

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА**

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Отчет по качеству расшифровки РСА структуры 4OGS
зеленого флуоресцентного белка
с мутациями S205A и T203V**

Выполнил студент 4-го курса
Пушкарев Сергей Владимирович
в рамках курса структурной биоинформатики

Москва, 2020 г.

Содержание

Содержание	2
Аннотация	2
Введение	3
Результаты и обсуждение	5
Общая информация о модели	5
Состав комплекса	5
История записи PDB	5
Авторы записи PDB	5
Статья с исследованием, в рамках которого получили структуру	5
Метод решения фазовой проблемы	6
Число измеренных рефлексов, разрешение и полнота. Диапазон разрешений структурных факторов	6
Параметры кристаллографической ячейки и кристаллографическая группа	7
Некристаллографические симметрии в асимметрической ячейке	7
Значения показателей качества модели в целом	7
R-фактор и Rfree	7
Карты Рамачандрана	7
Другие индикаторы	9
Маргинальные остатки	10
Анализ пяти избранных маргинальных остатков	10
Ser86, цепь B	10
Val 193, цепь A	13
Asn 170, цепь A	15
Glu 90, цепь B	15
CRO 66, цепь A	16
Сравнение модели PBD и PDB REDO	17
Заключение	18
Благодарности	18
Список литературы	18

Аннотация

В этом отчете рассматривается структура 4OGS. Кратко сформулирована биологическая задача, в связи с которой была получена структура. Также приведены общие данные о модели, показатели ее качества, полученные из разных источников и веб-сервисов.

Рассмотрены маргинальные остатки и коротко произведено сравнение оригинальной структуры со структурой, полученной по протоколу PDB REDO.

Введение

Зеленый флуоресцентный белок (GFP) представляет собой β -бочонок, состоящий из 11 β -тяжей (рис. 1) [1]. В центре β -бочонка располагается флуорофор, спонтанно образующийся из трех аминокислот: Ser65, Tyr66 и Gly67 в присутствии кислорода (рис. 2) [1].

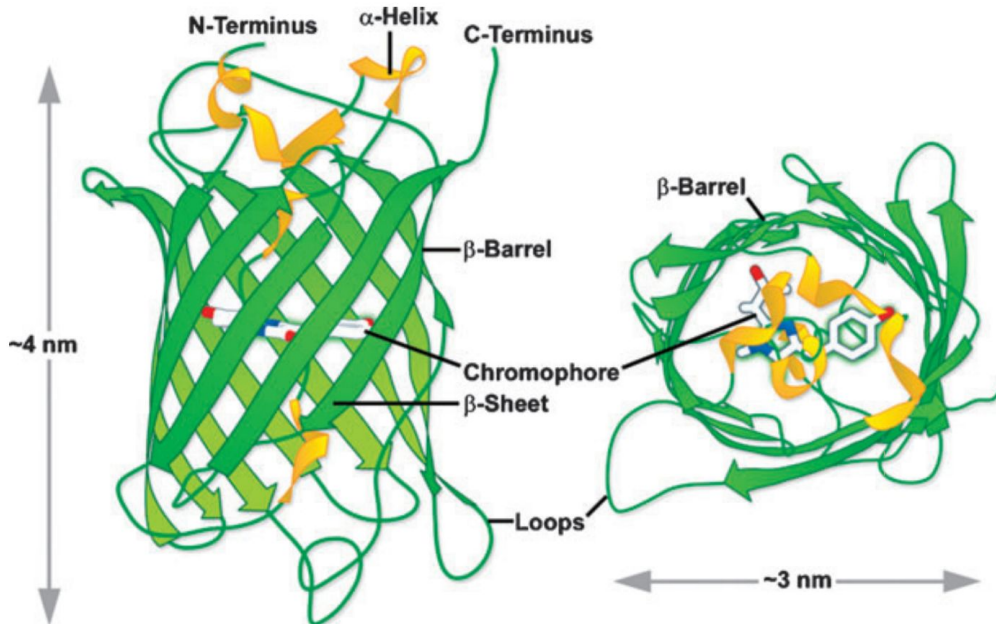


Рисунок 1. Структура GFP. Взято из [1].

Благодаря тому, что для формирования флуорофора необходимо только присутствие кислорода, а N- и C-концы белка свободно находятся на поверхности β -бочонка, GFP представляет из себя идеальный инструмент для создания флуоресцентно меченых fusion-белков [1]. Не будет преувеличением сказать, что открытие и последующее изучение GFP произвели революцию в визуализации биологических процессов на клеточном уровне [1].

Фотофизические процессы, в частности флуоресценция, очень чувствительны к окружению возбужденной молекулы. Мутируя последовательность GFP, можно изменять оптические свойства флуорофора. На сегодняшний день с помощью мутагенеза на основе GFP было получено 186 [2, 3] флуоресцентных белков.

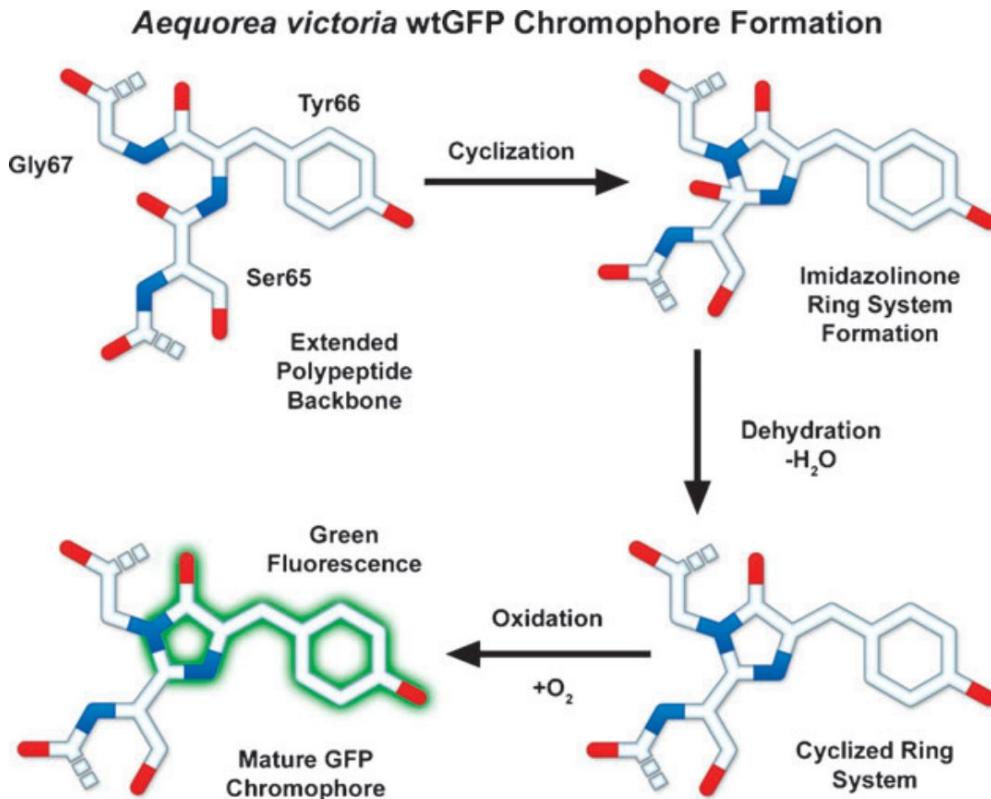


Рисунок 2. Созревание флуорофора GFP. Взято из [1].

