

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Отчёт по качеству РСА расшифровки структуры
комплекса NEC вируса герпеса первого типа (PDB:
4ZXS)**

Работа студента 4-го курса
Кан Марины Петровны

Москва

2020 г.

Оглавление

Список сокращений	2
Аннотация	3
Введение.....	3
Результаты и обсуждение.....	3
Общая информация о модели.....	3
Значения индикаторов качества модели в целом	4
14 Маргинальных остатков	6
Анализ 5 маргинальных остатков	7
Сравнение модели из PDB с моделью из PDB_redo	12
Заключение	13

Список сокращений

HSV-1 – Вирус герпеса первого типа;

NEC - nuclear egress complex, комплекс выхода из ядра;

RSR - Real-space R-value, пространственный R-фактор;

RSRZ - Real-space R-value Z-score, Z-score пространственного R-фактора;

Аннотация

При помощи различных сервисов была получена и систематизирована информация, необходимая для оценки качества модели структуры 4ZXS, были выявлены и рассмотрены детально некоторые маргинальные остатки, а также проведено сравнение модели PDB с моделью из PDB_redo.

Введение

Вирусы герпеса способны инфицировать почти всех позвоночных и некоторых моллюсков в том числе человека. Вирус герпеса первого типа (HSV-1) вызывает у человека такие заболевания, как герпетический стоматит, губной герпес и генитальный герпес, но обычно находится в латентном состоянии и проявляется при иммунодефиците. Особенностью герпесвирусов является то, что те упаковываются в капсид в ядре, но размер капсида не позволяет свободно проходить через ядерные поры, поэтому вирусы герпеса используют особый способ отпочковывания. Связываясь со внутренней ядерной мембраной, вирус образует перинуклеарные частицы, которые впоследствии связываются со внешней ядерной мембраной, высвобождая капсиды в цитоплазму. В этом процессе принимает участие nuclear egress complex (NEC), состоящий из двух консервативных вирусных белков, UL31 и UL34. Чтобы разобраться механизме NEC-опосредованного отпочковывания, авторы определили кристаллические структуры NEC [1]. В данном отчёте будет изучено качество расшифровки комплекса NEC (PDB 4ZXS).

Результаты и обсуждение

Общая информация о модели

Структуру расшифровали в 2015 году учёные из Бостона: Janna M. Bigalke и Ekaterina E. Heldwein [1]. Комплекс NEC состоит из двух белков, которые в свою очередь состоят из двух цепей. Белок UL34 состоит из цепей А и В, а белок UL31 из цепей С и D (Рис. 1).

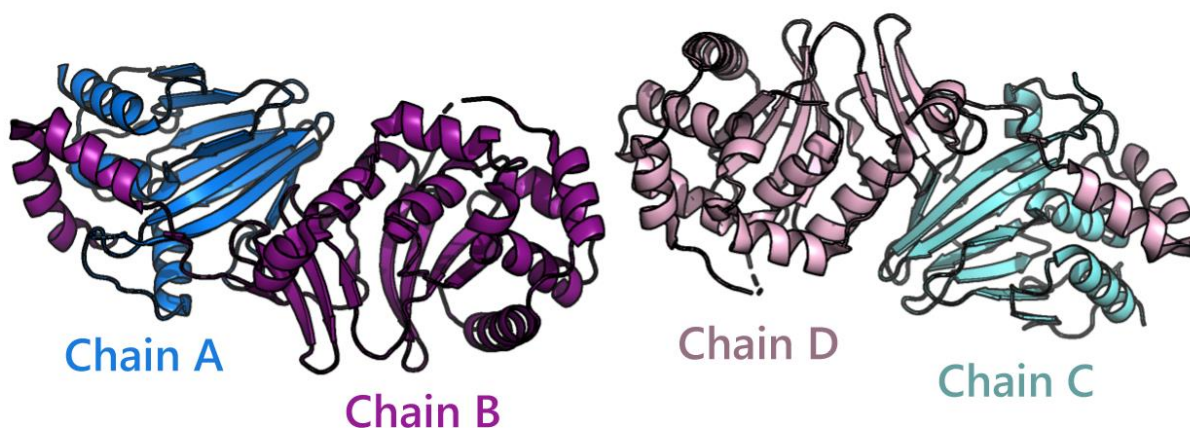


Рисунок 1. Структура комплекса NEC (PDB: 4ZXS) в виде Cartoon. Раскраска по цепям.

Таблица 1. Общая информация о модели 4ZXS [2].

Метод решения фазовой проблемы	Молекулярное замещение
Разрешение	2.77 Å
Число измеренных рефлексов	27433

Полнота набора структурных факторов	99.77%	
Диапазон разрешений структурных факторов	2.772 - 47.86	
Кристаллографическая группа	P 6	
Наличие некристаллографических симметрий в асимметрической ячейке	Нет	
Параметры кристаллографической ячейки	Длина, Å	Угол, °
	a = 110.529	$\alpha = 90$
	b = 110.529	$\beta = 90$
	c = 155.85	$\gamma = 120$

Значения индикаторов качества модели в целом

Качество структуры можно оценить при помощи различных параметров. В их числе R-фактор и R-free, характеризующие соответствие модели экспериментальным данным. R-free строится так же, как и R-фактор, с той разницей, что набор измеренных рефлексов разбивается на два набора – контрольный и рабочий. Это необходимо, чтобы избежать переоптимизации. Считается, что R-free <20% - хорошее значение, R-free >40% - плохое значение. Если разность R-фактора и R-free по модулю больше 10%, то имеет место переоптимизация.

