

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

Отчёт по качеству расшифровки структуры алкогольдегидрогеназы из
Pyrobaculum aerophilum (PDB ID 4JBI)
методом рентгеноструктурного анализа

Буянова Софья,
студентка 4 курса

Москва, 2016г

Аннотация

В отчёте рассмотрена общая информация о ферменте алкогольдегидрогеназа, некоторые индикаторы качества модели структуры 4JBI фермента, а также проведён анализ маргинальных аминокислотных остатков.

Введение

Актуальность

Архея *Ruobaculum aerophilum*, из которой был выделен описываемый фермент, принадлежит специфической группе архей, называемой *Crenarchaeota*. Они метаболически "универсальны", так как могут использовать органические и неорганические субстраты как для аэробного, так и анаэробного дыхания. Изучение метаболических путей и таких процессов, как транскрипция, трансляция архей (причем именно *crenarchaea* являются предками современных эукариот), занимает центральное место в понимании эволюции эукариотической клетки. Изучение белков архей поможет нам больше знать об эволюции эукариот¹.

Так же для этой алкогольдегидрогеназы показана ее термостабильность, поэтому ее структура может пролить свет на стабильность белков.

Общие сведения

Алькогольдегидрогеназы широко распространены во всех живых организмах (археи, бактерии, грибы, растения и животные). Они принадлежат к дегидрогеназам /редуктазы надсемейства (МДР), класс аферментов, ответственных за катализ всех биологических окислительно-восстановительных реакции. Алькогольдегидрогеназы бывают трех классов, в зависимости от длины цепи. Изучаемый мной фермент *RuAeADHII* принадлежит к 1 классу - средних по длине.

У алкогольдегидрогеназ известно несколько доменов - НАДФ-связывающий домен, а также каталитический домен, который связывает ион цинка с каталитической функцией и структурный цинк связывающий сайт. Короткоцепочечные дегидрогеназы не содержат цинк-связывающего домена, зато у них есть нуклеотидсвязывающий домен, чья укладка известна, как укладка Россмана. Интересно, что N-концевой участок каталитического домена похож на GroEL-связывающую мобильную петлю GroES шаперонинов из *Escherichia coli*. Изображение фермента приведено на рис. 1 и рис. 2.

Алькогольдегидрогеназа катализирует реакцию:



В настоящей работе рассматривалась модель структуры фермента 4JBI². Структура этого фермента была получена Vitale et. al., потому что предыдущие исследования *RuAeADHII* показали, что у нее отсутствует активность для большинства стандартных субстратов, используемых для тестирования активности алкогольдегидрогена. Более того, выравнивание аминокислотной последовательности *RuAeADHII* с другими, хорошо изученными, ADHs, показали, что у *RuAeADHII* отсутствуют ключевые остатки, участвующие в каталитически важном связывании Zn^{2+} , а так же остатки, напрямую участвующие в катализе.

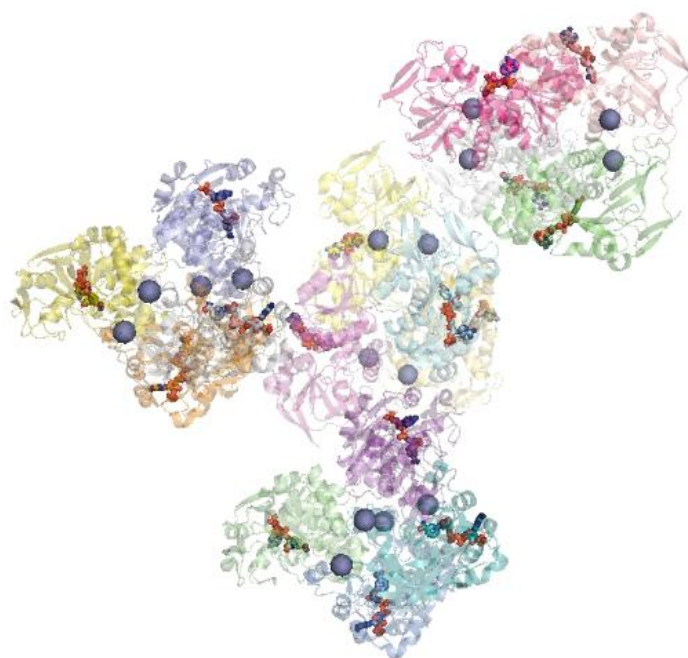


Рис. 1. Изображение фермента алкогольдегидрогеназы из археи *Pyrobaculum aerophilum*. Цепи показаны разными цветами, ионы цинка - синими сферами, NADP - в модели sticks, покрашенные по типу атомов.

