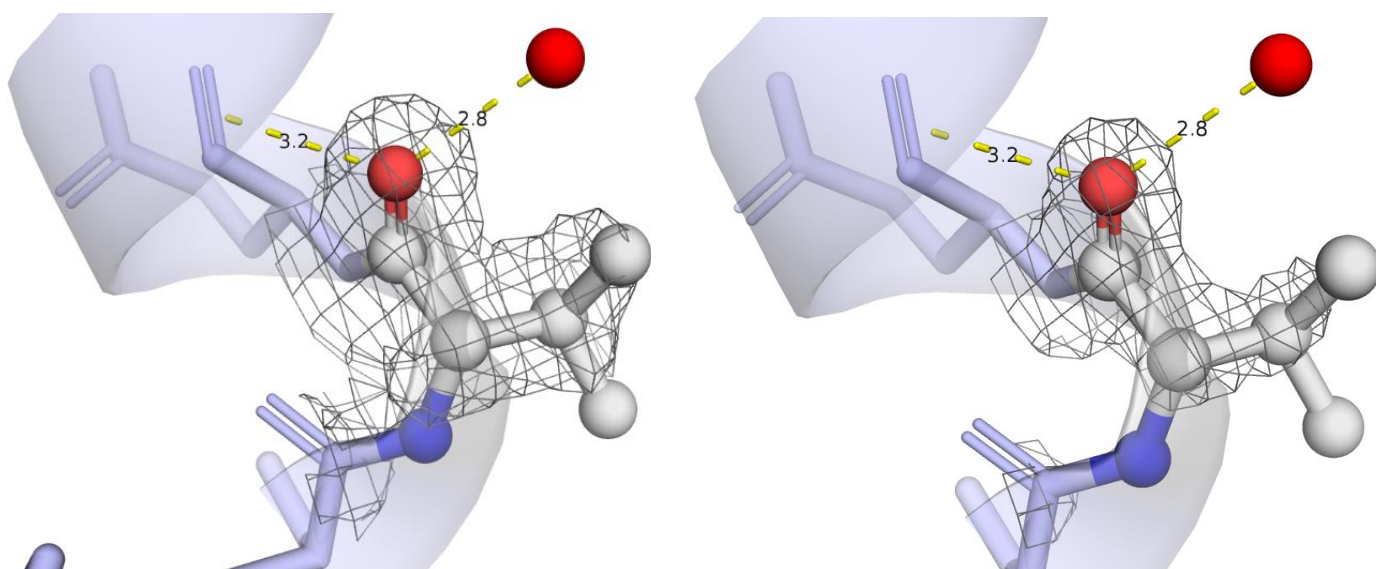


Рисунок 9. Визуализация функции электронной плотности в области Val4 на уровне подрезки 0,5  $\sigma$  (слева) и 1  $\sigma$  (справа). Молекула воды обозначена красной сферой.

Итак, очевидно, что электронная плотность вписана не самым лучшим образом, что связано с ее низкими значениями, так как на уровне подрезки 0,5 есть непокрытые участки (Рисунок 9). Однако ключевыми критериями, которые позволяют с большой уверенностью утверждать, что данный остаток является маргиналом – это его торсионные углы, как остова, так и боковой цепи. Случаи, когда остаток имеет конформацию боковой цепи, отличную от разрешенных можно иногда объяснить уникальными взаимодействиями (водородные связи, стэкинг-взаимодействия, взаимодействия с ионами металла) между соседними остатками в асимметрической ячейке или остатками, которые находятся в месте контакта между соседними молекулами в кристалле. Однако в данном случае не было обнаружено никаких весомых причин, объясняющих такое конформационное положение этого остатка. Дополнительным свидетельством маргинальности, служит информация о перекрывании ванн-дер-ваальсовых радиусов атома HG12 валина и атома O в NOH1140 (0,539 Å).

**Вывод:** из-за того, что значения электронной плотности для данного остатка получились очень низкими, то произошла ошибка при расшифровке, с чем связано плохое совпадение модели и электронной плотности, а также энергетически невыгодные значения торсионных углов.

## Lys26

Таблица 5. Характеристики остатка Lys26.