В случае со структурой 4ZXS, имеем значение R-free = 26.47%, R-фактор = 21.95%. Качество модели не дотягивает до хорошего всего на 6.47%. Разность R-фактора и R-free = 4.52%, что говорит о том, что авторы избежали переоптимизации.

Кроме того, существуют показатели качества отдельных остатков: RSR, комфортность окружения остатка, торсионные углы (карты Рамачандрана и ротамеры боковых цепей). Маргиналами называют остатки с нетипичными значениями перечисленных выше показателей.

Маргиналов по RSRZ = 4.7%, что хуже, чем в структурах с похожим разрешением, маргиналов по ротамерам всего 2.7%, что лучше, чем в структурах со схожим разрешением. Маргиналы по карте Рамачандрана отсутствуют в данной структуре (Рис. 2), но это обсудим позже.

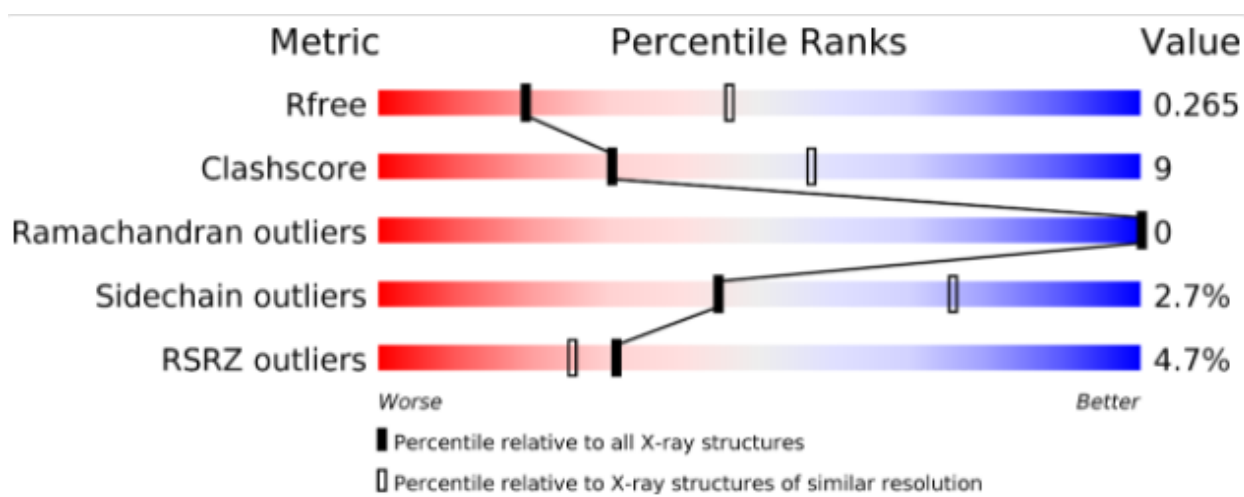


Рисунок 2. Показатели качества модели структуры 4ZXS из базы данных PDB [2].

Сервис MolProbity позволяет получить дополнительные характеристики качества структуры (Рис 3):

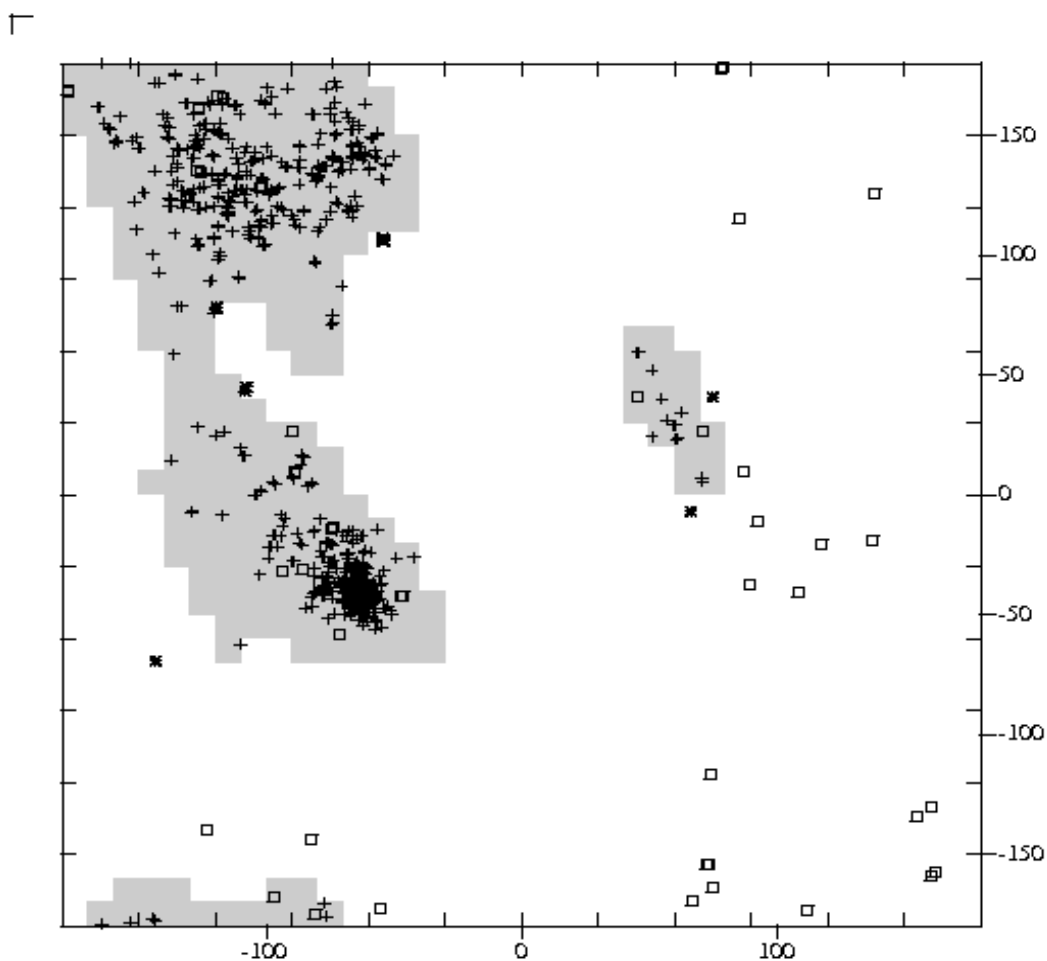
- "Clashscore" – число недопустимых наложений атомов на 1000.
- "Poor rotamers" — число остатков с маргинальными по отклонению от ротамеров боковыми цепями.
- "Favored rotamers" — число и процент остатков в предпочитаемой области по отклонению от ротамеров боковыми цепями.
- "Ramachandran outliers" — полные маргиналы по карте Рамачандрана;
- "Ramachandran favored" — число и процент остатков в предпочитаемой области;

