

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ им. М.В.ЛОМОНОСОВА
ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

Отчет о качестве РСА структуры белка PDB ID: 4U63

Авдюнина Полина

401 группа

Москва
2018г.

Содержание

Аннотация	2
Введение	2
Результаты	3
Общая информация о модели	3
Значения индикаторов качества модели в целом	3
R-фактор	3
R-free	4
Перекрытие Ван-дер-Ваальсовых радиусов атомов (Clashscore)	4
Карта Рамачандрана	5
Ротамеры (Sidechain outliers)	5
RSR outliers	5
Анализ маргинальных остатков	5
Lys3 chain A	6
Leu91 chain B	6
Glu158 chain A	7
Lys170 chain A	7
Lys2 chain A	8
Сравнение модели из PDB с моделью из PDB-REDO [3]	9
Заключение	10
Список литературы	10

Аннотация

В работе был произведен анализ качества расшифровки структуры PDB ID: 4R11. Описаны

основные показатели качества модели и их соответствие общепринятым значениям.

Рассмотрено несколько маргинальных аминокислотных остатков.

Введение

Исследуемый белок является бактериальной ДНК-фотолиазой, выделенной из бактерии *Agrobacterium fabrum* (PhrA).

ДНК-фотолиаза является одним из ферментов репарации ДНК, активация которого происходит под действием видимого света. ДНК-фотолиаза удаляет фотопродукты и пиримидин-пиримидиновые димеры, образующиеся в молекуле ДНК под действием коротковолнового УФ-излучения. Процесс, в котором участвует фермент, называется фотореактивацией. Такие фотореактивирующие ферменты имеются у бактерий и низших эукариотических организмов, но в клетках млекопитающих они не обнаружены.

Структура была получена методом рентгеноструктурного анализа в работе 2015 года с разрешением 1.67 Å [1]. PhrA содержит 5,10-метенилтетрагидрофолат (МТНФ) в качестве антенного хромофора с уникальным сайтом связывания, а также FAD в качестве кофактора. Строение белковой глобулы можно увидеть на рис. 1.

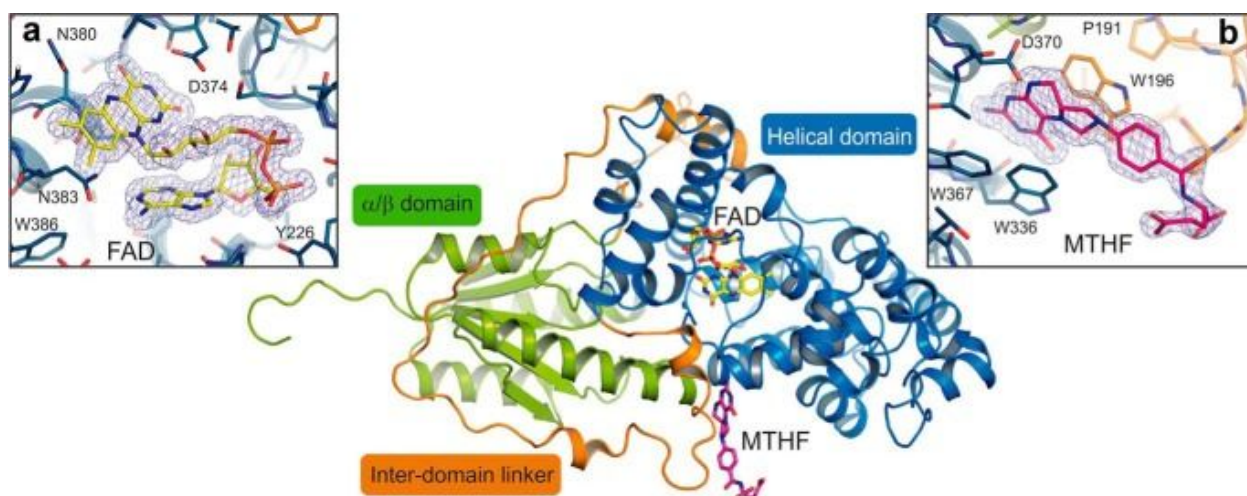


Рис.1 Визуализация исследуемой структуры 4U63

На картинке методом визуализации *ribbon* изображены α/β -домен (N-концевой домен; зеленый) и спиральный домен (каталитический домен; синий), связанный длинным междоменным линкером (оранжевый). Кофакторы МТНФ (окрашены в розовый цвет) и FAD (окрашены в желтый цвет) представлены в виде модели палочек. Сайты связывания FAD / МТНФ можно увидеть на вставках, лиганды изображены методом *stick* с отображенным соответствующим участком карты электронной плотности ($\sigma = 1.2$).

