

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

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

Отчёт по расшифровке структуры Zlez

Демкина А. О.

**Москва
2017г.**

СОДЕРЖАНИЕ

1 ВВЕДЕНИЕ	3
2 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ	4
2.1 Общая информация о модели	4
2.2 Индикаторы качества модели	5
2.3 Маргинальные остатки	8
2.4 Анализ маргинальных остатков и гетеромолекул.	8
2.4.1 Thr 180	8
2.4.2 276 Met	9
2.4.3 Анализ молекулы 4-(2-гидроксиэтил)-1-пиперазин этансульфоновой кислоты	9
2.4.4 303 Arg	10
2.4.5 304 Asn	10
2.5 Сравнение PDB и PDB_redo	11
3 ЗАКЛЮЧЕНИЕ	13
4 СПИСОК ЛИТЕРАТУРЫ	14

1 ВВЕДЕНИЕ

Зlez - это структура фермента β -лактамазы (рисунок 1) из *Oceanobacillus iheyensis*. По классификации ферментов это фермент класса гидролаз с ЕС номером: [3.5.2.6]. Функция данного фермента заключается в расщеплении β -лактамного кольца (рисунок 2), входящего в состав некоторых классов антибиотиков, например, антибиотиков класса пенициллинов. Соответственно, данный белок отвечает за устойчивость к антибиотиком.

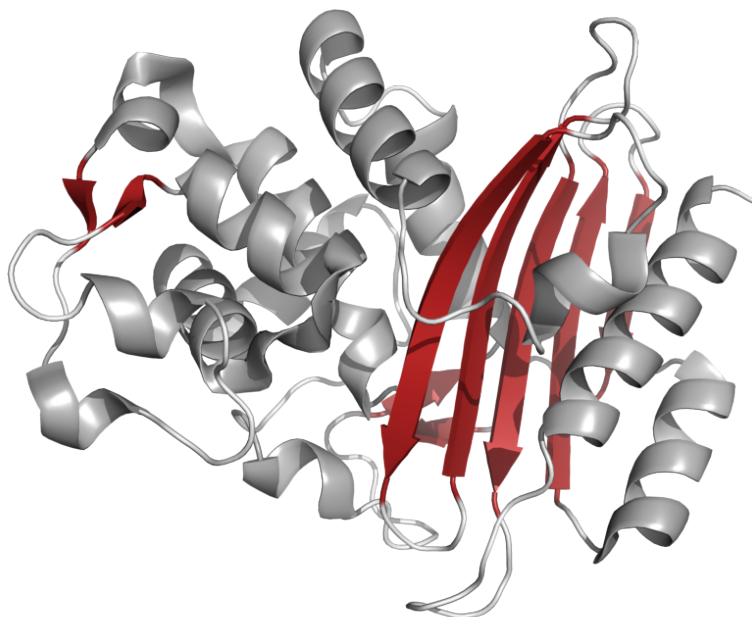


Рисунок 1. Структуры β -лактамазы Zlez

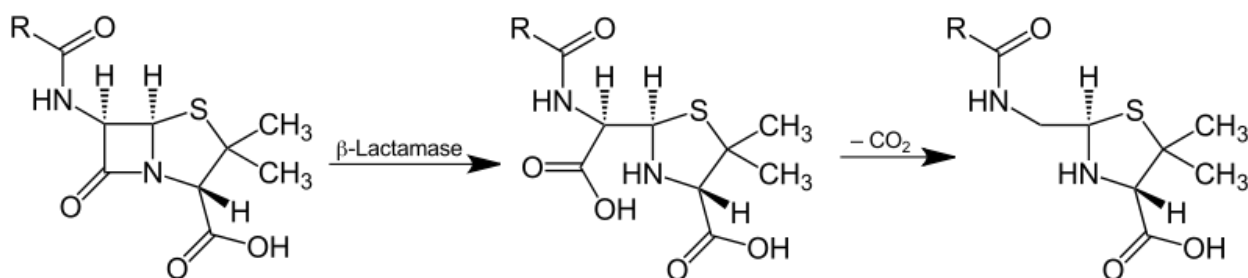


Рисунок 2. Реакция, катализируемая β -лактамазой, изображения взято с <https://en.wikipedia.org/wiki/Beta-lactamase>

Бактерия *Oceanobacillus iheyensis* является галотолерантной бактерией, обитающей на глубине более 1050 м в Тихом океане и данная β -лактамаза представляет интерес для изучения из-за того, что она развивалась, видимо, достаточно независимо от других ферментов данного класса и способна работать при большой концентрации соли.