В статье, где была получена рассматриваемая в отчете структура [4], авторы исследовали эффект мутаций T203V/S205A на перенос протона в возбужденном состоянии (ППВС) флуорофора. В отличие от мутанта T203V/S205V, для которого в прошлом было показано отсутствие ППВС [5], спектр флуоресценции T203V/S205A указывает на наличие ППВС, что неожиданно, учитывая близость физико-химических свойств аланина и валина.

Для разрешения этого противоречия авторы, среди прочих методов, прибегли к рентгеноструктурному анализу мутанта T203V/S205A с последующим использованием полученной структуры в качестве входных данных молекулярно-динамической симуляции.

Полученная структура депонирована в PDB [6] с кодом 4OGS [7]. О ней и пойдет речь в данном отчете.

Результаты и обсуждение

Общая информация о модели

Помимо самих координат структуры, запись в PDB содержит большое количество метаданных, характеризующих все этапы получения структуры (получение кристаллов, PCA-эксперимент, решение фазовой проблемы, уточнение модели),

состав полученной структуры, а также показатели качества структуры и т.д. В этом разделе описана общая информация о модели и то, как она была получена.

Состав комплекса

Из соответствующей web-страницы в PDB [7], а также из самого файла .pdb структуры получил информацию о составе модели.

В структуре присутствуют две пептидные цепи А и В. Вместо остатков 65-67 в обеих цепях указан один модифицированный остаток б6, соответствующий хромофору.

Помимо двух белковых цепей имеется также один ион Cl⁻ и 155 молекул воды.

Примечательно, что в оригинальной статье в таблице 1 под заголовком “refinement” указано 157 молекул воды и 3521 атомов белка, в то время как в разделе “experimental” [8] страницы PDB в разделе “refinement” указано 155 молекул воды и 3519 атомов белка.

Возможно, авторы представили в статье не окончательные данные. В самом файле .pdb присутствует 155 молекул воды, а также 3480 полей ATOM и 42 поля HETATM, принадлежащих хромофору. Суммарно атомов белка получается 3522, однако, с учетом двух альтернативных конформаций трех атомов серина 30 в цепи А получается 3519 уникальных атомов белка.

Таким образом, можно заключить, что в случае 4OGS web-страница PDB достовернее описывает содержимое файла .pdb, чем авторы соответствующей статьи.

История записи PDB

Из раздела “Entry History” web-страницы структуры в PDB [7] получил дату депозита структуры: 2014-01-16 и дату релиза структуры: 2014-05-14.

Авторы записи PDB

В разделе “Entry History” web-страницы структуры в PDB [7] нашел авторов записи PDB: Remington, S.J., Saif, M.

Статья с исследованием, в рамках которого получили структуру

С web-страницы структуры [7] в разделе “Literature” нашел DOI статьи: 10.1039/c4cp00311j.

Сама статья:

Wineman-Fisher V. et al. Insight into the structure and the mechanism of the slow proton transfer in the GFP double mutant T203V/S205A //Physical Chemistry Chemical Physics. – 2014. – Т. 16. – №. 23. – С. 11211-11223.

Метод решения фазовой проблемы

Согласно статье [4], для решения фазовой проблемы использовался метод молекулярного замещения. Фазы были рассчитаны по структуре мутантного GFP S205V (цепь А), депонированной в PDB под кодом 2QLE [9].

Число измеренных рефлексов, разрешение и полнота. Диапазон разрешений структурных факторов

Согласно разделу “Experimental” записи в PDB [8], было измерено 27588 рефлексов (Number Reflections (Observed)) из 27559 теоретически возможных (Number Reflections

(All)). Полнота этого набора рефлексов высока: 99.9%. Диапазон разрешений: 2.22-50 Å.

Из файла структурных факторов получили, что по какой-то причине при уточнении модели авторы решили сократить диапазон разрешений с 2.22-50 Å до 2.21-15 Å. В результате число рефлексов уменьшилось до 23186.

Подозрительным кажется то, что максимальное разрешение подмножества всех рефлексов лучше, чем глобальное максимальное разрешение.

Из того же файла структурных факторов узнаем, что из этих 23186 рефлексов только 15202 удовлетворяют некоему “observation criterion” [10]. О том, что это за критерий, можно косвенно догадываться по статье авторов структуры [4]: в ней указано, что рефлексов, с интенсивностью больше в два раза, чем значение шума ($I > 2\sigma(I)$) 14 381. Это число не совпадает с 15202 и с 15270 – суммой work и free рефлексов, указанных в разделе “Refinement” [8]. Возможно, 14 381 – это только про work рефлексы. Их в разделе “Refinement” на PDB 14492, снова не совпадение.

Все это выглядит очень странно и непонятно. Рефлексы, не прошедшие “observation criterion” имеют “x” в поле “status”.

Так или иначе, от 23186 рефлексов осталось около 14-15 тысяч. Кажется странным производить фурье-синтез электронной плотности по всем 23186, если в уточнении будет использоваться только 14-15 тысяч. Исходя из этого, я предполагаю, что для фурье-синтеза реально использовалось как раз только 14-15 тысяч рефлексов. В этой связи, полнота полученного набора структурных факторов составляет, согласно validation report, около **65.5%**, что фантастически меньше 95%, на которые было рекомендовано ориентироваться на лекциях и семинарах.

Кажется, что авторы просто выкинули треть данных по какой-то неуказанной причине, не понизив при этом разрешение (большинство “x” рефлексов находятся в конце файла структурных факторов, то есть соответствуют рефлексам с высоким разрешением). Возможно, они хотели таким образом оставить разрешение своей структуры высоким, ведь на него смотрят чаще, чем на полноту. Несовпадение данных статьи, страницы pdb и файла структурных факторов оставляет очень неприятное впечатление.

Параметры кристаллографической ячейки и кристаллографическая группа

Кристаллографическая ячейка задается тремя длинами ребер a, b, c и тремя углами α , β , γ между ними. Все возможные симметрии ячейки формируют группу симметрии. Согласно wwPDB Validation Report [11], a,b,c, α , β , γ кристаллографической ячейки имеют значения 86.61 Å, 86.61 Å, 119.23 Å, 90.00°, 90.00°, 90.00° соответственно. Ячейка имеет группу симметрии P 41 21 2.

Некристаллографические симметрии в асимметрической ячейке

Согласно документации формата mtcif на wwPDB [12], данные о некристаллографических симметриях должны содержаться в поле “struct_ncs_oper”, которого в mtcif файле для нашей структуры нет. Исходя из этого, заключаю, что некристаллографических симметрий нет.

Значения показателей качества модели в целом

В этом разделе рассмотрены показатели, характеризующие качество всей модели, полученные как из страницы структуры в PDB, так и с помощью других сервисов.