Результаты

Общая информация о модели

Таблица 1. Основные параметры структуры 4U63

Параметр	Значение
Состав комплекса	Одна субъединица, 1 молекула FAD, 1 молекула МТНФ
Год	2015
Авторы	Patrick Scheerer, Fan Zhang, Jacqueline Kalms, David von Stetten, Norbert Krauß, Inga Oberpichler and Tilman Lamparter
Разрешение	1.67 Å
Метод разрешения фазовой проблемы	MIR
Количество измеренных рефлексов	84 703
Полнота набора структурных факторов	100.0
Диапазон разрешений структурных факторов	34.40-1.67
Параметры элементарной ячейки	$a = b = 81.87 \text{ Å}$, $c = 195.95 \text{ Å}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$

Значения индикаторов качества модели в целом

Основными индикаторами качества модели считаются R фактор и R-free. Также качество структуры можно оценить по доле маргинальных остатков по разным критериям, таким как карта Рамачандрана, перекрывание Ван-дер-Ваальсовых радиусов атомов (clashscore), длин и углов связей, ротаметров.

R-фактор

R-фактор является оценкой того насколько расшифрованная структура способна предсказать наблюдаемые результаты. Хорошими значениями для этого параметра считаются показатели меньше 0.25; R-фактор рассматриваемой структуры равен 0.180, т.е. Полученная модель хорошо предсказывает результаты рентгеноструктурного эксперимента.

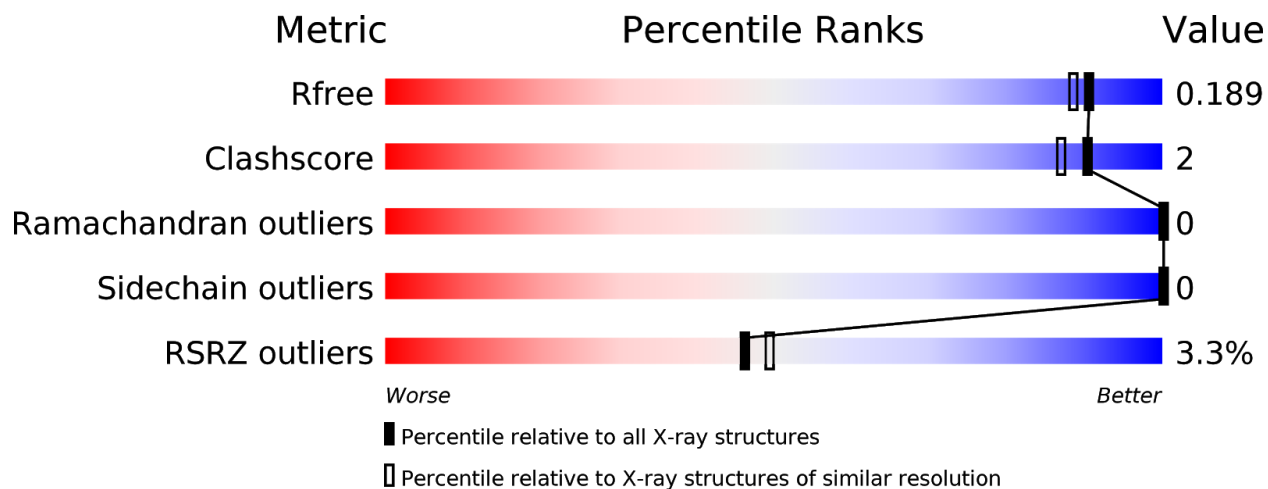


Рис.2 Основные показатели качества модели (черным – относительно всех PCA моделей; белым – относительно PCA моделей близкого разрешения).

