

# Отчёт

по качеству расшифровки структуры беіо

Выполнил студент 4 курса Сергей Маргасюк

Преподаватели:

А. В. Алексеевский

С. А. Спирин

В. Ю. Лунин

## Аннотация

В этой работе из нескольких баз данных собраны сведения о структуре беіо и выполнена оценка качества расшифровки структуры в целом и отдельных остатков.

## Введение

беіо – структура лёд-связывающего белка (IBP) EfcIBP из метагенома бактериальных симбионтов антарктической инфузории *Euplotes focardii*. Цель расшифровки – найти структурные особенности, позволяющие этим белкам связывать лёд [1]. В результате расшифровки структуры и докинга показано, что белок отличается от других IBP наличием двух связывающих лёд поверхностей.

## Результаты и обсуждение

### Общая информация о модели

Расшифровку структуры выполнили авторы работы [1] (Nardini, M., Mangiagalli, M., Nardone, V., Bar Dolev, M., Vena, V.F., Sarusi, G., Braslavsky, I., Lotti, M.) в 2017 году.

Структура включает цепь белка и лиганды (сульфат-ион и глицерин) [2]. В эксперименте измерены 183713 рефлексов с разрешениями 0.84Å – 46.23Å, выбрано разрешение структуры 0.84Å (полнота набора рефлексов 94.4%). Пространственная группа кристалла –  $P 2_12_12_1$ , некристаллографические симметрии ячейки отсутствуют. Параметры кристаллографической ячейки: **a**: 45.482 Å; **b**: 50.717 Å; **c**: 92.452; Å **α**: 90° ;**β**: 90°; **γ**: 90° [3].

Фазовая проблема решена методом молекулярного замещения из структуры белка CoIAFP (PDB ID 3WP9) [1].

### Качество модели в целом

Построенная при расшифровке модель хорошо описывает экспериментальные данные:  $R = 0.11$ ,  $R_{free}=0.12$ . Отметим, что при этом сервис EDS [4] не способен

построить модель с R-value, отличающимся менее, чем на 5% от опубликованного. На рисунке 1 приведена карта Рамачандрана из выдачи сервиса MolProbity [5]:

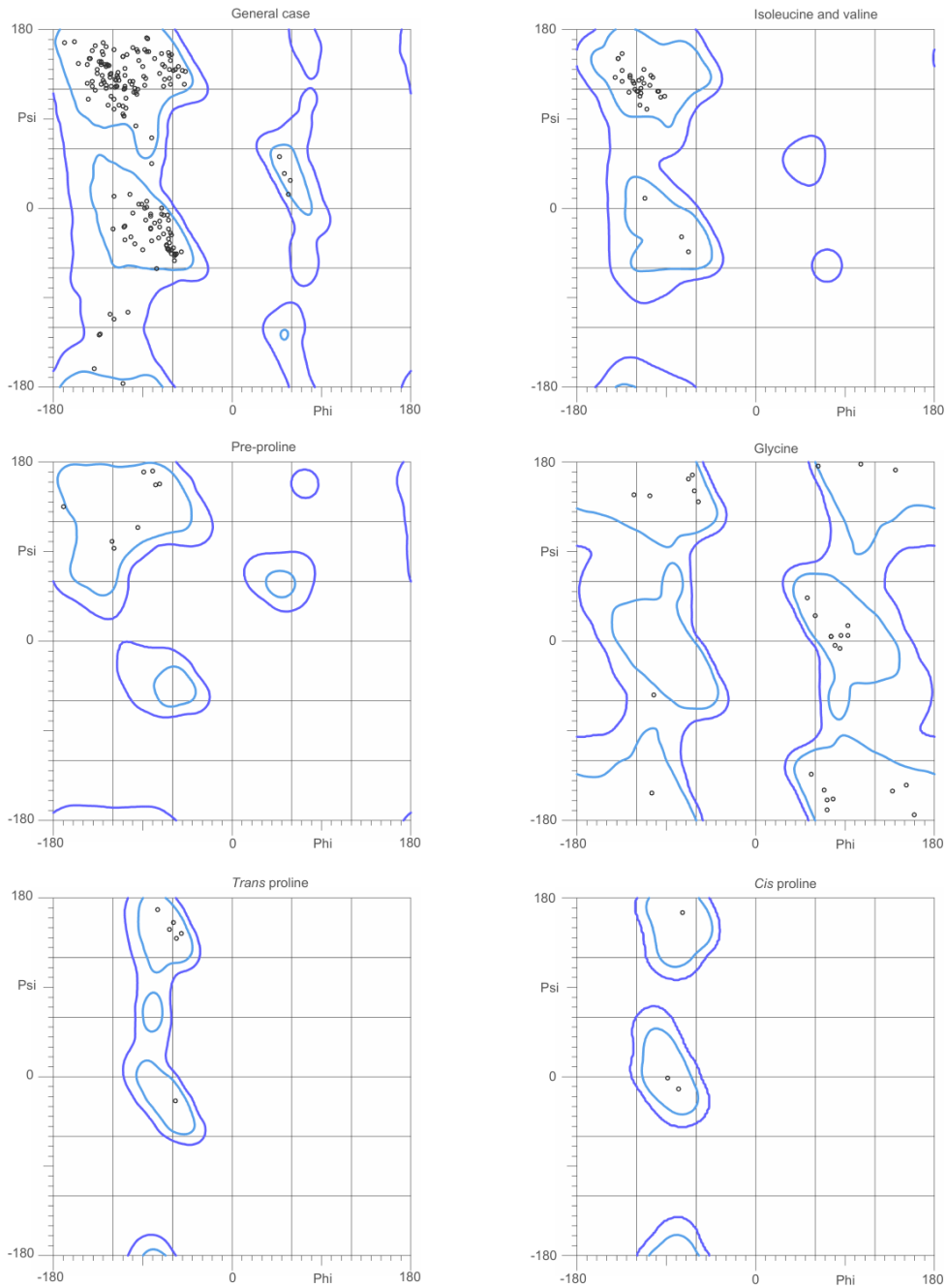


Рисунок 1. Карта Рамачандрана структуры be10. Фиолетовым цветом показаны допустимые области, синим – предпочитаемые (по данным MolProbity)

Все аминокислотные остатки находятся в допустимых областях (т.е. нет маргинальных остатков), 96% – в предпочитаемых. Таким образом, структура несколько отклоняется от обучающей выборки MolProbity (в предпочитаемой области лежит 98% остатков обучающей выборки). Общий показатель MolProbity (соответствие структуры параметрам, полученным из обучающей выборки) – 1.63 (51 перцентиль): вероятно, это обусловлено тем, что белок функционирует в условиях, сильно отличающихся от стандартных (низкая температура).

## Маргинальные остатки

В рассматриваемой модели нет маргинальных остатков по критериям: запрещенная область карты Рамачандрана, неблагоприятная конформация боковой цепи. В таблице 1 представлен список некоторых маргинальных остатков (все остатки с RSRZ > 2 и некоторые остатки, в которых есть атомы, аномально (менее, чем на сумму Ван-дер-Ваальсовых радиусов) сближенные с другими атомами).

Таблица 1. Некоторые маргинальные остатки структуры бею [6]

№	Номер	Причина отнесения к маргинальным
1	HIS259	RSRZ = 3.7 > 2 (высокий RSRZ значит, что ЭП модели для остатка слабо соответствует экспериментальной), пересечение Ван-дер-Ваальсова радиуса атома HD2 с атомом O молекулы HON644 (0.45 Å)
2	THR38	RSRZ = 3.6 > 2
3	HIS258	RSRZ = 3.6 > 2
4	THR208	пересечение Ван-дер-Ваальсовых радиусов атома HG21[B] с атомом O молекулы HON475 (0.64 Å) и атома HG23[B] с атомом O молекулы HON453 (0.68 Å)
5	THR114	пересечение Ван-дер-Ваальсовых радиусов атома HG23 с атомом H32 молекулы GOL301 (0.51 Å), атома HA с атомом H11 молекулы GOL301 (0.45 Å), атома OG1 с атомом H32 молекулы GOL301 (0.41 Å)

Рассмотрим подробнее эти остатки.

Остаток His259 находится на С-конце цепи белка, поэтому относительно подвижен. В кристалле в эксперименте его положение не фиксировано, и пики электронной плотности оказываются в нескольких позициях, и их величина меньше. В модели представлены два варианта положения этого остатка, но они, на мой взгляд, не описывают все конформации, получившиеся в эксперименте (в этом случае пики электронной плотности были бы выше, примерно в два раза ниже пиков прочих атомов, для которых положение определено); таким образом, модель плохо описывает электронную плотность (поэтому у остатка высокий RSRZ) и слабо отражает реальное расположение боковой цепи (поэтому возможно сильное сближение с соседними атомами). Модель и поверхности уровня электронной плотности ( $1\sigma$  и  $2\sigma$ ) изображены на рисунке 2.