Clashscore, all atoms:	9.69	98 th percentile* (N=143, 2.77Å ± 0.25Å)	
Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Poor rotamers	22	3.30%	Goal: <0.3%
Favored rotamers	595	89.34%	Goal: >98%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	758	97.06%	Goal: >98%

Рисунок 3. Показатели качества модели структуры 4ZXS, полученные с помощью сервиса MolProbity [3].

Согласно сервису, маргиналы по карте Рамачандрана отсутствуют, в предпочитаемой области находится 97.06%, что очень близко к идеалу, к 98%. С ротамерами ситуация хуже – их 3.3%, в наилучшей ситуации их число не должно превышать 0.3%, а предпочитаемых ротамеров на 10% меньше от идеальной ситуации в 98%.

Вернёмся к маргиналам по карте Рамачандрана. Сервис EDS [4] выявил 9 маргиналов (Рис. 4) в отличие от MolProbity, который не выявил ни одного. Видимо некоторые остатки, находящиеся в области favoured по версии MolProbity, считаются маргиналами по версии EDS.



X = PHI mapped to [-180,180>

Y = PSI mapped to [-180,180>

PDB file : /pub/db/eds/sfd/4zxs/pdb4zxs.ent

Glycines (open squares): 67 ; Start/end residues : 22

D-amino acids : 0 ; Residues with missing atoms : 0

Residues in Ramachandran plot checked : 714 out of 803

In core regions (plus signs): 705 ; Outliers (asterisks): 9

Percentage outliers: 1.3

An average ≤ 2.0 Å model has ~0.5% outliers

See: Kleywegt, G.J. and Jones, T.A. (1996). Structure 4, 1395-1400.

Рисунок 4. Карта Рамачандрана, полученная при помощи сервиса EDS [4].

14 Маргинальных остатков

В таблице 2 представлены маргинальные остатки и критерий, по которому они были отобраны. Под перекрытием понимается перекрытие Ван-дер-Ваальсовых радиусов. Маргиналами по RSRZ считаются остатки с $RSRZ > 2$. Всего выявлено 4 остатка с возможной инверсией цепи, 38 маргиналов по RSRZ, 18 остатков с недопустимыми ротамерами и 9 маргиналов по значению торсионного угла на карте Рамачандрана по версии сервиса EDS.

Таблица 2. Маргинальные остатки модели 4ZXS, отобранные по различным критериям.

Остаток	Критерий маргинальности
Chain A, PRO-160	1) Перекрытие 0.69Å HA и A 159 MET HE3 2) Cis-конформация
Chain B, CYS-122	1) Перекрытие 0.65Å HB3 и B 125 CYS HB2 2) Маргинал по углу связи CA-C-N, 4.0 σ
Chain B, ASP-199	1) Перекрытие 0.60Å HB3 и B 201 ARG HH21 2) Недопустимый ротамер
Chain B, ARG-201	1) Перекрытие 0.71Å HB2 и B 218 ILE HB 2) Недопустимый ротамер
Chain B, HIS-246	1) Перекрытие 0.41Å CE1 и B 95 PRO HG3 2) Недопустимый ротамер 3) Подозрение на инверсию боковой цепи
Chain C, PRO-103	1) Перекрытие 0.71Å HD3 и C 101 HIS O
Chain D, VAL-84	1) Перекрытие 0.94Å HG11 и C 172 GLU HG2
Chain D, ASN-105	1) Перекрытие 0.44Å OD1 и D 102 VAL HG22 2) Недопустимый ротамер
Chain D, ASP-199	1) Перекрытие 0.90Å OD1 и D 193 ARG NH1 2) Недопустимый ротамер
Chain D, HIS-246	1) Недопустимый ротамер
Chain A, ALA-174	1) RSRZ = 6.5
Chain B, ALA-133	1) RSRZ = 5.6
Chain A, HIS-104	1) RSRZ = 3.2 2) Недопустимый ротамер
Chain C, ALA-174	1) Перекрытие 0.56Å HB3 и C 173 ASP O 2) Недопустимое расположение на карте Рамачандрана по версии сервиса EDS

