

**Отчет о качестве структуры ГТФазы RhebL1 (PDB-код  
3OES), полученной с помощью рентгеноструктурного  
анализа (РСА).**

*Отчет студента 4-го курса  
Абдрахманова Алибека*



Москва

2015 г.

# Оглавление

<b>Аннотация</b> .....	3
<b>Введение</b> .....	3
<b>Результаты и обсуждение</b> .....	4
Общая информация о модели .....	4
Значения индикаторов качества модели в целом .....	5
Анализ маргинальных остатков .....	12
Сравнение модели из PDB с моделью из PDB_redo.....	15
<b>Заключение</b> .....	16
<b>Литература</b> .....	16

## Аннотация

В данном отчете приведены некоторые показатели качества структуры Зоес. Показаны различные параметры, такие как: R-free, R, параметры симметрии, температурный фактор для некоторых аминокислотных остатков, карты Рамачандрана, с помощью которых было показано, что некоторые аминокислотные остатки являются маргинальными и ряд других параметров, полученных с разных серверов.

## Введение

Белок, изучаемый в данном отчете – человеческая ГТФаза RhebL1 с PDB-кодом 3OES. Относится к классу гидролаз [3.6.5.2]. Связывает и гидролизует молекулу ГТФ. Представляет собой мономер (цепь А), длиной 201 аминокислотных остатков (рисунок 1). Белок связывает 2 лиганда: ион магния (2+) и ГНФ (гуанилатфосфоаминофосфоновую кислоту). Экспрессионная система, из которой был получен белок – бактерия *Escherichia coli*.

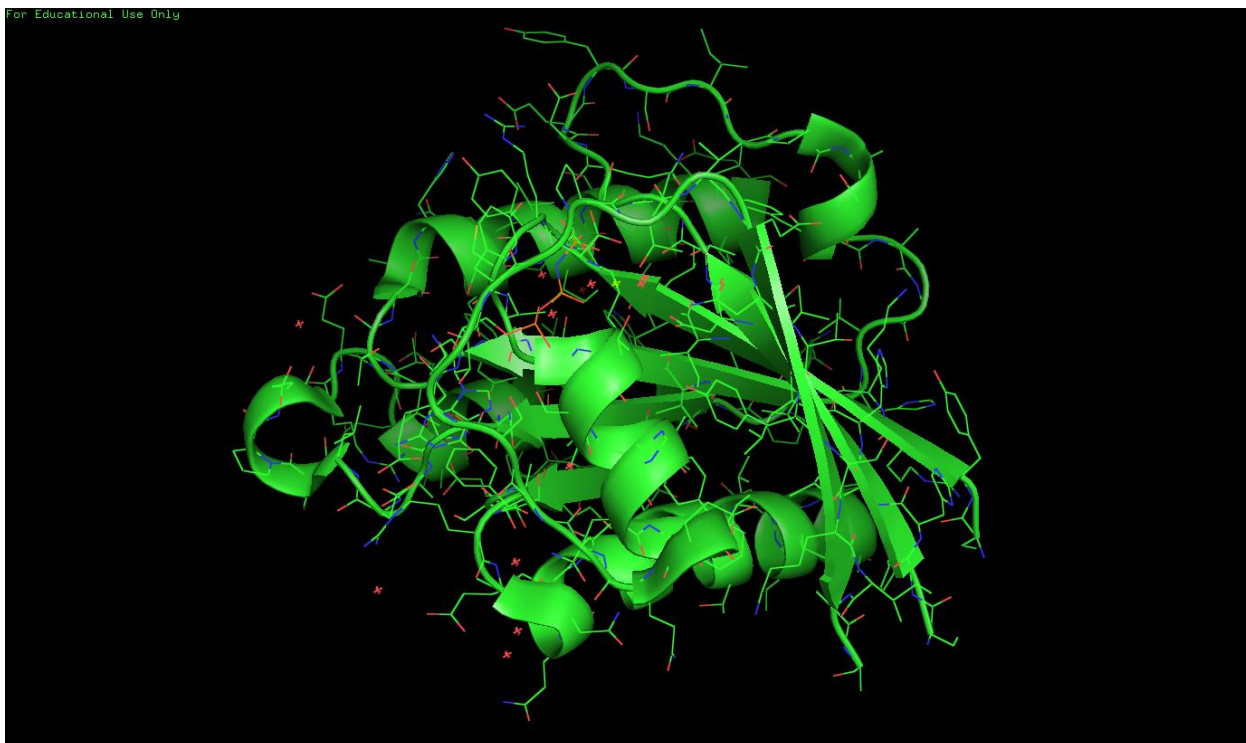


Рисунок 1. Структура ГТФазы RhebL1 (3OES), визуализированная средствами PyMol.

Данная ГТФаза вызвала интерес, поскольку является ярким представителем семейства Ras-ГТФаз, которая является важным компонентом в NF-каппа В – метаболическом пути. Известно, что этот метаболический путь является одним из ключевых в иммунном ответе организма на бактериальное заражение [1].