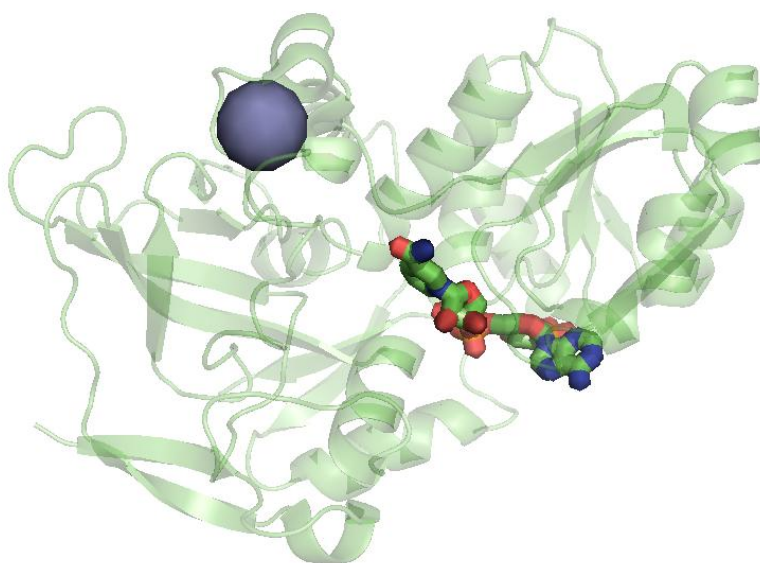


Рис. 2. Изображение цепи А фермента алкогольдегидрогеназы из археи *Pyrobaculum aerophilum*. Ионы цинка изображены синими сферами, NADP - в модели sticks, покрашенные по типу атомов.

Результаты и обсуждения

Общая информация о модели

Модель структуры 4JBI была получена в 2013 году группой Annalisa Vitale, Natasha Thorne, Scott Lovell, Kevin P. Bataille, Xin Hu, Min Shen, Sabato D'Auria, Douglas S. Auld². Общая информация о модели систематизирована в таблице 1.

Таблица 1. Общая информация о модели 4JBI

| | |
|--|---|
| Разрешение | 2.35 Å |
| Число измеренных рефлексов | 254725 |
| Число использованных рефлексов | 254638 |
| Метод решения фазовой проблемы | MOLECULAR REPLACEMENT - молекулярное замещение |
| Кристаллографическая группа | P 1 21 1 |
| Полнота набора структурных факторов | 99.89 |
| Диапазон разрешений структурных факторов | 2.35-44.589 |
| Параметры кристаллографической ячейки | Длина (Å): a = 138.13 b = 169.26 c = 149.15 Угол (°): $\alpha = 90$ $\beta = 116.25$ $\gamma = 90$ |

Индикаторы качества модели в целом и по отдельным аминокислотным остаткам

Для того, чтобы понимать, насколько построенная модель соответствует действительности, необходимо всегда обращать внимание на индикаторы качества построенной модели. Это можно сделать с помощью сервиса MolProbity³. Показатели качества модели для 4JBI приведены в табл. 2.

Таблица 2. Индикаторы качества модели

| | Значение | Процент | Пороговый уровень |
|---|----------|---------|--|
| R-фактор | 0.1639 | | |
| R_free | 0.2133 | | <20% |
| RSR | 0.189 | 18.9 | <20% |
| Clashscore | 2.61 | | 100 th percentile* (N = 335, 2.35Å ± 0.25Å) |
| Clash score - это число недопустимых наложений (> 0.4 Å) атомов на 1000 | | | |

| | | | |
|--|-------------|--------|---|
| Маргиналы по карте Рамачандрана | 10 | 0.19% | Goal: <0.05% |
| Число остатков в предпочитаемой области карты | 5160 | 97.43% | Goal: >98% |
| Валентные углы, существенно отклоняющиеся от теории | 167 / 55353 | 0.30% | Goal: <0.1% |
| Cis пролины: | 48 / 304 | 15.79% | Expected: ≤ 1 per chain, or $\leq 5\%$ |
| Ковалентные связи, существенно отклоняющиеся от теории (twisted peptide) | 1 / 5312 | 0.02% | Goal: 0 |