Анализ 5 маргинальных остатков

Остаток: Цепь A, ALA-174

Критерий маргинальности: RSRZ = 6.5

RSRZ показывает, насколько хорошо остаток вписывается в электронную плотность. На уровне подрезки 0.5 углерод боковой цепи аланина и кислород остова остаётся всё еще не покрытым электронной плотностью. Остаток расположен на C-конце белка наряду с ASP-173, который тоже является маргиналом по RSRZ (RSRZ = 4.4). Заметим, что GLU-172

уже неплохо вписывается в свою электронную плотность, хотя хочется его немного подвинуть ниже (Рис. 5).

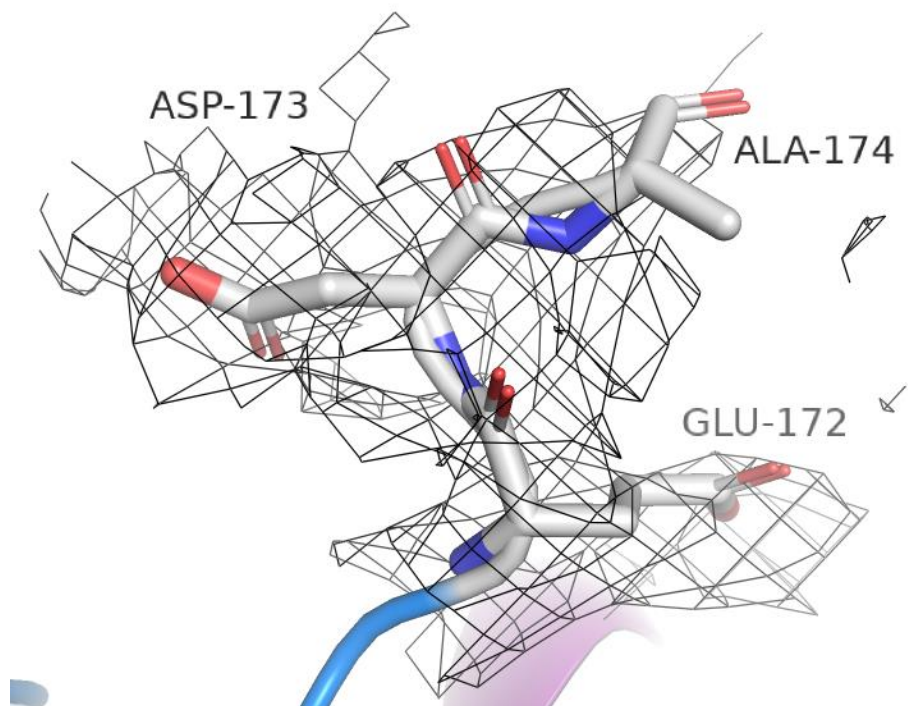


Рисунок 5. Маргинальный ALA-174 в окружении. Уровень подрезки 0.5, carve = 2.5. Концы белка обычно удаётся хуже разрешить из-за высокой термической подвижности, которая приводит к неопределённости в положении. Я бы сказала, что авторы погорячились, определив таким образом координаты этого остатка. На мой взгляд, лучше было бы оставить этот остаток unmodelled или попытаться вписать его как-то по-другому. Таким образом, данный маргинал – ошибка расшифровки.

Остаток: Цепь А, HIS-104

Критерий маргинальности: RSRZ = 3.2 и недопустимый ротамер.

Видно, что электронная плотность кольца гистидина немного съехала (Рис. 6). На уровне подрезки 1 лучше видно, что электронную плотность кольца лучше расположить чуть правее (Рис. 7). Замечу, что электронная плотность остова читается лучше, но следующий ASN-105 хуже вписывается в свою плотность, а 106 остаток вообще не разрешён и отсутствует в этой структуре. HIS-104 расположен в длинной петле, которые, как известно, хуже разрешаются и имеют высокую термическую подвижность в отличие от более структурированных частей белка. Таким образом, данный маргинал – ошибка расшифровки. Отдельно отмечу, что для данного остатка RSCC, показывающий соответствие положения электронной плотности улучшился до довольно хороших (0.9) в модели PDB-redo.

