

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

**Отчет по качеству RSA расшифровки
структуры РНК-шаперона Hfq (3SB2)**

Выполнила студентка

4 курса ФББ МГУ

Антоненкова Юлия

Москва

2021

Оглавление

Введение.....	3
Результаты.....	3
Общая информация о модели.....	3
Значения индикаторов качества модели в целом.....	4
Маргинальные остатки.....	6
Анализ 5 маргинальных остатков.....	7
Сравнение модели из PDB с моделью из PDB_redo.....	10
Заключение.....	11
Список литературы.....	11

Введение

РНК-шаперон Hfq – комплекс, состоящий из 6 одинаковых цепей, организованных в кольцо. Он участвует в репликацию фага Q β и является важным посттрансляционным регулятором нескольких типов РНК, влияя на множество функций в клетке. Этот белок принадлежит бетапротеобактерии *Herbaspirillum seropedicae SmR1* и нарабатывался с помощью *E.coli*. Белок консервативен, каждая его цепь состоит из одной альфа-спирали и пяти бела-тяжей. Судя по расчетам, полученная структура 3SB2 совпадает со структурой комплекса в растворе.

Результаты

Общая информация о модели

Рассматриваемая структура (см. рисунок 1) [4] состоит из 6 цепей, организованных в кольцо, в комплексе с глицерином. Структура была выложена в 2012 году группой Kadowaki, M.A. и др., ими же в 2011 году была опубликована статья [1], описывающая структурные особенности комплекса.

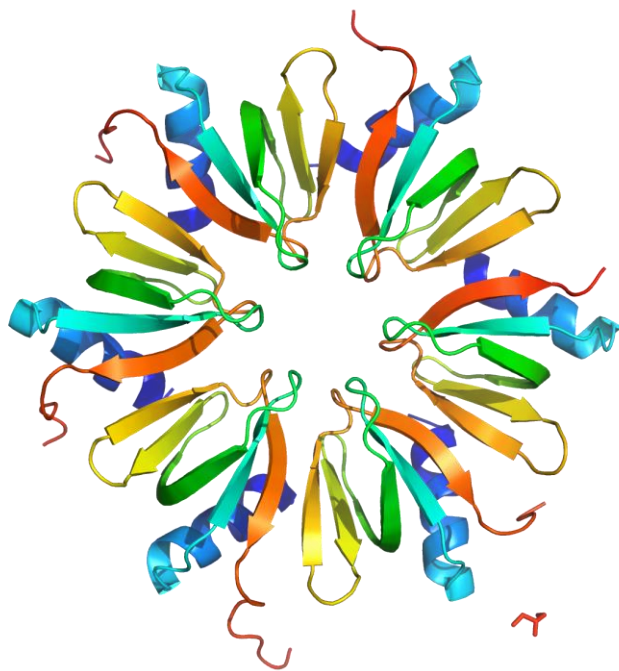


Рисунок 1. Общий вид структуры 3sb2.

Фазовая проблема решалась с помощью метода молекулярного замещения. Авторам удалось измерить 13197 рефлексов из 13664 с разрешением от 2,63 Å до 58,78 Å, полнота

набора структурных факторов составляет 97.3 %, а разрешение всей структуры – 2,63 Å. Параметры кристаллографической ячейки (длины сторон и углы) представлены в таблице. Кристаллографическая группа структуры - P 1 21 1.

Таблица 1. Параметры кристаллографической ячейки.

Длина (Å)	Угол (°)
a = 59.639	$\alpha = 90$
b = 65.323	$\beta = 105.47$
c = 60.987	$\gamma = 90$

Значения индикаторов качества модели в целом

По данным PDB, показатели качества модели следующие:

R_{free} : 0.280

R_{work} : 0.199

$R_{observed}$: 0.203

Показатель R-Value Free довольно велик, это говорит о не очень хорошем качестве модели, а заметная разница между R_{work} и R_{free} – о возможной небольшой переоптимизации модели. Заметим, что из 5-ти параметров, отраженных на рисунке 2, только один (маргинальные по боковым цепям остатки) показывает хорошее качество модели относительно других моделей в PDB.