R-фактор и R_{free}

Согласно странице структуры в PDB, R-фактор по всем рефлексам (R-value Observed) составил 0.187, R-work – 0.181, R-free – 0.306. Налицо довольно большая разница (> 0.1) между R-work и R-free, которая может свидетельствовать об оверфиттинге.

Авторы статьи комментируют это высокое значение R-free как “следствие высокой мозаичности кристалла и плохого качества данных”. Возможно, авторы из-за этого плохого качества данных и были вынуждены выкинуть треть рефлексов.

Карты Рамачандрана

Для построения карт Рамачандрана использовался сервис MolProbity [13].

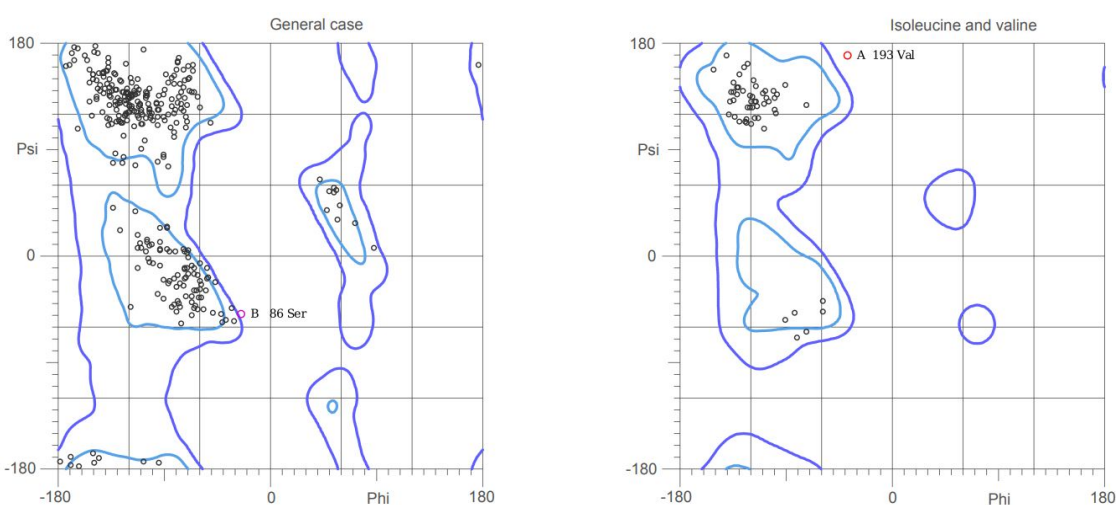


Рисунок 3. Карты Рамачандрана для общего случая (все остатки, кроме глицина, пролина и препролиновых) и для изолейцина и валина.

Из рисунка 3 видно, что большинство остатков находятся в разрешенных областях. Видим двух аутлаеров: Ser86 цепи B и Val 193 цепи A.

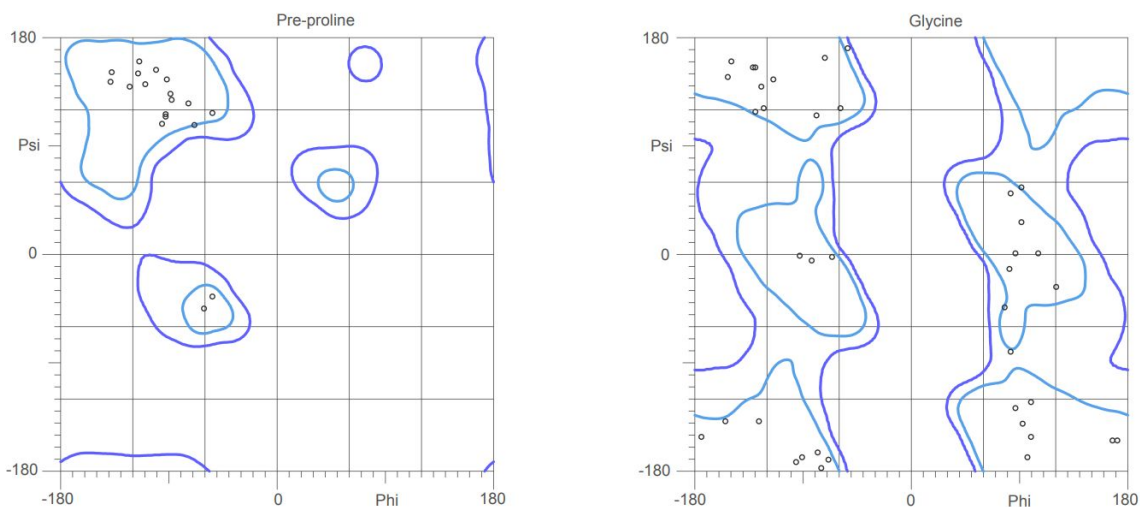


Рисунок 4. Карты Рамачандрана для препролиновых остатков (слева) и глицинов (справа).

Здесь заметных отклонений не найдено.

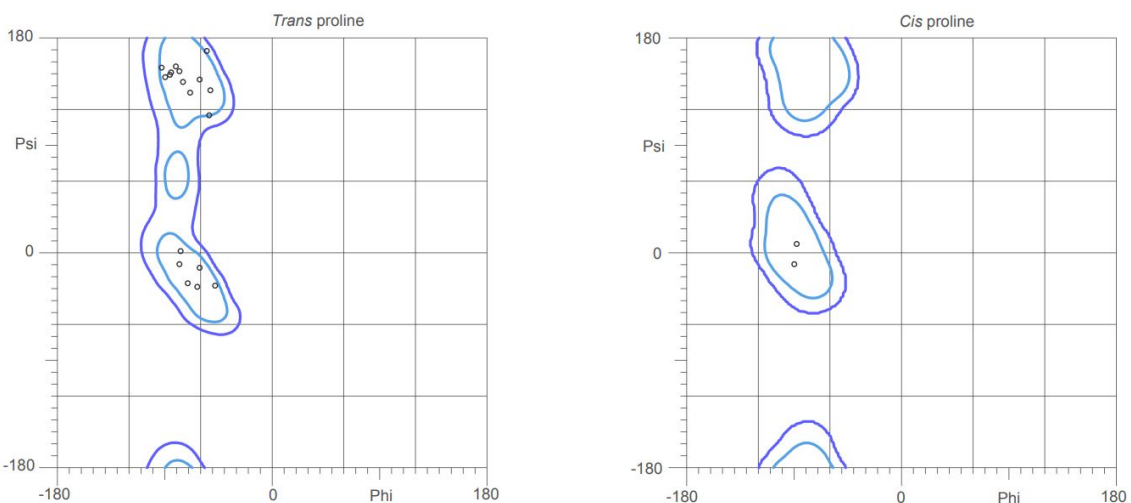


Рисунок 5. Карты Рамачандрана для транс-пролинов (слева) и цис-пролинов (справа).

Как видно из рисунка 5, В анализируемой структуре имеется два цис-пролина. Помимо карт, MolProbity подвел итог: 93.9% (416/443) всех остатков находятся в благоприятных регионах (98%) регионах. 99.5% (441/443) всех остатков находятся в разрешенных (>99.8%) регионах. Тем не менее, в статистике MolProbity по структуре показатель "Rama distribution Z-score" отмечен зеленым, что говорит о том, что в целом распределение торсионных углов остова хорошее.