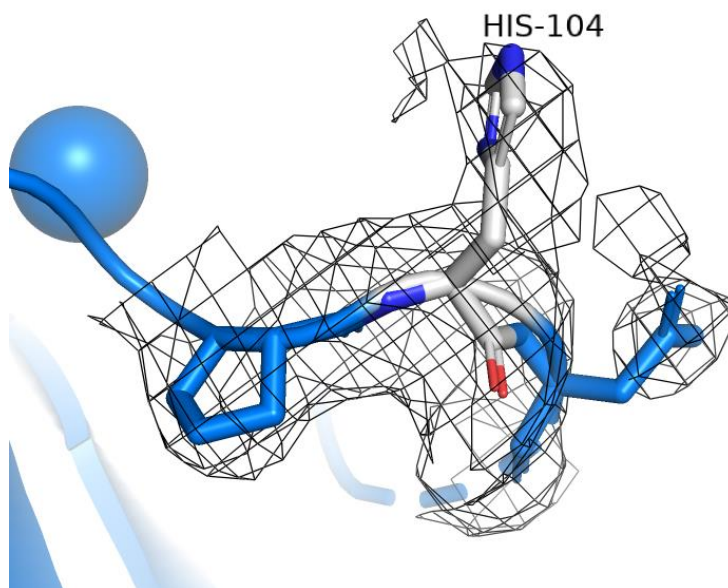


Рисунок 6. Маргинальный HIS-104 в окружении. Соседние остатки показаны в виде sticks: PRO-103, ASN-105. Уровень подрезки 0.5, carve = 2.5.

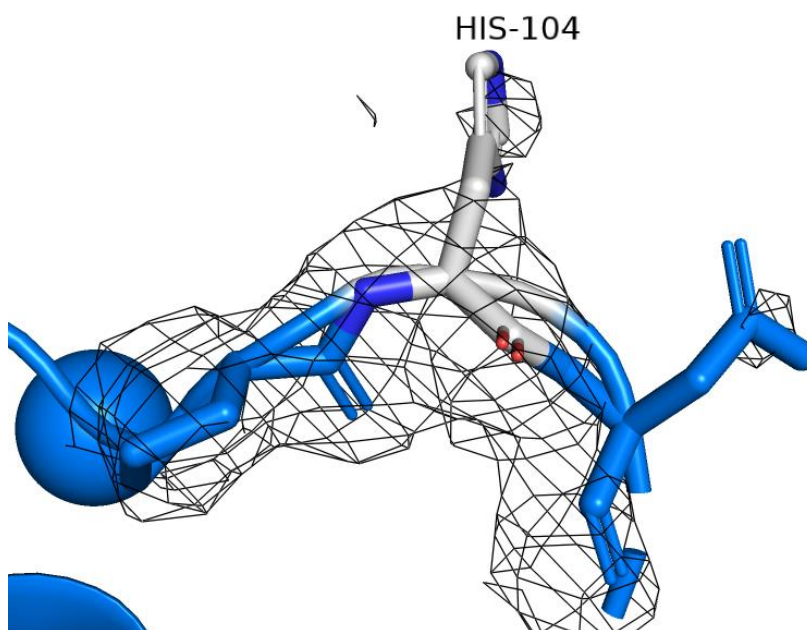


Рисунок 7. Маргинальный HIS-104 покрашен по цветам по умолчанию в окружении. Уровень подрезки 1, carve = 2.5.

Остаток: Цепь В, CYS-122

Критерий маргинальности:

- 1) Перекрывание 0.65\AA HB3 и В 125 CYS HB2
- 2) Маргинал по углу связи CA-C-N, 4.0σ

Перекрывание водородов скорее всего вызвано образованием водородных связей с водой. Угол маргинальный, но в целом всё очень хорошо вписывается в электронную плотность (Рис. 8). Таким образом, считаю этот остаток - особенностью структуры, а не ошибкой расшифровки..

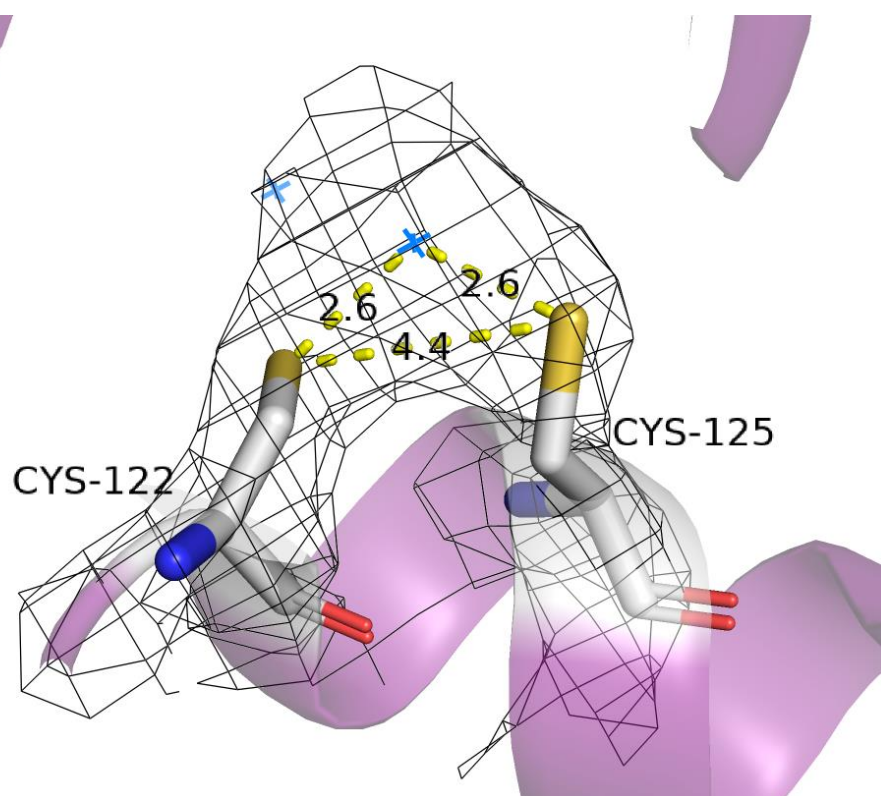


Рисунок 8. Маргинальный CYS-122 в окружении. Вода показана синей «снежинкой». Кроме того, здесь показано расстояние между соседними цистеинами в 4.4 Å. Уровень подрезки 0.5, carve = 2.5.