Для данной структуры R-фактор равен 0.1639. Rfree равняется 0.2133, а (R-free - R) равняется 0.0494. Хорошие значения: $R_{free} < 20\%$, в нашем случае это значение немного превышает. $(R_{free} - R) > 10\%$ могло бы настораживать в отношении переоптимизации (overfitting), но в нашей случае у нас значение $< 10\%$, значит у нас нет переоптимизации модели.

Real Space R (RSR) позволяет найти маргиналов по соответствию боковой цепи “экспериментальной” электронной плотности. Сумма берется по узлам пространственной решетки в окружении группы атомов. Хорошие значения: $RSR < 20\%$. Так как сервер EDS анализирует цепи поотдельности, то рассмотрим только цепь А. Гистограмма значений RSR для всех остатков цепи А структуры 4JBI приведена на рисунке 3.

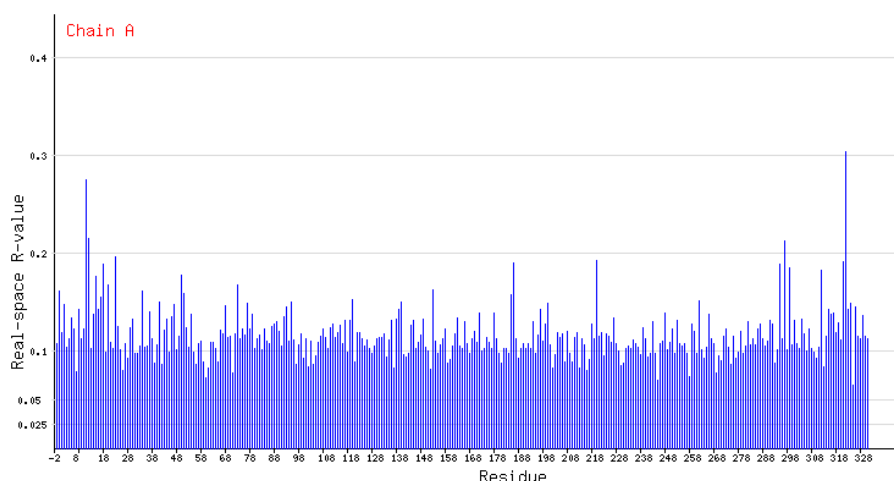


Рис. 3. Гистограмма значений RSR для всех остатков цепи А структуры 4JBI. По оси Y отложены значения RSR, по оси X отложены номера остатков в цепи А. Значения, которые превышают 0.2 считаются маргинальными.

Z-score - относительный R-фактор. Z-score вычисляется относительно всех остатков такого же типа из структур с примерно таким же разрешением. Высокие положительные значения $Z > 2$ свидетельствуют о том, что остаток плохо вписан в электронную плотность по сравнению с другими структурами с тем же разрешением (=>маргинал). Если RSR плохой, а RSR-Z – хороший, то значит координаты атомов расшифрованы плохо, но не хуже, чем в других подобных структурах. Значения z-score для цепи А приведены на рисунке 4.

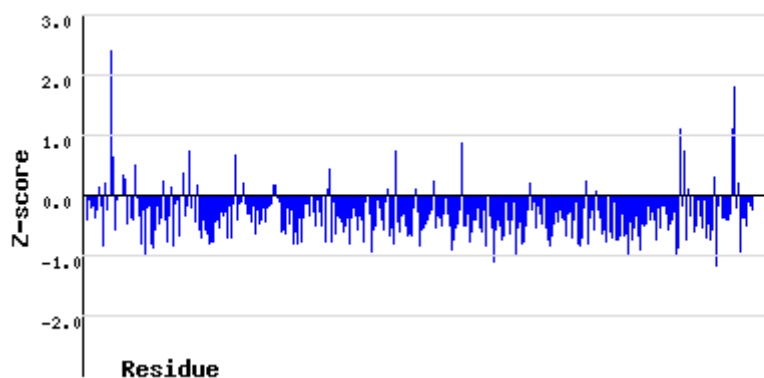


Рисунок 4. Значения z-score для цепи А. По оси Y отложены значения z-score, по оси X отложены номера остатков в цепи А. Значения, которые превышают 2.0 считаются маргинальными.