2 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

2.1 Общая информация о модели

Белок закристаллизован в виде мономера, связанного с 4-(2-гидроксиэтил)-1-пиперазин этансульфоновой кислотой и ионами Cl^{1-} и Ca^{2+} . Также в структуре присутствует много молекул воды. Авторы работы [1]: Toth, M., Smith, C., Frase, H., Mobashery, S., Vakulenko, S.. Для решения фазовой проблемы ими был применён метод молекулярного замещения с использование программы molrep [2]. Ниже рассмотрим несколько характеристик модели, взятых с сайта rcsb.org [3].

Общее количество рефлексов: 84295, число рефлексов, использованных для определения структуры: 84272. Разрешение данных очень хорошее: 1.25Å, смоделированы в том числе и молекулы воды. Диапазон разрешений для структурных факторов равен: 1.25Å - 41.855Å.

Полнота набора структурных факторов: 99.9 %. Основной причиной неполноты структурных факторов является отсутствие измерений по некоторым направлениям.

Описание параметров кристаллографической ячейки:

Поле записи о кристаллографической симметрии в pdb-файле, CRYST1, содержит следующую информацию:

```
CRYST1 55.980 55.980 165.890 90.00 90.00 120.00 P 31 2 1 6
```

Первые три значения соответствуют длинам направляющих векторов ячейки кристалла, следующие три - углы между ними, дальше идёт название кристаллографической группы (P 31 2 1) и, последним, число молекул в ячейке - 6. Данный тип кристаллографической ячейки относится к простому типу и является тригональной ячейкой, для которой разрешены следующие 6 видов операций симметрии:

- x, y, z
- $\bar{y}, x - y, \frac{1}{3} + z$
- $\bar{x} + y, \bar{x}, \frac{2}{3} + z$
- x, y, \bar{z}
- $x - y, \bar{y}, \frac{2}{3} - z$
- $\bar{x}, \bar{x} + y, \frac{1}{3} - z$

Ассиметрической единицей является мономер белка. За некристаллографические симметрии в файле pdb отвечает поле MTRIXn, которого нет в pdb-файле для белка 3lez. Восстановленный кристалл белка показан на рисунке 3

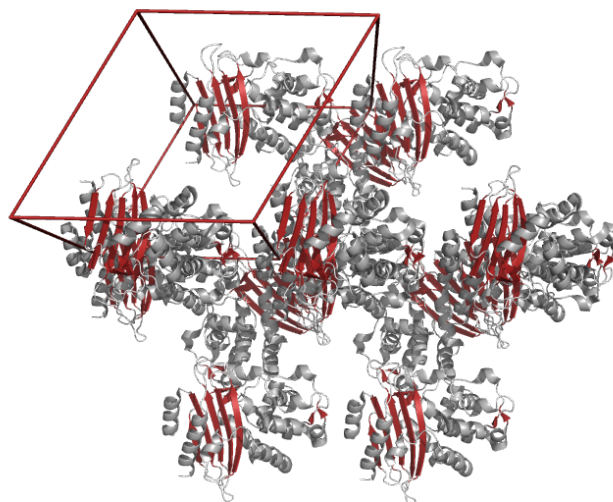


Рисунок 3. Восстановленная структура кристалла 3lez

2.2 Индикаторы качества модели

Для оценки качества модели есть несколько разных критериев. Ниже мы разберём некоторые из них. Основные значения таких оценок приведены на сайте PDB (рисунок 4), более подробный отчёт можно скачать там же.

