

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

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Оценка качества расшифровки структуры белка гидроксинитриллиазы
Baliospermum montanum (3WWP) методом рентгеноструктурного анализа**

***Отчет студентки 4-го курса
Байкузиной Полины Георгиевны***

**Москва
2017 г.**

АННОТАЦИЯ

В работе был проведен анализ качества расшифровки структуры белка гидроксинитриллизасы балиоспермума горного (PDB ID 3WWP) методом рентгеноструктурного анализа, были рассмотрены как показатели качества структуры в целом, так и для отдельных остатков.

ВВЕДЕНИЕ

Гидроксинитриллизасы (HNL) – это ферменты класса лиаз, которые в общем случае катализируют обратимую реакцию отщепления синильной кислоты от циангидринов с образованием кетонов или альдегидов (рис.1).

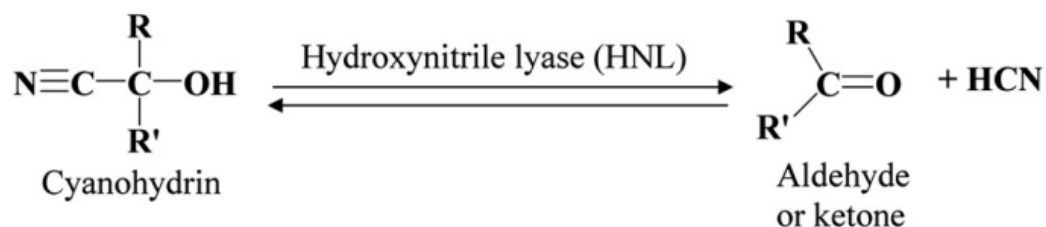


Рисунок 1. Реакция, катализируемая гидроксинитриллиазами [1].

Гидроксинитриллизасы встречаются в основном у растений, а также у бактерий и некоторых членистоногих. С помощью цианогенеза растения защищаются от животных, поедающих растения, и патогенных микроорганизмов.

Гидроксинитриллизасы обладают исключительной энантиоселективностью, различают R- и S-селективные HNL. Именно благодаря этому свойству гидроксинитриллизасы широко используются в биотехнологии. В данной работе рассматривается S-селективная гидроксинитриллиаза.

S-HNL (EC 4.1.2.47) имеют α/β -гидролазную укладку и представляют собой димер, состоящий из идентичных субъединиц (рис.2).

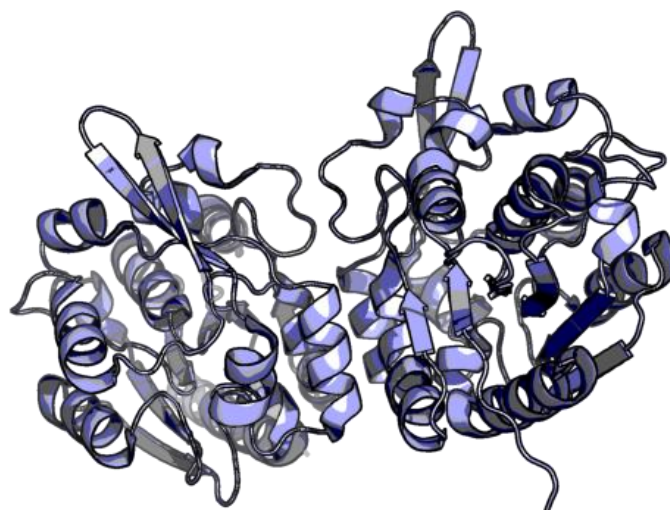


Рисунок 2. Гидроксинитриллиаза, структура 3WWP.

Структура 3WWP была получена японскими учеными Shogo Nakano, Mohammad Dadashipour, Yasuhisa Asano в 2014 году. Помимо данной структуры была получена структура 3WVO с разрешением 2.55Å. Гидроксинитриллиазы могут использовать большое количество субстратов, при этом сохраняя стереоселективность в большинстве случаев. В своем исследовании авторы анализировали структуру, динамику активного сайта и обнаружили, что гидрофобные остатки в активном сайте образуют гидрофобные контакты с субстратом и боковые цепи этих остатков подвижны. Авторы пришли к выводу, что подвижность и размещение гидрофобных остатков важны для широкой субстратной специфичности S-гидроксинитриллиз [1].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Общая информация о модели

Структура 3WWP была получена методом рентгеноструктурного анализа и опубликована на сайте PDB в 2014 году учеными Shogo Nakano, Mohammad Dadashipour, Yasuhisa Asano, также была опубликована статья «Structural and functional analysis of hydroxynitrile lyase from *Baliospermum montanum* with crystal structure, molecular dynamics and enzyme kinetics», содержащая информацию об исследовании структуры.

Ассиметрическая единица 3WWP содержит шесть субъединиц. Помимо белковых субъединиц модель содержит молекулы цитрата, этандиола, сульфат-ионы и молекулы воды.

Краткая информация о структуре приведена в таблице 1.

Таблица 1. Общая информация о модели.

Разрешение, Å	1.90
Диапазон разрешений структурных факторов, Å	45.99 – 1.90
Полнота набора структурных факторов, %	99.8
Число измеренных рефлексов	181508
Кристаллографическая группа	C 2 2 2
Параметры кристаллографической ячейки	$a = 191.962\text{Å}$, $b = 261.58\text{Å}$, $c = 91.987\text{Å}$ $\alpha = 90.00$, $\beta = 90.00$, $\gamma = 90.00$

Для решения фазовой проблемы был использован метод молекулярного замещения с использованием структуры 1 EB8 из организма *Manihot esculenta* с разрешением 2.1 Å.

Значения индикаторов качества модели в целом

Значения индикаторов качества модели в целом были получены с помощью сервиса EDS [2].

R-фактор отражает соответствие модели и эксперимента. Значение R-фактора составляет 0,173. Хорошими значениями считаются показатели меньше 0,25, поэтому можно считать, что полученная модель хорошо соответствует экспериментальным данным.