Показатель	Значение
RSR	0,078
Z-score	-0,8
Температурный фактор B	18,780
Конформация боковой цепи	Предпочтительная область
Комфортность окружения	-5.589
Положение	На поверхности белка

Этот остаток программа отнесла к маргинальным по параметру комфортности окружения. Т.е. это остаток должен располагаться так, чтобы образовались более энергетически выгодные контакты внутри белка.

В нашем случае неудобным оказался лизин 26, который находится на поверхности белка. (распространенный случай, когда остаток, находящийся на поверхности глобулы по данному критерию определяется как маргинальный). Однако если восстановить соседние молекулы в кристалле, то можно увидеть, что лизин стабилизирован довольно большим количеством водородных связей с остатками из своей же цепи, молекулами воды и остатками соседней молекулы (Рисунок 10).

Электронная плотность хорошо вписана, что позволяет заключить, что атомы остатка вписаны правильно. Об этом же можно заключить взглянув на значения RSR и z-score.



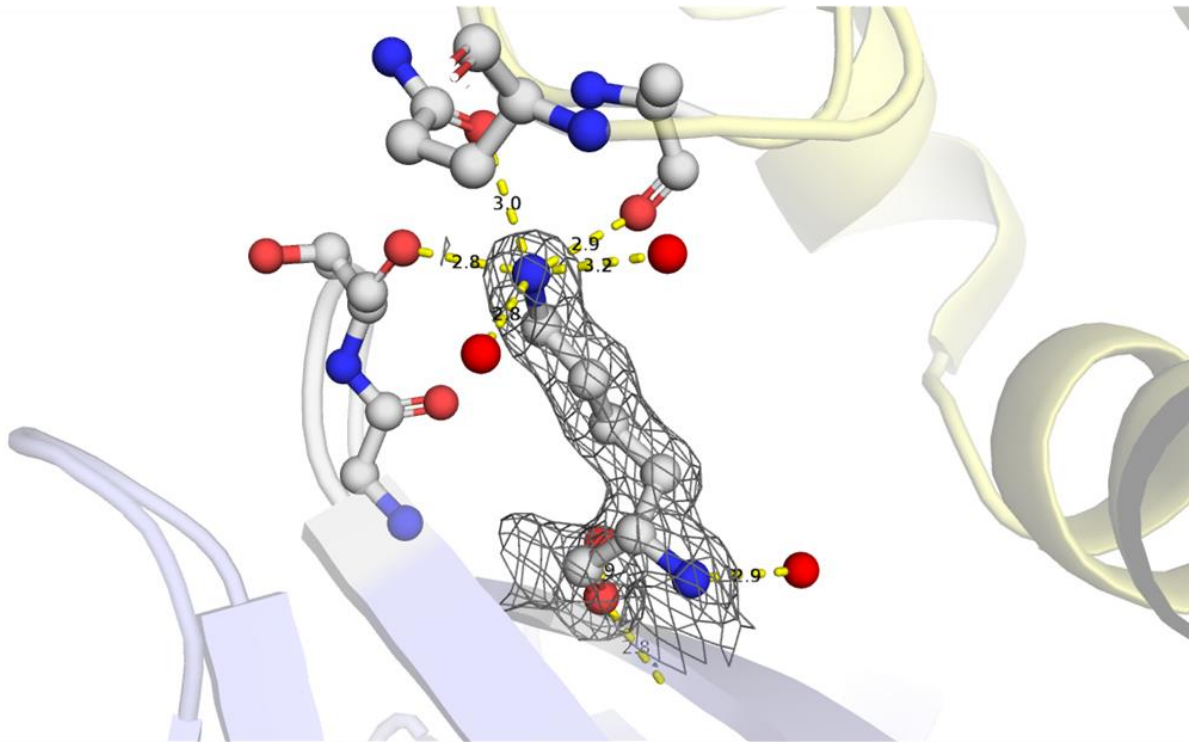


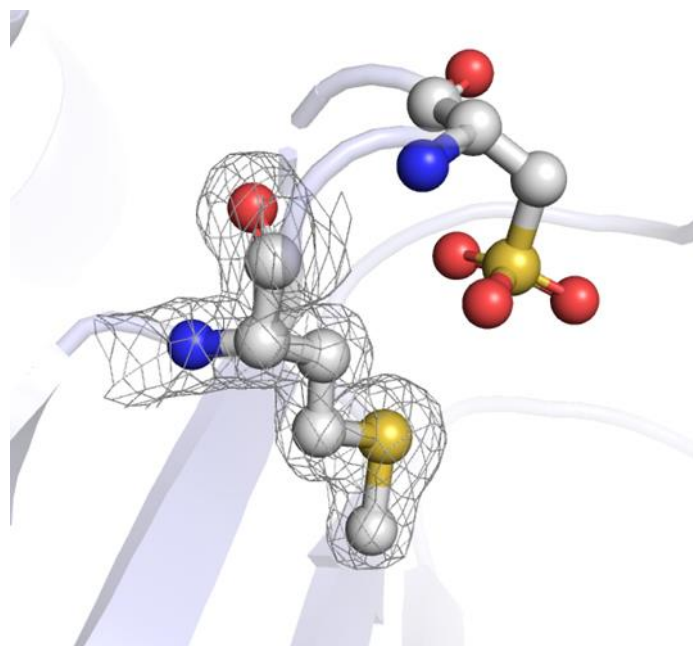
Рисунок 10.. Многочисленные полярные контакты Lys26. Разные молекулы белка обозначены светло-голубым и светло-желтым цветом. Также визуализирована электронная плотность на уровне подрезки 1  $\sigma$ .