Также маргинальность остатка можно оценивать с помощью карты Рамачандрана, на горизонтальной оси которой откладывается угол Phi, а на вертикальной - угол Psi. Карты Рамачандрана, полученные с помощью сервиса MolProbity, вы можете найти на рис. 5.

Из рис. 5 видно, что в модели 4JVI есть несколько маргинальных остатков, например 77 валин или 111 глицин. Такое представление, как карта Рамачандрана принято для визуальной оценки аминокислотных остатков на предмет их маргинальности или преобладания торсионных углов. Для того, чтобы рассмотреть каждый остаток в отдельности, можно воспользоваться текстовой выдачей MolProbity.

MolProbity Ramachandran analysis

4jbi.pdb, model 1

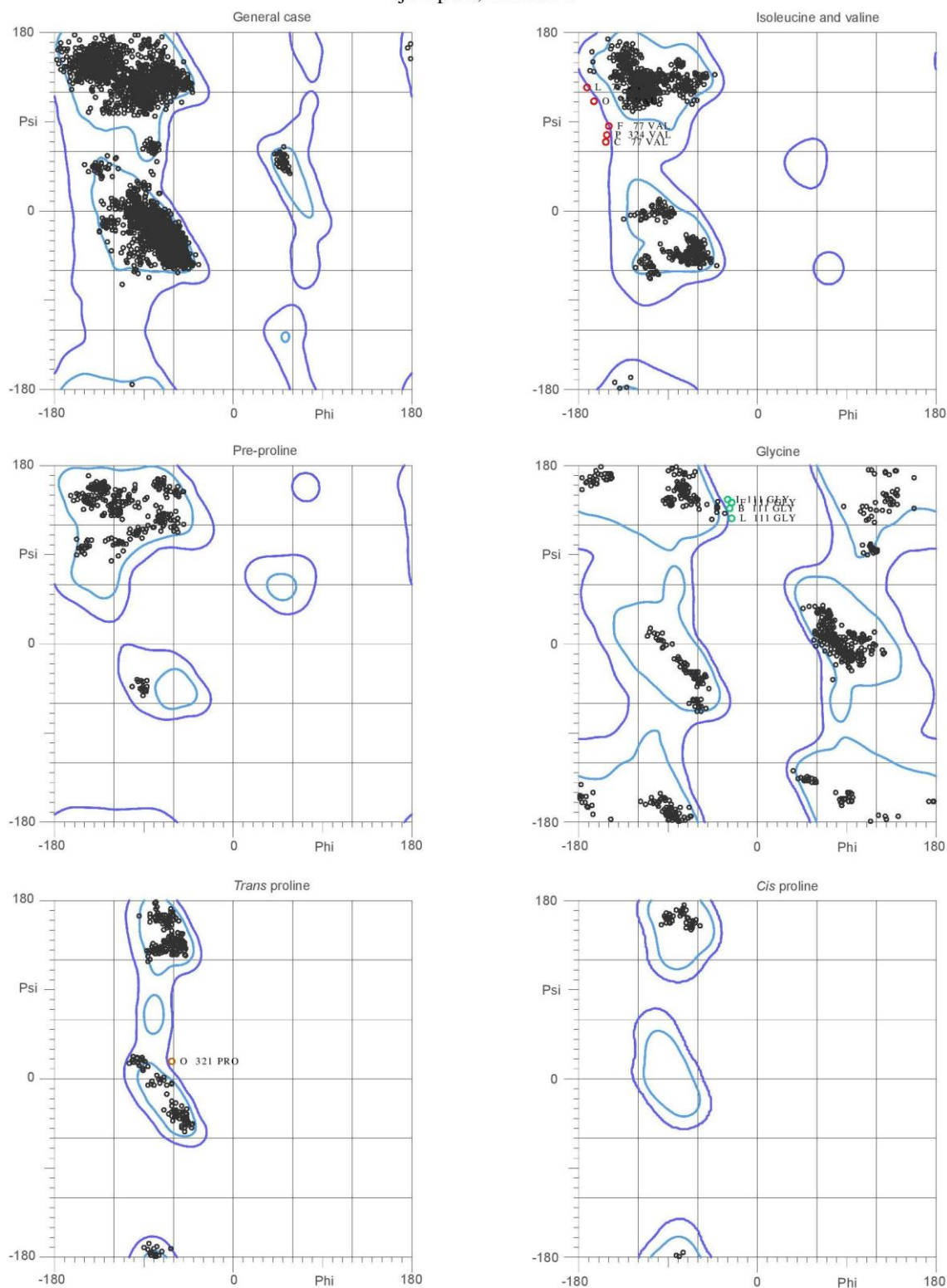


Рис.5. - Карты Рамачандрана для всех аминокислотных остатков, полученные с помощью сервиса MolProbity. Отдельно показаны карты Рамачандрана для изолейцина и валина, глицина, пре-пролина, транс и цис-пролина. Маргинальные по карте Рамачандрана остатки на картах для отдельных аминокислот выделены цветом и подписаны.

Анализ маргинальных остатков

Рассмотрим теперь отдельные аминокислотные остатки на предмет их маргинальности. Список маргинальных по различным критерия остатков приведен в табл. 3

Таблица 3. Список маргинальных остатков. Ротамеры - остатки, маргинальные по положению боковой цепи

| Аминокислота | Цепь | Номер остатка | RSR | Ramachandran outlier | Clashscore >0.4 | Twisted peptide | Ротамер |
|--------------|------|---------------|-------|----------------------|-----------------|-----------------|---------|
| ASN | P | 304 | 0.342 | | | | |
| GLU | K | 20 | 0.338 | | | | |
| VAL | P | 38 | 0.294 | | 0.56Å | | |
| ARG | P | 2 | 0.267 | | | | |
| VAL | P | 324 | 0.265 | Да | | | |
| VAL | O | 324 | 0.210 | | | Да | |
| SER | P | 320 | 0.128 | | | | Да |

Будем работать с маргинальными остатками приведенными в таблице 3.С помощью следующего скрипта рассчитаем функцию электронной плотности и рассмотрим каждый их указанных в таблице 3 остатков.

```
load 4JBI.pdb
```

```
load 4jbi.omap, 4jbi_map
```

```
hide lines
```

```
hide (resn HOH)
```

```
isomesh 1_0_map, 4jbi_map, 1.0, backbone, 2.5, carve=2.5
```

```
bg_color white
```

```
set mesh_color, grey50
```

```
zoom resi 304 and chain P
```