R-free вычисляется точно так же, как R-фактор, но по относительно небольшой заранее отобранной случайной выборке экспериментальных данных (рефлексов), не используемых при оптимизации модели (в данном случае 5% от полученных данных). Таким образом, R-free является контролем переоптимизации модели – правильно оптимизированная модель должна предсказывать одинаково хорошо все наблюдаемые результаты (а не

только те, по которым она оптимизировалась). В данном случае R-free составил 0.189, это значение лучше, чем у большинства РСА структур подобного разрешения (Рис. 2), а также очень близко к границе “хороших” значений. Это значит что переоптимизации модели не наблюдается.

Перекрытие Ван-дер-Ваальсовых радиусов атомов (Clashscore)

Этот параметр отражает число стерических перекрытий (Ван-дер-Ваальсовы радиусы соседних атомов перекрываются более чем на $0,4 \text{ \AA}$) в расчете на 1000 атомов. Всего наша структура содержит 482 аминокислотных остатка. Значение clashscore для рассматриваемой структуры - 2.0.

Карта Рамачандрана

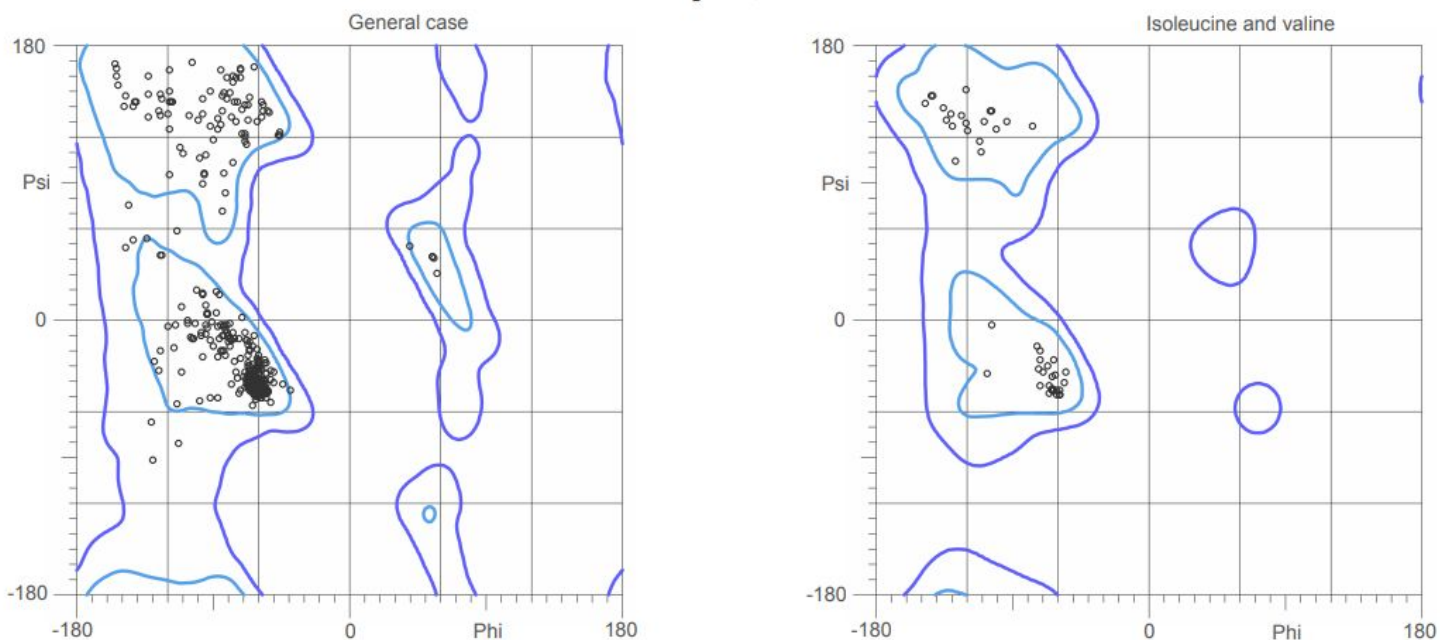


Рис. 3. Карты Рамачандрана для рассматриваемого белка

Карта Рамачандрана позволяет оценить, есть ли маргиналы по торсионным углам ϕ и ψ .
На

рисунке 3 представлена карта Рамачандрана для всего белка, построенная программой MolProbity [2]. В данном случае нет ни одного остатка в запрещенной области, а в предпочтительных регионах находятся 98% (510/522) остатков.

