**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**имени М.В.ЛОМОНОСОВА**

##### ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Отчет о качестве расшифровки структуры белка 1W5D методом рентгеноструктурного анализа**

**Работу выполнила:**

**Елисеева Ю.И.**

**402 группа**

**Москва, 2014 г.**

**Оглавление**

1. **Аннотация…………………………………………………………………………….....3**

**Использованные сокращения………………………………………………………...3**

1. **Введение………………………………………………………………………………….4**
2. **Результаты и обсуждение……………………………………………………………...5**
   1. **Общая информация о модели……………………………………………………..5**
   2. **Значения индикаторов качества модели в целом……………………………...5**
   3. **Маргинальные остатки……………………………………………………………8**
      1. **Список нескольких маргинальных остатков…………………………………..8**
      2. **Детальный анализ нескольких маргинальных остатков………..……9**
   4. **Сравнение модели из PDB с моделью из PDB\_redo…………………………..12**
3. **Выводы…………………………………………………………………………………13**
4. **Использованные источники…………………………………………………………14**
5. **Аннотация**

В данном отчете я попробовала продемонстрировать анализ структуры 1W5D [2]. Рассматривались разные индикаторы для того, чтобы сделать выводы о качестве 1W5D. Соответственно, целью являлось показать недостатки в этой структуре.

**Использованные сокращения:**

B- factor – температурный фактор

PBP - пенициллин-связывающий домен

PBPs-C1 - пенициллин-связывающий домен подкласса C1

RSR (real-space R factor) – пространственный R factor

1. **Введение**

Геном Bacillus subtilis, кодирует 16 пенициллин-связывающих доменов (PBPs), вовлеченных в синтез и/или ремоделирование пептидогликана в течение сложного жизненного цикла этой спорулирующий Грам-положительной палочковидной бактерии. PBP4a (кодируемый геном dacC) – это низкомолекулярное вещество, ясно показывающее in vitro DD-карбоксипептидазную активность. В 2007 году была определена кристаллическая структура его одного и в комплексе с пептидом (D-α-aminopymelyl-ε-D-alanyl-D-alanine), который скрывает терминальный C- конец “Bacillus peptidoglycan stem peptide” [1].

PBP4a кристаллизовался в группу симметрии кристаллической решётки P3212 с 1 молекулой в асимметрической ячейке.

Actinomadura R39 DD - пептидаза использовалась в качестве модели в методе молекулярного замещения.

PBP4a структура содержит 3 домена, 1 из которых сохраняет типичную структуру и активный сайт пенициллин-связывающих протеинов. Этот домен состоит из 5-виткового антипараллельного β-листа, покрытого с 2-х сторон α-спиралями. Между консервативными мотивами 1 и 2 находится вставка из 221 остатка, формирующая домены II и III, при этом домен III вставлен во II. Каждый из 3-х доменов имеет

β-лист, ориентированный направо к остальным.

Длинная неструктурированная петля (130-156) присутствует в домене II. Она хорошо упаковывается в пенициллин-связывающем домене, с остатками Pro142, Asp145, Tyr150, которые являются частью специфического кармана на дне активного сайта.

На противоположной стороне домена II серия остатков лизина (Lys83, Lys85, Lys86, Lys114, Lys119, Lys122 и Lys265) формирует положительно заряженную поверхность на стороне обратной пенициллин-связывающему домену. Эти лизины не консервативны в первичной структуре, и эта особенность может быть уникальной для B. subtilis PBP4a.

Структура кристалла PBP4a и его комплекса сообщила информацию о специфичной ферментативной активности PBPs-C1, которая описана как DD - карбоксипептидазная,

DD – транспептидазная и DD – эндопептидазная [1].

1. **Результаты и обсуждение**
   1. **Общая информация о модели [2]**

В данной работе использовался метод молекулярного замещения для решения фазовой проблемы [1].

Число измеренных рефлексов составляет 34 427.

Число рефлексов с силой сигнала, превышающей стандартное отклонение более чем в три раза равен 33 638. Это составляет 97,7% от всех.

Разрешение: 2.10 Å.

