

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

им. М.В. ЛОМОНОСОВА

Факультет биоинженерии и биоинформатики

Отчет по качеству расшифровки методом PCA структуры ЗВМХ

Студента 4 курса

Бикметова Д.В.

МОСКВА, 2016

Оглавление

Аннотация	3
Введение	3
Результаты и обсуждение.....	3
1. Общая информация о модели	3
2. Значения индикаторов качества модели в целом	4
3. Список маргинальных остатков.....	8
4. Анализ отдельных маргинальных остатков	8
A: HIS 222	8
A: ASN 312.....	9
B: LYS 373	10
B: ALA 26	11
A: ARG 640.....	11
5. Сравнение модели из PDB с моделью из PDB_redo.....	12
Заключение	14
Список литературы	15

Аннотация

Данная работа посвящена изучению качества трехмерной модели с PDB-идентификатором 3VMX, которая принадлежит бета-гексаминидазе *Bacillus subtilis*. Здесь содержится краткое описание исследуемого белка и его модели. Кроме того, рассматриваются различные индикаторы качества модели. Особое внимание уделено маргинальным остаткам. В отчете приведен список некоторых из них, а отдельные маргинальные остатки проанализированы более подробно. Также имеется сравнение структуры 3VMX из PDB и PDB_redo.

Введение

Бета-гексаминидаза (EC 3.2.1.52)-фермент, катализирующий гидролиз гликозидной связи между последним N-ацетил-D-гексозамином и остатком уксусной кислоты в молекуле поли-N-ацетил- β -D-гексозамина. Эта реакция у бактерий имеет место в пути рециркуляции пептидогликана. Таким образом, рассматриваемый фермент играет важную роль в биосинтезе клеточной стенки. Белок локализован в клеточной мембране. Фермент функционирует в виде мономера. Количество аминокислотных остатков - 642, масса составляет 70580 Да. Белок содержит два домена. Фермент принадлежит к 3 семейству гликозидаз.

3VMX- одна из первых полученных моделей двухдоменной бета-N-ацетилглюкозамидазы. Расшифровка структуры привела к открытию механизма действия таких ферментов, Asp-His диады.

Результаты и обсуждение

1. Общая информация о модели

Модель 3VMX содержит 5 различных типов молекул:

1. Две идентичные белковые цепи, обозначенные А и В. Каждая цепь содержит 617 аминокислотных остатков.
2. Ацетат ион, 3 молекулы в цепи А и 4 в В.
3. Ион натрия, по 2 в каждой цепи.

4. 1-этоксис-2-(2-этоксисэтоксис)этан (P4G), 2 молекулы в цепи А и 1 в цепи В
5. Вода, 739 и 649 молекул в цепях А и В соответственно.

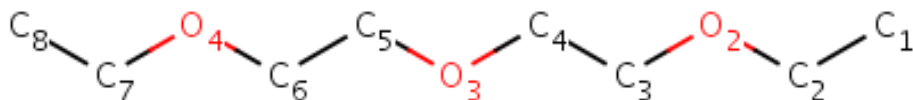


Рис.1. Структурная формула лиганда 1-этоксис-2-(2-этоксисэтоксис)этан (P4G).

Модель 3VMX была размещена в базе данных PDB 13 декабря 2007 года. Автор модели-Fischer, S. Фазовая проблема решалась с помощью молекулярного замещения. В качестве отправной точки служили известные структуры с идентификаторами 1EX1, 1EX, 1J8V, 1LQ2 и 1X38. Ниже представлены характеристики экспериментальных данных.

Таблица 1. Характеристики экспериментальных данных.

Диапазон разрешений структурных факторов	42.83 - 1.40 Å
Разрешение	1.40 Å
Полнота набора структурных факторов	90.9 %
Группа симметрии	P 1
Параметры кристаллографической ячейки	a=58.38 Å, b=73.08 Å, c=83.57 Å alpha=79.79, beta=69.61, gamma=88.25
Число использованных рефлексов	227962 (из 250923 измеренных)

В асимметрической ячейке некристаллографической симметрии не наблюдается.