R-free служит контролем переоптимизации модели. В данном случае значение R-free составляет 0,199. Хорошими значениями считаются показатели меньше 0,2. Кроме этого, оценивают разницу R и R-free, которая должна быть меньше 10%. В нашем случае R-free - R = 0,026, что меньше 0,1. Это говорит о том, что переоптимизации не произошло.

Для построения карты Рамачандрана был использован сервис MolProbity [3]. Карта Рамачандрана позволяет найти маргиналов по конформации остова, основываясь на значениях торсионных углов ϕ и ψ . MolProbity строит карты для шести разных случаев: общая карта, для изолейцинов и валинов, остатков перед пролином, для глицинов, транс- и цис-пролинов. Полученные карты приведены на рис. 3.

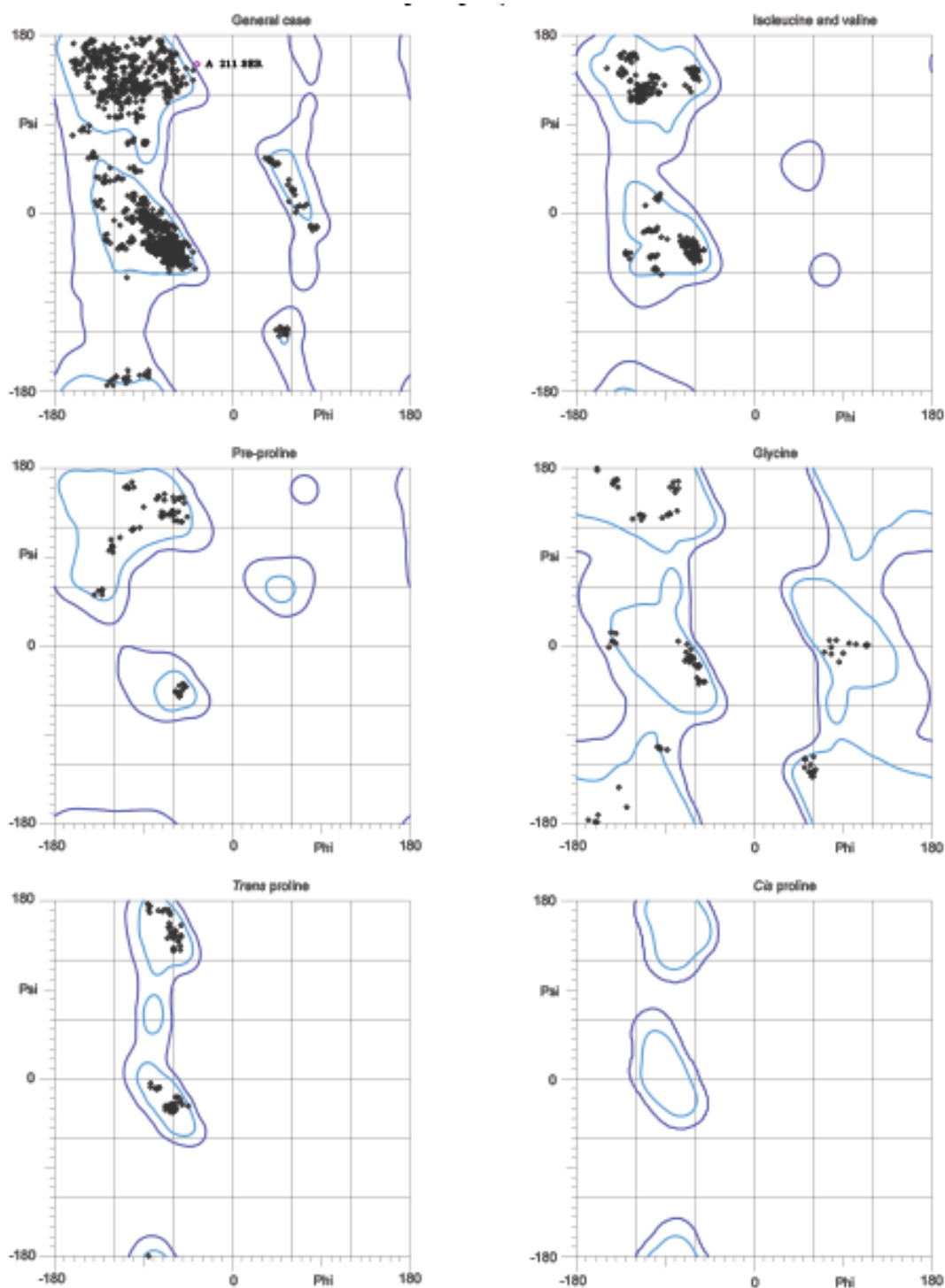


Рисунок 3. Карты Рамачандрана, полученные с помощью сервиса MolProbity. Голубыми линиями выделены предпочитаемые области, синим – разрешенные.

В запрещенной области оказался только один остаток, это Ser211 из цепи А (0,06%). MolProbity анализировал 1556 остатков, всего остатков в структуре 1568. 1475 остатков (94,8%) попали предпочитаемую область, 1555 остатков (99,9%) попали в разрешенную область.

Сервис MolProbity выдает информацию о плохих ротамерах. Ротамеры – это боковые цепи в типичных для данного типа остатка конформациях. Они описываются наборами допустимых углов χ . У каждого типа остатков свое количество ротамеров.

Структура содержит 72 плохих ротамера (5,31%). По оценке сервиса их должно быть не более 0,3%. Этот показатель сильно влияет на качество расшифровки структуры.

Пространственный R-фактор (RSR) характеризует насколько модель группы атомов соответствует экспериментальной электронной плотности. Также RSR позволяет искать маргинальные остатки. Плохими считаются показатели больше 20%.

Z-score является относительной оценкой RSR, показывает насколько RSR для остатка отличается от среднего RSR для такого же типа остатков в структурах PDB с похожим разрешением. Значения Z-score > 2 свидетельствуют о том, что остаток плохо вписан в электронную плотность. На рис. 4 представлены значения Z-score для каждого остатка каждой цепи. На рис. 5 представлено положение маргинальных остатков по RSR (Z-score > 2).