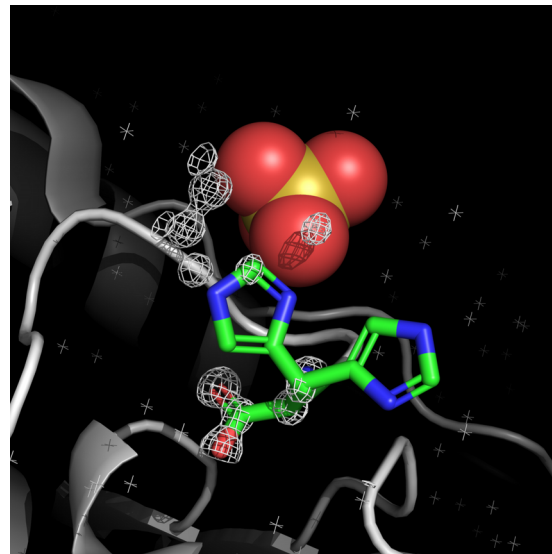
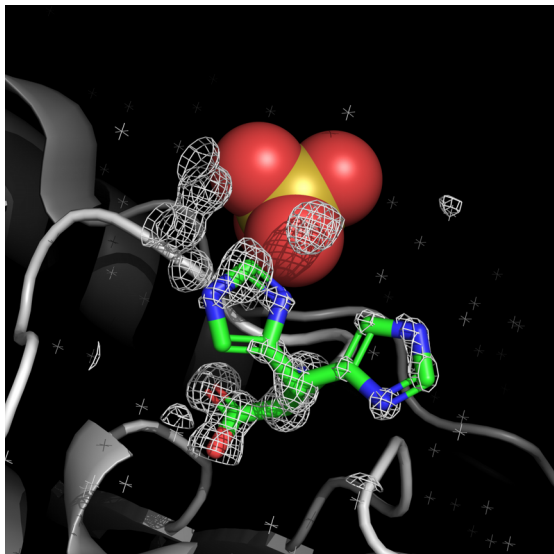


Рисунок 2. Остаток His259 и поверхности уровня электронной плоскости  $1\sigma$  (слева) и  $2\sigma$  (справа)

Остаток Thr38 находится на N-конце цепи, его структура, аналогично предыдущему остатку, восстановлена плохо. Модель здесь не предлагает несколько вариантов положения, но пики электронной плотности для этих атомов слабые (так, их не видно на уровне подрезки  $2\sigma$ , в то время как центры атомов соседнего остатка легко угадываются из поверхностей уровня). Модель и поверхности уровня электронной плотности ( $1\sigma$  и  $2\sigma$ ) изображены на рисунке 3.

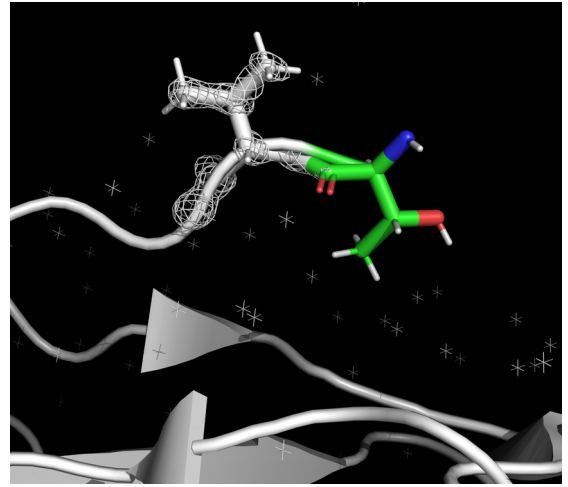
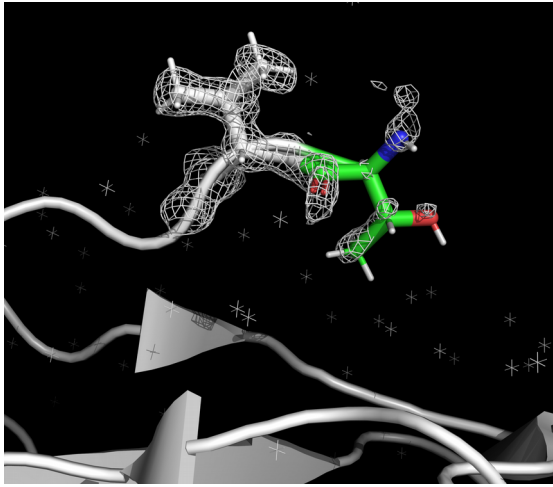