2. Значения индикаторов качества модели в целом

Значения R-фактора для данной модели следующие:

R-Free: 0.166

R: 0.127

Это очень хорошие показатели, свидетельствующие о хорошем качестве модели.

Структура была проанализирована с помощью сервиса MolProbity. Для анализа к структуре были добавлены атомы водорода и выполнены предложенные сервисом инверсии остатков:

A:312 ASN

B:68 GLN

B:374 ASN

Также была построена карта Рамачандрана, представленная ниже (Рис.2.). Согласно полученным данным, все остатки лежат в разрешенных областях и 97,5% лежит в предпочтительных. MolProbity также предоставляет некоторые характеристики качества структуры, указанные в таблице ниже.

Таблица 2. Характеристики качества модели по результатам сервиса MolProbity.

All-Atom Contacts	Clashscore, all atoms:	2.1		99 th percentile * (N=480, 1.40Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	8	0.76%	Goal: <0.3%
	Favored rotamers	1020	97.14%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	1210	97.50%	Goal: >98%
	MolProbity score [^]	1.08		99 th percentile * (N=3363, 1.40Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	1 / 9807	0.01%	Goal: 0%
	Bad angles:	7 / 13249	0.05%	Goal: <0.1%

Peptide	Cis Prolines:	8 / 64	12.50%	Expected: ≤ 1 per chain, or $\leq 5\%$
Omegas	Cis nonProlines:	2 / 1176	0.17%	Goal: $< 0.05\%$