## Результаты и обсуждение

### Общая информация о модели

Комплекс 3OES расшифрован в 2010 году группой ученых: Nedyalkova, L., Zhong, N., Tempel, W., Tong, Y., Shen, L., Loppnau, P., Arrowsmith, C.H., Edwards, A.M., Bountra, C., Weigelt, J., Vochkarev, A., Park, H., Structural Genomics Consortium. Белок состоит из одной цепи А, длиной 201 а.о. Для решения фазовой проблемы был использован метод молекулярного замещения, в качестве гомологичных структур взяты структуры из семейства RAS-ГТФаз.

На рисунке 2 дан график, с характеристиками структуры с сайта PDB.

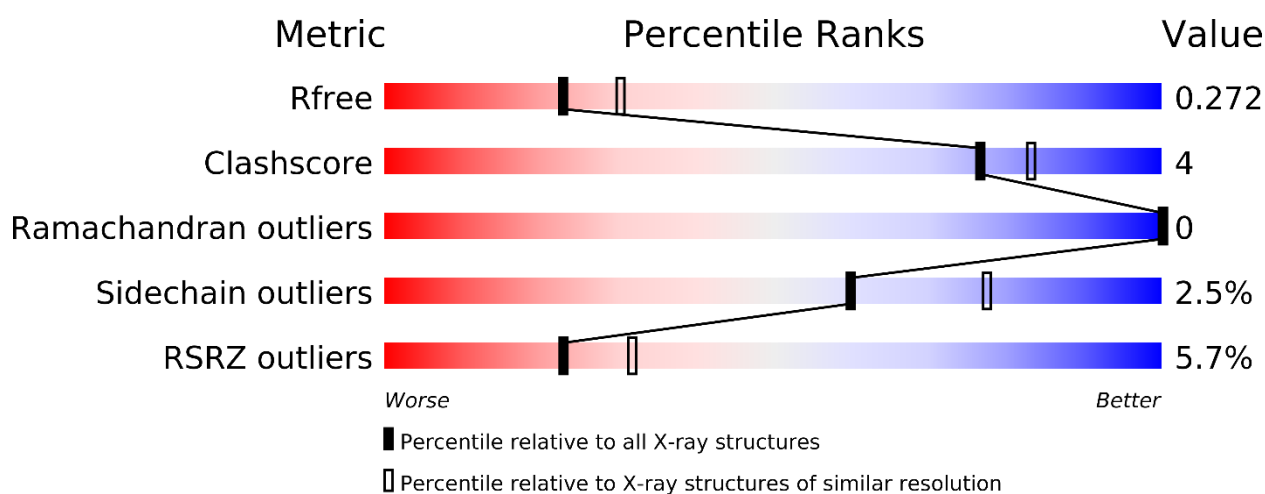


Рисунок 2. График с характеристиками качества структуры 3OES. Ниже описано значение этих характеристик.

Основная информация о структуре 3OES (ссылка на источник приведена не в конце, а здесь, для удобства

[http://ftp.wwpdb.org/pub/pdb/validation\\_reports/oe/3oes/3oes\\_full\\_validation.pdf](http://ftp.wwpdb.org/pub/pdb/validation_reports/oe/3oes/3oes_full_validation.pdf) ):

- Разрешение структуры: 2.3 Å
- Диапазон разрешения структуры: 19.66-2.30 Å
- Атомов в структуре: 1241 (с лигандами), 1195 (без лигандов)
- Всего 8810 структурных факторов (all), из них измеренных 8810 (obs)
- 8391 факторов использовано для оптимизации модели
- 419 факторов использовались в качестве R-free (5%)
- Значение R-фактора: 0.212
- Значение R-free: 0.271
- Полнота набора структурных факторов: 99.6%

- RSR (со стандартным отклонением): 0.154 (0.055)
- Параметры кристаллографической ячейки:

```
464 CRYST1 72.665 126.745 41.356 90.00 90.00 90.00 C 2 2 21 8
```

Первые 3 числа - длины направляющих векторов (72.665, 126.745, 41.356), затем 3 числа - углы между этими векторами, все равны 90°. Затем идет кристаллографическая группа симметрии - C 2 2 21. И количество молекул в ячейке (8).

### Значения индикаторов качества модели в целом

- R-фактор – показывает соответствие между моделью и экспериментальными данными (структурными факторами). Рассчитывается, как:

$$R = \frac{\sum_{hkl} |F_{hkl}^{calc} - F_{hkl}^{obs}|}{\sum_{hkl} F_{hkl}^{obs}} * 100\%$$

Где  $F_{calc}(h,k,l)$  – рассчитанные модули структурных факторов,  $F_{obs}(h,k,l)$  – померенные в эксперименте. В формуле берется модуль, однако на самом деле числитель возводят в квадрат и минимизируют эту величину, это сделано для удобства. Чем меньше значение R-фактора, тем лучше. Условной границей, характеризующей хорошее качество структуры является значение 2.5. Поскольку значение R-фактора меньше 2.5, можно говорить о хорошем качестве расшифровки модели.

- R-free – рассчитывается по аналогичной формуле, однако не для всех структурных факторов, а для 5-10%, в моем случае 5%. Эти 5% являются контрольными, в идеале значения R-фактора и R-free должны быть близки.

Посчитана разница между значениями R-фактора и R-free для проверки того, что не была допущена переоптимизация модели, разница составляет 0.059 (или 5.9%). Можно заметить, что разница достаточно большая, что возможно говорит о переоптимизации модели. На графике с рисунка 2 значение рулетки для R-free находится в красной области.