Ротамеры (Sidechain outliers)

Данный параметр учитывает конформацию боковых цепей аминокислот и позволяет найти

маргинальные остатки по этому критерию. Ротамерами называются боковые цепи в предпочитаемых для данного остатка положениях, которые описываются наборами допустимых значений торсионных углов χ_i . В данной структуре маргинальных (запрещённых) остатков по положению боковых цепей - 6, что составляет 0.14 % от всех остатков. По оценке же MolProbity таких остатков должно быть меньше 0.3%, поэтому их меньше оптимального.

RSR outliers

Параметр RSR характеризует, насколько хорошо остаток в модели вписывается в экспериментальную электронную плотность. Обычно оценивается Z-score (RSRZ) - насколько RSR для остатка в данной модели отличается от среднего значения RSR по структурам из PDB близкого разрешения. Если Z-score > 2, считается, что остаток маргинальный по RSR и не очень хорошо описывается экспериментальными данными. RSR оценивается для каждого остатка отдельно и может служить мерой локального качества. Для оценки качества структуры в целом можно посмотреть на долю маргинальных остатков по RSR: в данной структуре их 4.4% (21/482). Это весьма средние значения.

Анализ маргинальных остатков

Маргинальных остатков по картам Рамачандрана получено не было, поэтому они отсутствуют в таблице. 5 Выбранных остатков перечислены в таблице 2.

Таблица 2. Примеры маргинальных остатков

Остаток	Критерий, по которому остаток маргинальный
TRP196, chain A	Clashscore (0.595 Å)
LYS355, chain A	Clashscore (0.552 Å)
LEU215, chain A	Clashscore (0.453 Å)
PRO233, chain A	Clashscore (0.429 Å)
PHE436, chain A	Clashscore (0.419 Å)

TRP196, chain A

Очень большой clashscore 0.595 Å.

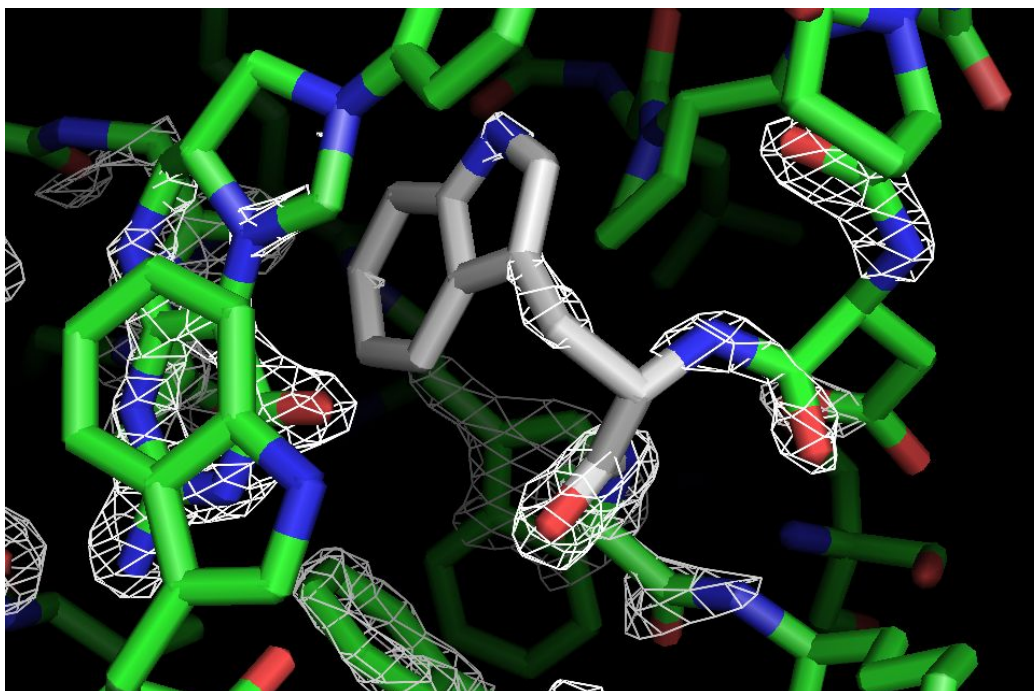


Рис 4. Изображение TRP196

LYS355, chain A

Очень большой clashscore 0.552 Å.