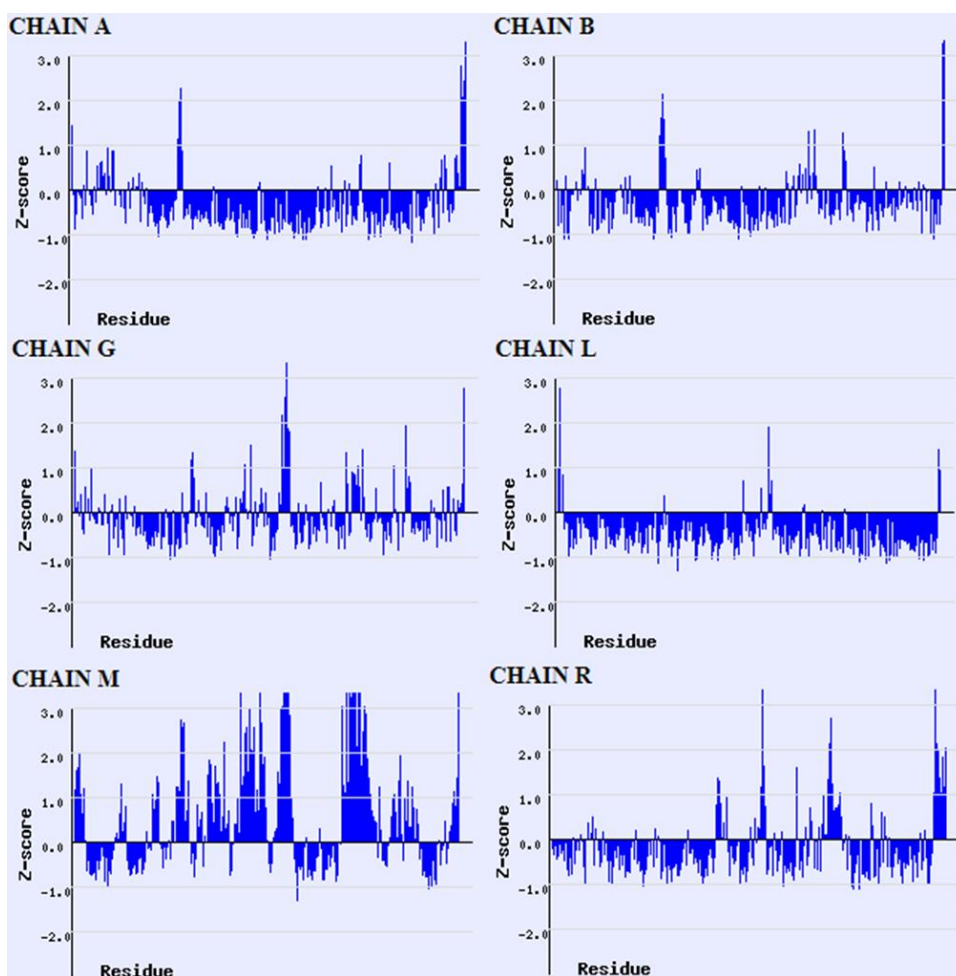


Рисунок 4. Значения Z-score для каждого остатка.

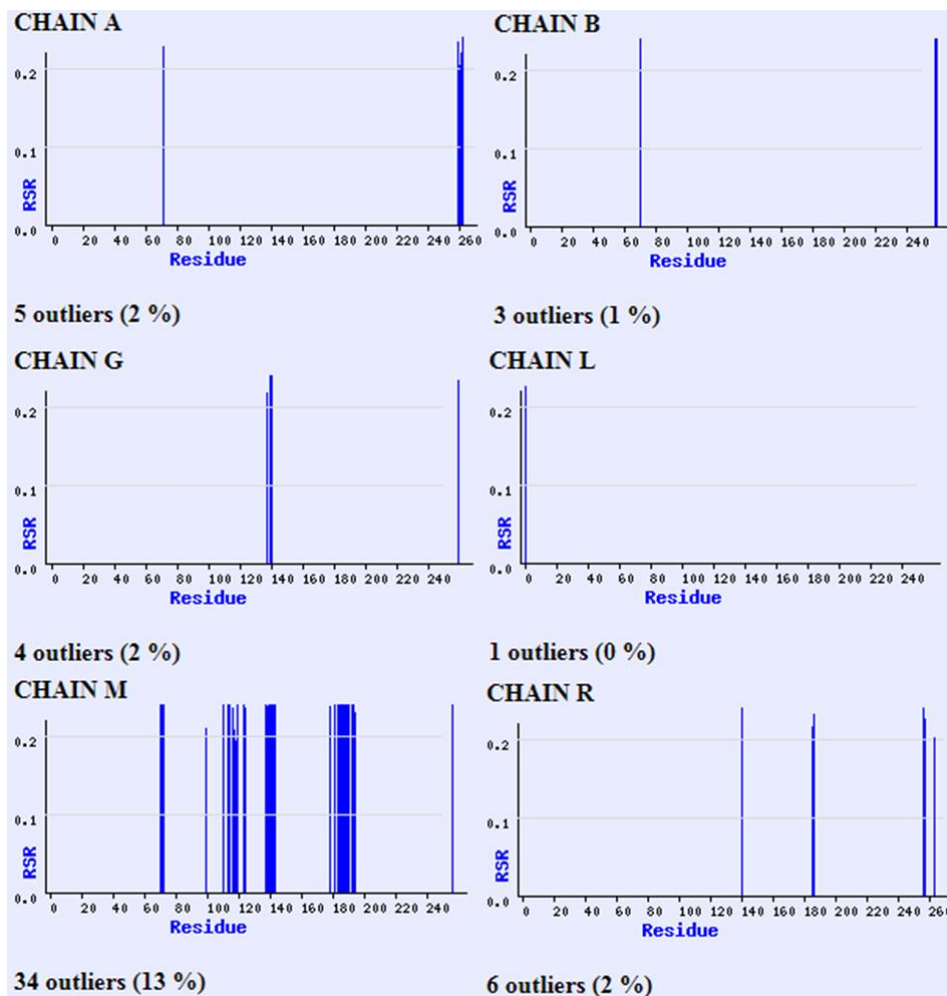


Рисунок 5. Положение маргинальных остатков и их значение RSR.