Рисунок 3. Остаток Trp38 и поверхности уровня электронной плоскости  $1\sigma$  (слева) и  $2\sigma$  (справа). Для сравнения изображен также остаток Val39 (белая стержневая модель)

Остаток His259 также расположен в слабо структурированном участке молекулы белка (второй с С-конца), поэтому слабо “покрыт” электронной плотностью. Модель и поверхность уровня электронной плотности  $2\sigma$  изображена на рисунке 4.

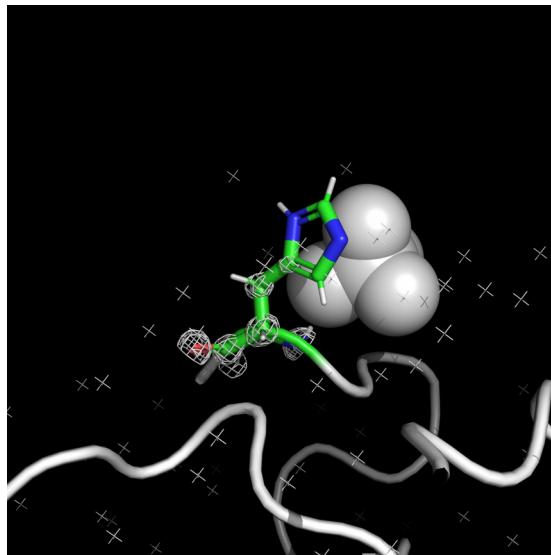
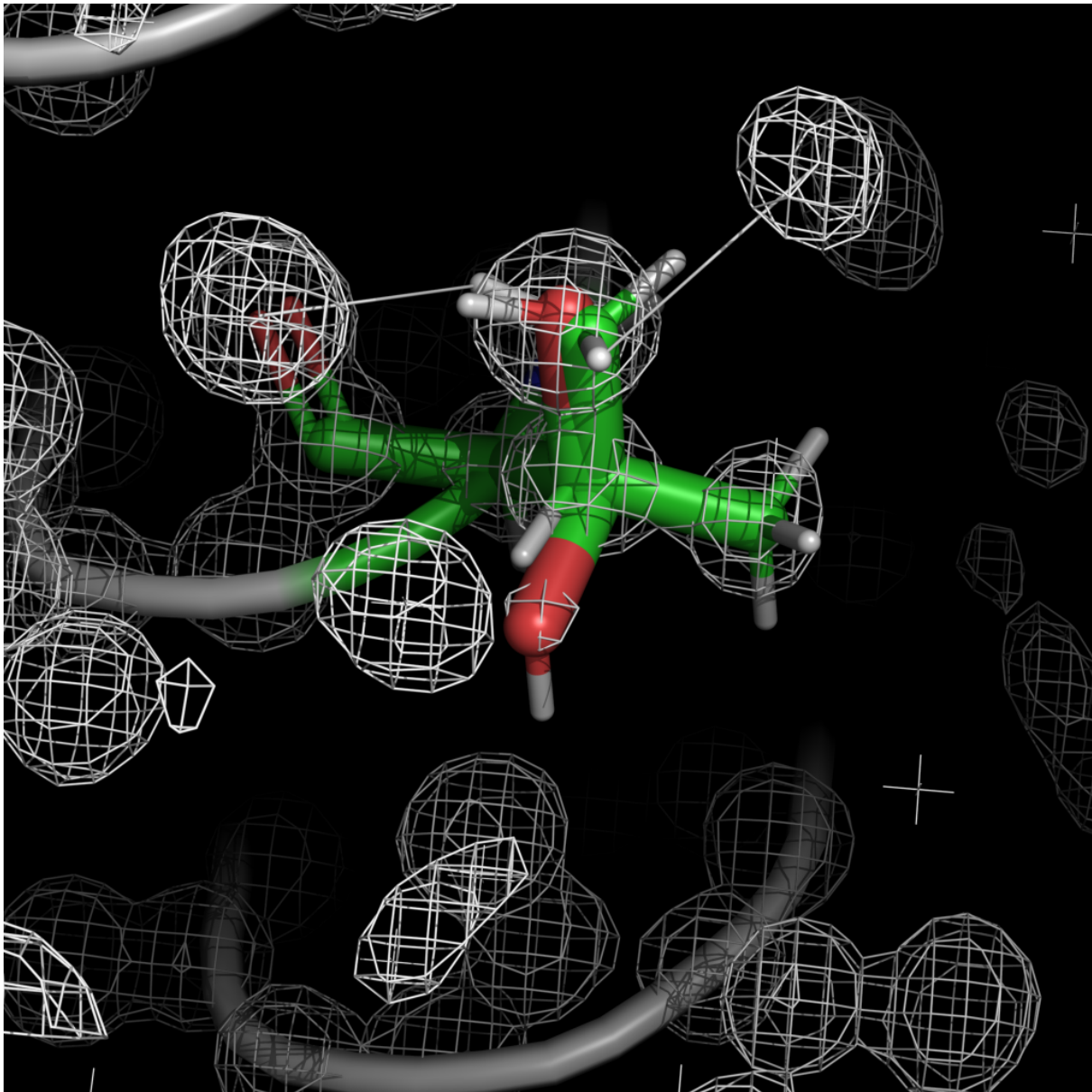