Остаток: Цепь В, HIS-246

Критерий маргинальности:

- 1) Перекрытие 0.41Å CE1 и В 95 PRO HG3
- 2) Недопустимый ротамер
- 3) Подозрение на инверсию боковой цепи

Положение гистидина часто расшифровывают неверно, т.к. азот и углерод имеют схожее число электронов, поэтому их сложно различить. Определить положение кольца гистидина можно по косвенным признакам – посмотрев на окружение. Сервис MolProbity обозначил этот остаток, как маргинальный, как раз по подозрению в инверсии боковой цепи. В ближайшем окружении этого остатка только вода, минимальное расстояние от гистидина до воды составляет 3.1 Å, причём это расстояние от атома углерода (Рис. 9). 3.1 Å достаточно, чтобы образовать водородную связь, но углерод недостаточно электроотрицательный, чтобы образовывать такие связи. Поэтому здесь скорее всего атом азота и инверсия боковой цепи, причём заметим, что тут довольно оптимальная геометрия водородной связи. Гистидин выступает в качестве донора водородной связи.

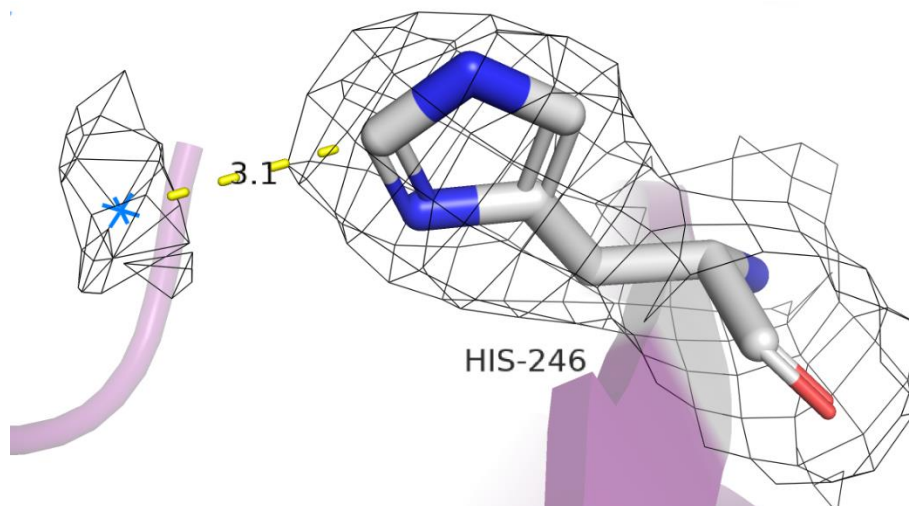


Рисунок 9. Маргинальный HIS-246 в оригинальной структуре, хранящейся в PDB, вода показана «снежинкой». Жёлтым показано минимально возможное расстояние между водой и гистидином. Уровень подрезки 0.5, carve = 2.

Кроме того, у гистидина наблюдается небольшое перекрытие водородов с пролином (Рис. 10), а также недопустимый ротамер HIS-246. Считаю, что этот аминокислотный остаток довольно неплохо вписывается в свою электронную плотность, поэтому ротамер – это особенность данной модели. Что касается перекрытия водородов, то разрешение модели не позволяет отличить отдельные атомы, поэтому перекрытия водородов можно считать ошибкой расшифровки.

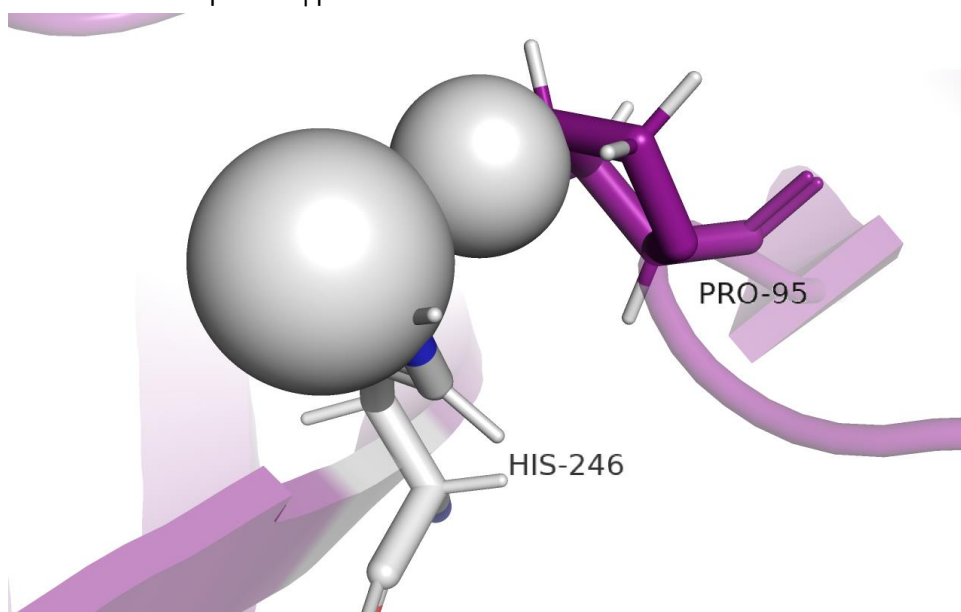


Рисунок 10. Маргинальный HIS-246 перекрывается с водородом PRO-95. Перекрывающиеся атомы показаны в виде sphere.