По рисункам 4-5 видно, что цепь М разрешена хуже всего, в ней 34 остатка с Z-score > 2, в то время как в остальных цепях в среднем по 4 таких остатка.

Анализ маргинальных остатков

В таблице 2 приведена информация о 12 аминокислотных остатках, являющихся маргинальными по одному и более критериям.

Таблица 2. Информация о некоторых маргинальных остатках в структуре 3WWP.

Остаток	Карта Рамачандрана	Ротамеры	RSR	C β - конформация ($>0.25\text{\AA}$)	Валентные углы	Недопустимые наложения атомов	Не транс- и не цис-
Val124A		+			+		
Thr142A		+		+		+	
Ile163A		+		+	+	+	
Arg199A		+			+	+	
Ser211A	+						
Leu260A		+	+			+	
His103G		+					
Leu106G		+				+	
Thr142L		+		+		+	+
Lys72M			+			+	
Glu123M			+				
Leu183M		+	+				
Thr185R		+	+			+	
Asn135R				+			
Ala263R						+	+

Анализ 5 маргинальных остатков

Далее более подробно будут рассмотрены 5 маргинальных остатков.

1. Arg199A

Данный остаток является маргинальным по трем параметрам: плохой ротамер, неправильный валентный угол, атом CD перекрывается с атомом кислорода из молекулы воды.

На рис.6 показан остаток аргинина с обозначенными углами χ . С помощью MolProbity были получены следующие значения углов: 281.9°, 123.2°, 68.9°, 265.6°. Полученное значение для χ_2 не наблюдается у аргининов (рис.7).

Рисунок 6. Arg199A, желтым цветом обозначены углы χ_1 , χ_2 , χ_3 , χ_4

9

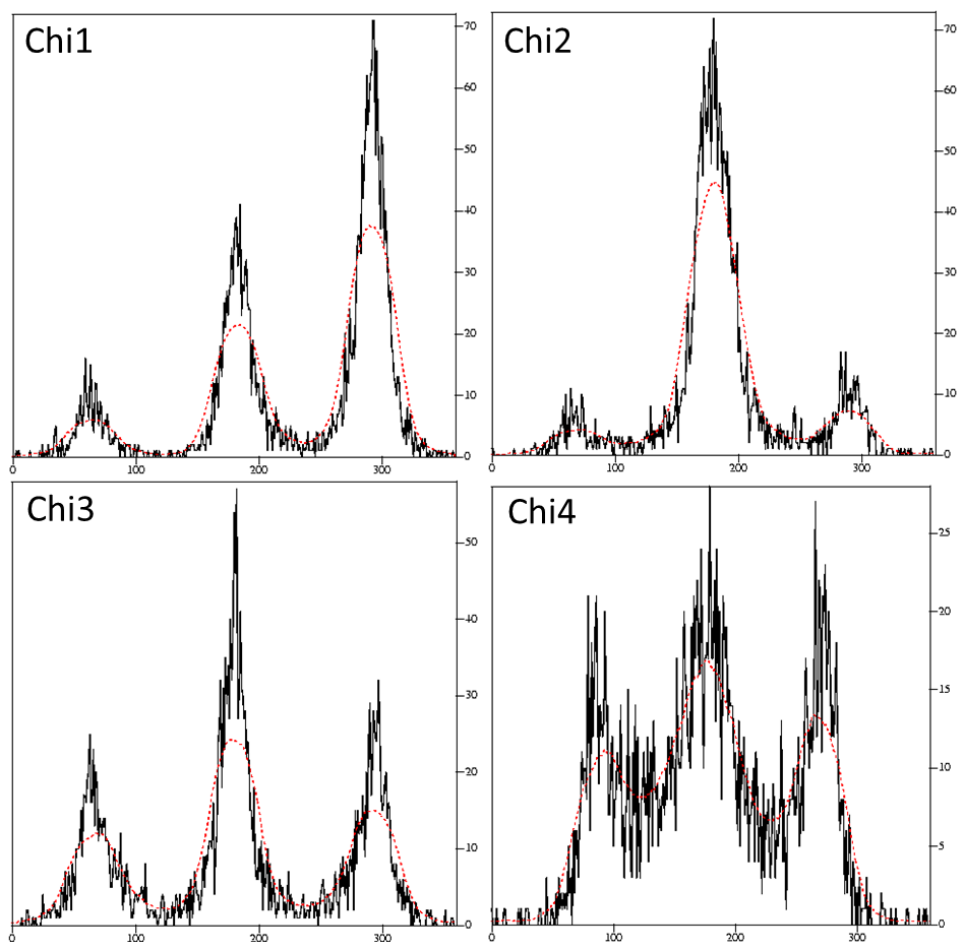


Рисунок 7. Распределение конформаций углов χ для аргинина. Получено с помощью сервиса Rotamers revisited [4]. В случае отрицательного значения нужно прибавить 360° .

На рис. 8 отмечен проблемный валентный угол CD-NE-CZ (отклонение 4.3σ).

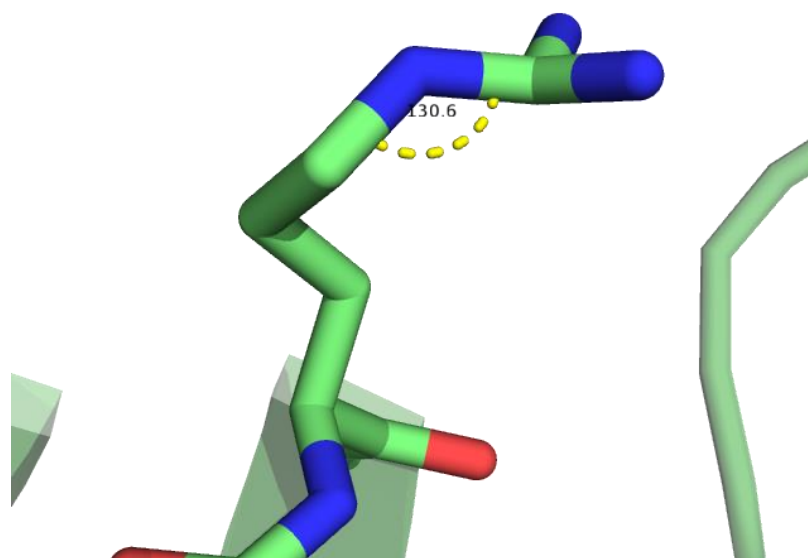


Рисунок 8. Arg199A, желтым цветом обозначен проблемный валентный угол.