**Вывод:** данный остаток было определен как маргинальный только по параметру комфортности окружения. Однако, видимо, в расчет не брались контакты с соседней молекулой, в следствие чего был сделан вывод, что окружение неблагоприятное. Однако можно с уверенностью сказать, что этот остаток не является маргиналом, так как у него наблюдается хорошее соответствие электронной плотности модели и эксперимента, а также предпочтительные конформации как остова, так и боковой цепи.

#### Met 82

Таблица 5. Характеристики остатка Met82.

Показатель	Значение
RSR	0,072
Z-score	-0,46
Температурный фактор B	16,450
Конформация боковой цепи	Предпочтительная область
Комфортность окружения	-5.346
Положение	Внутри глобулы белка



*Рисунок 11. Неблагоприятное окружение метионина 82 внутри белка. Электронная плотность представлена на уровне подрезки  $\sigma 1$ .*

Данный остаток также попал в разряд маргинальных из-за своего неблагоприятного окружения. Еще одним доводом выбрать его для анализа было тот факт, что Met82 находится внутри глобулы белка. Оказалось, что рядом с этим остатком находится модифицированный цистеин, т.е. остаток цистеинилсульфоновой кислоты, который, как можно заметить, чрезвычайно полярный и создает то самое неблагоприятное окружение (Рисунок 11). Это хороший пример как внутренние модификации остатков могут повлиять на вычисление этого параметра.

**Вывод:** Данный остаток не является маргинальным, так как электронная плотность хорошо вписана в атомы модели. Однако имеет место неблагоприятное окружение, обусловленное присутствием модифицированного остатка цистеина 84 рядом с метионином 82.