Рисунок 4. Остаток His158 и поверхность уровня электронной плоскости  $2\sigma$ .

На рисунке 5 изображена модель и поверхность уровня электронной плотности  $2\sigma$  вокруг остатка Thr208. Остаток не закреплен жестко в структуре белка (находится в петле): для него получены 2 конформации, при этом одна из них (В – на рис. 5 в этой

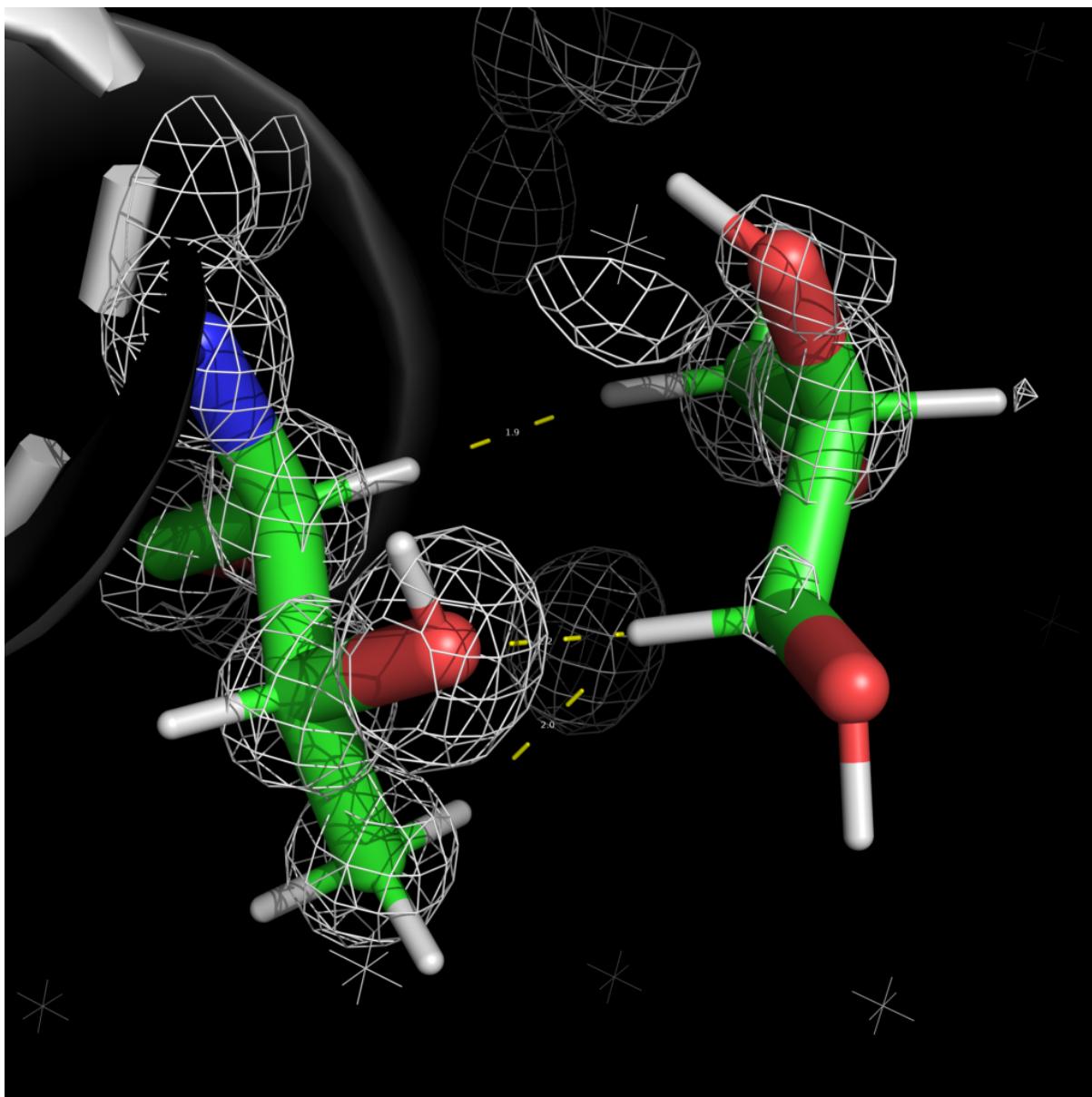
конформации вверх направлен атом углерода) менее заселенная (это видно, например, из того, что пик электронной плотности в направлении вниз-влево, соответствующий атому О в конформации В, слабый). Молекулы воды разрешены единственным способом, по видимости, соответствующим более заселенной конформации белка, А, поэтому наблюдается аномальное сближение в модели атомов конформации В с окружающими молекулами воды.