- Clashscore отображает число неблагоприятных стерических перекрываний атомов на 1000, оно равно 4.
- Следующий важный параметр для оценки качества структуры - пространственный R-фактор, который позволяет провести соответствие

между моделью групп атомов и экспериментальной электронной плотностью. Рассчитывается по следующей формуле:

$$RSR = \frac{\sum_{A \in L} |\rho_{\text{ЭКСП}} - \rho_{\text{МОДЕЛЬ}}|}{\sum_{A \in L} \rho_{\text{ЭКСП}}} [ \cdot 100\% ]$$

На сервере **EDS** есть графики, показывающие этот параметр для структуры Зоес, причем приведены его значения отдельно для полипептидной цепи, отдельно для молекул воды и лигандов. На рисунке 3 продемонстрирована карта значений RSR для цепи А, состоящей из 157 а.о. Из неё видно, что есть остатки с  $RSR > 0.3$  (30%), что не очень хорошо, к таким остаткам относятся:

- ✓ Tyr 67 (0.324)
- ✓ Arg 111 (0.352)
- ✓ Arg 169 (0.407)

Следует отметить, что большинство остатков превышает планку в 10%, которая является показателем хорошей расшифровки структуры.

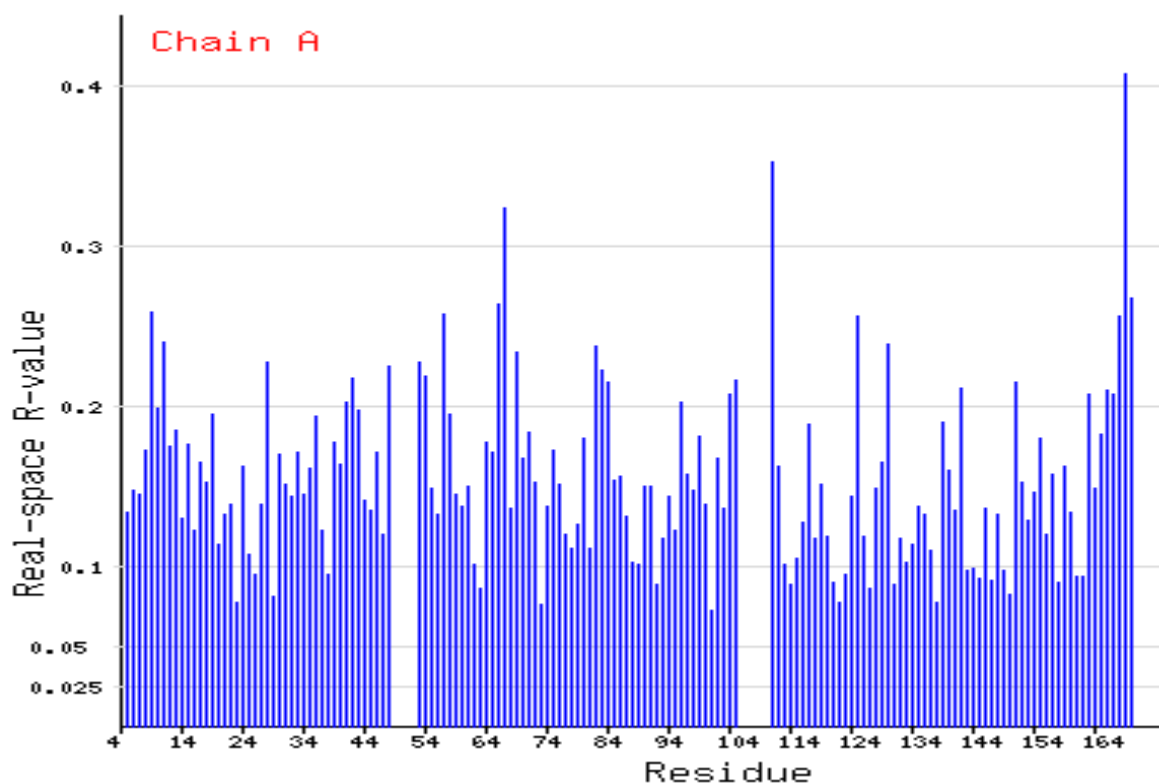
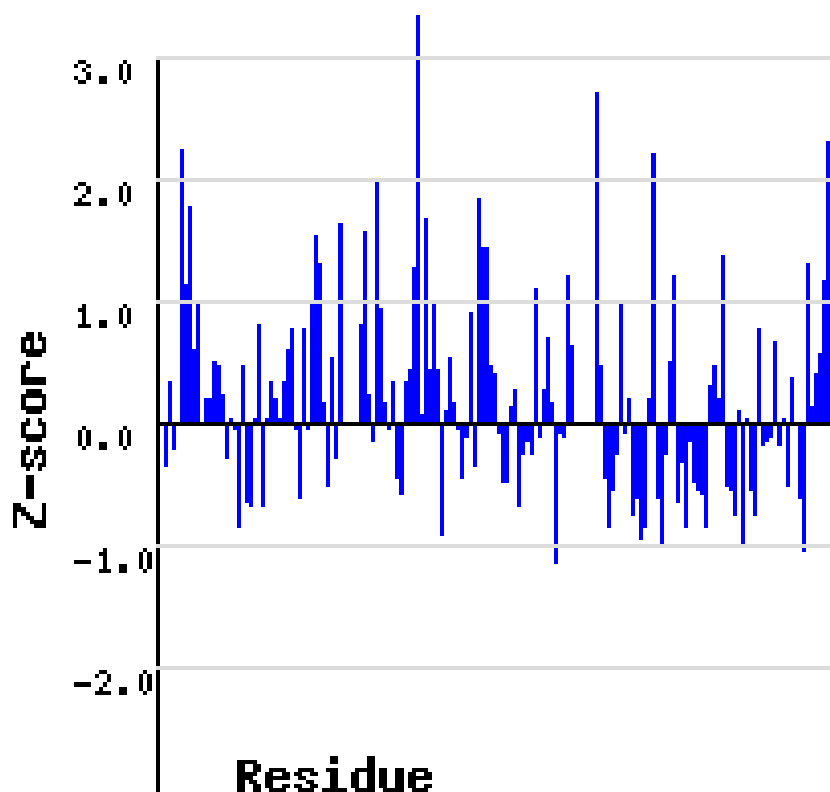