Минимальное и максимальное разрешение для использованных рефлексов составляет 58.38 и 2.10 Å соответственно.

Полнота набора рефлексов равна 97.7 %.

Параметры элементарной ячейки: a = 67.41 Å, b = 67.41 Å, c = 228.46 Å; α = 90.00 °, β = 90.00 °, γ = 120.00°.

Группа симметрии: P (примитивная ячейка) 32 1 2.

* 1. **Значения индикаторов качества модели в целом**

R-фактор: 0.229

R\_free: 0.268

R\_free – R-фактор = 3.9% (<10%) – переоптимизации не наблюдается.

Число маргинальных остатков по карте Рамачандрана равно 7. Ниже приведены соответствующие аминокислоты и их углы (phi,psi):

82 LEU (172.4, 118.8)

85 LYS (-168.4, -28.1)

181 GLY (62.9, -34.4)

183 GLU (164.6, 144.1)

190 PRO (-48.9, -84.7)

191 LYS (163.1, 108.0)

431 LYS (-46.8, -178.8)

Сама карта Рамачандрана приведена ниже (рис.1.)

92,8 % всех остатков находятся в предпочитаемой области, а 98,5% в разрешенной.

Это довольно хорошо, что видно, и из самой карты. Остатки, которые не попали в эти области, по-видимому, являются маргинальными. Более подробное их рассмотрение будет далее.

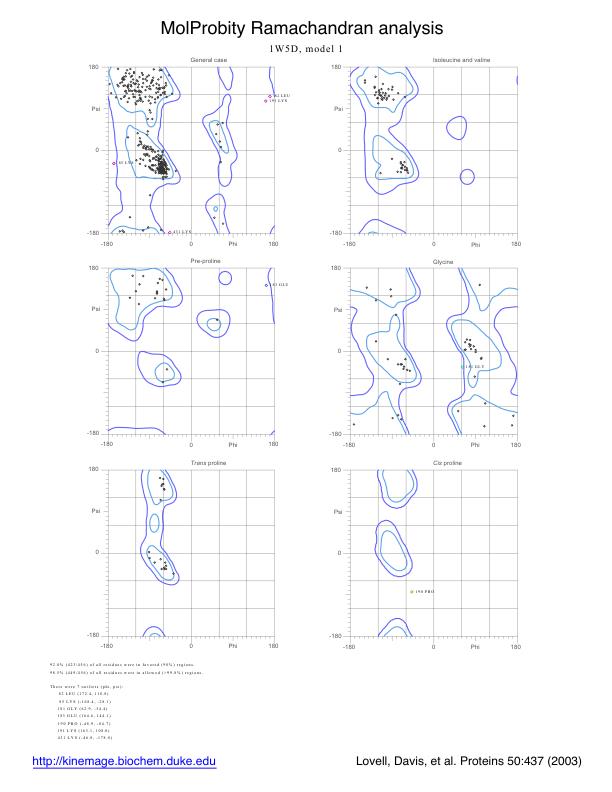


Рис.1. Карта Рамачандрана изучаемой структуры [3]

Также рассмотрела некоторые другие индикаторы модели [4].

Например, RSR (пространственный R-factor), который при норме равен от 10 до 20%. А у моей модели средний RSR составляет 22,9%, что не очень хорошо. Значит, многие атомы плохо вписываются в построенную электронную плотность. Общее распределение RSR-фактора для структуры 1W5D представлено на рисунке 2 (для каждого остатка показано соответствующее ему значение RSR) . На этом же рисунке видно несколько выдающихся пиков, судя по всему, это маргинальные остатки, речь о них пойдет дальше.