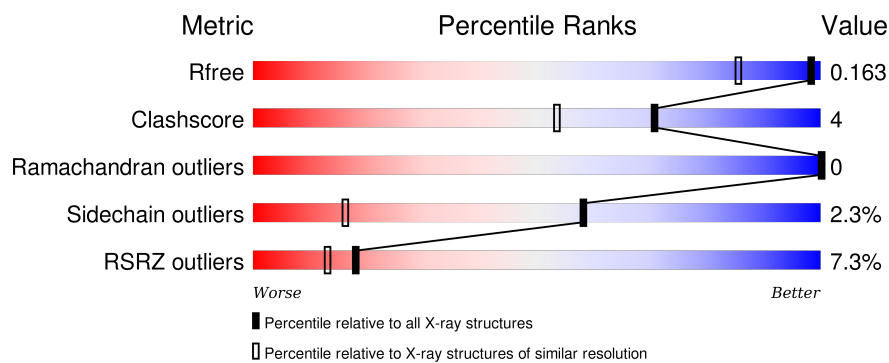


Рисунок 4. Краткий отчёт о качестве структуры PDB

Давайте разберём некоторые из этих критериев подробнее.

R-фактор описывает соответствие построенной модели экспериментальным данным и вычисляется по следующей формуле: $R = \frac{\sum_{hkl} |F_{hkl}^{calc} - F_{hkl}^{obs}|}{\sum_{hkl} F_{hkl}^{obs}}$, где F_{hkl}^{obs} и F_{hkl}^{calc} - рассчитанные и экспериментально полученные модули структурных факторов соответственно. Значение R-фактора уменьшается в процессе оптимизации, то есть чем меньше его значение, тем лучше оптимизирована модель. Для данной модели значение R-фактора равно 0.154, что является хорошим показателем, однако более точным критерием является так называемый R-free фактор, который вычисляется по той же формуле, что и R - фактор, однако для его подсчёта

используется набор структурных факторов, случайно выбранных из исходного набора и не участвующих в оптимизации модели. Для данной структуры R-free равен 0.174, что может говорить о лёгкой переоптимизации модели. Однако значение R-free сильно меньше, чем среднее значение данного фактора для структур с таким разрешением.

Такой параметр, как Clashscore показывает число стерических перекрываний атомов и для данной модели равно 4.

Следующий показатель - Ramachandran outliers - показывает число остатков, чьи ψ и ϕ углы выходят за пределы разрешённых значений, согласно картам Рамачандрана. Согласно pdb, все аминокислотные остатки находятся в допустимых областях карты Рамачандрана. Используя сервис MolProbity (<http://molprobity.biochem.duke.edu/>) можно получить чуть более подробную информацию. Если посмотреть на выдачу этого сервиса (рис 5), то можно увидеть, что не все остатки находятся в разрешённых, но не благоприятных областях, но процент последних всё-равно высок: 96.27 %.

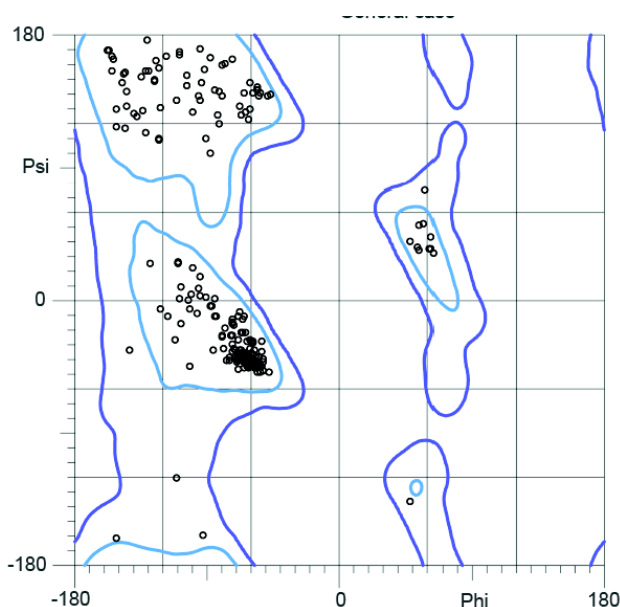


Рисунок 5. Карта Рамачандрана для всех аминокислотных остатков из структуры 3lez

Число $C\beta$; - атомов, отклоняющихся от "идеального" положения равно 0, число валентных связей, отклоняющихся от ожидаемых равно 3, причём одно из них относится к гетеромолекуле и не является остатком белка, а углов 0.

RSR - пространственный R-фактор, позволяет посмотреть соответствие рассчитанной по модели электронной плотности и реальной. Ниже, на рисунке 6, приведён график RSR для аминокислотных остатков белка. Как видно из этого графика, есть 10 остатков со значениями выше 20%.