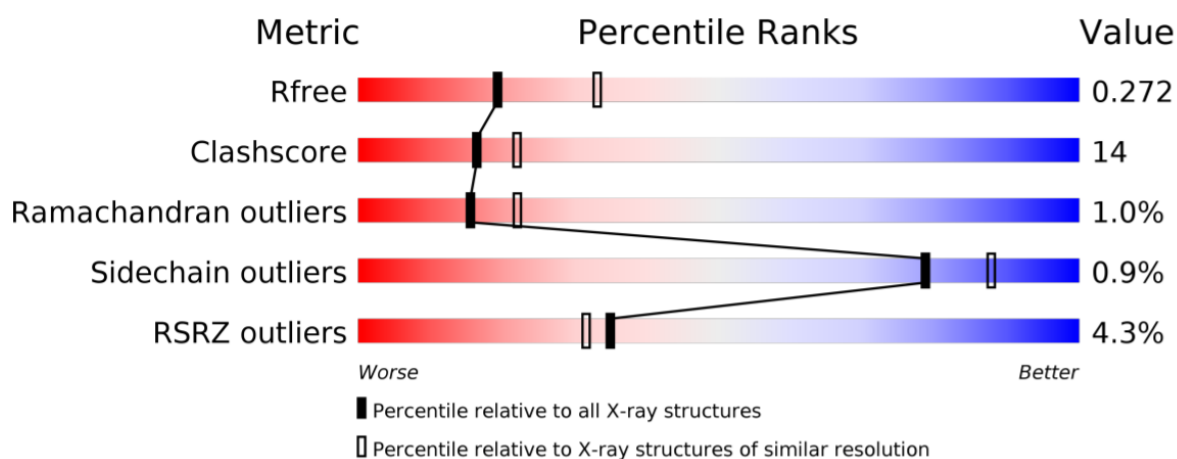


Рисунок 2. Показатели качества структуры по данным PDB.

С помощью сервиса MolProbity [2] были получены дополнительные параметры модели и карты Рамачандрана. Так, в структуре найдено слишком большое количество (2,48%) неротамеров, а количество предпочтительных ротамеров достаточно мало (92,26%). Количество недопустимых наложений атомов (1,5%) среднее: эта модель находится в 51-м перцентиле. В модели не обнаружено связей и углов, значительно отклоняющихся от теоретических данных, а также остатков с большим отклонением атома C β и цис-пролинов. По картам Рамачандрана 89,8% остатков находятся в предпочтительных областях (что довольно мало), 99,0% - в разрешенных областях. Также выявлено 4 маргинальных остатка (1,05%), они отмечены на рисунке 3.