Другие индикаторы

Из других индикаторов я рассмотрел статистику MolProbity и RSRZ outliers.

Таблица 1. Статистика MolProbity.

Summary statistics

All-Atom Contacts	Clashscore, all atoms:	9.41		90 th percentile* (N=457, 2.21Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	19	5.29%	Goal: <0.3%
	Favored rotamers	294	81.89%	Goal: >98%
	Ramachandran outliers	2	0.45%	Goal: <0.05%
	Ramachandran favored	415	93.89%	Goal: >98%
	Rama distribution Z-score	-1.51 ± 0.36		Goal: abs(Z score) < 2
	MolProbity score [^]	2.45		55 th percentile* (N=10183, 2.21Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	3 / 3563	0.08%	Goal: 0%
Peptide Omegas	Bad angles:	0 / 4834	0.00%	Goal: <0.1%
	Cis Prolines:	2 / 20	10.00%	Expected: ≤1 per chain, or ≤5%
Additional validations	Chiral volume outliers	0/529		
	Waters with clashes	12/155	7.74%	See UnDowser table for details

Общей мерой качества структуры по clashscore (перекрывания vdW-радиусов), качеству ротамеров и распределению остовных торсионных углов является MolProbity score. Исходя из него, наша структура чуть выше среднего среди 10000 структур с подобным разрешением.

Clashscore у структуры достойный, однако по много плохих ротамеров и неидеальная карта Рамачандрана.

С web-страницы PDB извлекли следующие показатели:

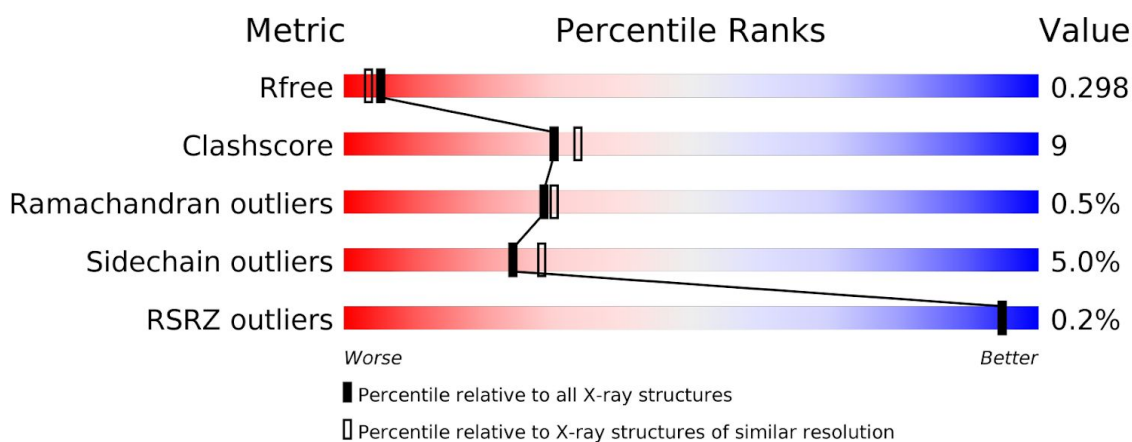


Рисунок 6. Показатели качества структуры.

Здесь, помимо четырех вышеперечисленных параметров есть RSRZ outliers, характеризующий сколько остатков в нашей структуре описываются хуже, чем в среднем по структурам с похожим разрешением.

Видим, что таких остатков очень мало, что можно трактовать как то, что модель хорошо вписана в наблюдаемую электронную плотность (ЭП) в том смысле, что теоретические значения ЭП атомов модели хорошо совпадают с наблюдаемыми. При прекрасном соответствии остатков плотности получаем остальные метрики ниже среднего. В виду вышеназванных проблем с полнотой набора рефлексов кажется очень смелым доверять такой плотности. То есть, то, что в нее что-то хорошо вписали не много говорит об адекватности вписываемой модели. Мы увидим это на примере анализа одного из маргиналов.

Маргинальные остатки

Ниже в таблице 2 сведены некоторые маргинальные остатки и критерии, по которым они являются маргинальными. Остатки были найдены с помощью MolProbity и/или указанные в wwPDB Validation Report [11].

Таблица 2. Некоторые маргинальные остатки.

Остаток, цепь	Критерий маргинальности
Ser86, B	Аутлаер на карте Рамачандрана
Val 193, A	Аутлаер на карте Рамачандрана
Asn 170, A	Подозрение на инверсию амидной группы
Asn 170, B	Подозрение на инверсию амидной группы
Asn 149, A	Подозрение на инверсию амидной группы
Glu 90, B	Нереалистичная длина связи CB-CG (1.82 против 1.52 в идеальном случае)
Tyr 145, A	Неротамерный радикал
Glu 34, B	Неротамерный радикал
CRO 66, A	Аутлаер по геометрии (множество связей и углов)
Gly 4, B	RSRZ аутлаер

Анализ пяти избранных маргинальных остатков

1. Ser86, цепь B

Серин 86 является маргиналом по положению на карте Рамачандрана. Однако, видно, что он находится не прямо в середине запрещенной области, а всего лишь чуть-чуть выходя за границу разрешенной области.

Рассмотрим сначала положение остатка в структуре, возможно, его маргинальность объясняется какой-то функциональной значимостью.



Рисунок 7. Положение Ser 86 в структуре.

Как можно видеть, Ser 86 расположен достаточно близко к петле, которая проходит сквозь β -бочонок и образует хромофор. Возможно, такая структура довольно нетипична и маргинальность остатка объясняется этим? Рассмотрим подробнее.

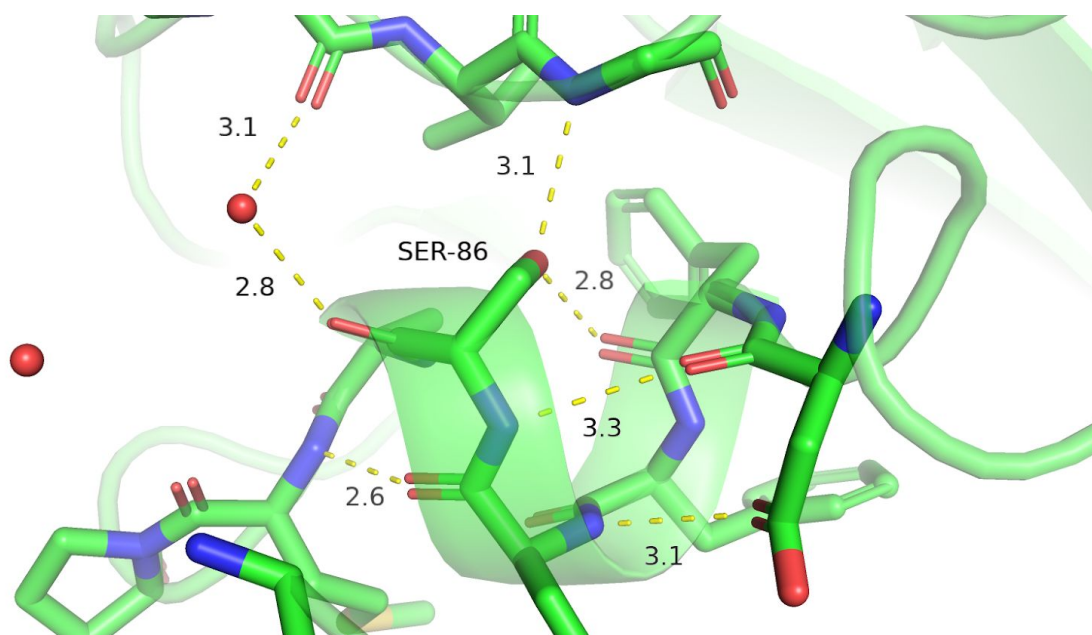


Рисунок 8. Окружение Ser 86.