При анализе электронной плотности вокруг данного остатка заметно, что остаток не достаточно хорошо вписывается. Возможно, это неточность расшифровки, и корректировка угла χ_2 и валентного угла привела бы остаток в более удачную конформацию.

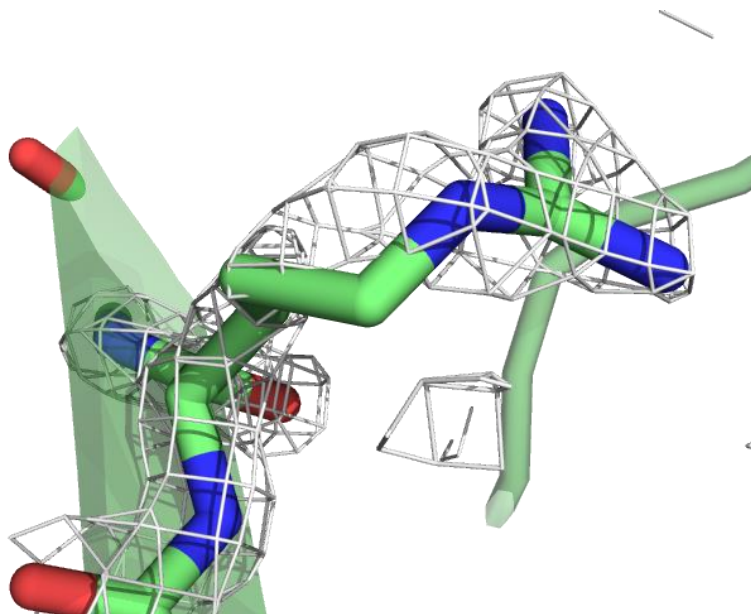


Рисунок 9. Электронная плотность вокруг остатка Arg199A на уровне подрезки 3σ .

2. Ser211A

Этот остаток является единственным маргиналом по карте Рамачандрана. На рис. 10 отмечены торсионные углы ϕ и ψ .

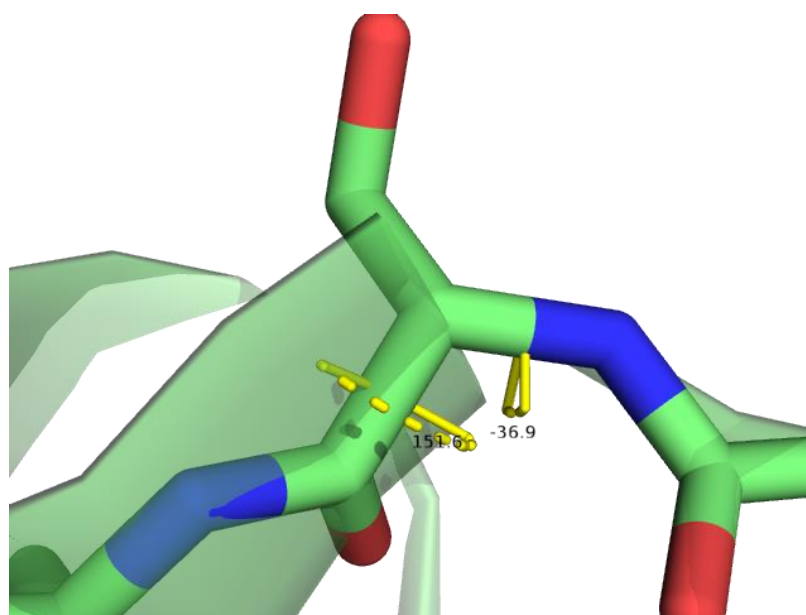


Рисунок 10. Ser211A, желтым цветом обозначены углы ϕ и ψ .

На рис. 11 показана функция электронной плотности для остатка. Атом кислорода не вписывается в электронную плотность, хочется его развернуть. Вероятно, это также ошибка расшифровки.

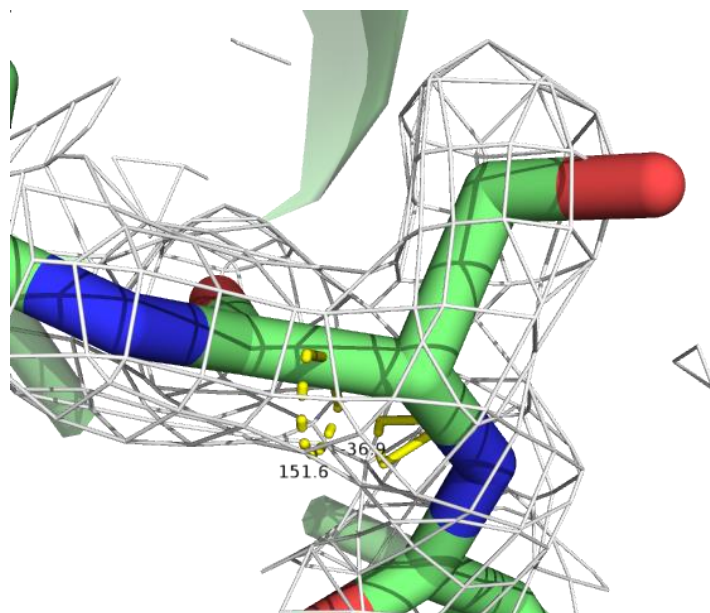


Рисунок 11. Электронная плотность вокруг остатка Ser211A на уровне подрезки 2 σ . Желтым цветом обозначены углы ϕ и ψ .

3. Leu260A

На рис. 12 обозначены углы χ_1 и χ_2 . Согласно MolProbity углы равняются 257.8° и 89.5°. Из рис.13 видно, что для лейцинов не характерно такое значение угла χ_1 .