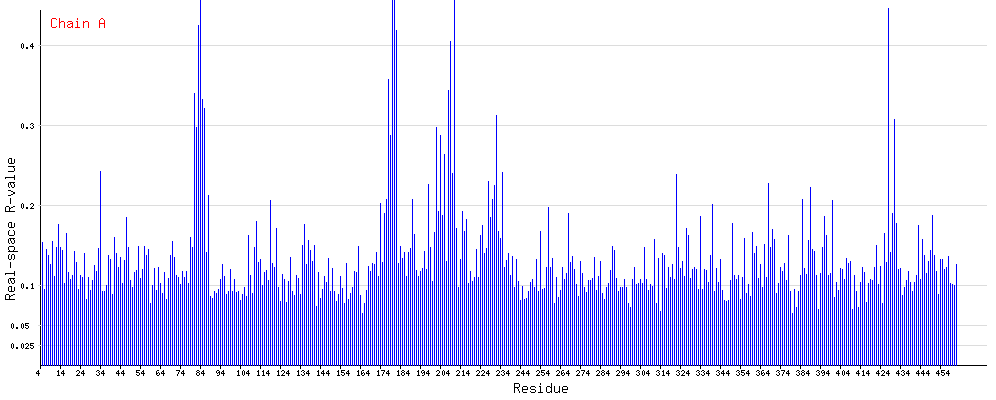


Рис.2. RSR изучаемой структуры [4]

Еще один немаловажный индикатор качества модели – это температурный фактор, отвечающий за ‘размазанность’ функции электронной плотности, соответствующей определенному атому. Хорошо, когда он принимает значения меньше 20. В данном случае средний B-фактор равен 22,9, и видно, что для некоторых остатков он принимает значения больше 60! Это нехорошо, значит, эти атомы сильно колеблются, и мы не можем их хорошо идентифицировать в пространстве.

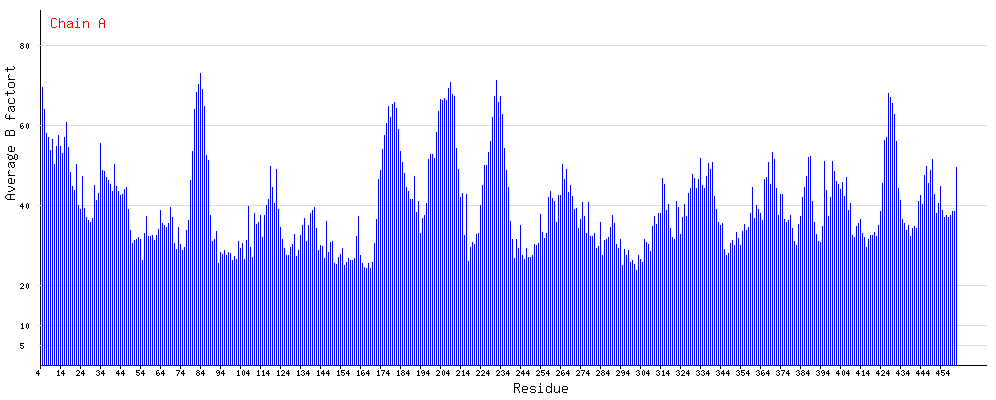


Рис. 3. Распределение значений температурного (В) фактора для атомов структуры [4]

Ещё рассмотрела Z-score,который можно посчитать так: Z-score=(RSR-<RSR>)/sigma.

Z-score считается для отдельного атома, и в хорошем случае он должен быть меньше 2, здесь средний Z-score равен 0.13. Это неплохо. Но опять видны остатки (рис.4.), у которых этот параметр превышает 2, всего таких остатков 4.37% от всех.

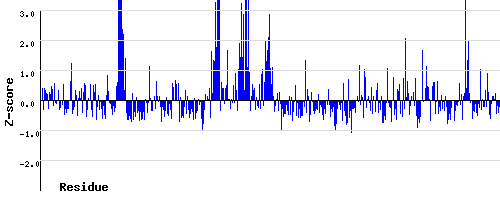


Рис. 4. Z-score остатков изучаемой структуры [4]

Еще 1 параметр, на который стоит обратить внимание, это “Significant regions”. Это ‘значимые регионы’ с большим RSR. Некоторые из этих пиков будут рассмотрены далее.