Исходя из рисунка 8, Ser 86 участвует в сети водородных связей, которые могут быть важны для поддержания структуры. Посмотрим на электронную плотность.

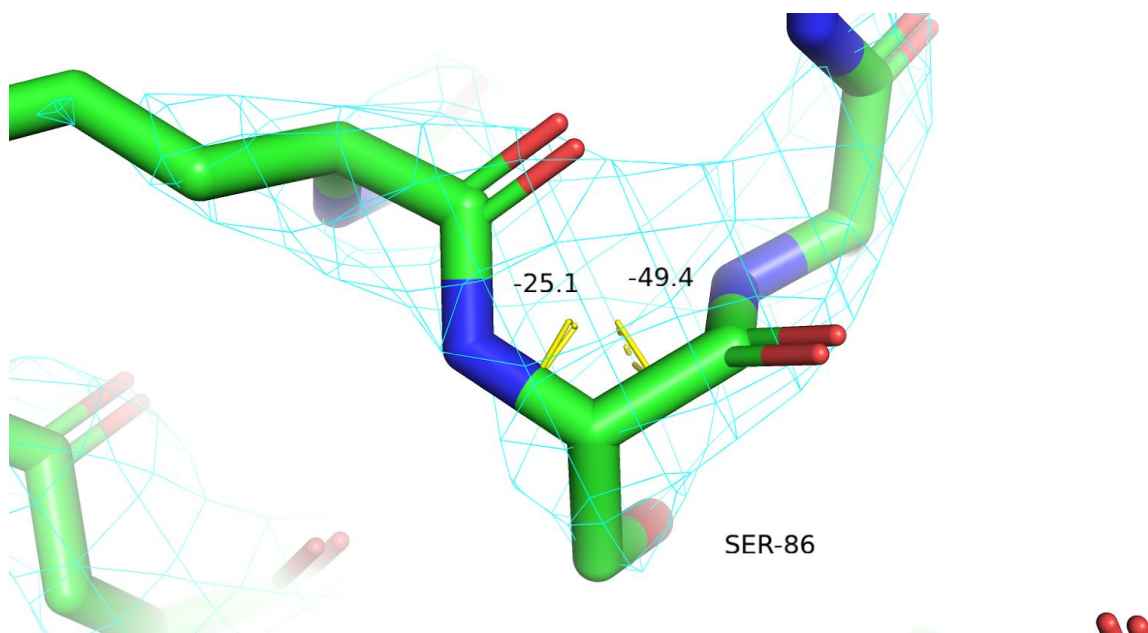


Рисунок 9. Электронная плотность на уровне подрезки 2 сигма, подписаны торсионные углы остова.

Плотность довольно расплывчатая, отдельные атомы не видны. Руководствуясь только этой структурой не очевидно, имеем мы ошибку в интерпретации плотности или остаток имеет какое-то важное функциональное значение.

Обратимся к структуре 2QLE, которая использовалась для решения фазовой проблемы, возьмем цепь А, которая использовалась авторами. Структура имеет разрешение 1.59 Å, возможно, она поможет нам решить судьбу Ser 86.

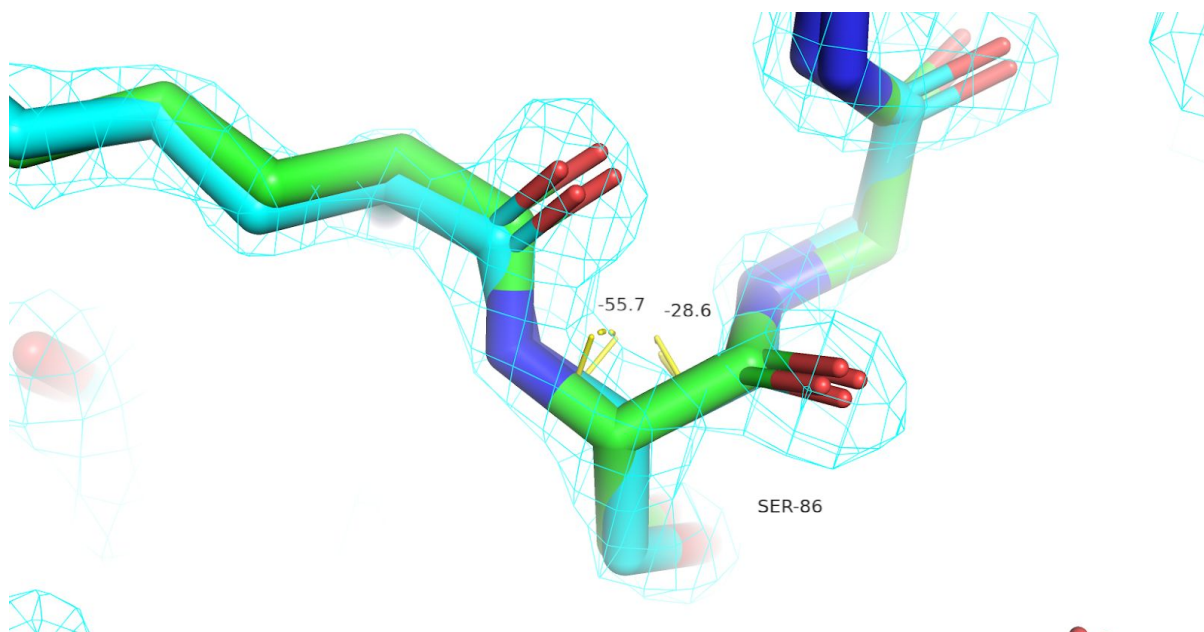


Рисунок 10. Наложенные 4ogs и 2qle. Электронная плотность 2qle на уровне подрезки 2 сигма. Атомы 2qle обозначены бирюзовым, атомы 4ogs – зеленым. Указаны торсионные углы для Ser 86 из цепи А 2qle.

Как видно из рисунка 10, структуры похожи в рассматриваемом участке. Лучшее качество плотности 2qle придает уверенности в использовании ее как референса. Исходя из структуры 2qle, Ser 86 может выполнять свою функцию в структуре и иметь при этом идеальные значения торсионных углов остова (-55.5 и -28.6 попадает в оптимальный регион карты Рамачандрана).

Таким образом, Ser 86 цепи В структуры 4ogs является аутлаером вследствие низкого качества электронной плотности и/или ее ошибочной интерпретации.

2. Val 193, цепь А

Аналогично для Val 193 из цепи А посмотрим его расположение в структуре белка.