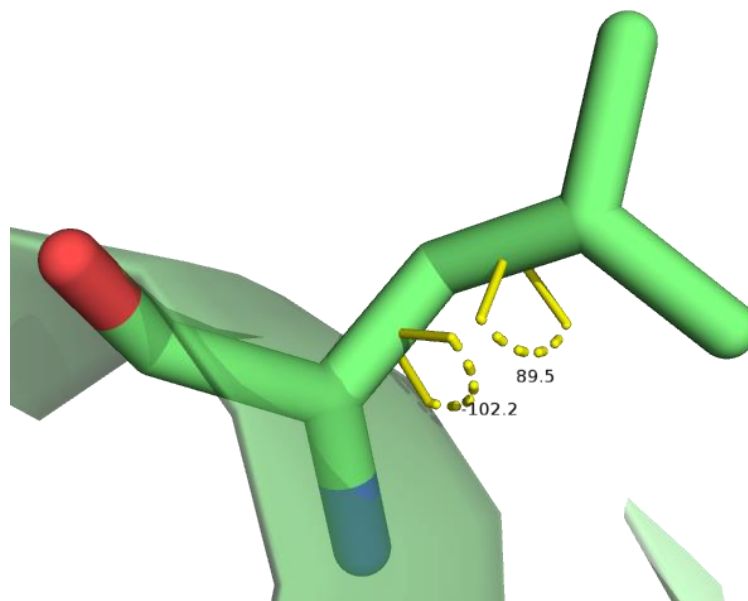


Рисунок 12. Leu260A, желтым цветом обозначены углы χ_1 и χ_2 .

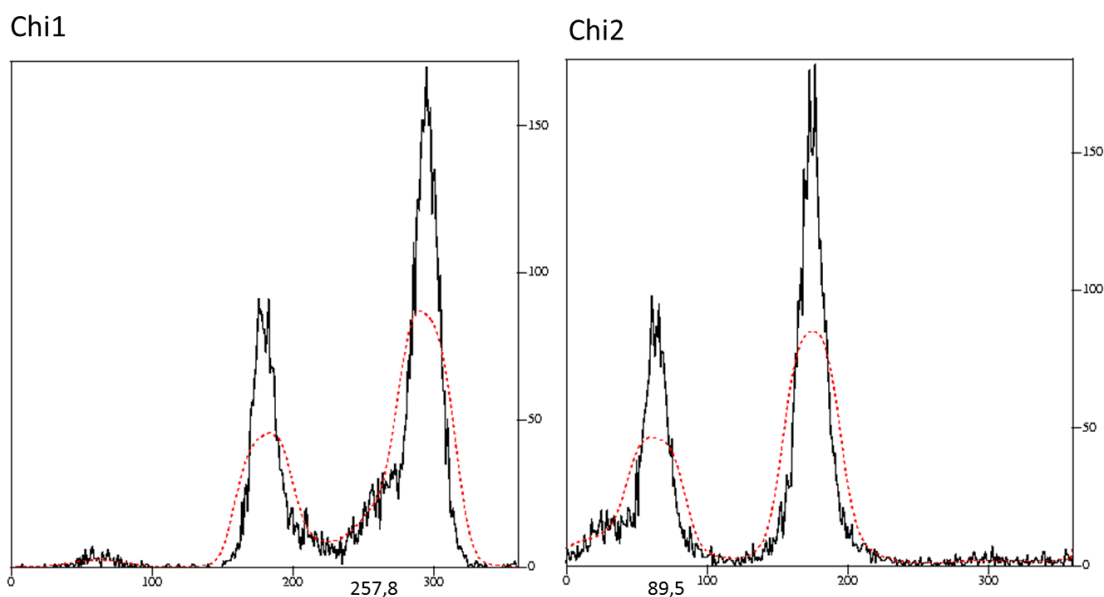


Рисунок 13. Распределение конформаций углов χ для лейцина.

Кроме этого, остаток является маргинальным по RSR. Но значения RSR и Z-score не сильно превышают допустимые значения, 0,204 и 2,08, соответственно. Остаток не полностью вписывается в функцию электронной плотности.

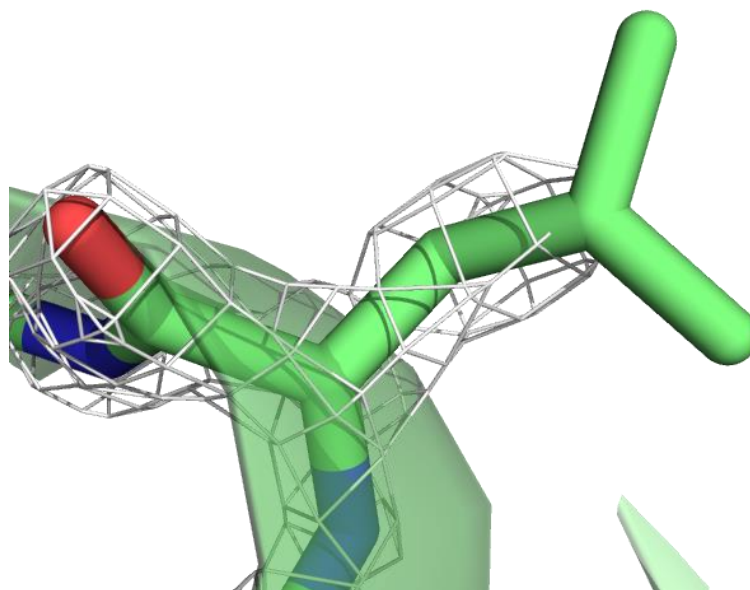


Рисунок 14. Электронная плотность вокруг остатка Leu260A на уровне подрезки 1.5 σ .

4. Thr142L

Данный остаток является плохим ротамером, перекрываются атомы, находится в конформации не цис- и не транс-. На рис. 15 отмечен угол χ , значение которого совсем не характерно для треонинов (рис. 16).