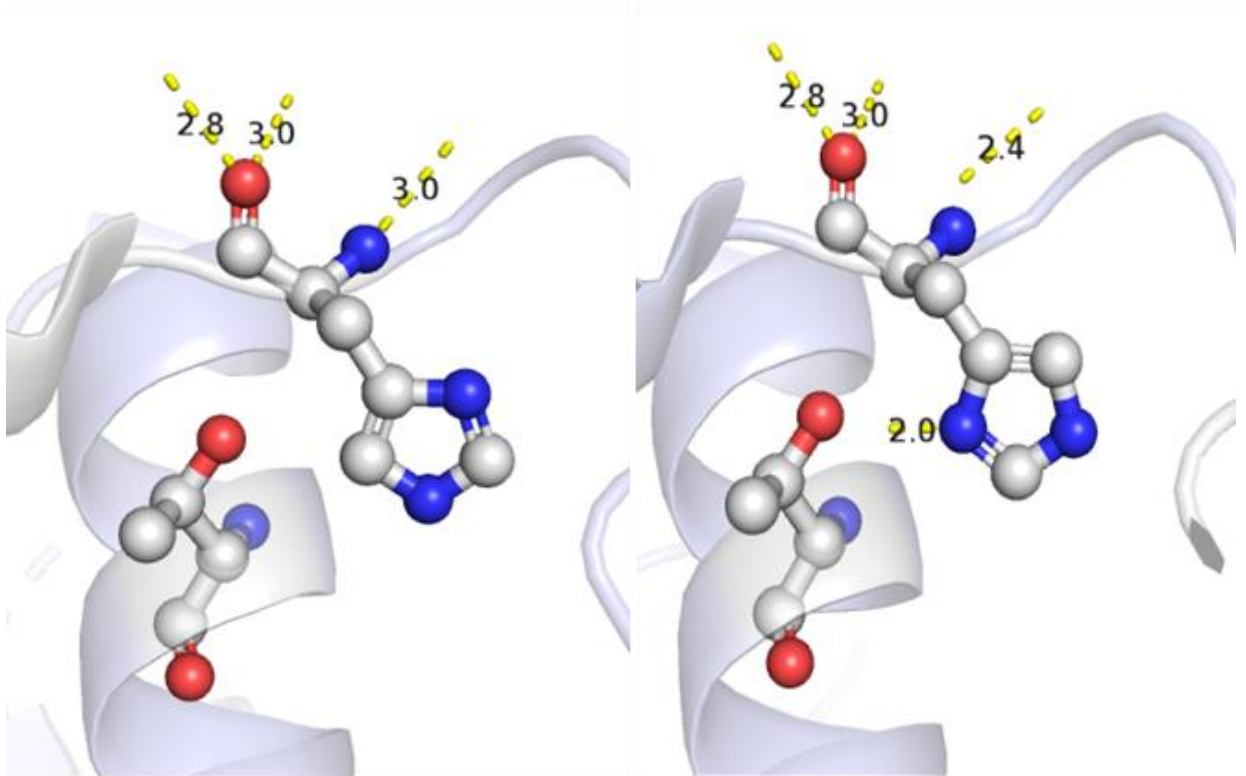


Рисунок 12 . (Слева) остаток His132 и остаток Thr55, (справа) остаток His 132 остаток Thr55 после инверсии. Желтым пунктиром обозначена водородная связь, цифрами – длина связи.