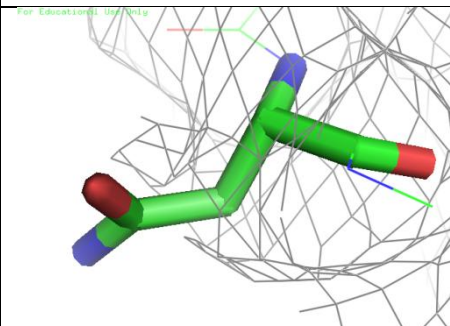
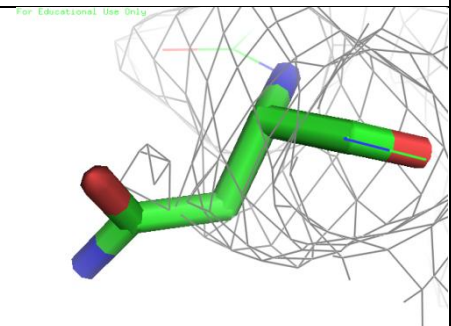
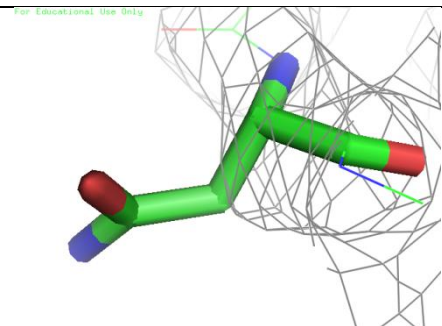
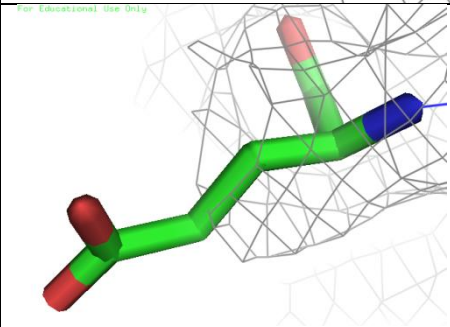
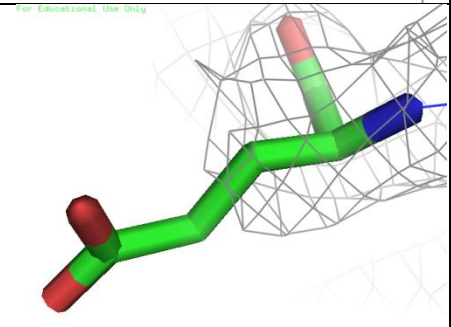
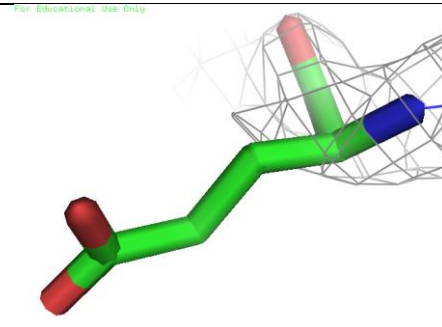