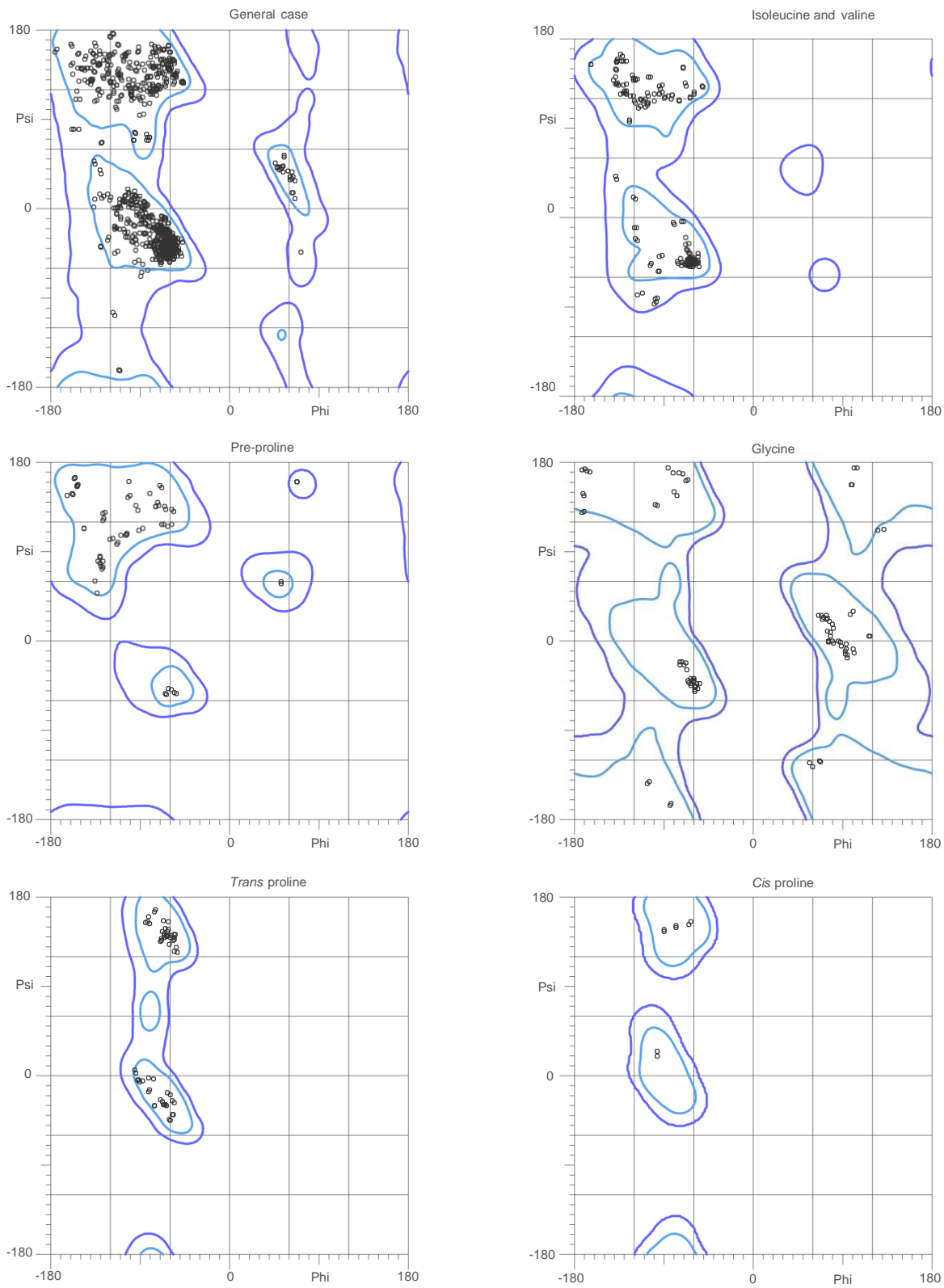


Рис.2. Карта Рамачандрана для структуры 3BMX. Получена при помощи сервиса MolProbity.

3. Список маргинальных остатков

На основе данных MolProbity, EDS и Full wwPDB X-ray Structure Validation Report была составлена таблица нескольких маргинальных остатков структуры ЗВМХ.

Таблица 3. Некоторые маргинальные остатки. Жирным начертанием выделены остатки, более подробно рассмотренные в следующем разделе.

Остаток	Причина маргинальности
A 222 HIS B 222 HIS	Cis-конформация
B 268 THR	Выброс по длине связи (worst is CB--OG1: 4.5 σ) и углу (worst is OG1-CB-CG2: 4.9 σ)
A 495 VAL A 560 MET A 638 THR A 640 ARG B 374 ASN B 495 VAL B 554 ASN B 577[A] LEU B 577[B] LEU	Отклонение боковой цепи от ротамеров
A 643 ACT	Лиганд ацетат, длина связи
B LYS 373 B ALA 26	Высокий RSR

4. Анализ отдельных маргинальных остатков

Для более подробного анализа были выбраны пять различных маргинальных остатков. При упоминании выше они выделены жирным

A: HIS 222

Данный остаток в обоих цепях белка присутствует в Cis-конформации. Это ненормальное состояние для гистидина. Однако, у этого остатка хороший показатель RSR (менее 0,1). Кроме того, он отлично вписывается в электронную плотность. Ниже

приведено изображение восстановленной электронной плотности при уровне подрезки 2. Таким образом, по всей видимости, Cis-конформация данного остатка является особенностью структуры исследуемого белка, а не недостатком модели.

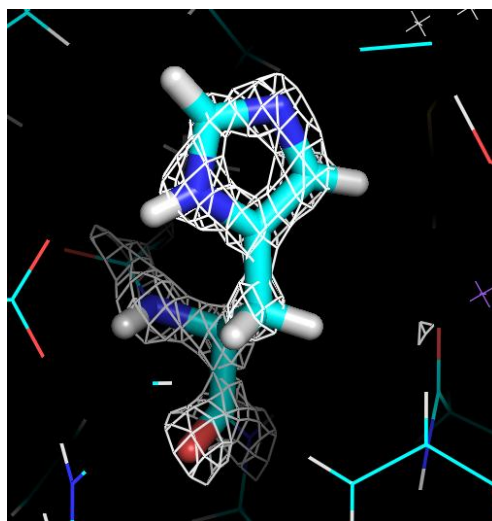


Рис.3. Остаток A:HIS 222. Белым изображена ЭП, уровень подрезки 2. Раскраска в соответствии с типом атомов.