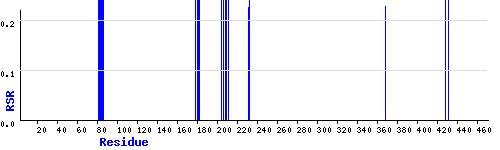


Рис. 5. Significant regions изучаемой структуры [4]

* 1. **Маргинальные остатки**

**3.3.1. Список нескольких маргинальных остатков**

|  |  |
| --- | --- |
| № | Показатели |
| Asp145 | Перекрывание с Thr 21(HG22) |
| Tyr150 | Перекрывание с Thr 21(HG21) |
| Lys 191 | Перекрывание с A2163 HOH O; вне разрешенной области на карте Рамачандрана. |
| Thr 81 | Плохое значение RSR (высокий) |
| Leu 82 | Плохое значение RSR (высокий) |
| Lys 83 | Плохое значение RSR (высокий) |
| Gly 84 | Плохое значение RSR (высокий) |
| Lys 85 | Плохое значение RSR (высокий); вне разрешенной области на карте Рамачандрана |
| Lys 86 | Плохое значение RSR (высокий) |
| Gly 181 | Вне разрешенной области на карте Рамачандрана |
| Gly 183 | Вне разрешенной области на карте Рамачандрана |
| Pro 190 | Вне разрешенной области на карте Рамачандрана |
| Lys 431 | Вне разрешенной области на карте Рамачандрана |

Таблица 1. Список нескольких маргинальных остатков [3,4]

* + 1. **Детальный анализ нескольких маргинальных остатков**

В первую очередь интересовал анализ тех остатков, роль которых обсуждалась в статье [1]. Например, Pro142, Asp145, Tyr150, которые являются частью специфического кармана на дне активного сайта.

Посмотрела на их параметры[3,4]: у Pro 142 всё хорошо, а вот 2 других, похоже, маргиналы:

Asp 145: clash 0.788Å, O with A 148 THR HG22

Tyr150: clash 0.481Å, CE2 with A 148 THR HG21

Посмотрела на эти остатки в PyMol [5]:

Tyr150: MolProbity [4] выдает, что у него происходит перекрывание с Thr 21. Выделила Tyr 150 циановым цветом, оранжевым обозначила атомы, между которыми должно происходить перекрывание.

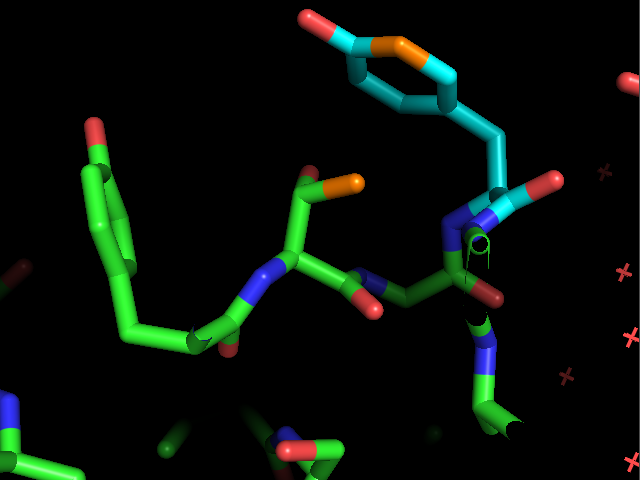


Рис. 6. Tyr150: перекрывание 0.481Å, CE2 с A 148 THR HG21

Asp 145: аналогично. Смотрим перекрывание между Asp 145(желтый) и Thr 148. Атомы, между которыми должно происходить перекрывание выделены оранжевым цветом.

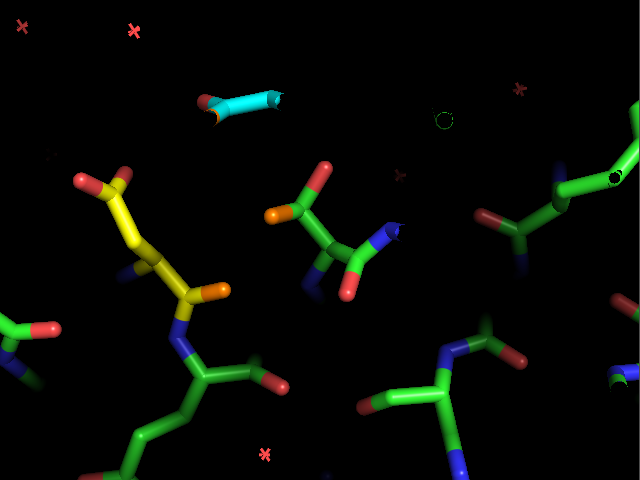


Рис.7. Asp 145: перекрывание 0.788Å, O с A 148 THR HG22

Также рассмотрела Lys 191, у него проблемы с несколькими параметрами:

Lys 191(Clash: 0.599Å HE2 with A2163 HOH O; Ramachandran: OUTLIER (0%) General / 163.1,108.0 [3].

У Lys 191 должно быть перекрывание между HE2 и водой A2163 HOH O (выделены оранжевым). И еще у этого остатка, видимо, неправильная укладка боковой цепи (что следует из карты Рамачандрана).

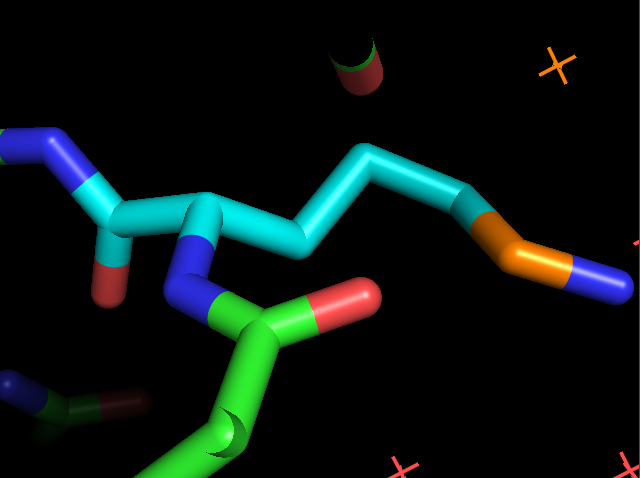


Рис. 8. Lys 191(перекрывание: 0.599Å HE2 с A2163 HOH O; Ramachandran: OUTLIER (0%) General / 163.1,108.0

Для остатков 81-86 плохое RSR. Посмотрела для них электронную плотность. Использовала уровень срезки 2 сигма. Как видно, атомы плохо поместились в электронную плотность. Также у этих остатков плохой z-score (больше 2). И температурный фактор больше 60! Значит, это маргиналы.

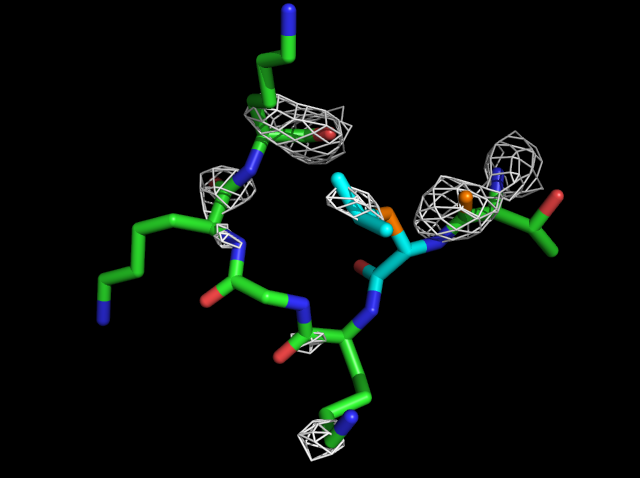


Рис. 9. Электронная плотность 81-86 остатков.

Также построила электронную плотность с уровнем подрезки 2 сигма, для остатков из статьи. Видно, что они хорошо в нее вписываются (рисунок 10).