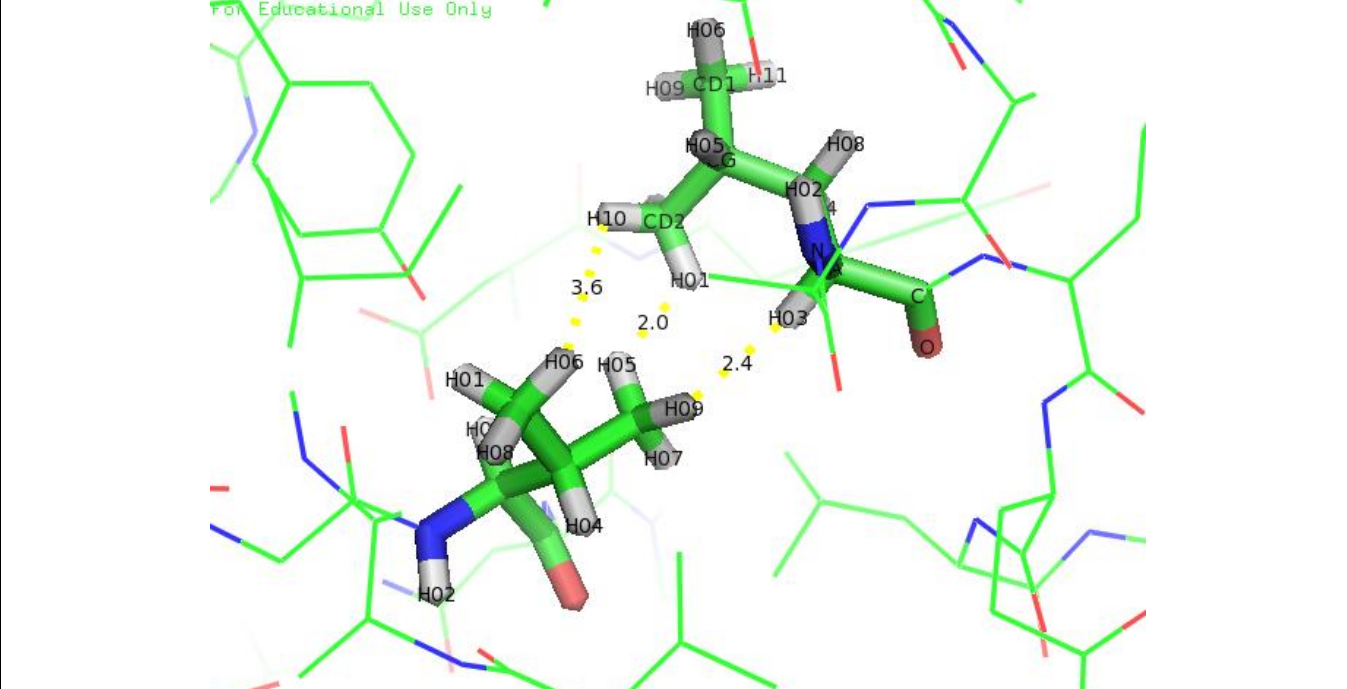
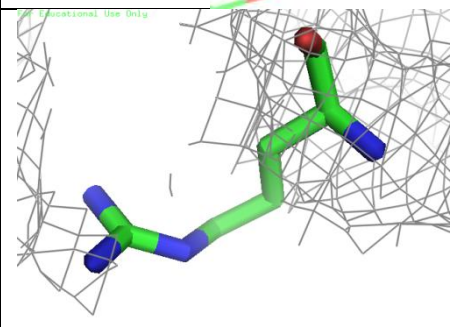
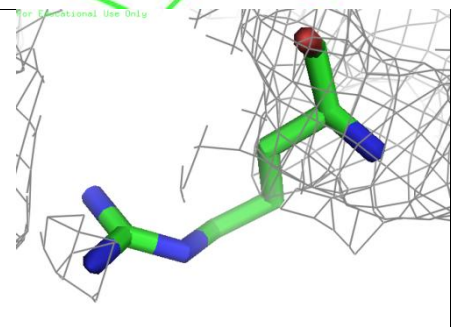
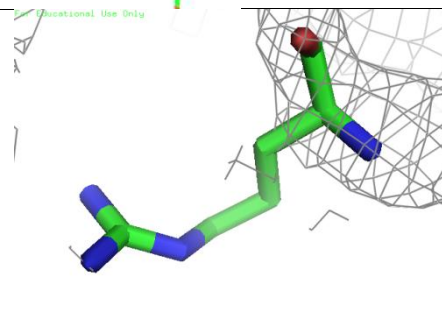
Изображение электронной плотности при разных уровнях подрезки (0.25, 0.5, 1) для остатков представлены в таблице 4.