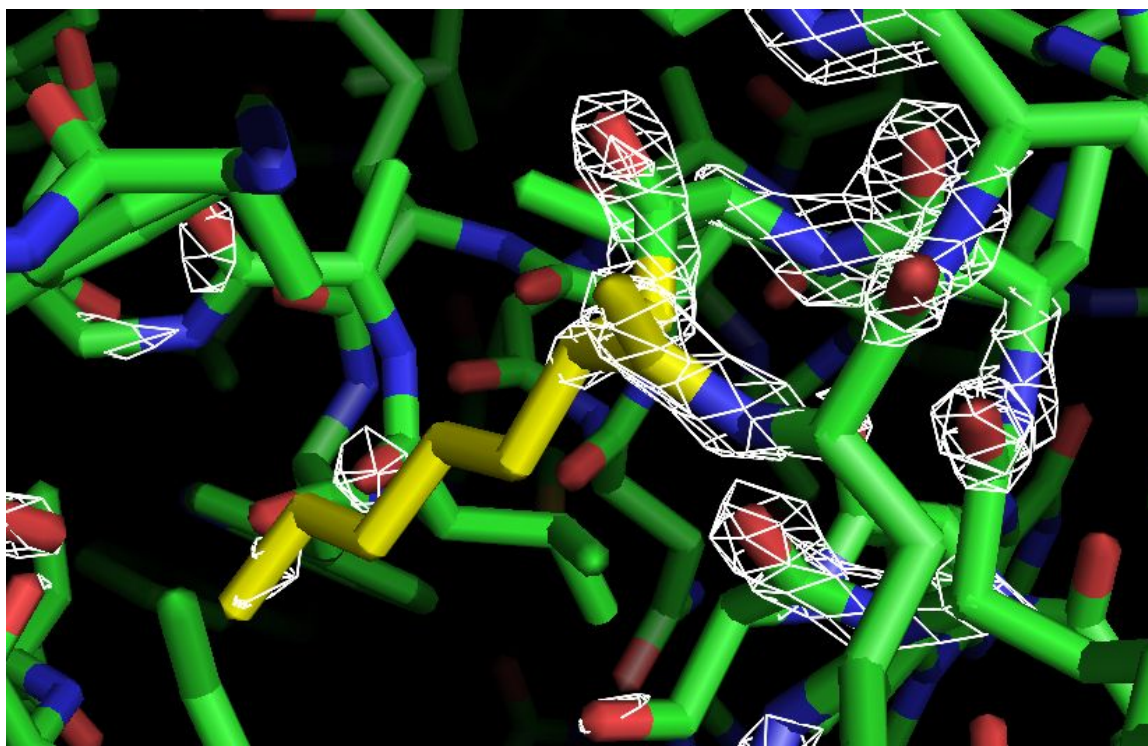


Рис 5. Изображение LYS355

LEU215, chain A

Очень большой clashscore 0.453 Å.