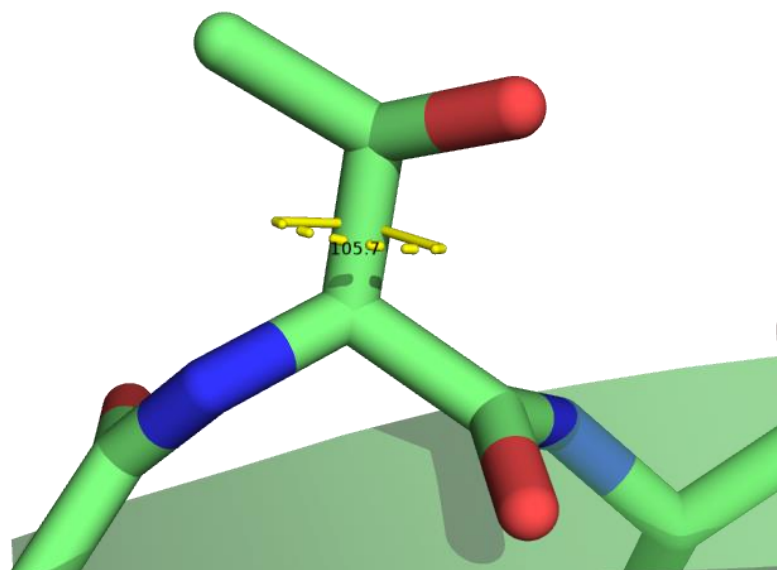


Рисунок 15. Thr142L, желтым цветом обозначен угол χ .

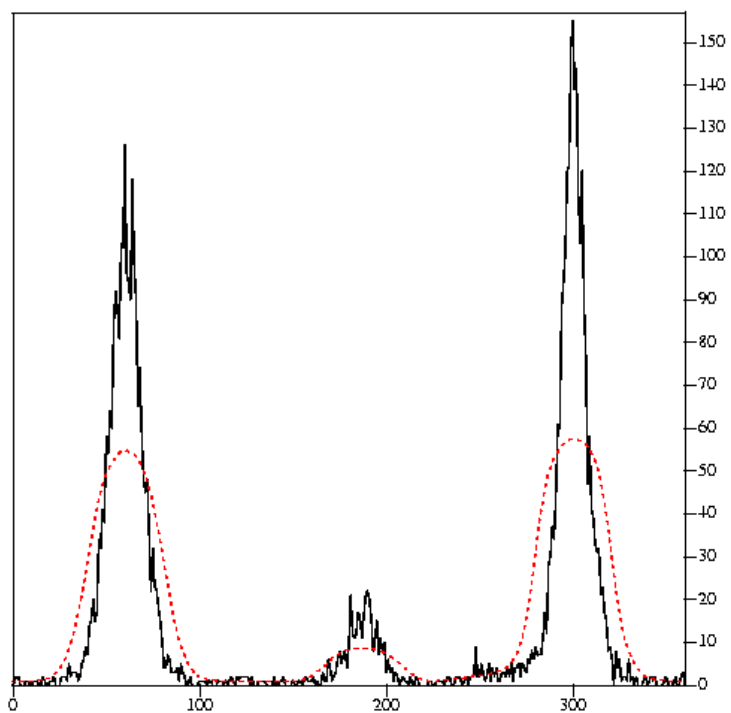


Рисунок 16. Распределение конформаций углов χ для треонина.

Данный остаток находится в конформации не цис- и не транс-, угол ω составляет -170.3° , что не характерно для белков. Обычно угол ω принимает значения 180° , редко 0° .

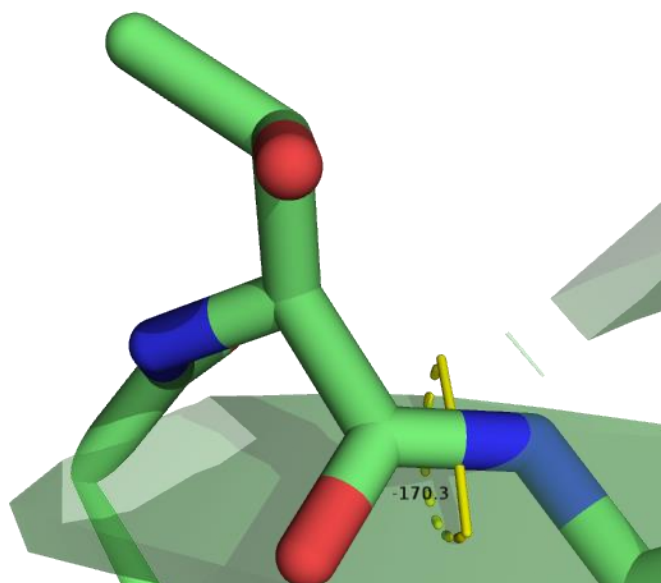


Рисунок 17. Thr142L, желтым цветом показан угол ω .

На рис. 18 показана функция электронной плотности для остатка. Остаток вписывается в электронную плотность, но не совсем так, как надо. Вероятно, это неточность расшифровки, и можно было скорректировать угол χ .

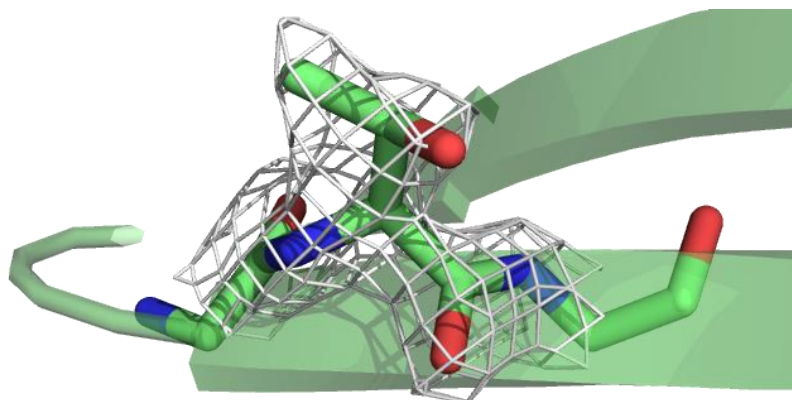


Рисунок 18. Электронная плотность вокруг остатка Thr142L на уровне подрезки 1.5 σ .

5. Glu123M

Glu123M является маргиналом по RSR. Значение RSR для него составляет 0,436, Z-score 4,2. Остаток очень плохо вписывается в экспериментальную функцию электронной плотности (рис. 19).