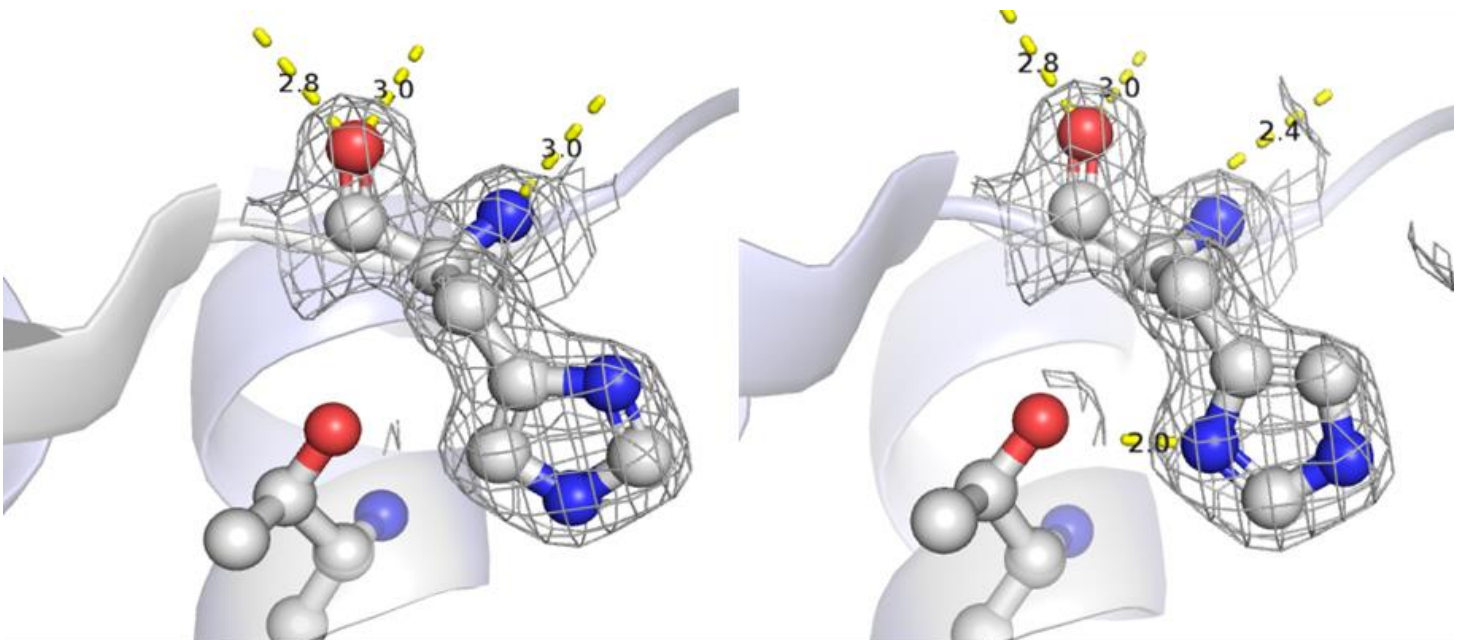


Рисунок 13 . (Слева) остаток His132 и остаток Thr55, (справа) остаток His 132 остаток Thr55 после инверсии. Желтым пунктиром обозначена водородная связь, цифрами – длина связи. Электронная плотность визуализирована на уровне подрезки  $\sigma 1$

Для структуры 1IE0 была зафиксирована инверсия боковых групп у остатка His132. Сравнение положение остатков приведено на Рисунке 12. После проведения инверсии, остаток гистидина 132 образует водородную связь с остатком рядом лежащего треонина55. Если взглянуть на Рисунок 13, то можно заметить, что электронная плотность не отличается для остатка гистидина до и после инверсии. Поэтому при расшифровке структуры была допущена ошибка. Это заключение согласуется также с отчётом в базе данных PDBREPORT. Поэтому данный остаток является маргинальным.

**Вывод:** Изображение остатка His132 (Рисунок 12) и электронной плотности вокруг него (Рисунок 13) в двух вариантах (с инверсией и без) позволяет сделать вывод о том, что его маргинальность связана с ошибкой расшифровки. Ошибка – следствие схожести распределения электронной плотности у остатка гистидина.

*Таблица 6. Сводная таблица по результатам детальной проверки потенциальных маргинальных остатков.*

Остаток	Комментарий	Итог
Ser3	В начале цепи, высокая подвижность, плохо вписана электронная плотность	Маргинал
Val4		Маргинал
Lys26	Неблагоприятное окружение из-за положения остатка на поверхности белка. Его многочисленные контакты свидетельствуют о правильности расшифровки	Не маргинал
Met82	Неблагоприятное окружение не следствие ошибки, а результат присутствия модифицированного остатка	Не маргинал
His132	Требуется инверсия, так как образуется водородная связь	Маргинал

## PDB-REDO

PDB\_redo [7] оптимизирует структуру полностью автоматизированными методами и предоставляет сравнение полученного результата с исходной структурой. В таблице 7 представлено сравнение некоторых параметров структур 1IE0 оригинальной и оптимизированной. Полученная оптимизированная структура была дана на вход серверу Molprobity. На рисунке 14 приведено изображение совмещения моделей 1IE0 из PDB и PDB\_REDO.