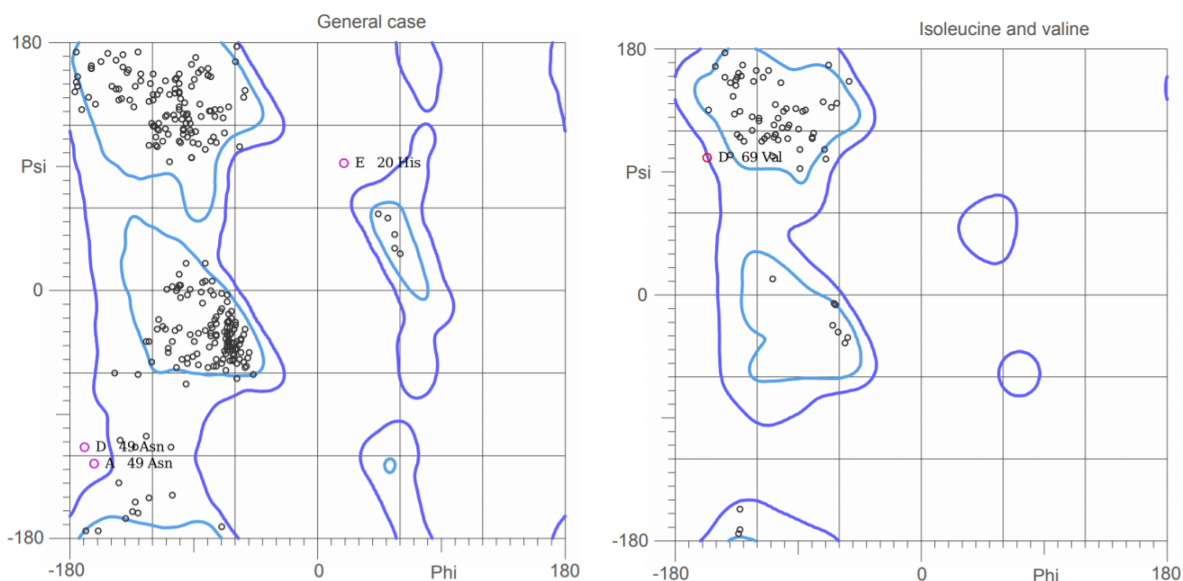


Рисунок 3. Карты Рамачандрана для большинства остатков (слева) и для изолейцина и валина (справа). Кругами обозначены остатки модели. Розовыми кругами обозначены маргинальные остатки: Asn49 из двух цепей, His20, Val69.

Далее рассмотрим параметр RSR (Real-space R-Value). Его можно получить с помощью сервиса EDS [3]. На рисунке 4 приведены распределения RSR в двух цепях. Визуально видно несколько маргинальных остатков, имеющих относительно большое значение RSR.

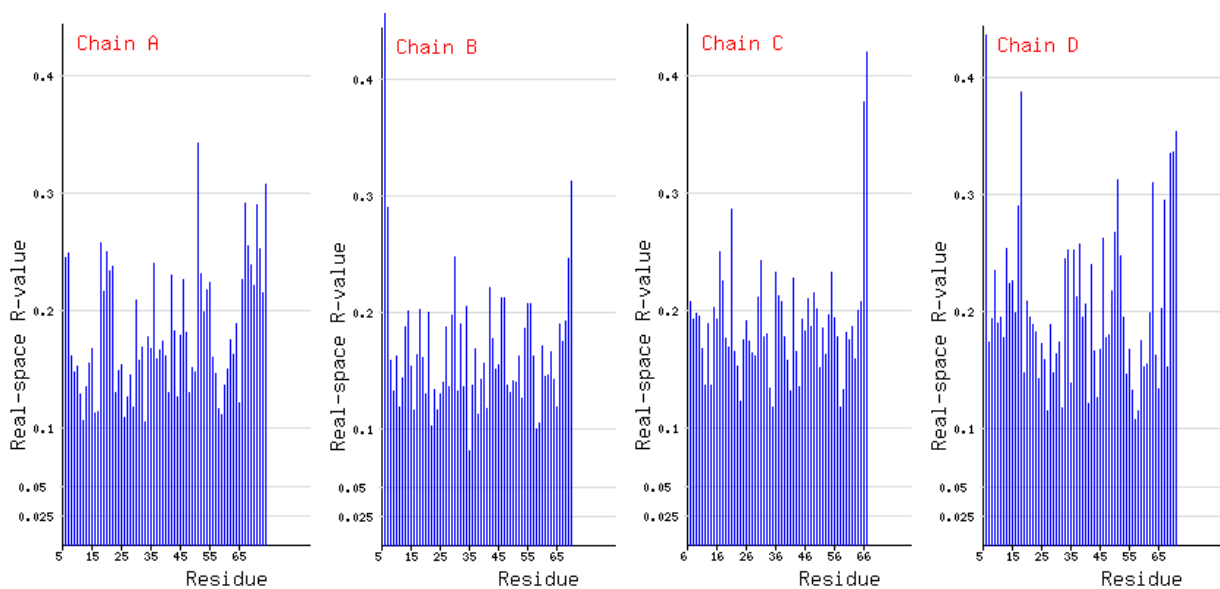


Рисунок 4. Значения параметра RSR для остатков четырех цепей (A, B, C, D). Большие значения соответствуют худшему соответствию электронной плотности модели экспериментальной электронной плотности.