RSR-Z - это оценка RSR для отдельных остатков, основанная на сравнении со средним RSR для таких же остатков из других структур из PDB со сходным разрешением. На рисунке 7 видно, что есть несколько остатков со значением Z-RSR больше 2, что может служить сигналом о том, что они плохо вписаны в электронную плотность.

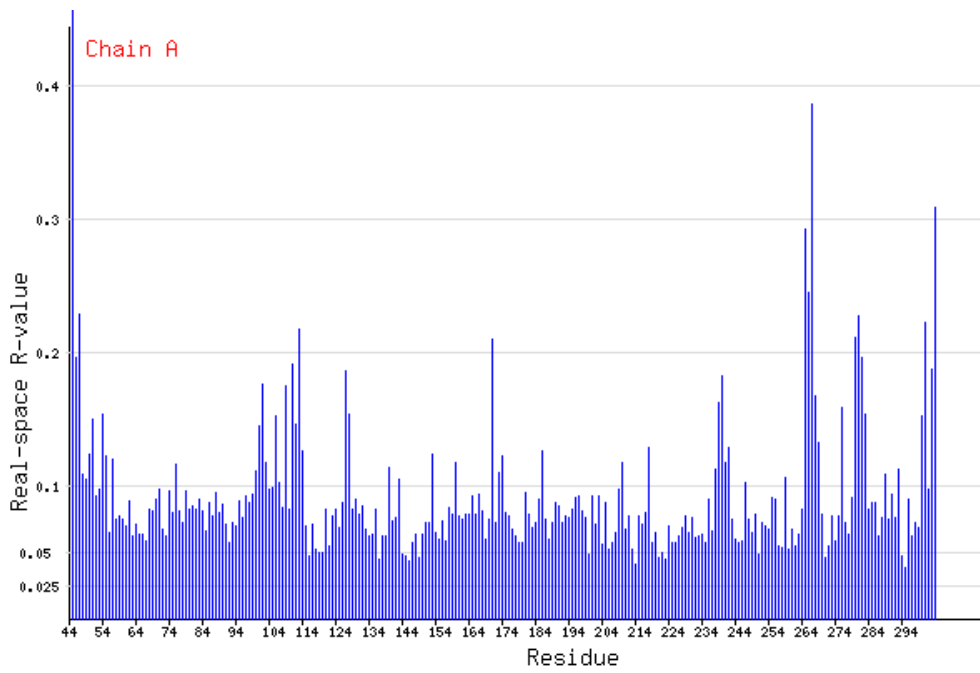


Рисунок 6. График значений RSR для аминокислотных остатков 3lez

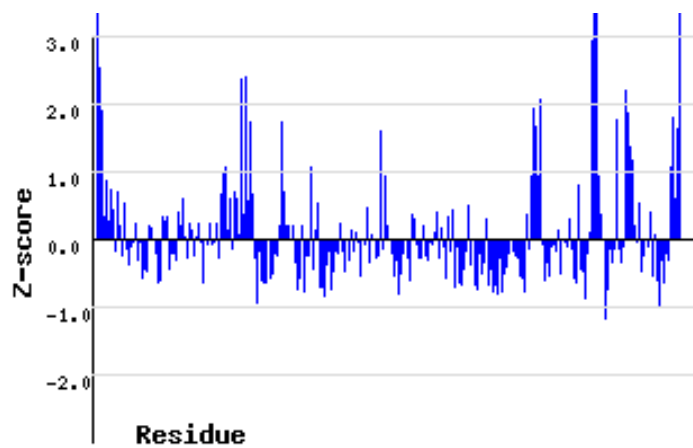


Рисунок 7. График значений RSR-Z для аминокислотных остатков 3lez

Дальше была проверена комфортность окружения остатков с помощью сервиса WHAT CHECK (<http://swift.cmbi.ru.nl/servers/html/index.html>), по результатам работы сервиса вышло, что только 2 остатка имеют неблагоприятное окружение, мы рассмотрим их дальше.