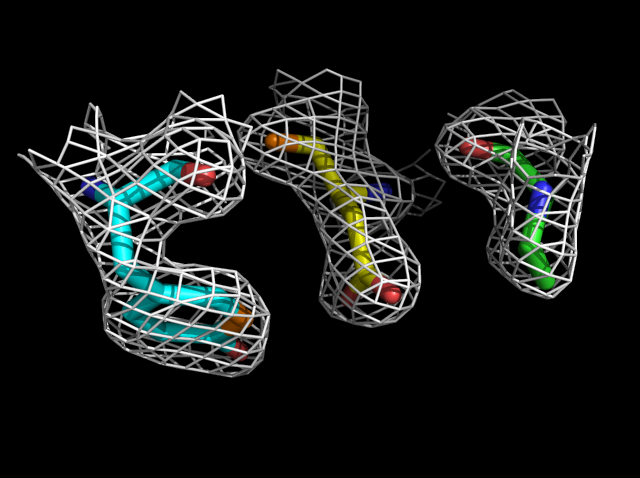


Рис. 10. Электронная плотность у аминокислот: Pro142 (зеленый), Asp145 (желтый), и Tyr150 (циановый)

Также в статье [1] упоминается серия остатков лизина (Lys83, Lys85, Lys86,Lys114, Lys119, Lys122 и Lys265), посмотрела на их электронную плотность.

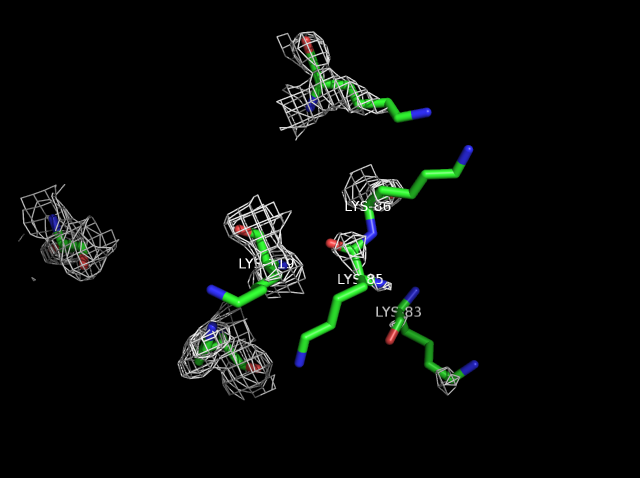


Рис. 11. Электронная плотность у остатков лизина

Видно, что некоторые остатки плохо вписались: Lys 83,Lys 85, Lys 86, Lys 119. И действительно про остатки 81-86, мы знаем, что они маргиналы. А у Lys 119 RSR = 20, 6.

* 1. **Сравнение модели из PDB с моделью из PDB\_redo**

PDB\_redo служит для оптимизации структуры. Сравнила некоторые данные из PDB для 1W5D с PDB\_redo. Видно, что они улучшились:

|  |  |  |
| --- | --- | --- |
| Параметр | PDB | PDB\_redo |
| R-Value | 0.224 | 0.1908 |
| R-free | 0.268 | 0.2216 |
| Ramachandran plot appearance | -1.341 | -0.176 |
| Chi-1/Chi-2 rotamer normality | -2.325 | -1.381 |

Таблица. 2. Данные для 1W5D из PDB и из PDB\_redo

1. **Выводы**

В целом, я думаю, качество структуры 1W5D является нормальным со средним разрешением (2.1 Å). Также структура неплохо оптимизирована. Конечно, есть и недостатки: довольно много остатков вне разрешенной области на карте Рамачандрана, многие атомы плохо вписываются в построенную электронную плотность (высокий RSR), многие атомы плохо локализованы в пространстве (большие значения температурного фактора).

1. **Использованные источники:**
2. Sauvage, E., Duez, C., Herman, R., Kerff, F., Petrella, S., Anderson, J.W., Adediran, S.A., Pratt, R.F., Frere, J.M., Charlier, P.(2007) J. Mol. Biol., 371, 528–539.

Crystal structure of the Bacillus subtilis penicillin-binding protein 4a, and its complex with a peptidoglycan mimetic peptide.

1. Сервис PDB (<http://www.rcsb.org/pdb/home/home.do>)
2. Сервис MolProbity (<http://molprobity.biochem.duke.edu/> )
3. Сервер EDS (<http://eds.bmc.uu.se/eds/index.html> )
4. Программа PyMol v-1.7.2.1 (<http://www.pymol.org/>)