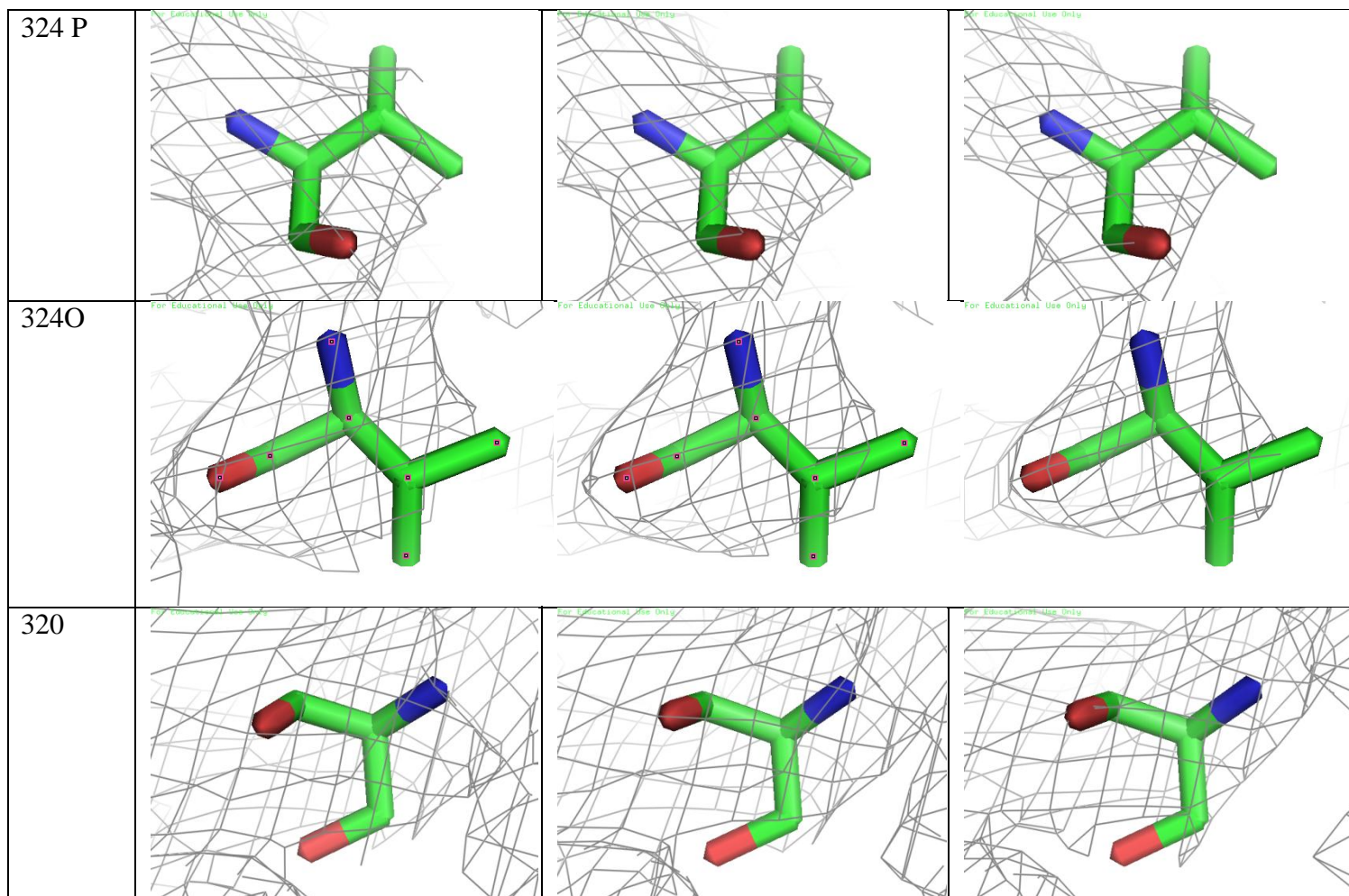
При этом непонятно, почему программа MolProbity вдруг посчитала, что между 38 и 315 остатками цепи P есть стерическое затруднение (3 столбец табл. 4), расстояния между атомами вполне приемлемы, и больше, чем их Ван-дер-Ваальсовы радиусы.