A: ASN 312

Это один из тех остатков, боковую цепь которых сервис MolProbity предложил инвертировать. Его RSR достаточно хорош и составляет 0.115. Других отклонений от нормы не наблюдается. Если визуализировать функцию его электронной плотности видно, что по форме поверхности не удастся определить, должна ли быть инверсия. Водородных связей, стабилизирующих этот остаток в структуре обнаружено не было. К сожалению, по данному набору экспериментальных данных установить правомерность инверсии затруднительно.

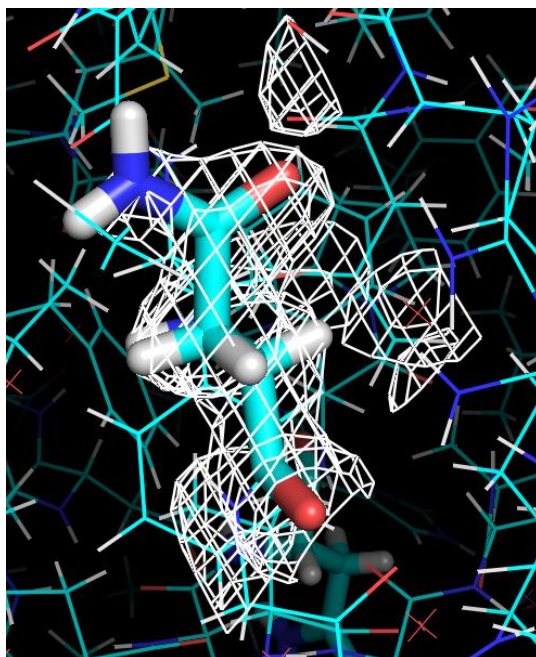


Рис.4. Остаток А: ASN 312. Белым изображена ЭП, уровень подрезки 1. Остаток инвертирован сервисом MolProbity. Раскраска в соответствии с типом атомов.

В: LYS 373

Сервис MolProbity не дает информации по данному остатку. Однако он привлекает внимание тем, что имеет самый большой во все модели RSR, равный 0,41. Визуализация электронной плотности этого остатка затруднительна даже при небольших уровнях подрезки. Даже на уровне подрезки 0,5, облако ЭП маленькое и разорванное. Возможно, координаты этого остатка ошибочны, однако при таком наборе экспериментальных данных установить его точные координаты невозможно.

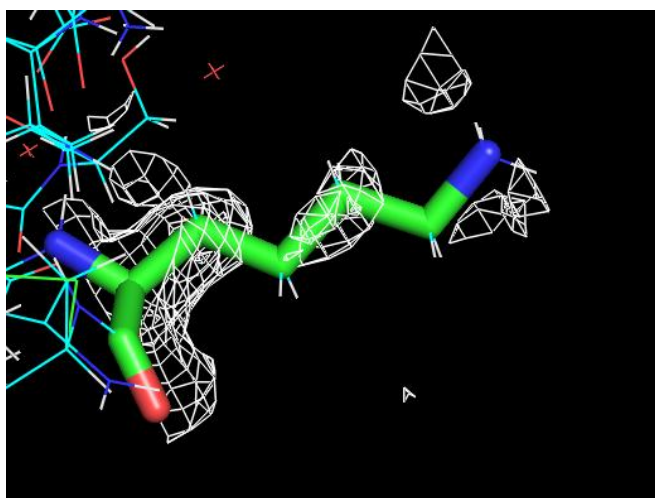


Рис.5. Остаток В: LYS 373. Белым изображена ЭП, уровень подрезки 0,5. Раскраска в соответствии с типом атомов.