Рисунок 3. RSR для 157 а.о. цепи А структуры ЗОЕС

- Еще одним параметром оценки качества является Z-Score (RSR-Z).  
Рассчитывается, как:

$$Z\text{-score} = (\text{RSR} - \langle \text{RSR} \rangle) / \sigma$$

Это параметр, позволяющий сравнивать RSR остатка модели со средним RSR в выборке моделей PDB с таким же разрешением. Если Z-score какого-либо остатка > 2, остаток вписывается в электронную плотность значительно хуже, чем остаток в структуре с таким же разрешением. На рисунке 4 представлен график со значениями Z-Score для цепи A, состоящей из 157 а.о.



*Рисунок 4. Z-Score для 157 а.о. цепи A структуры 3OES*

Именно из этих предпосылок, утверждается, что а.о. с Z-score > 2, является маргинальным. Таких остатков 7 (видно из рисунка 4), а именно:

- ✓ Val 9
- ✓ Tyr 67 (>3)
- ✓ Arg 111
- ✓ Pro 125
- ✓ Ala 168
- ✓ Arg 169 (>3)

✓ Val 170

Следует заметить, что все остатки, выбивающиеся по RSR, также выбиваются и по Z-score (Tyr 67, Arg 111 и Arg 169), что, наверняка, говорит о их маргинальности. Для подтверждения этого высказывания построил карту Рамачандрана для белка RhebL1 с помощью сервера **MolProbity**.

- Еще один параметр, по которому проверялось качество структуры ЗОЕС – температурный фактор (B). Он показывает то, насколько далеко могут отклоняться определенные а.о. при тепловых колебаниях. Имеет смысл среднеквадратичного отклонения, поэтому значения этого параметра выше 49 показывают то, что температурные колебания вносят большой вклад. На рисунке 5 показан график.

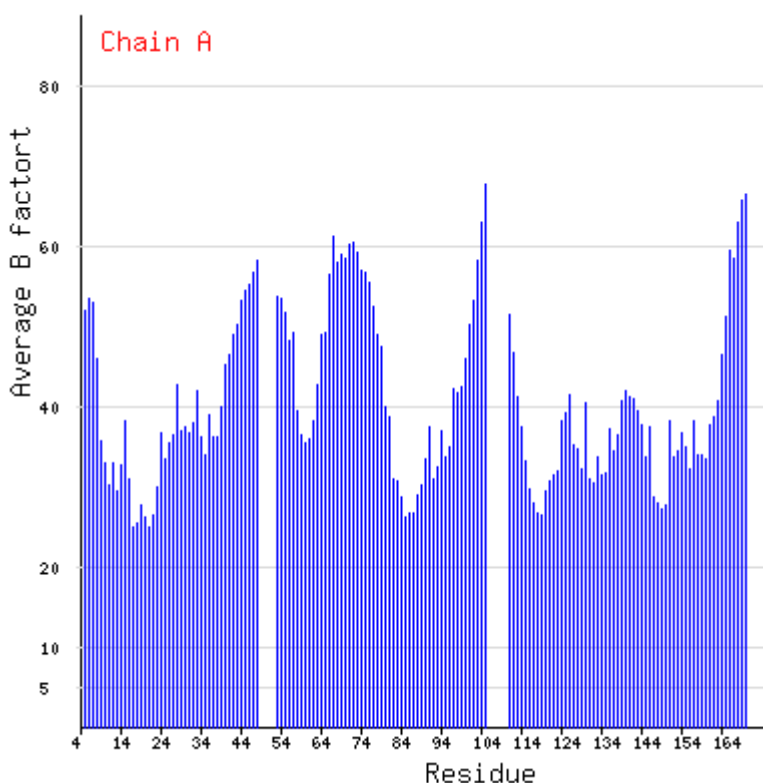


Рисунок 5. Значения температурного фактора - B

Остатки, имеющие завышенный B-фактор:

- ✓ Glu 105 (68)
- ✓ His 104 (63)
- ✓ Ala 168 (63)
- ✓ Arg 169 (65.7)
- ✓ Val 170 (66.5)



✓ Tyr 67 (61.2)

- Наконец, как уже говорилось, с помощью сервера **MolProbity**, была построена карта Рамачандрана (представлена на рисунке 6), а также таблица 1.