2.3 Маргинальные остатки

Таблица 1

Таблица маргинальных остатков

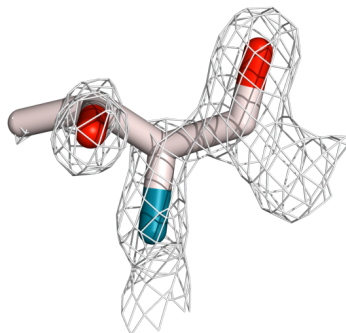
#	Остаток	Причина отнесения к маргинальным
1	128 Glu	Неблагоприятное окружение, параметр < -3
2	182 Leu	Неблагоприятное окружение, значение параметра ≈ 2.5
3	102 Glu	Poor rotamer, outlier, $\chi = 16.7, 269.2, 240.5$
4	143 Ser	Poor rotamer, outlier, $\chi = 221.9$
5	267 Asp	Poor rotamer, outlier, $\chi = 231.5, 277.5$
6	303 Arg	Poor rotamer, outlier, $\chi = 286.2, 270.8, 122.9, 162.8$
7	180 Thr	Cis-конформация
8	276 MET	Длина N-C связи
9	304 Asn	Значение Z-score > 4

2.4 Анализ маргинальных остатков и гетеромолекул

В этом разделе давайте рассмотрим подробнее некоторые остатки, которые мы считали "маргинальными" выше.

2.4.1 Thr 180

Согласно сервису EDS данный остаток находится в цис-конформации, что вызывает подозрения в его маргинальности. Однако, если посмотреть на то, как данный остаток вписывается в электронную плотность (см. рисунок 8), то можно увидеть, что даже при большом уровне подрезки (значение $\sigma = 3$), остаток прекрасно вписывается в неё, соответственно это может действительно быть необычная конформация аминокислотного остатка в белке.

Рисунок 8. остаток 180 Thr с электронной плотностью, значение $\sigma = 3$

2.4.2 276 Met

Данный остаток, согласно сервису EDS имеет слишком большую длину связи. Если посмотреть на этот остаток в структуре (рисунок 9), то видно, что возможно, проблемы были вызваны наличием двух состояний этого остатка в pdb-файле. Согласно карте электронных плотностей, вероятнее всего только одно состояние, так как электронная плотность покрывает полностью только одно из них, особенно при большом уровне подрезки, но, видимо, наличие альтернативного варианта, ухудшает вписывание остатков.

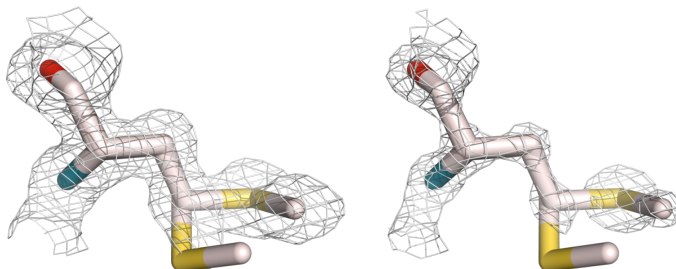


Рисунок 9. остаток 276 Met с электронной плотностью, значение $\sigma = 1$ на изображении слева и 3 изображении справа

2.4.3 Анализ молекулы 4-(2-гидроксиэтил)-1-пиперазин этансульфоновой кислоты

Молекула данной кислоты была закристаллизована вместе с белком, что означает, что её контакт с белком был довольно прочным. Если посмотреть на то, как она вписывается в электронную плотность (рисунок 10), то можно увидеть, что даже при небольшом уровне подрезки (1σ), покрытие электронной плотностью у молекулы достаточно слабое, но при этом есть участок молекулы (концевой фосфат), который хорошо покрыт электронной плотностью даже при уровне обрезки в 3σ , что может означать, что данный участок молекулы прочно связан с белком. В действительности, согласно структуре 3lez, фосфатная часть молекулы соприкасается с белком, в то время как остальная часть молекулы "висит" в растворе белка.