Остальные остатки довольно логично классифицированы, как маргинальные. Как правило, совершенно не понятно не только, как расположен в белке этот остаток, но и какой там остаток вообще

Таблица 4

Изображение электронной плотности маргинальных остатков

| Остаток | Подрезка | | |
|---------|--|--|---|
| | 0.25 | 0.5 | 1 |
| 304 |  |  |  |
| 20 |  |  |  |
| 38 |  | | |
| 2 |  |  |  |



Сравнение модели из PDB с моделью из PDB_redo

Было проведено сравнение модели 4JBI из PDB с оптимизированной моделью из PDB_REDO. Значения некоторых параметров для оценки качества двух моделей приведены в таблице 9. Значения R-фактора ухудшилось, а R-free улучшились, и разница их также уменьшилась. Значит, новая модель чуть лучше подогнана под экспериментальные данные. Но при этом появились полярные атомы, которые не образуют водородных связей. Так же по данным PDB_redo значительно улучшились показатели для 25 остатков и при этом ухудшились для 214. Полную информацию можете посмотреть на странице http://www.cmbi.ru.nl/pdb_redo/jb/4jbi/index.html.

| | PDB | PDB_redo |
|------------------------------------|--------|----------|
| R | 0.1610 | 0.1873 |
| R-free | 0.2130 | 0.2053 |
| R-free-R | 0.052 | 0.018 |
| Несвязанные доноры/акцепторы | 0 | 303 |
| Количество стерических затруднений | 394 | 169 |

Заключение

4JBI - это модель среднего разрешения. В целом качество её разрешения и можно оценить как от среднего до плохого на основании индикаторов качества. Почти все показатели, указанные в отчете превышают пороговые значение. Хотя мне кажется это вполне логично для такого большого белка, как алкогольдегидрогеназа. Сервер PDB_REDO улучшил такие показатели модели, как R-фактор и Rfree, но при этом появилась куча других проблем. Да и в целом улучшение показателей для 25 остатков и при этом ухудшение для 214 остатков говорит о посредственном улучшении модели.

1. Vitale, A., Rosso, F., Barbarisi, A., Labella, T. & Auria, S. D. Properties and evolution of an alcohol dehydrogenase from the Crenarchaeota *Pyrobaculum aerophilum*. *Gene* 461, 26–31 (2010).
2. Vitale, A. *et al.* Physicochemical Characterization of a Thermostable Alcohol Dehydrogenase from *Pyrobaculum aerophilum*. 8, (2013).
3. <http://molprobit.biochem.duke.edu/index.php>
4. EDS <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=4jbi>
5. PDB <http://www.rcsb.org/pdb/explore/explore.do?structureId=4jbi>
6. PDB_redo www.cmbi.ru.nl/pdb_redo/ai/4jbi/index.html