Также 16 остатков отмечены в EDS как значительные. Критерий для этого отбора таков: $RSR > (\langle RSR \rangle + 2.00 \cdot \sigma)$. Больше всего таких остатков в цепи E (7 штук). Самый частый остаток – Gln6, он встречается в 3-х цепях из 6-ти.

На основе RSR подсчитывается $Z\text{-score} = (RSR - \langle RSR \rangle) / \sigma$. Будем считать маргинальными остатки с $Z\text{-score} > 2$. Маргинальные остатки по этому признаку полностью совпадают с маргинальными остатками по предыдущему критерию.

Маргинальные остатки

По параметрам, рассматриваемым выше, были выбраны несколько маргинальных остатков. Данные по этим параметрам для каждого остатка приведены в таблице 2.

Таблица 2. Некоторые маргинальные остатки с указанием параметров, по которым они маргинальны.

Остаток	Параметр(-ы)
Gln6 B D F	Z-score
His20 E	Карта Рамачандрана
Asn49 A D	Карта Рамачандрана
Thr50	Отклонение CA (A-F), Z-score (E), clash (D)
Val51 A	Неротамер
Ala66 C E	Z-score
Arg67	Отклонение CA (A), Z-score (C), clash (F)
Val69 D	Карта Рамачандрана, Z-score

Анализ 5 маргинальных остатков

1. Val69, цепь D

Этот остаток находится на самом конце цепи, в этом месте нет вторичной структуры, этим, вероятно, можно объяснить плохую электронную плотность вокруг остатка. Маргинальность остатка по карте Рамачандрана можно объяснить подвижностью петли и неполнотой данных: с такой электронной плотностью, наверное, сложно было корректно вписать атомы. Авторы не стали придумывать расположение боковой цепи остатка в отсутствие электронной плотности в этом месте, это может объяснить большой Z-score. В этом случае маргинальность остатка – скорее следствие плохих экспериментальных данных в этом месте и (или) их плохой расшифровки.

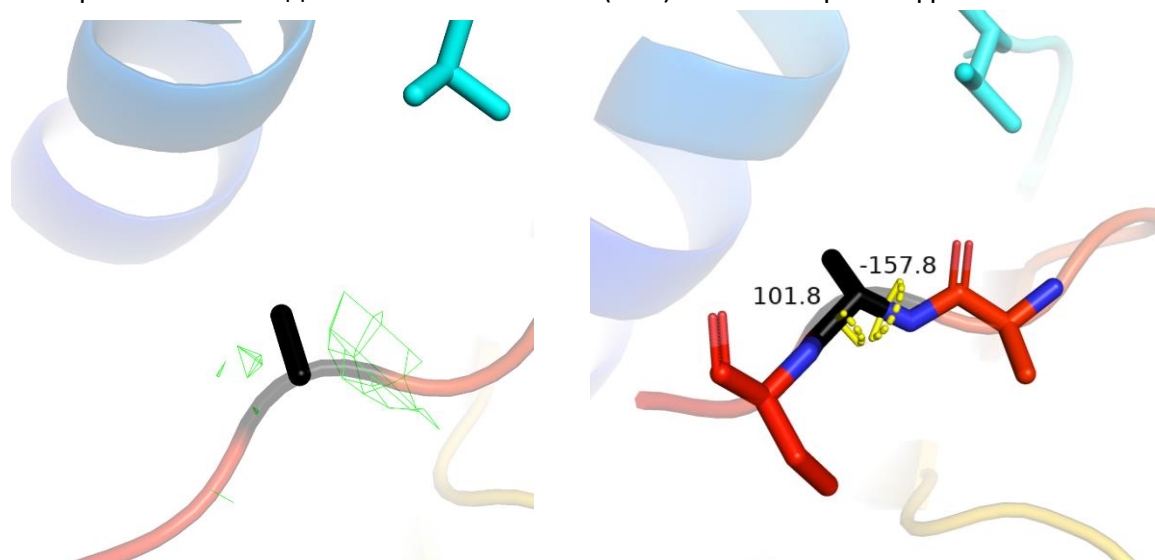


Рисунок 5. Концевой участок цепи D. Валин 69 показан черным. Слева показана электронная плотность на уровне подрезки 1. Справа – торсионные углы остова.