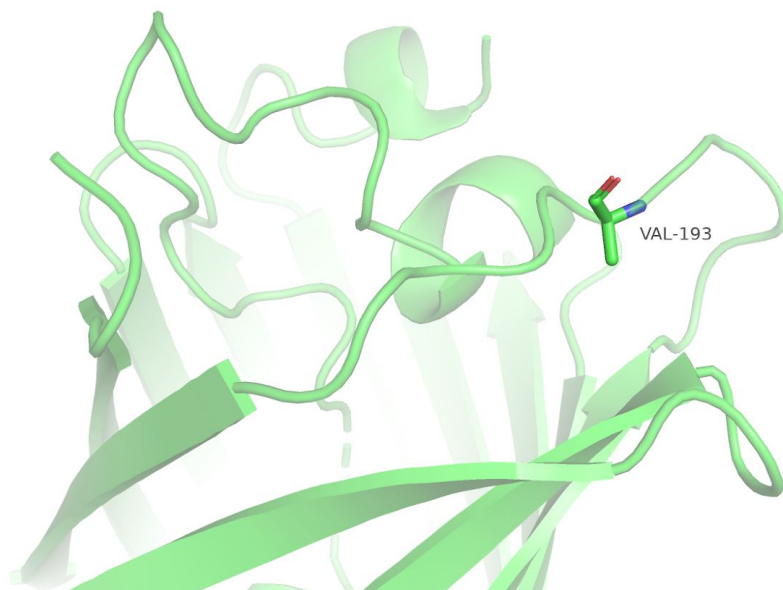


Рисунок 11. Положение Val 193 в структуре.

По сравнению с Ser 86, Val 193 находится в петле, экспонированной в раствора и просто соединяющей два β -тяжа. Следовательно, к гипотезе о его большой функциональной значимости относиться следует с большим недоверием.

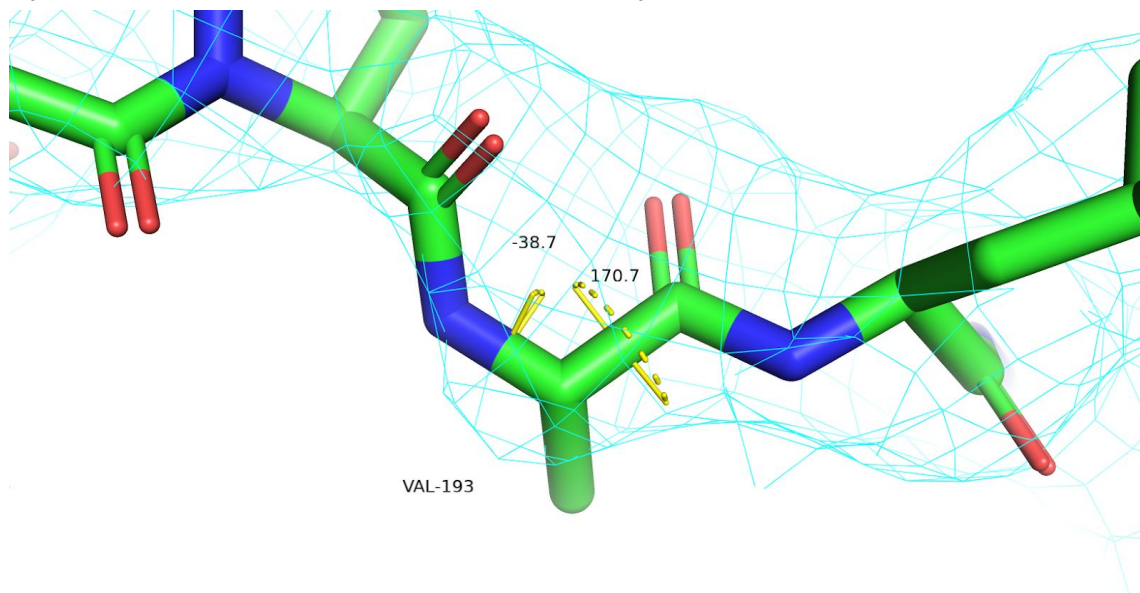


Рисунок 12. Электронная плотность вокруг Val 193 на уровне подрезки 1 сигма. Указаны торсионные углы остова.

Видим близость двух карбонильных кислородов, расстояние между которыми 2.7 Å. Крайне неправдоподобно.

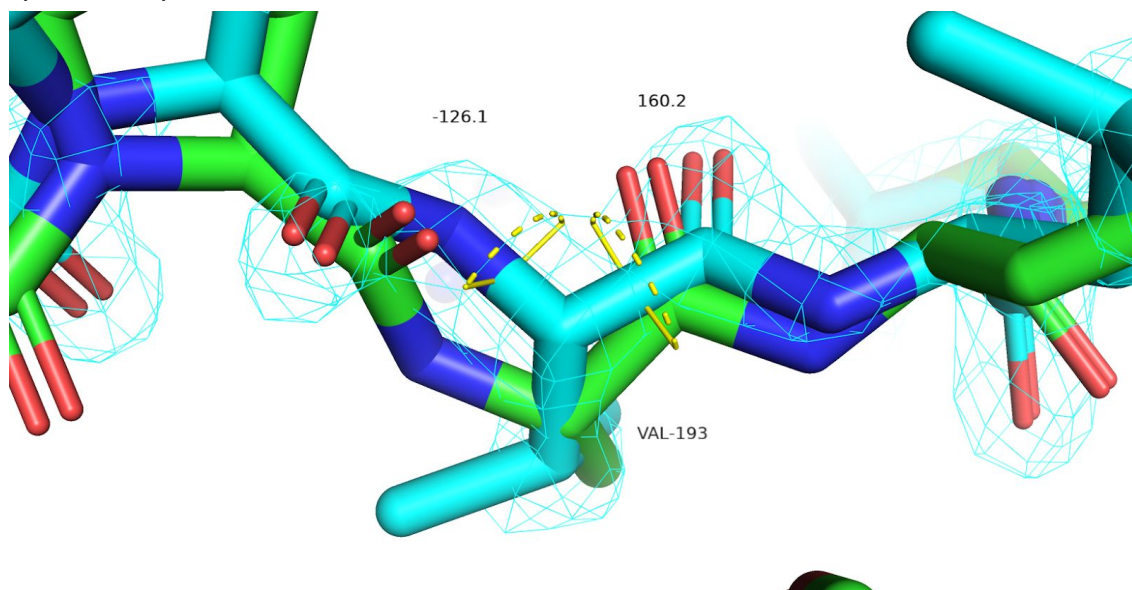


Рисунок 13. Наложенные 4ogs и 2qlе. Электронная плотность 2qlе на уровне подрезки 2 сигма. Атомы 2qlе обозначены бирюзовым, атомы 4ogs – зеленым. Указаны торсионные углы для Val 193 из цепи A 2qlе.

В референсной структуре (2qlе) оба торсионных угла лежат в оптимальной области карты Рамачандрана.

Вердикт аналогичный Ser 86: Val 193 цепи A структуры 4ogs является аутлаером вследствие низкого качества электронной плотности и/или ее ошибочной интерпретации.