Остаток: Цепь А, PRO-160

Критерий маргинальности:

- 1) Перекрытие 0.69\AA HA и A 159 MET HE3
- 2) Cis-конформация

На странице практикума написано рассмотреть остатки, находящиеся в цис-конформации. Около 10% пролинов могут находиться в белках в цис-конформации, поэтому в данном

случае это особенность структуры, учитывая, что пролин неплохо вписывается в свою электронную плотность. Далее наблюдаются перекрытия водородов (Рис. 11).

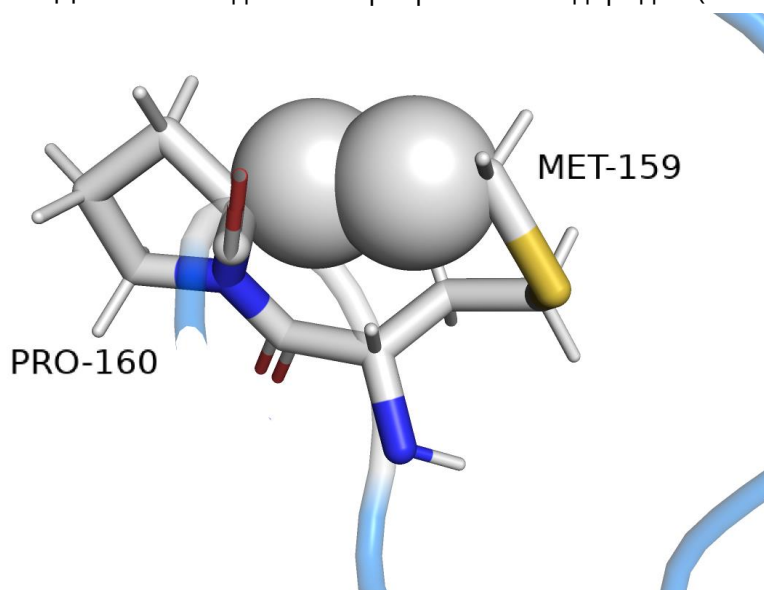


Рисунок 11. Маргинальный PRO-160 и соседний остаток MET-159 показаны в виде sticks. Наблюдаем перекрывание водородов.

Перекрытие скорее всего вызвано тем, что соседний метионин скорее всего располагается как-то по-другому, потому что он не очень хорошо вписывается в свою электронную плотность (Рис. 12). Таким образом, пролин не маргинал, а особенность расшифровки данной структуры.

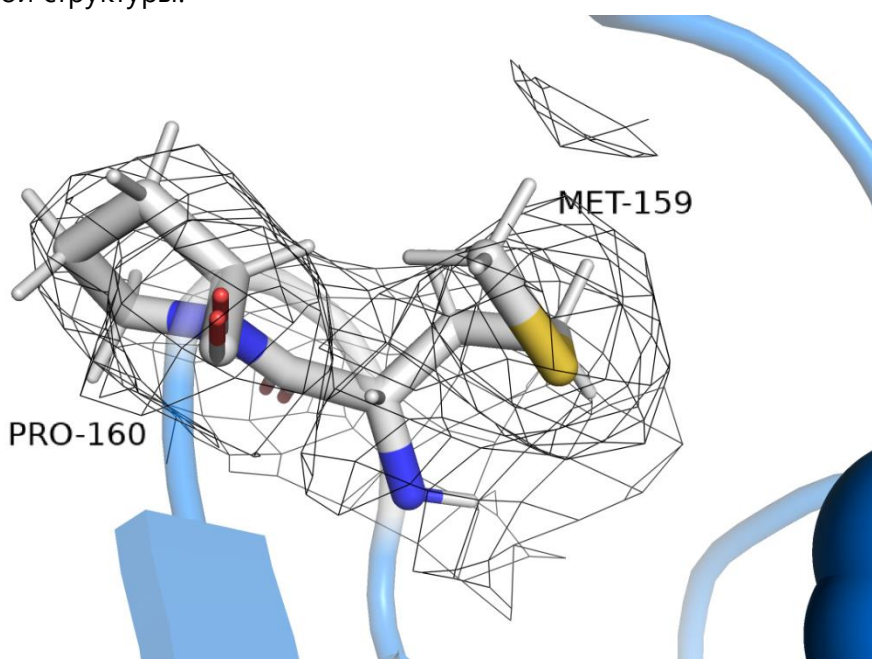


Рисунок 12. Маргинальный PRO-160. Уровень подрезки 0.5, carve = 2.