В: ALA 26

Этот остаток также отличается высоким RSR (0.301). Кроме того, это первый по счету остаток в модели. MolProbity также не дал информации о нем. При визуализации ЭП возникли те же проблемы, как и в случае предыдущего рассмотренного аминокислотного остатка. Тем не менее, при визуализации с уровнем подрезки 0,5 видно, что атомы плохо вписываются в электронную плотность. Возможно, это свидетельствует об ошибке в расшифровке, хотя судить по визуализации при подрезке 0,5 сложно. Таким образом, определить правильные координаты остатка также трудно при данном наборе экспериментальных данных. Вероятно, трудности с экспериментальным определением положения этого остатка в пространстве вызваны тем обстоятельством, что он первый в структуре. Например, эксперименту могут мешать сильные его колебания. В пользу этой точки зрения говорит высокий температурный фактор 38.

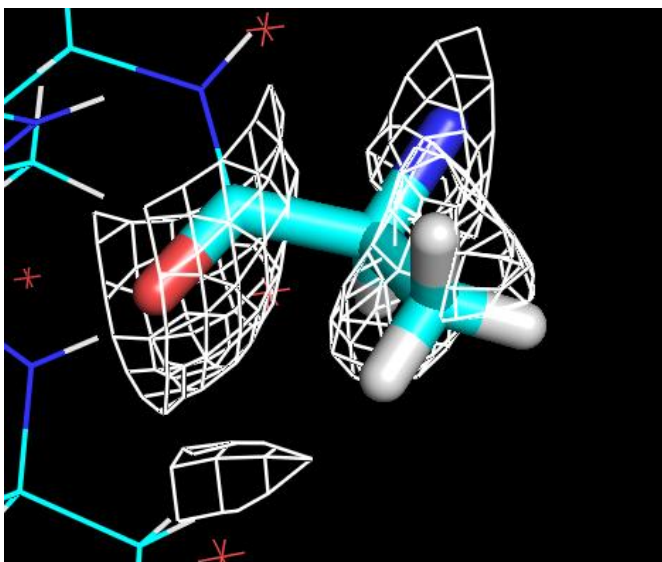


Рис.5. Остаток В: ALA 26. Белым изображена ЭП, уровень подрезки 0,5. Раскраска в соответствии с типом атомов.

А: ARG 640

Остаток имеет допустимый пространственный R фактор 0,186. Он был выбран маргинальным по следующим причинам:

- остаток на 0.68\AA перекрывается с остатком А: THR 638.
- Углы связей боковой цепи отклоняются от ротационных (chi angles: 105.3,218.1,290.1,215.1)

Как показала визуализация, этот остаток очень плохо вписывается в электронную плотность. Скорее всего, его координаты расшифрованы неверно. Возможно, это вызвано тем, что он расположен очень близко к концу структуры, где расшифровка всегда хуже. Может быть, имеет место вклад колебаний. Температурный фактор всех остатков конца структуры несколько выше, чем у остальных.

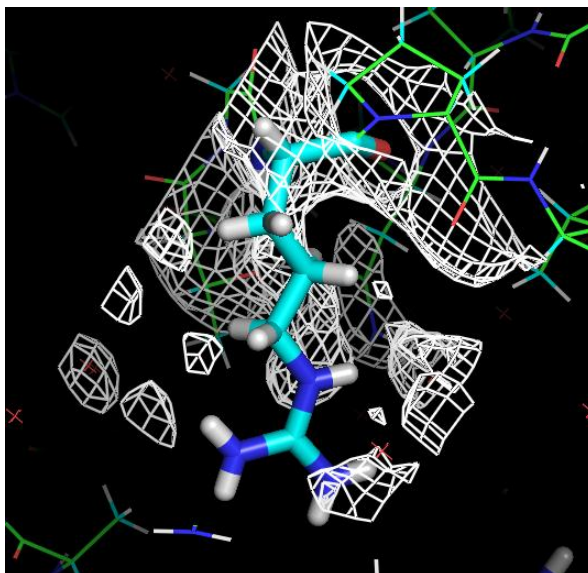


Рис.5. Остаток A: ARG 640. Белым изображена ЭП, уровень подрезки 0,5. Раскраска в соответствии с типом атомов.