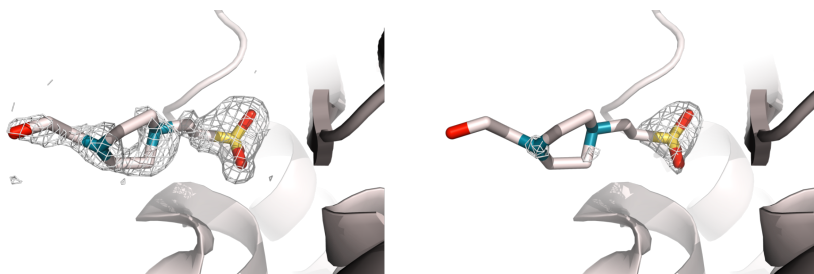


Рисунок 10. Молекула 4-(2-гидроксиэтил)-1-пиперазин этансульфоновой кислоты с электронной плотностью, значение $\sigma = 1$ на изображении слева и 2 изображении справа

2.4.4 303 Arg

Данный остаток, согласно сервису EDS, является маргиналом из-за того, что углы χ имеют значения, выходящие за допустимые (286.2, 270.8, 122.9, 162.8). Если посмотреть на карту его электронной плотности (рисунок 11), то можно увидеть, что он вписан в неё неправильно и углы в аминокислотном радикале должны быть другими, более того, у данного остатка полностью отсутствует покрытие уже при значении $\sigma = 2$, так что можно сказать, что данный остаток является маргиналом.

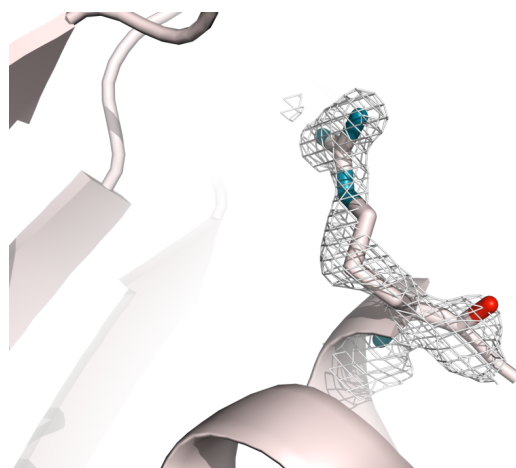


Рисунок 11. Остаток 303 Arg с электронной плотностью, значение $\sigma = 1$

2.4.5 304 Asn

Данный остаток имеет очень большое значение RSR-Z = 4.44. Это С-концевой остаток цепи белка, который смотрит в раствор, соответственно можно предположить его большую подвижность. Если посмотреть на карту его электронной плотности, то она покрывает только часть его и существует только при уровне обрезки = 1 (рисунок 12), так что, вероятнее всего, можно сказать, что этот остаток является маргинальным.

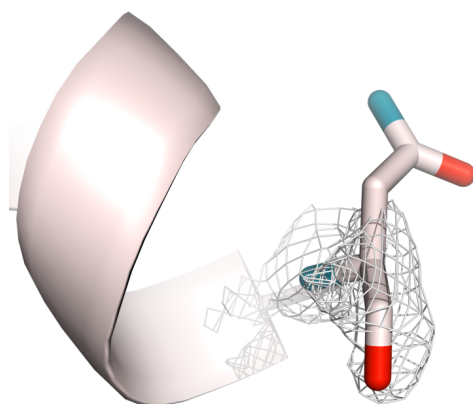


Рисунок 12. Остаток 304 Asn с электронной плотностью, значение $\sigma = 1$

2.5 Сравнение PDB и PDB_redo

Сервис PDB redo позволяет оптимизировать структуру, взятую из pdb, используя различные параметры уточнения, выбирая между изо- и анизотропным B-фактором, перестраивая конформации ротамеров, вращая боковые цепи остатков для оптимизации сети водородных связей, пересматривая молекулы воды и конформации лигандов.

Для данной структуры pdb redo не дало сильных изменений в конформации белка и аминокислотных остатков, только алгоритм PyMol по распознавания вторичной структуры у структур из pdb redo относит к β -листам намного меньше остатков (рисунок 13).