Сравнение модели из PDB с моделью из PDB_redo

База данных PDB-redo хранит оптимизированные структуры. Для нашей модели произошло улучшение по многим параметрам. R-free для оптимизированной модели составил 17.35%, теперь нашу модель можно считать хорошей, а модель из PDB мы охарактеризовали, как удовлетворительную. Качество ротамеров, хотя и незначительно, но

тоже улучшилось, число остатков, которые лучше вписываются в электронную плотность составило 665. Единственный параметр, по которому произошло ухудшение, это торсионные углы (Рис. 13). Мы получили 9 маргинальных остатков, расположенных в запрещённой зоне карты Рамачандрана.

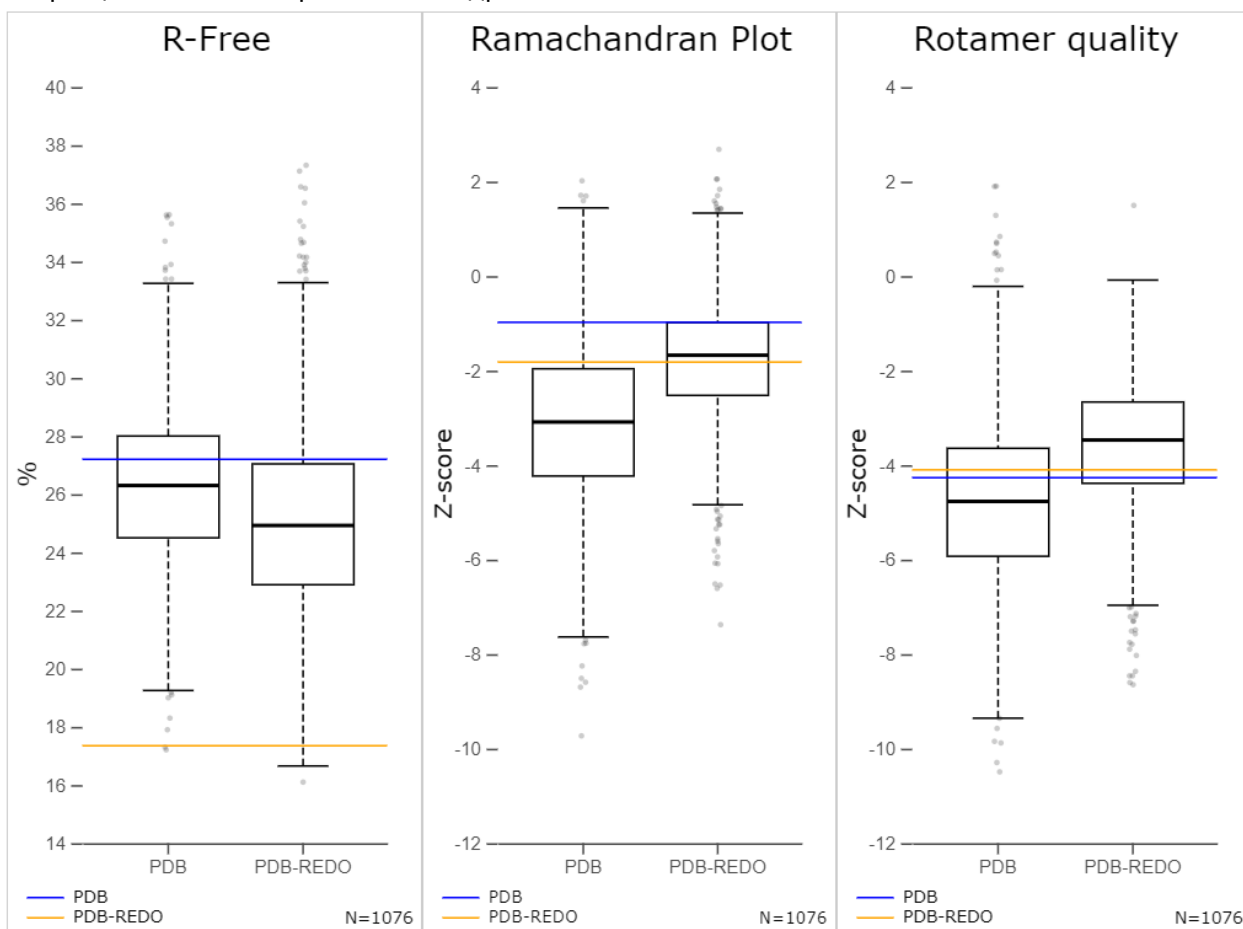


Рисунок 13. Модель PDB, полученная из базы данных PDB-redo.

Заключение

Качество расшифровки данной структуры я считаю удовлетворительным, но для разрешения 2.7 Å это очень хороший результат.

Список литературы

- [1] J. M. Bigalke and E. E. Heldwein, "Structural basis of membrane budding by the nuclear egress complex of herpesviruses," *EMBO J.*, vol. 34, no. 23, pp. 2921–2936, Dec. 2015, doi: 10.15252/emj.201592359.
- [2] Запись PDB: <https://www.rcsb.org/structure/4ZXS>
- [3] Сервис MolProbity: <http://molprobity.biochem.duke.edu/>
- [4] Сервис EDS: <http://www.ebi.ac.uk/pdbe/eds>
- [5] PDB-redo: <https://pdb-redo.eu/db/4zxs>