5. Сравнение модели из PDB с моделью из PDB_redo.

Было проведено сравнение 3BMX с моделью, полученной из PDB_redo. В таблице ниже представлено сопоставление общих параметров моделей. В то время как R-free почти не изменился, выросли значения R и Z факторов.

Таблица 4. Сравнение общих показателей качества моделей из PDB и PDB_redo.

	PDB	PDB_redo
R	0.1282	0.1372
R-free	0.1650	0.1649
Z(R-free)	11.80	20.40

Структура из PDB_redo также была проанализирована с помощью сервиса MolProbity. Ниже представлены полученные результаты. Как видно из таблицы, многие показатели

почти не изменились. Отметим, что уменьшилось число перекрывающихся атомов, Clashscore сократился вдвое. Также незначительно сократилось число остатков, имеющих отклонения боковой цепи от ротамеров. Но в то же время, появился еще один остаток с выпадающей длиной связи.

All-Atom Contacts	Clashscore, all atoms:	1.12		99 th percentile* (N=480, 1.40Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	8	0.76%	Goal: <0.3%
	Favored rotamers	1028	97.90%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	1211	97.58%	Goal: >98%
	MolProbity score [^]	0.91		100 th percentile* (N=3363, 1.40Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.17%	Goal: 0
	Bad bonds:	2 / 9807	0.02%	Goal: 0%
Bad angles:	2 / 13249	0.02%	Goal: <0.1%	
Peptide Omeegas	Cis Prolines:	8 / 64	12.50%	Expected: ≤1 per chain, or ≤5%
	Cis nonProlines:	2 / 1176	0.17%	Goal: <0.05%

Средствами PyMol данные модели были совмещены. В целом, укладка не изменилась. Имеются лишь небольшие изменения, касающиеся отдельных остатков. При этом, остатки, поменявшие свои координаты, расположены, в основном, на концах элементов вторичной структуры.