All-Atom Contacts	Clashscore, all atoms:	3.8		99 <sup>th</sup> percentile* (N=355, 2.30Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	5	4.13%	Goal: <0.3%
	Favored rotamers	111	91.74%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	148	98.01%	Goal: >98%
	MolProbity score <sup>^</sup>	1.64		98 <sup>th</sup> percentile* (N=8909, 2.30Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 1252	0.00%	Goal: 0%
	Bad angles:	0 / 1705	0.00%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 4	0.00%	Expected: ≤1 per chain, or ≤5%

*Таблица 1. Данные, полученные сервером **MolProbity**. Красным помечены параметры, которые не удовлетворяют условию хорошей модели.*

Что означают значения в полях таблицы (о которых не было ранее сказано):

- 1) "Poor rotamers" - число остатков с маргинальными по отклонению от ротамеров боковыми цепями. 4.13%, когда в хорошей модели должно быть до 0.3%. Видно, что таких остатков 5, а именно:
  - ✓ Arg 15
  - ✓ Ile 69
  - ✓ Ser 73
  - ✓ Leu 103
  - ✓ Lys 136
- 2) "Favored rotamers" – число и процент ротамеров в предпочитаемой области - 91.74%, когда в хорошей модели должно быть > 98%.
- 3) MolProbity score - интегральная оценка структуры по данным этого сервиса (складывается из нескольких параметров)
- 4) C-beta deviations > 0.25Å - число C-beta с неприемлемым отклонением от ожидаемого положения, видно, что по этому критерию всё хорошо.
- 5) Bad backbone bonds - число ковалентных связей, существенно отклоняющихся от теории (типично - 0)
- 6) Bad backbone angles - число валентных углов, существенно отклоняющихся от теории (типично - 0)
- 7) "Ramachandran outliers" - полные маргиналы по карте Рамачандрана, лежат вне допустимой области; такие встречаются в 0.05% остатков при высоком разрешении. В данной структуре такие а.о. не наблюдаются.

- 8) "Ramachandran favored" - число и процент остатков в предпочитаемой области; в идеале таких должно быть > 98%. Видно, что для данной структуры это значение 98.01%.