2. Val51, цепь A

Этот остаток расположен в петле между тяжами бета-листа. Здесь тоже нет достаточной электронной плотности, но авторы все-таки решили вписать сюда валин. Скорее всего, этим объясняется то, что эта форма валина не является ротамером. Вероятно, маргинальность этого остатка – следствие плохих экспериментальных данных в этом месте и их плохой расшифровки.

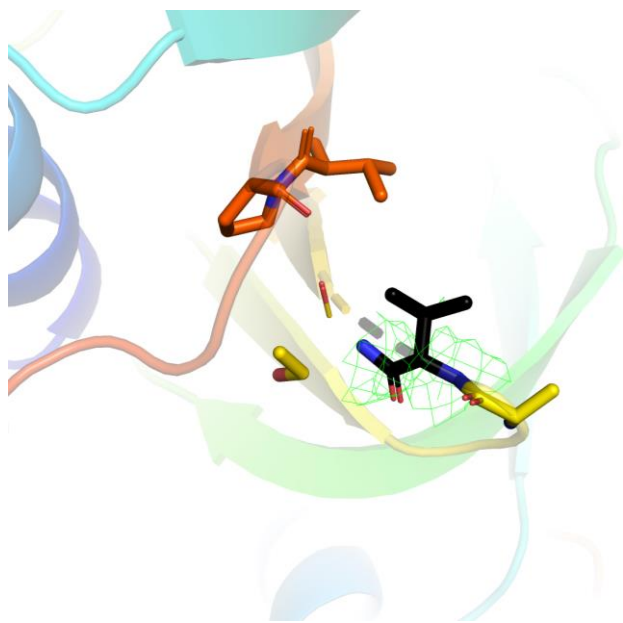


Рисунок 6. Участок цепи A, петля между бета-тяжами. Черный остаток – валин 51. Для него показана электронная плотность на уровне подрезки 1.

3. Gln6, цепь D

По рисунку 7 видно, что, аналогично, нет достаточной электронной плотности, чтобы корректно вписать остаток, поэтому электронная плотность модели не соответствует экспериментальной. Здесь сложности возникли, скорее всего, из-за подвижности остатка вследствие его расположения: это самый первый остаток этой цепи в модели, к тому же расположен он на поверхности комплекса. Можно предположить, что маргинальность остатка – следствие недостаточности экспериментальных данных.

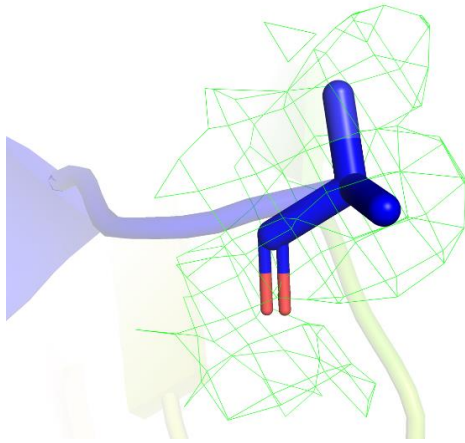


Рисунок 7. Участок цепи D с глутамином 6. Для него показана электронная плотность на уровне подрезки 0.5.

4. Thr50

Из рисунка 8 видно, что электронная плотность не соответствует расположению атомов в боковой цепи остатка. Скорее всего, экспериментальная электронная плотность не соответствует действительности вследствие подвижности участка, поэтому что нет других ротамера или остатка, которые бы подходили значительно лучше. Маргинальность остатка – следствие плохих экспериментальных данных на этом участке.