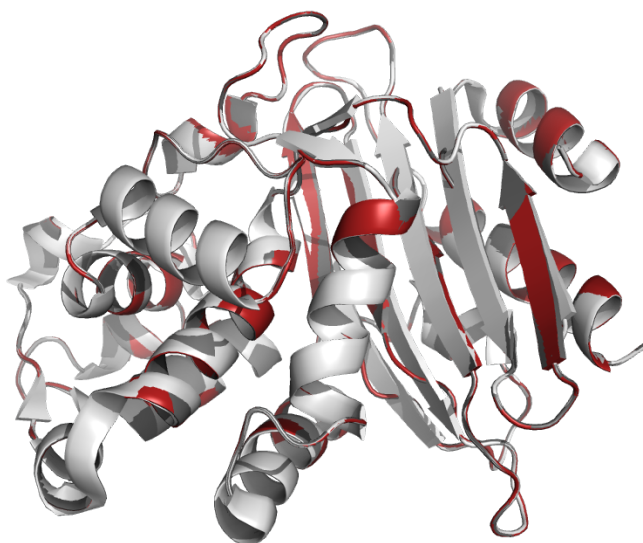


Рисунок 13. Сравнение исходной структуры (белый) со структурами, полученными с помощью pdb redo: conservative (серый) и fully-optimised (красный)

Согласно собственному отчёту pdb redo, он не произвёл значительного улучшения конформации у большинства аминокислотных остатков, значительное улучшение наблюдается у C-концевых остатков 301, 303 и 304, и у большого количества молекул воды, хотя положение 5 молекул воды претерпело значительное ухудшение. Если рассматривать изменения в конформации аминокислотных остатков, рассматриваемых в разделе выше, то там есть некоторые изменения (рисунок 14). Оба варианта Met 276 несколько сблизилась в структуре, видимо, чтобы второй вариант также лучше соответствовал экспериментально полученной электронной плотности. Концевые остатки Arg 303 и Asn 304, внешне, не претерпели сильных изменений, в том числе остались прежние неправильные значения углов χ у Arg, хотя, согласно данным pdb redo, конформация данных остатков была значительно улучшена.

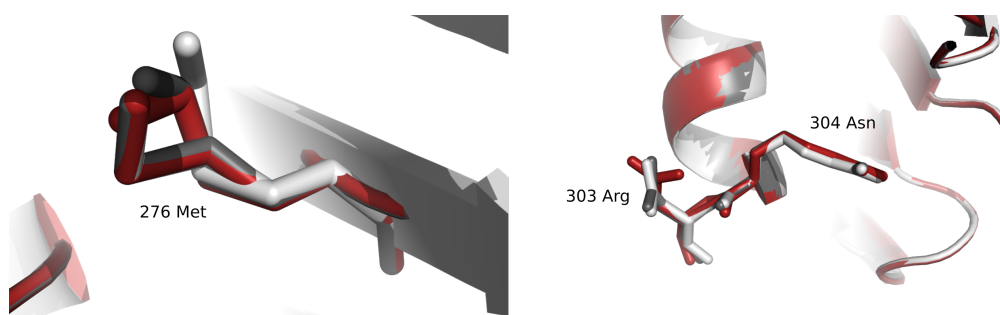


Рисунок 14. Сравнение расположения остатков 276, 303 и 304 в исходной структуре (белый) и структурах, полученных с помощью pdb redo: conservative (серый) и fully-optimised (красный)

3 ЗАКЛЮЧЕНИЕ

Данная структура определена очень хорошо, среди структур с таким же разрешением она имеет куда меньшей R-free фактор, само разрешение в 1.25 также является хорошим показателем. Отсутствуют остатки в запрещённых областях карты Рамачандрана, а также почти нет остатков с неблагоприятным окружением (есть 2 возможных остатка, но при более близком рассмотрении становится понятно что, скорее всего, это ошибочное утверждение), а pdb redo не внёс никаких серьёзных изменений в структуру.

Экспериментальных данных достаточно для точного установления положения атомов не только белка, но и молекул воды. Есть только небольшие проблемы на С-концевом участке белковой цепи, видимо, из-за наибольшей её подвижности в структуре.

В целом, можно сделать вывод, что данную модель можно считать очень точной и соответствующей экспериментальным данным.

4 СПИСОК ЛИТЕРАТУРЫ

1. Toth, M., Smith, C., Frase, H., Mobashery, S., and Vakulenko, S. (2010) An antibiotic-resistance enzyme from a deep-sea bacterium, *J. Am. Chem. Soc.*, **132**, 816–823.
2. Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
3. Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J.D., and Zardecki, C. (2002) The Protein Data Bank, *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 899–907.