*Рисунок 5. Остаток Thr208 и поверхность уровня электронной плоскости  $2\sigma$ . Тонкой линией связаны атомы, Ван-дер-Ваальсовы радиусы которых пересекаются в модели.*

На рисунке 6 изображена модель и поверхность уровня электронной плотности  $2\sigma$  вокруг остатка Thr114. В отличие от прочих рассмотренных маргинальных остатков, он находится внутри  $\alpha$ -спирали и хорошо разрешен. Для него найдены очень близкие

контакты с молекулой глицерина (показаны пунктиром на рис. 6). По видимости, это происходит из-за того, что глицерин здесь подвижен (так, ближайшие к “камере” на рис. 6 атомы углерода и кислорода моделированы по очень слабым пикам электронной плотности), и модель плохо изображает реальное положение молекулы.



*Рисунок 6. Остаток Thr114 и поверхность уровня электронной плоскости 2σ. Пунктирной линией связаны атомы, Ван-дер-Ваальсовы радиусы которых пересекаются в модели.*

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

Перестроение модели сервисом PDB\_redo слабо улучшает качество модели: на рисунке 7 приведены основные показатели качества этих моделей



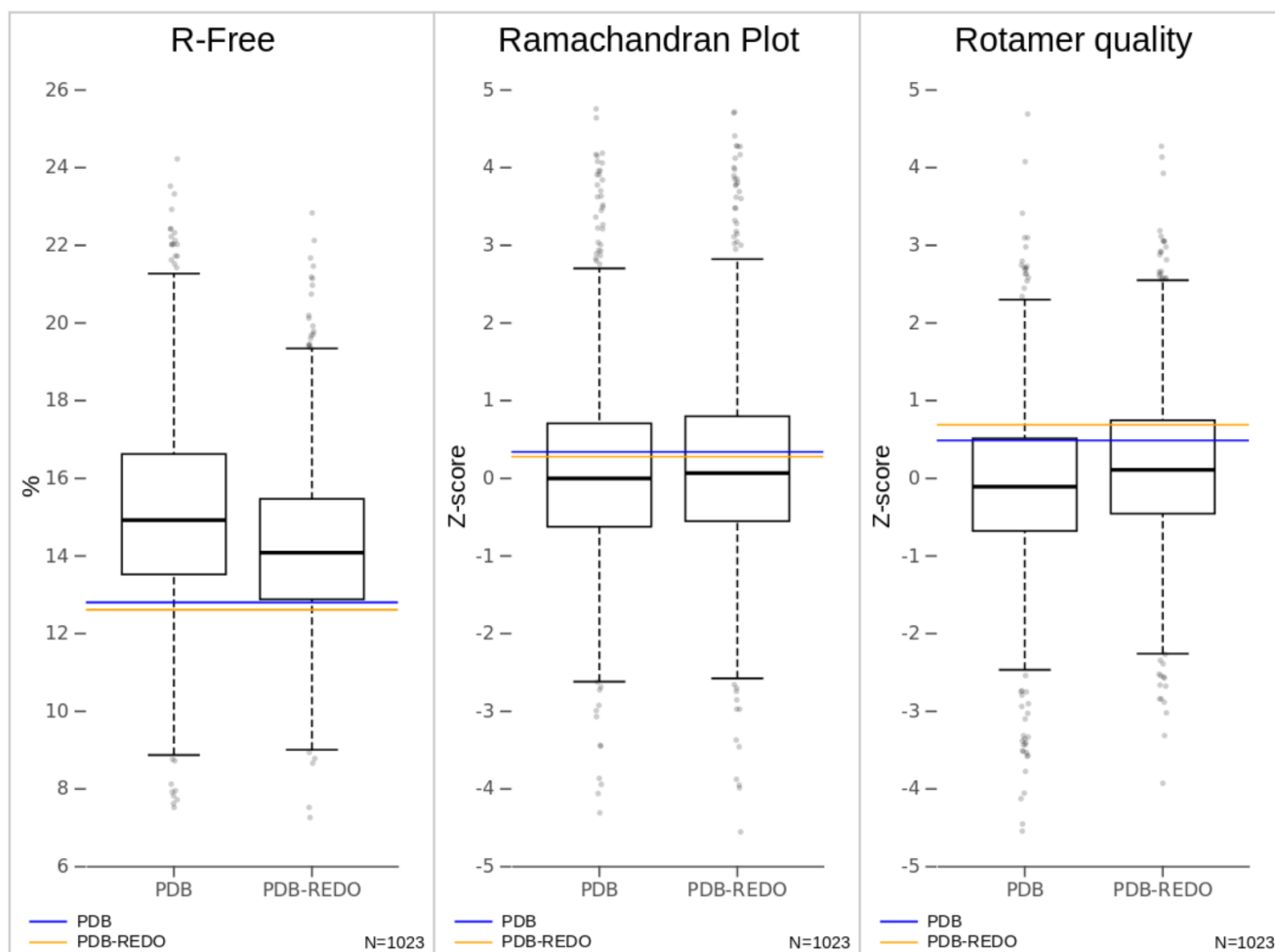


Рисунок 6. Сравнение модели beio из PDB и PDB\_redo. Для построения boxplot использованы данные 1023 моделей аналогичного разрешения.

Приведенное улучшение основных показателей приводит к ухудшению “физического реализма” модели: повышается количество пересечений атомов (96 перцентиль качества в PDB, 87 – в PDB\_redo) и уменьшается количество водородных связей (94 перцентиль качества в PDB, 88 – в PDB\_redo).