3. Asn 170, цепь A

Для аспарагина 170 цепи A есть подозрение на перепутанные атомы N и O амидной группы.

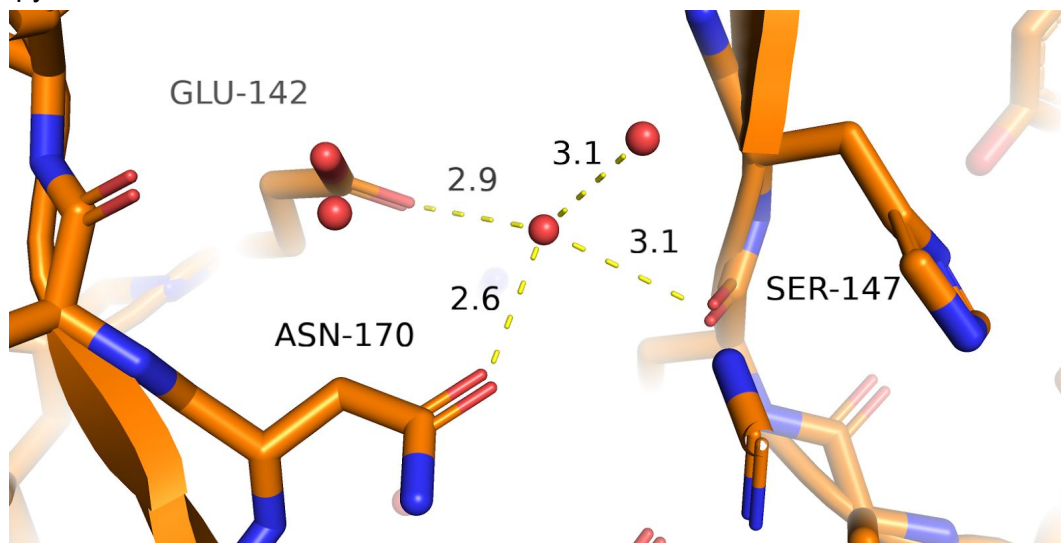


Рисунок 14. Окружение Asn 170.

При подробном рассмотрении окрестностей остатка видим воду рядом с как минимум тремя (учитывая кислород Asn 170) акцепторами водородной связи. Такая ситуация маловероятна, поскольку молекула воды может быть донором водородной связи только дважды.

Вывод: боковая цепь действительно была вписана неправильно.

4. Glu 90, цепь В

Связи С-С длиной около 1.8 ангстрем описаны [14], однако это, безусловно, аномалия, существующая только в виду сильной стерической загруженности участвующих атомов углерода. Подобной загруженности в атомах глутамата белка, конечно, не наблюдается. Потому, это явно ошибка интерпретации ЭП.

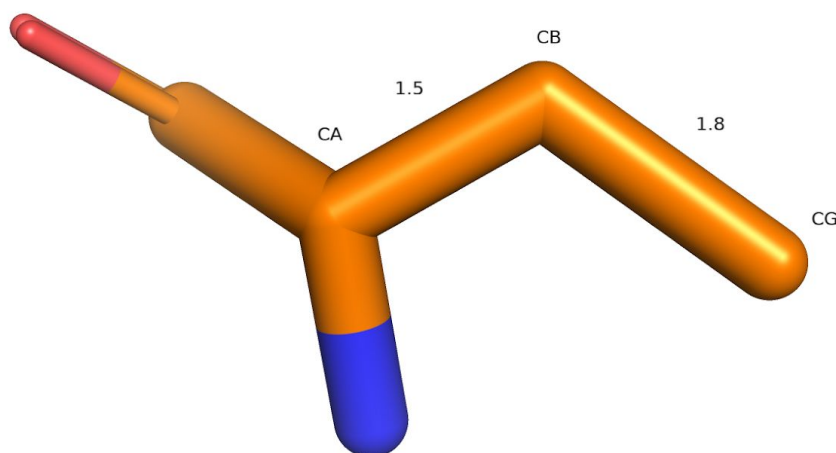


Рисунок 15. Glu 90 цепи В. Аномально длинная связь Cb-Cg. Карбоксильная группа не разрешена.

Этот остаток является ярким примером того, что вписывая атомы в некачественную ЭП, мы можем получить неправильную геометрию.

Интересно также то, что по RSRZ validation report [11] не считает этот остаток маргиналом. Ведь RSR оценивает только соответствие геометрии электронной плотности. Если ЭП плохая и геометрия такая же плохая (ЭП растягивает какую-то связь и мы вписываем эту связь длинной), то такой остаток будет иметь немаргинальные значения RSR.

Наверное поэтому для структур существует столько разных критериев качества. Каждый критерий проверяет адекватность только в каком-то одном смысле и может пропускать неадекватные по другим критериям остатки.

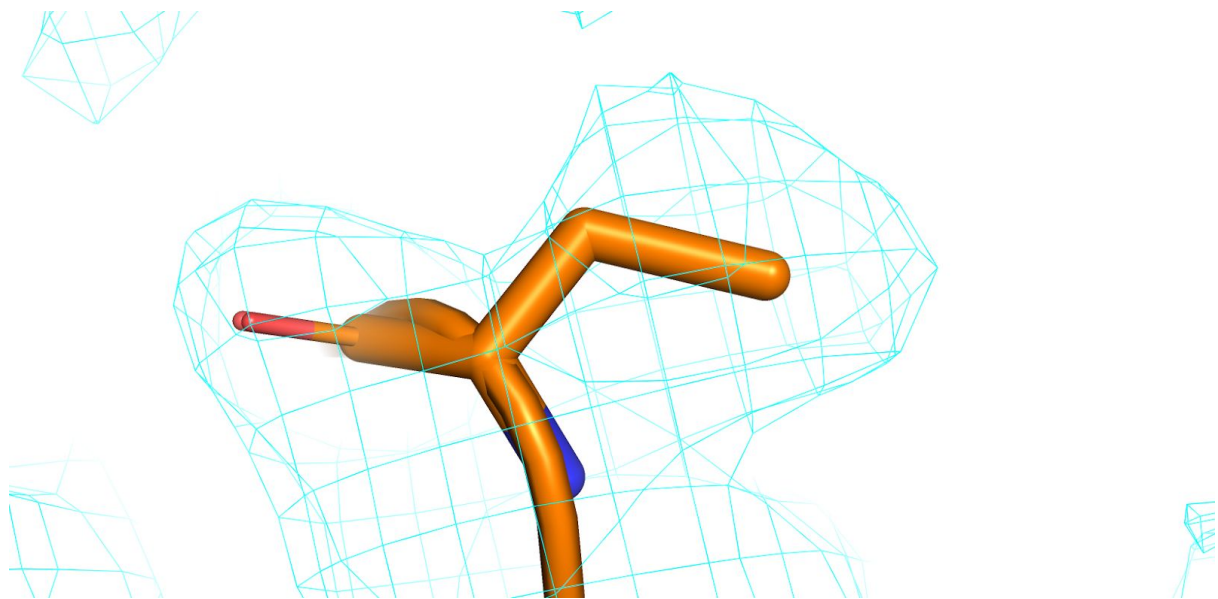


Рисунок 16. Электронная плотность вокруг радикала Glu 90 цепи В на уровне подрезки 0.5 сигма.