# MolProbity Ramachandran analysis

3oesFH\_reg.pdb, model 1

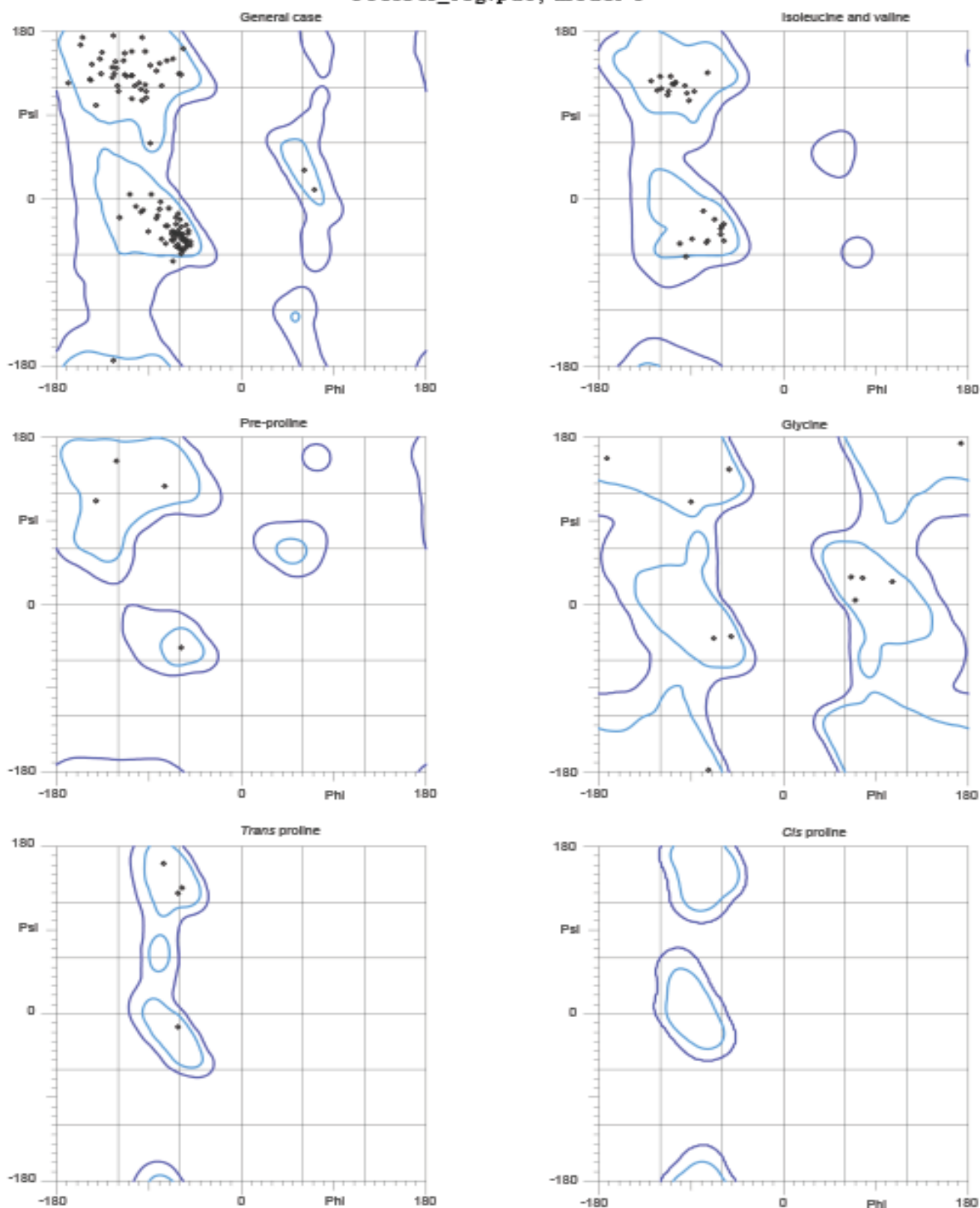
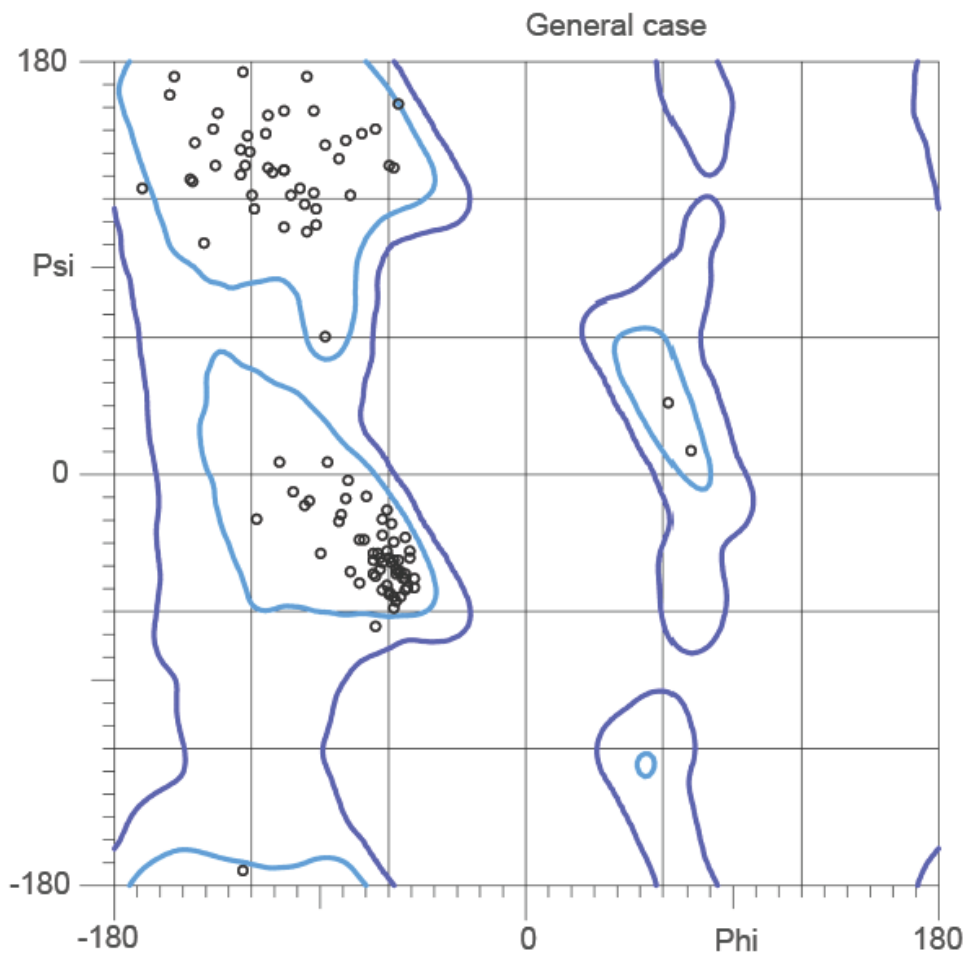


Рисунок 6. Карта Рамачандрана, полученная сервером **MolProbity**. На карте *General case* изображены все остатки. Она отдельно, более наглядно, показана на рисунке 6.



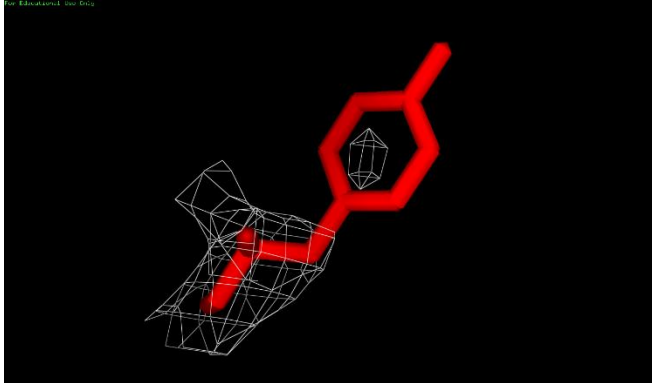
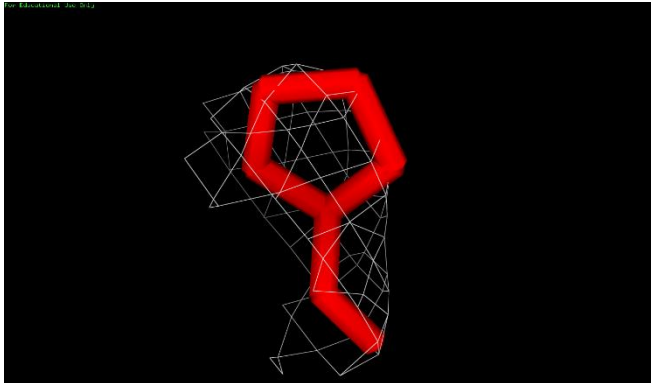
*Рисунок 7. Та же карта, но более хорошего качества для всех остатков.*