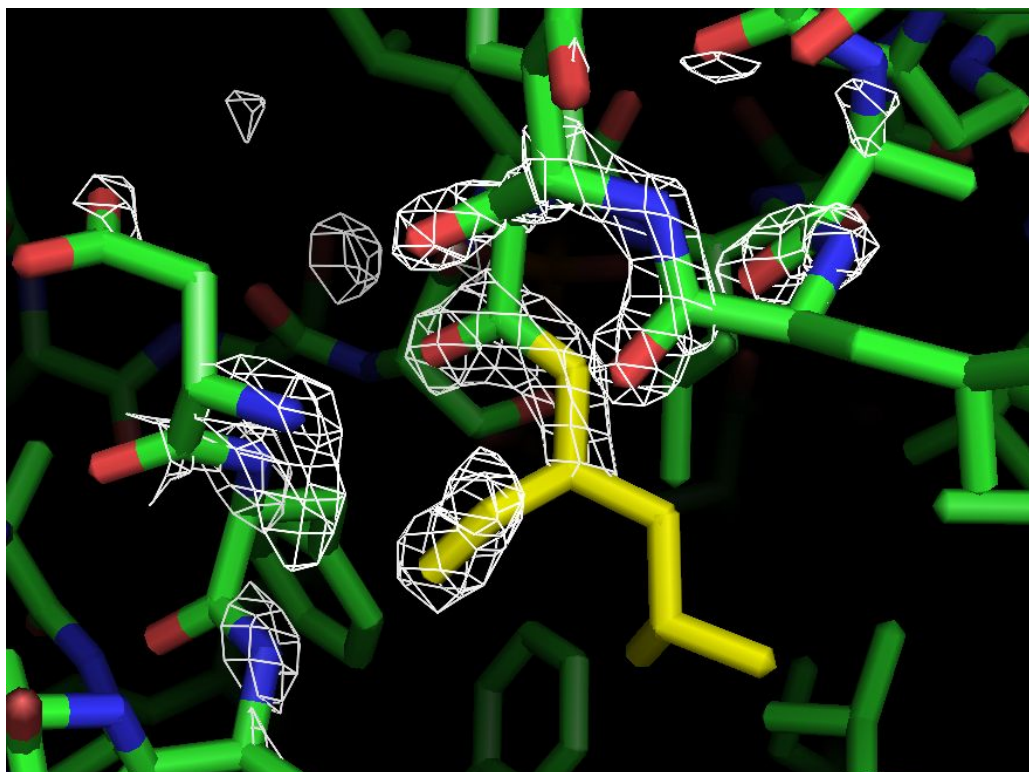


Рис 6. Изображение LEU215

PRO233, chain A

Очень большой clashscore 0.429 Å

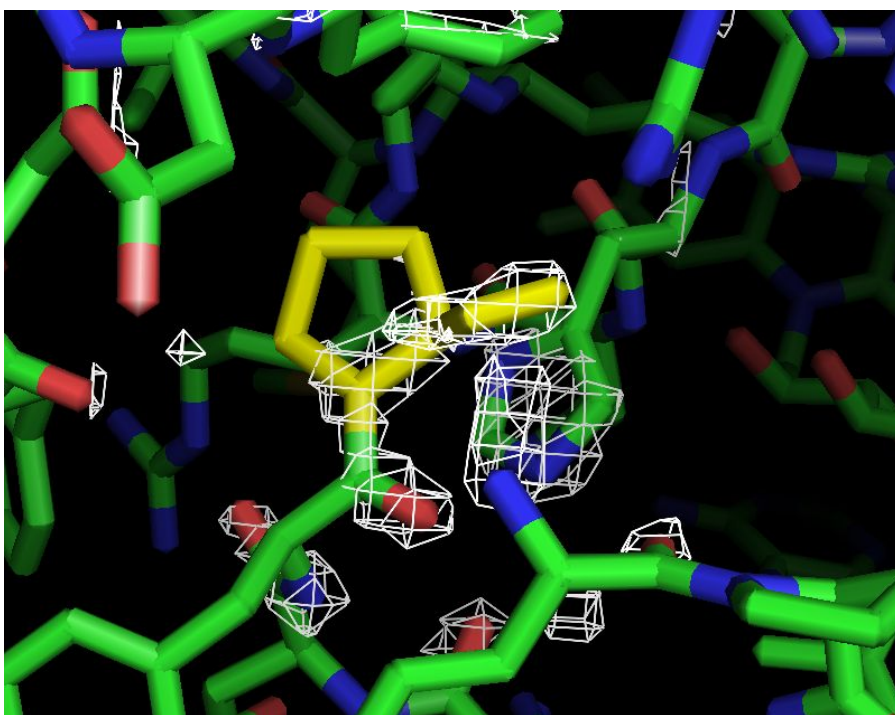


Рис 7. Изображение PRO233

PHE436, chain A

Очень большой clashscore 0.419 Å

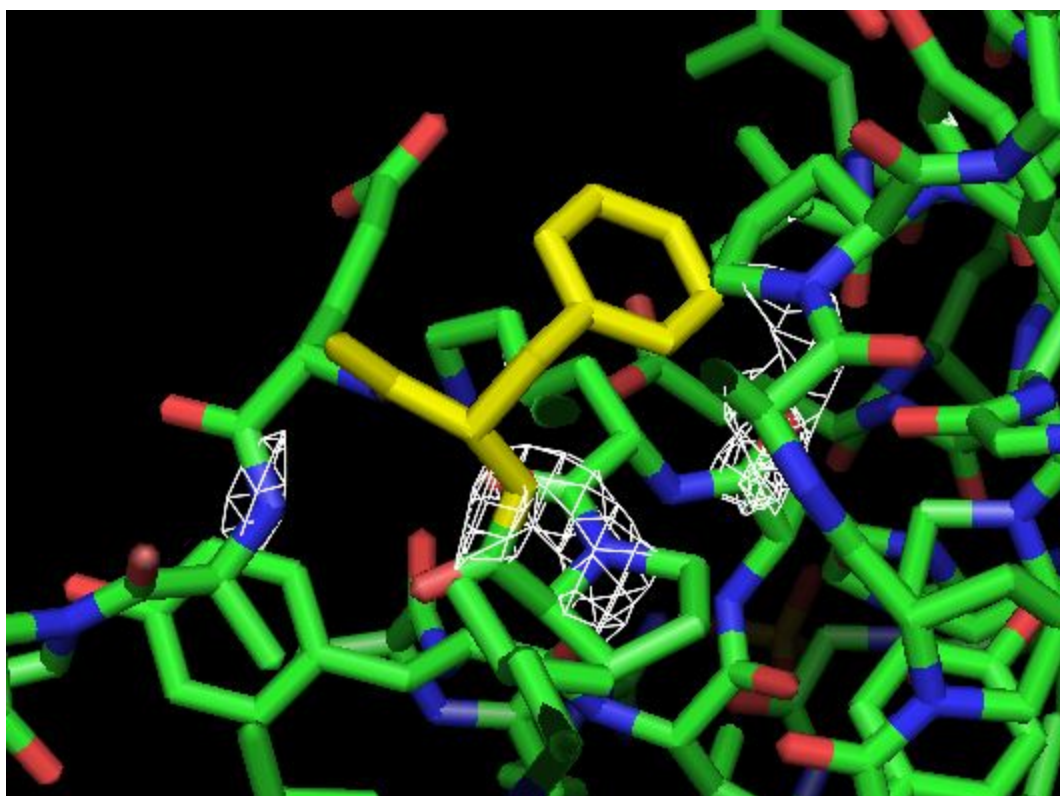


Рис 7. Изображение PHE436

Сравнение модели из PDB с моделью из PDB-REDO [3]

Как видно из таблицы 3, в PDB-REDO не удалось улучшить модель по многим параметрам, особенно стоит отметить разницу между R и R-free ч. Множество остатков было хуже вписано в электронную плотность, однако были убраны некоторые молекулы воды и 6 аминокислотных остатков были лучше вписаны в электронную плотность.

Таблица 3. Сравнение модели из PDB с моделью из PDB-REDO

Параметр	PDB	PDB-REDO
R-фактор	0,1570	0,1784
R-free	0,1812	0,2083
Residues fitting density better	0	6
Ramachandran plot appearance	86	81
Rotamer normality	93	70
Residues fitting density worse	0	128

Заключение

Структура белка 4U63 расшифрована достаточно качественно по меркам средней структуры из PDB. Она имеет разрешение 1.6 Å и её показатели лучше, чем у большинства структур такого же разрешения. По данным из PDB_redo можно заключить, что структура могла быть оптимизирована гораздо лучше, особенно если посмотреть на R-free, ротаметры и вписанные остатки. Рассмотренные маргиналы в основном были получены из-за ошибок расшифровки.

Список литературы

1. Scheerer, Patrick, et al. "The class III cyclobutane pyrimidine dimer photolyase structure reveals a new antenna chromophore binding site and alternative photoreduction pathways." *Journal of Biological Chemistry* 290.18 (2015): 11504-11514.
2. <http://molprobity.biochem.duke.edu>
3. <https://pdb-redo.eu/db/4ri1>