5. CRO 66, цепь А

Согласно validation report, хромофор имеет много отличий относительно модельного соединения в Chemical Component Dictionary [15]: 6 “плохих” связей и 8 “плохих” углов. Нужно оговориться, что хромофор GFP может существовать как в виде нейтральной молекулы, так и в виде аниона. В этой связи, сравнение усредненной геометрии по кристаллу (где хромофор в окружении белка и имеет другую рКа) с геометрией из, например, раствора или с геометрией по банку PDB (где лежат разные белки с разной рКа хромофора) может быть не вполне корректным. Более того, просто рН при кристаллизации одного и того же белка должен приводить к разному соотношению двух форм.

Дополнительно проблемы создает возможность радиационного повреждения хромофора (захват хромофором электрона) [16], что усиливает гетерогенность молекул хромофора в кристалле.

Хромофор, безусловно, является “сердцем” GFP, и точная интерпретация его ЭП очень важна для каких-то выводов по структуре. Однако, в зависимости от эксперимента, гетерогенность молекул хромофора может вызывать отклонения от какого-либо референса или существующей статистики по PDB.

Например, в структуре 2qlе с разрешением 1.59 Å в цепи А, которую использовали для решения фазовой проблемы с 4ogs, наблюдается 6 “плохих” связей и 12 “плохих” углов. Более того, даже в структуре 6JGJ [17], которая является одной из наиболее хорошо разрешенных структур GFP (0.78 Å) наблюдается 3 “плохие” связи и 3 “плохих” угла.

Таким образом, в контексте GFP аутлаеры по структуре хромофора – это совсем нетривиальная вещь, которая вполне может не являться ошибкой в интерпретации плотности. Поэтому я не считаю, что в 4ogs CRO 66 является аутлаером.

Сравнение модели PBD и PDB REDO

Наша структура была найдена в банке PDB REDO [18]. В первую очередь поразило то, что аж у 117 остатков **ухудшилось** соответствие модели наблюдаемой плотности. Улучшилось же только у одного. На первый взгляд это совершенно контринтуитивно, однако, припоминая качество ЭП, можно понять, что именно это и должно произойти! В контексте этой структуры, лучше структура, которая хуже соответствует плохой плотности, но зато меньше отклоняется от здравого смысла (вроде той длинной связи в глутамате 90 или аулаеров карты Рамачандрана).

Таблица 3. Сравнение исходной и PDB REDO структур.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
<i>R</i>	0.1765	0.2391
<i>R-free</i>	0.3063	0.2944
<i>Bond length RMS Z-score</i>	0.643	0.475
<i>Bond angle RMS Z-score</i>	0.837	0.756
Model quality (raw scores percentiles)		
<i>Ramachandran plot appearance</i>	25	58
<i>Rotamer normality</i>	2	24
<i>Coarse packing</i>	N/A	N/A
<i>Fine packing</i>	2	13
<i>Bump severity</i>	34	81
<i>Hydrogen bond satisfaction</i>	43	28

Как можно видеть, R-value действительно повысился в REDO структуре, но зато все остальные показатели, кроме Hydrogen bond satisfaction, сильно улучшились.

Заключение

Подведу итог. Непонятные непосвященным манипуляции с количеством рефлексов и полнота 65.5% оставляют неприятное впечатление. Возможно, где-то просто люди промахнулись мимо клавиши, когда писали число рефлексов.

На примере аутолаеров карты Рамачандрана и длинной связи в Glu 90 было показано как плохая ЭП приводит к плохим результатам. В таких случаях при применении протокола PDB REDO наблюдается контринтуитивное падение соответствия модели электронной плотности на фоне улучшения большинства показателей качества.

Учитывая то, что структур GFP в PDB, скорее всего, очень много и есть даже структуры с субатомарным разрешением [17], я бы не стал использовать 4ogs для своей работы. Если бы был необходим мутант по тем же положениям, что и в 4ogs, возможно, надежнее было бы получить его чисто *in silico*, мутировав остатки в хорошей структуре и минимизировав окружение.

Несмотря на все претензии к структуре, свою биологическую задачу она выполнила. Общий фолд, хромофор и мутантные остатки, в которых были заинтересованы исследователи, были хорошо различимы.

Даже стартовав с такой структуры, с помощью минимизации энергии, которая всегда предшествует молекулярной динамике, неточности в геометрии были, скорее всего, устранены. В ходе симуляции авторы смогли предложить механизм ППВС, разрешающий противоречие между наличием ППВС в T203V/S205A и его отсутствием в T203V/S205V.

Благодарности

Хочу поблагодарить преподавателей курса структурной биоинформатики за наглядную и увлекательную подачу материала, а также за возможность послушать приглашенных специалистов в области разрешения структур.

Очень ценным также является формирование критического отношения к структурам в PDB. Курс значительно расширил представления о существующих инструментах работы со структурами и, в целом, о микроскопической “жизни” биомолекул.

Список литературы

1. Day R. N., Davidson M. W. The fluorescent protein palette: tools for cellular imaging //Chemical Society Reviews. – 2009. – Т. 38. – №. 10. – С. 2887-2921.
2. Lambert T. J. FPbase: a community-editable fluorescent protein database //Nature methods. – 2019. – Т. 16. – №. 4. – С. 277-278.
3. <https://www.fpbases.org/>
4. Wineman-Fisher V. et al. Insight into the structure and the mechanism of the slow proton transfer in the GFP double mutant T203V/S205A //Physical Chemistry Chemical Physics. – 2014. – Т. 16. – №. 23. – С. 11211-11223.
5. Shu X. et al. An alternative excited - state proton transfer pathway in green fluorescent protein variant S205V //Protein Science. – 2007. – Т. 16. – №. 12. – С. 2703-2710.
6. <https://www.rcsb.org/>
7. <https://www.rcsb.org/structure/4OGS>
8. <https://www.rcsb.org/experimental/4OGS>
9. <https://www.rcsb.org/structure/2QLE>
10. http://mmcif.wwpdb.org/dictionaries/mmcif_pdbx_v50.dic/items/_diffn_reflns.pdbx_number_obs.html
11. https://files.rcsb.org/pub/pdb/validation_reports/og/4ogs/4ogs_full_validation.pdf
12. <http://mmcif.wwpdb.org/pdbx-mmcif-home-page.html>
13. Williams C. J. et al. MolProbity: More and better reference data for improved all - atom structure validation //Protein Science. – 2018. – Т. 27. – №. 1. – С. 293-315.
14. Ishigaki Y. et al. Longest C–C single bond among neutral hydrocarbons with a bond length beyond 1.8 Å //Chem. – 2018. – Т. 4. – №. 4. – С. 795-806.
15. <https://www.ebi.ac.uk/pdbe-srv/pdbechem/chemicalCompound/show/CRO>
16. <https://link.springer.com/book/10.1007/978-1-4939-7000-1>
17. Takaba K. et al. Subatomic resolution X-ray structures of green fluorescent protein //IUCrJ. – 2019. – Т. 6. – №. 3. – С. 387-400.
18. <https://pdb-redo.eu/db/4ogs>