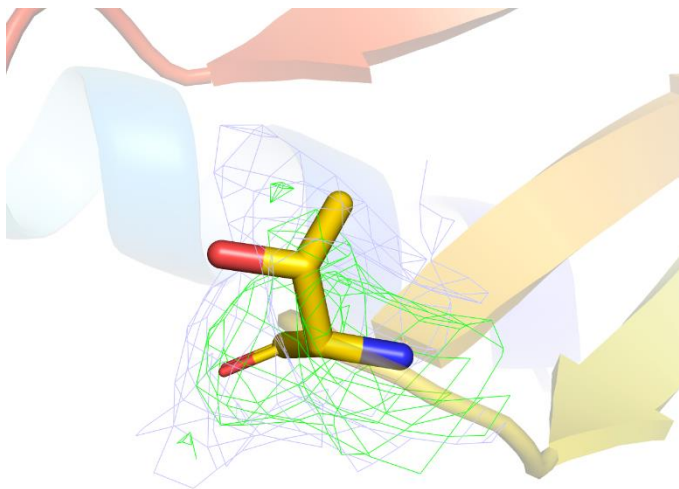


Рисунок 8. Участок цепи D с треонином 50. Электронная плотность показана на уровнях подрезки 0,5 (голубым) и 1 (зеленым).

5. His20, цепь E

Этот остаток расположен на карте Рамачандрана в неразрешенной области. Неестественность торсионных углов остова можно объяснить каким-нибудь взаимодействием в совокупности с гибкостью петли, на которой находится остаток.

Здесь, как видно на рисунке 9, рядом с рассматриваемым гистидином находится еще один гистидин. Не все его атомы есть на модели, поэтому невозможно определить, есть ли там значительное взаимодействие, но можно предположить, что там есть пиводородная связь. В этом случае это может быть особенностью структуры, а не ошибкой расшифровки.

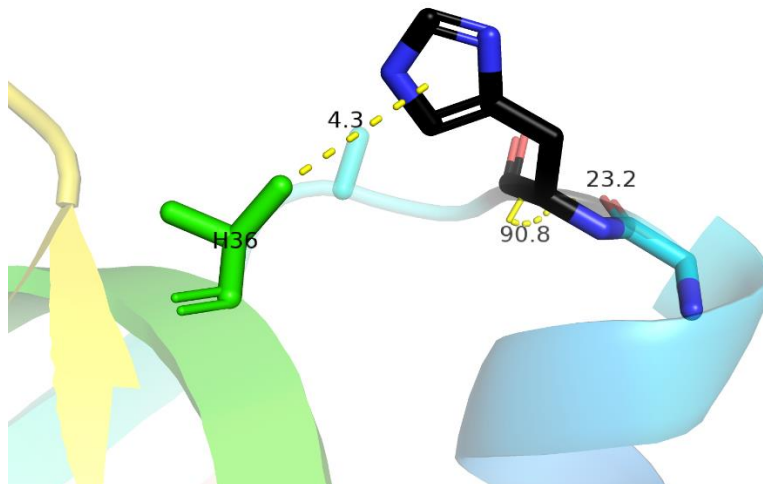


Рисунок 9. Участок цепи E с гистидином 20 (черный).

Сравнение модели из PDB с моделью из PDB_redo

Сервис PDB позволяет оптимизировать модель. В таблице 3 можно увидеть параметры сравнения моделей PDB и PDB_redo.