Из рисунков 6 и 7 видно, что только 2 аминокислотных остатка находятся за голубой линией, а именно: Arg 169 (на рисунке 6, под треугольником, координаты -66.42 и -67.35) и Glu 40 (слева вверху, координаты -168.94 и 125.53). Еще один остаток находится на голубой линии, это – Tyr 14 (координаты -56.71 и 162.1). По углам пси и фи, соответственно. Все остальные остатки находятся в пределах голубой линии.

Такая же работа была проделана для карт других остатков, но все они лежали внутри разрешенной области (внутри голубой линии).

#### Анализ маргинальных остатков

В результате проделанной работы появились некоторые кандидаты на роль маргинальных остатков, ниже приведена таблица 2, в которую внесены эти остатки, а также критерии по которым они отнесены в эту группу. Приведены 14 таких остатков. Для 5 из них (которые маргинальны сразу по нескольким критериям), ниже построена карта электронной плотности, а также некие комментарии.

N	Остаток	Почему маргинал?	
1	Val 9	1) Z-Score = 2.24 2) RSR = 0.259	
2	Arg 15	1) Poor rotamer (отклоняется угол: chi angles: 28,171.1,212.8,293)	
3	Tyr 67	1) Z-Score = 3.67 2) RSR = 0.324 3) B-factor = 61.2	
4	Ile 69	1) Poor rotamer (отклоняется угол: chi angles: 279.7,124.4) 2) RSR = 0,234	
5	Ser 73	1) Poor rotamer (отклоняется угол: chi angles: 25.2)	
6	Leu 103	1) Poor rotamer (отклоняется угол: chi angles: 228.7,64.3)	
7	His 104	1) B-factor = 63	
8	Glu 105	1) B-factor = 68	
9	Arg 111	1) Z-Score = 2.68 2) RSR = 0.352	
10	Pro 125	1) Z-Score = 2.2 2) RSR = 0.256 3) Перекрытие с Ser 124 (0.44Å)	
11	Lys 136	1) Poor rotamer (отклоняется угол: chi angles: 199.5,100.9,277.8,297.2) 2) Перекрытие с Glu 139 (0.44Å)	

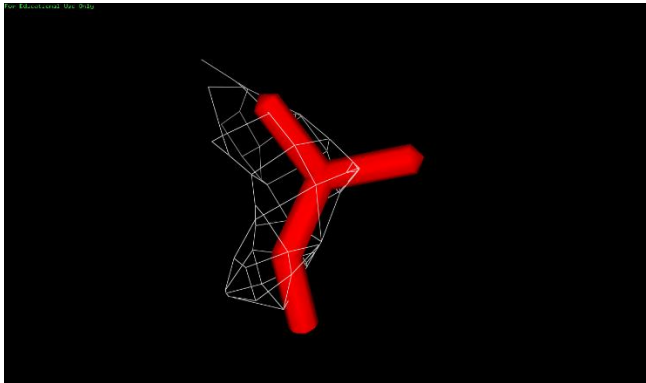
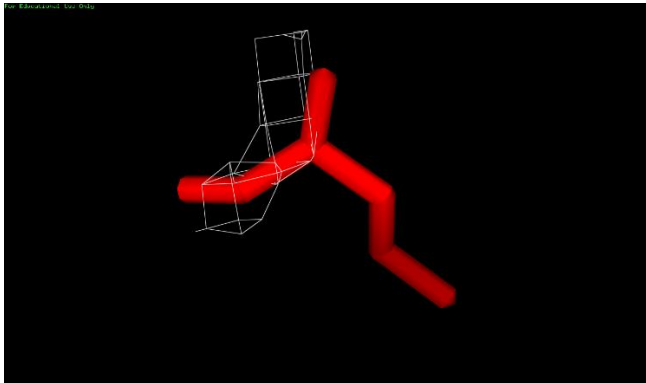
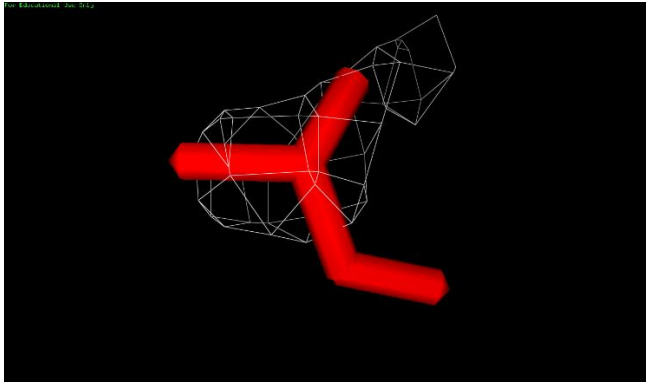
12	Ala 168	1) Z-Score = 2.31 2) RSR = 0.256 3) B-factor = 63	
13	Arg 169	1) Z-Score = 3.45 2) RSR = 0.407 3) B-factor = 65.7	
14	Val 170	1) Z-Score = 2.38 2) RSR = 0.267 3) B-factor = 66.5	

Таблица 2. В ней приведены 14 а.о., которые дали хотя бы 1 маргинальную характеристику. Для 5 а.о. которые маргинальны сразу по 3 признакам (в таблице выделены зеленым) построены карты электронных плотностей с уровнем подрезки = 1.5  $\sigma$  и carve = 1.5 Å.