Таблица 7. Сравнение параметров моделей PDB и PDB\_REDO

Параметр	PDB	PDB_REDO
R-factor	0.1688	0.1611
R-free	0.1974	0.1933
R-free – R-фактор	0.0286	0.0322
Outlier по карте Рамачандрона	1	0
Outlier по ротомерам	3	2
Остатки, требующие инверсии	1 His	1 Gln
MolProbity score	1,62	1.28
Clashscore	6,51	2.44

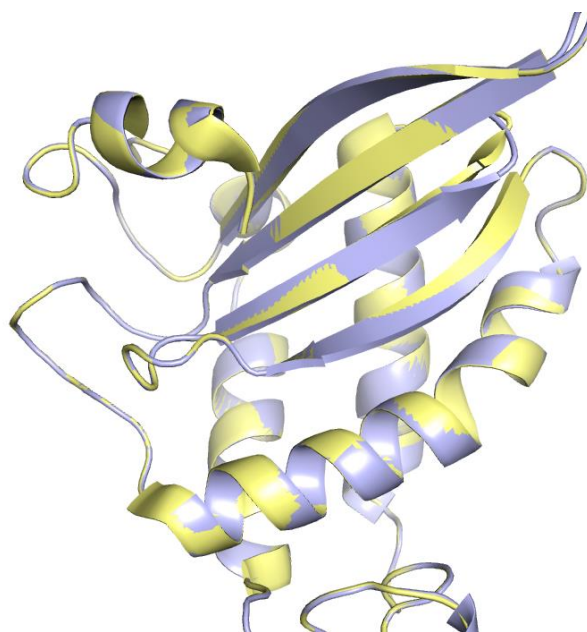
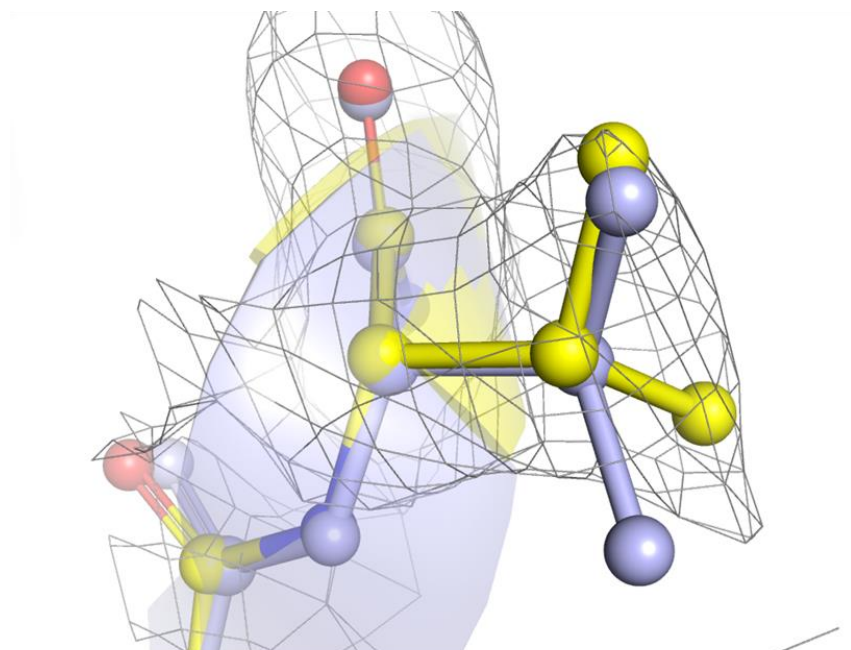


Рисунок 14. Совмещение моделей 1IE0 из PDB (серый) и PDB\_REDO (желтый)

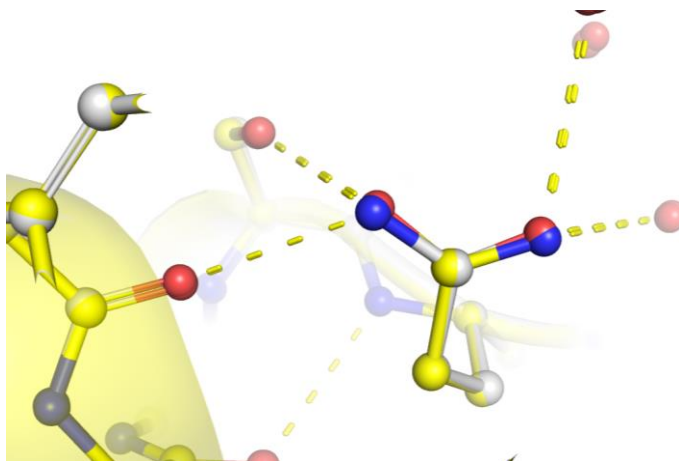
Как можно видеть и значение R-factor и значение R-free немного улучшились после оптимизации, однако их разница возросла, что может свидетельствовать о некоторой вероятности подгонки параметров. Оптимизация привела к уменьшению количества остатков, конформация которых находится в запрещенной области. Интересно, что те остатки, которые обсуждались в предыдущем разделе и которые были признаны маргиналами, но не по ошибке авторов структуры, а по причинам, связанным с физико-химическими особенностями, в структуре после оптимизации остались без изменений.