Таблица 3. Сравнение моделей PDB и PDB_redo.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
<i>R</i>	0,2032	0,2264
<i>R-free</i>	0,2845	0,2742
<i>Bond length RMS Z-score</i>	0,167	0,541
<i>Bond angle RMS Z-score</i>	0,433	0,840
Model quality (raw scores percentiles)		
<i>Ramachandran plot appearance</i>	7	59
<i>Rotamer normality</i>	31	32
<i>Coarse packing</i>	N/A	N/A
<i>Fine packing</i>	33	89
<i>Bump severity</i>	98	72
<i>Hydrogen bond satisfaction</i>	44	54
WHAT_CHECK	Report	Report

Параметры *R* и *R_{free}* почти не изменилась, хотя разница между ними стала меньше, что говорит о меньшей степени переоптимизации. Другие параметры, в основном, улучшились в

модели PDB_redo. Например, соответствие углов остова картам Рамачандрана улучшилось кардинально (рисунок 10) и заполнение водородных связей увеличилось. Единственный параметр, в котором модель PDB_redo уступает модели PDB – перекрывание ван-дер-ваальсовых сфер атомов. Также была достроена 31 боковая цепь остатков, изменен 1 ротамер, удалена 31 молекула воды. Тем не менее, только 5 остатков стали лучше соответствовать электронной плотности, а 23 – хуже.

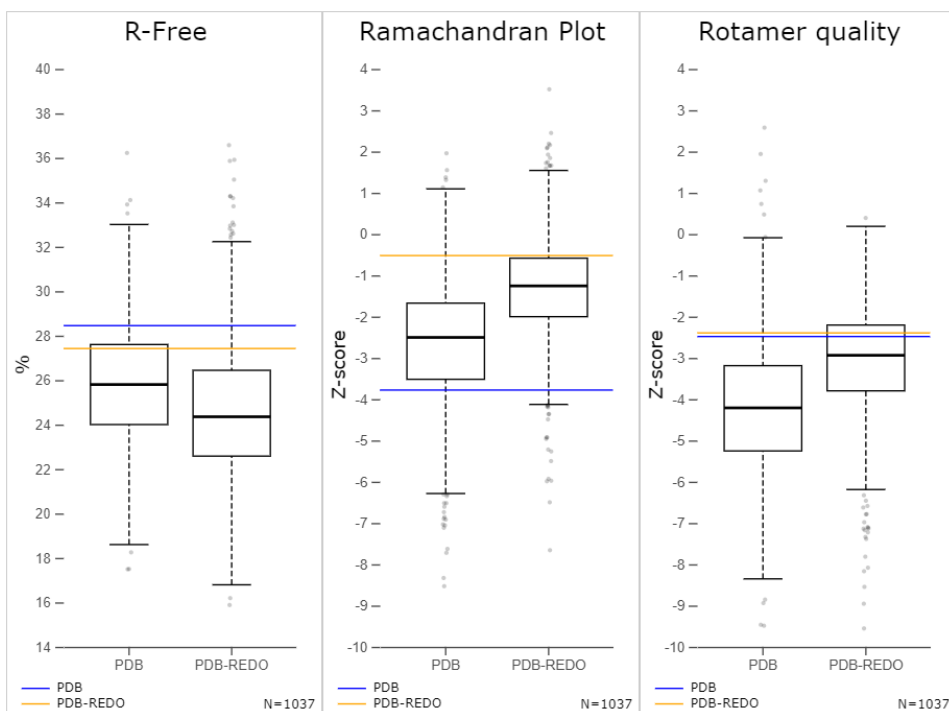


Рисунок 10. Качество модели в сравнении с моделями с похожим разрешением.

Заключение

В целом качество структуры достаточно плохое. Один из пяти рассмотренных маргинальных остатков с трудом может быть интересной особенностью структуры, остальные (как и многие другие маргинальные остатки), скорее всего, возникли из-за не очень хороших экспериментальных данных.

Список литературы

1. Kadowaki, M. A. S. *и др.* Structural characterization of the RNA chaperone Hfq from the nitrogen-fixing bacterium *Herbaspirillum seropedicae* SmR1. *Biochim. Biophys. Acta - Proteins Proteomics* **1824**, 359–365 (2012).
2. <http://molprobit.biochem.duke.edu/>
3. <http://eds.bmc.uu.se/eds/>
4. <https://www.rcsb.org/structure/3SB2>