Из таблицы 2 видно, что 4 маргинальных остатков плохо вписываются в карту электронной плотности с уровнем подрезки = 1.5  $\sigma$  и carve = 1.5 Å, это: Tyr 67, Ala 168, Arg

169 и Val 170. У этих а.о. наблюдаются повышенные параметры для температурного фактора, RSR и Z-Score. Поэтому их маргинальность является, скорее ошибкой расшифровки структуры.

Иначе дело обстоит с Pro 125, он достаточно хорошо вписался в карту электронной плотности, и имеет не настолько критические значения RSR и Z-Score, скорее всего, маргинальность этого остатка не является ошибкой расшифровки структуры, а является особенностью данного остатка.

### Сравнение модели из PDB с моделью из PDB\_redo

	PDB	PDB_redo
R-фактор	0.212	0.2011
R-free	0.271	0.2398
R-free - R-фактор	0.059	0.0387
Poor rotamers (число)	5	3
Poor rotamers (процент)	4.13	2.48
Favored rotamers (число)	111	114
Favored rotamers (процент)	91.74	94.21
Ramachandran outliers (число)	0	0
Ramachandran outliers (процент)	0	0
Ramachandran favored (число)	148	149
Ramachandran favored (процент)	98.01	98.68
Bad bonds (число)	0	7
Bad angles (число)	0	0

Из таблицы выше видно, что в целом, параметры улучшились. Единственным параметром, который стал хуже в PDB\_redo это – bad\_bonds, раньше их не было, а после оптимизации стало 7. На рисунке 8 показано, как PDB\_redo улучшил коэффициенты корреляции параметра RSR для некоторых остатков, однако, для остатков со слишком большими и выбивающимися значениями, ничего не изменилось.

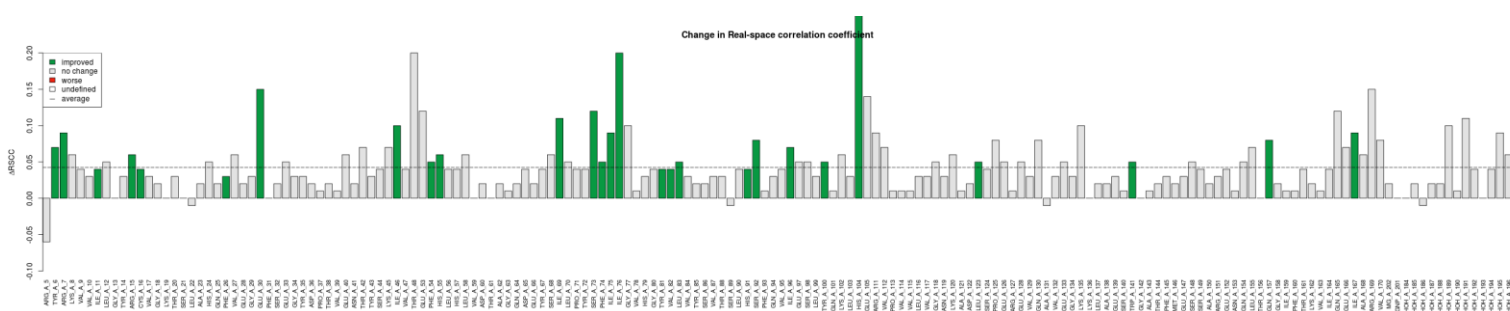


Рисунок 8. Изменение коэффициента корреляции параметра RSR.

[http://www.cmbi.ru.nl/pdb\\_redo/oe/3oes/index.html](http://www.cmbi.ru.nl/pdb_redo/oe/3oes/index.html)

## Заключение

Структура ЗОЕС среднего качества. Об этом свидетельствуют различные индикаторы качества модели. Разрешения структуры недостаточно для того, чтобы определить, где именно находятся многие атомы, это было показано на картах электронной плотности некоторых аминокислотных остатков. Однако, несмотря на это, на карте рамачандрана не было “выбивающихся” остатков, а также остатков, попавших в запрещенные области.

Новая модель, построенная PDB\_redo, несколько улучшила качество стурктуры, это также видно из различных индикаторов, описанных выше.

## Литература

1. J. Yuan, Y. Shan, X. Chen, W. Tang, K. Luo, J. Ni, B. Wan, and L. Yu, “Identification and characterization of RHEBL1, a novel member of Ras family, which activates transcriptional activities of NF-kappa B.,” *Mol. Biol. Rep.*, vol. 32, no. 4, pp. 205–14, Dec. 2005.