Чтобы точнее сравнить две структуры до и после оптимизации были рассмотрены два случая. Конечно же, было интересно. Что стало с маргинальным валином 4. Как видно на Рисунке 15, конформация его боковой цепи резко изменилась и теперь электронная плотность лучше вписана. Здесь однозначно оптимизация сделала свое дело.



*Рисунок 15. Остаток Val4 до (серый цвет) и после (желтый цвет) оптимизации. Электронная плотность визуализирована на уровне подрезки 0,5.*

Второй пример касается инверсии, которая все-таки присутствует даже после оптимизации, но теперь маргинал по этому параметру – это другой остаток. На рисунке 16 приведено изображение остатка из двух структур. Видно, что если повернуть боковую цепь на 180 градусов, то образуется новая водородная связь. Странно, что процедура оптимизации пропустила данный остаток.



*Рисунок 16. Остаток глутамина 147, для которого наблюдается инверсия после оптимизации.*

## **ВЫВОДЫ**

Структура 1IE0 обладает высоким разрешением 1,64 Å, однако расшифрована с некоторыми ошибками. В основном ошибки связаны с отсутствием электронной плотности на N-конце белка. Структура содержит достаточное количество маргинальных остатков, однако большинство из них находится на поверхности глобулы и обращены в раствор. Полная оптимизация с помощью PDB\_REDO позволила улучшить качество модели: уменьшились R-value и R-free, количество столкновений ван-дер-ваальсовых радиусов. Кроме того, отобранные ранее маргинальные остатки были исправлены. Остальные параметры изменились незначительно.

## **ЗАКЛЮЧЕНИЕ**

По результатам проведенного анализа можно сделать вывод, что качество структуры 1IE0 довольно хорошее, о чем свидетельствуют приведенные параметры оценки качества структуры. Только начальный участок цепи, из-за высокой подвижности недостаточно точно определен. Среди отобранных маргинальных остатков не было тех, кто непосредственно осуществляет катализ. Возможно, поэтому им не было уделено большего внимания в плане корректировки. В данной работе было показано, что не всегда аминокислотные остатки, описанные как маргинальные, являются таковыми. Все-таки ручная проверка необходима в некоторых случаях. Оптимизация модели помогает улучшить качество, однако тоже может быть неправа в отдельных моментах.

## СПИСОК ЛИТЕРАТУРЫ

- [1] A. Vendeville, K. Winzer, K. Heurlier, C. M. Tang, and K. R. Hardie 2005 Making 'sense' of metabolism: Autoinducer-2, LuxS and pathogenic bacteria *Nat. Rev. Microbiol.*
- [2] S. C. Belval, L. Gal, S. Margiewes, D. Garmyn, P. Piveteau, and J. Guzzo (2006) Assessment of the roles of LuxS, S-ribosyl homocysteine, and autoinducer 2 in cell attachment during biofilm formation by *Listeria monocytogenes* EGD-e *Appl. Environ. Microbiol.*
- [3] M. T. Hilgers and M. L. Ludwig (2001) Crystal structure of the quorum-sensing protein LuxS reveals a catalytic metal site. *Proc. Natl. Acad. Sci. U. S. A.*
- [4] <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1ie0>
- [5] <http://molprobity.biochem.duke.edu/>
- [6] [http://files.rcsb.org/pub/pdb/validation\\_reports/ie/1ie0/1ie0\\_full\\_validation.pdf](http://files.rcsb.org/pub/pdb/validation_reports/ie/1ie0/1ie0_full_validation.pdf)
- [7] [http://www.cmbi.ru.nl/pdb\\_redo/](http://www.cmbi.ru.nl/pdb_redo/)