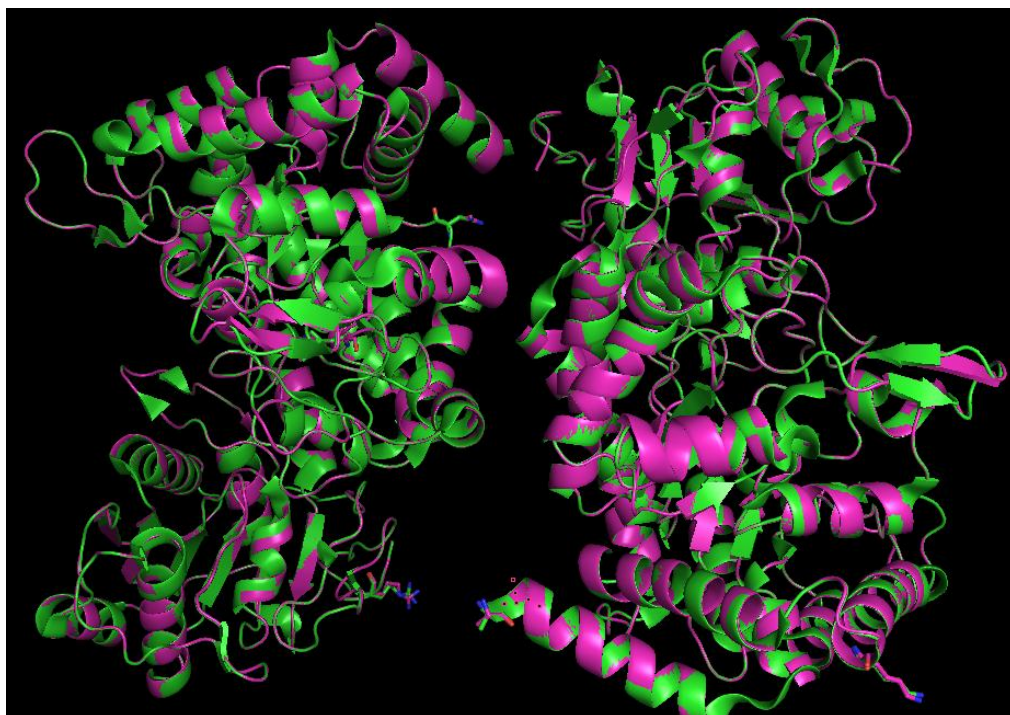


Рис. 6. Совмещение структур: зеленая-PDB, розовая-PDB_redo.

Отметим, что рассмотренный ранее неправильно расшифрованный остаток A: ARG 640 в новой модели изменил свое положение и теперь лучше соответствует экспериментальным данным.

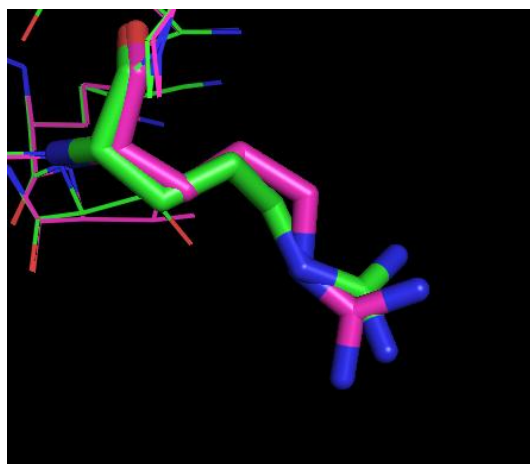


Рис.5. Совмещение остатков A: ARG 640: зеленый-PDB, розовый-PDB_redo.

Заключение

Исходя из проведенной работы можно заключить, что модель бета-гексаминидазы ЗВМХ весьма достойного качества. Она имеет отличные общие показатели и малое число

отклонений от правильной геометрии геометрии молекул. Некоторые из таких отклонений при этом являются структурными особенностями данного фермента. Расшифровка модели при данном наборе исходных данных очень хорошая, хотя имеет незначительные недочеты. Оптимизация модели заново в рамках проекта PDB_redo местами их исправляет. Модель из PDB_redo, тем не менее, очень слабо превосходит по качеству и без того довольно-таки хорошую исходную.

Список литературы

- [1] Litzinger S, Fischer S, Polzer P, Diederichs K, Welte W, Mayer C. Structural and Kinetic Analysis of Bacillus subtilis N-Acetylglucosaminidase Reveals a Unique Asp-His Dyad Mechanism. *The Journal of Biological Chemistry*. 2010;285(46):35675-35684. doi:10.1074/jbc.M110.131037.
- [2] Litzinger S, Duckworth A, Nitzsche K, Risinger C, Wittmann V, Mayer C. Muropeptide Rescue in Bacillus subtilis Involves Sequential Hydrolysis by β -N-Acetylglucosaminidase and N-Acetylmuramyl-L-Alanine Amidase . *Journal of Bacteriology*. 2010;192(12):3132-3143. doi:10.1128/JB.01256-09.
- [3] "RCSB PDB for 1XTM." [Online]. Available:
<http://www.rcsb.org/pdb/explore/explore.do?structureId=3bmx>
- [4] "EDS: 3bmx." [Online]. Available: <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=3bmx>
- [5] "Full wwPDB X-ray Structure Validation Report." [Online]. Available:
http://ftp.wwpdb.org/pub/pdb/validation_reports/xt/3bmx/3bmx_full_validation.pdf
- [6] "3bmx - PDB_REDO optimisation results." [Online]. Available:
http://www.cmbi.ru.nl/pdb_redo/xt/3bmx/index.html
- [7] "MolProbity." [Online]. Available:
<http://molprobity.biochem.duke.edu>