## Заключение

По моему мнению, высокое разрешение полученной структуры позволило авторам расшифровки получить очень качественную модель. Найденные неточности в модели часто также обусловлены тем, что разрешение высоко, и модель содержит информацию сразу о нескольких конформациях белка и лигандов. Возможно, с увеличением числа моделей высокого разрешения будут разработаны программы, способные с использованием представлений о возможных движениях молекулы дополнительно оптимизировать такие модели.

## Источники

[1] Mangiagalli, M. , Sarusi, G. , Kaleda, A. , Bar Dolev, M. , Nardone, V. , Vena, V. F., Braslavsky, I. , Lotti, M. and Nardini, M. (2018), Structure of a bacterial ice binding protein with two faces of interaction with ice. FEBS J, 285: 1653-1666. doi:[10.1111/febs.14434](https://doi.org/10.1111/febs.14434)

[2] RSCB PDB – 6EIO: Crystal structure of an ice binding protein from an Antarctic Biological Consortium

[3] PDB 6eio structure summary ‹ Protein Data Bank in Europe (PDBe) ‹ EMBL-EBI

[4] Electron Density Server at Uppsala University

[5] Vincent B. Chen, W. Bryan Arendall III, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson and David C. Richardson (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica D66: 12-21.

[6] Full wwPDB X-ray Structure Validation Report, [https://www.ebi.ac.uk/pdbe/entry-files/download/6eio\\_full\\_validation.pdf](https://www.ebi.ac.uk/pdbe/entry-files/download/6eio_full_validation.pdf)

[7] PDB-REDO: updated and optimised crystallographic structures