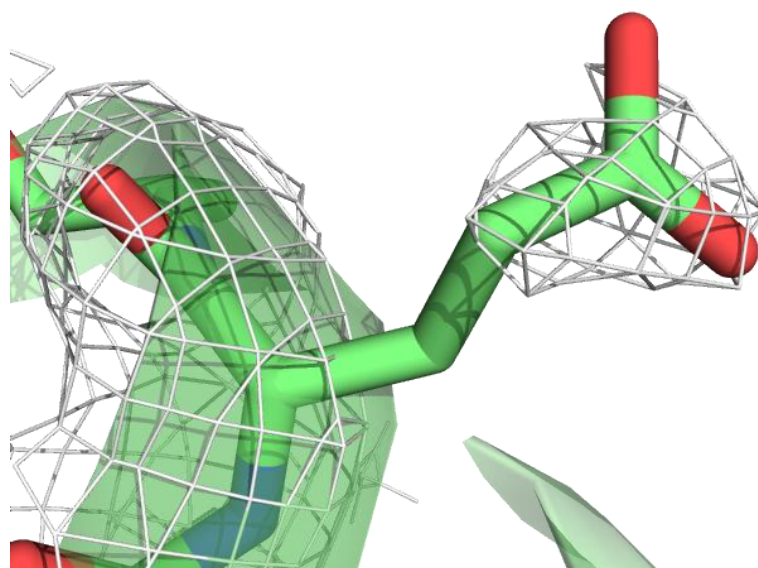


Рисунок 19. Электронная плотность вокруг остатка Glu123M на уровне подрезки 0.5σ .

Данный остаток согласно сервису EDS не является маргинальным в других цепях, но при визуализации обнаружилось, что не во всех цепях этот остаток хорошо вписывается в электронную плотность (рис. 20).

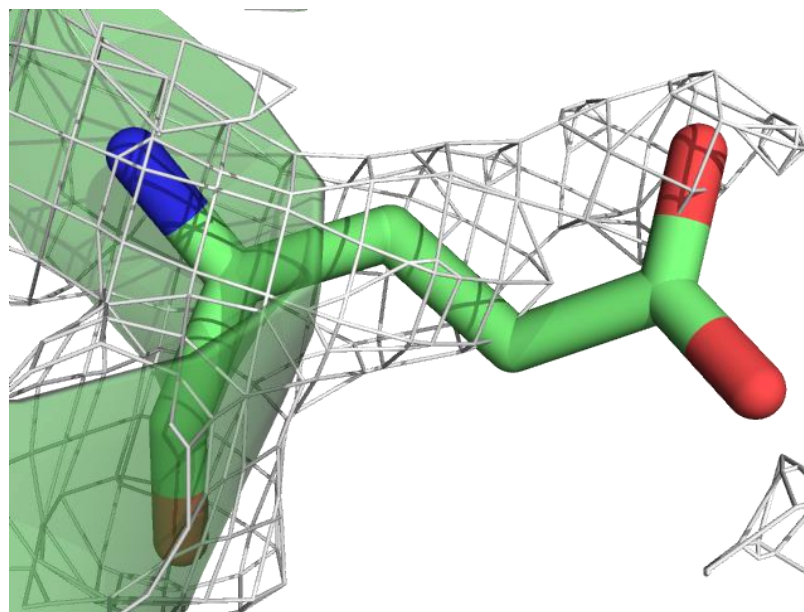


Рисунок 20. Электронная плотность вокруг остатка Glu123L на уровне подрезки 1σ .

Сравнение модели из PDB с моделью из PDB_REDO

Сервис PDB_REDO содержит оптимизированные версии записей PDB. В результате работы сервис представляет сравнение параметров оценки качества расшифровки для исходной структуры и структуры, полученной PDB_REDO (табл. 3). Зеленым цветом выделены параметры, значения которых улучшились для модели из PDB_REDO.

Таблица 3. Сравнение параметров оценки качества расшифровки структур из PDB и PDB_REDO.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
<i>R</i>	0.1595	0.1494
<i>R-free</i>	0.1983	0.1726
<i>Bond length RMS Z-score</i>	0.503	0.710
<i>Bond angle RMS Z-score</i>	0.708	0.840
Model quality (raw scores percentiles)		
<i>Ramachandran plot appearance</i>	36	64
<i>Rotamer normality</i>	41	77
<i>Coarse packing</i>	N/A	N/A
<i>Fine packing</i>	44	56
<i>Bump severity</i>	62	86
<i>Hydrogen bond satisfaction</i>	72	81

Из таблицы видно, что многие показатели улучшились. Для более наглядного представления в таблице 4 приведено сравнение параметров оценки качества для структур из PDB и PDB_REDO согласно MolProbity. Действительно, улучшились показатели по ротамерам, нет остатков в запрещенной области карты Рамачандрана. Но при этом заметно увеличилось количество плохих связей и углов.

Таблица 4. Сравнение параметров оценки качества расшифровки структур из PDB и PDB_REDO, данные получены с помощью сервиса MolProbity.

	PDB		PDB_REDO	
Poor rotamers	72	5.31%	14	1.03%
Favored rotamers	1248	92.04%	1309	96.53%
Ramachandran outliers	1	0.06%	0	0.00%
Ramachandran favored	1475	94.79%	1487	95.57%
MolProbity score [^]	2.36		1.11	
C β deviations >0.25Å	6	0.41%	5	0.34%
Bad bonds:	0 / 12983	0.00%	12 / 12983	0.09%
Bad angles:	13 / 17632	0.07%	25 / 17632	0.14%

ЗАКЛЮЧЕНИЕ

Можно заключить, что структура 3WWP расшифрована не очень хорошо, не смотря на то, что структура, в принципе, имеет хорошее разрешение (1.9Å) и большинство параметров оценки качества структуры находятся в пределах нормы. Данные из PDB_REDO показывают, что модель могла быть оптимизирована намного лучше.

СПИСОК ЛИТЕРАТУРЫ

1. Nakano S., Dadashipour M., Asano Y. Structural and functional analysis of hydroxynitrile lyase from *Baliospermum montanum* with crystal structure, molecular dynamics and enzyme kinetics // *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. – 2014. – Т. 1844. – №. 12. – С. 2059-2067.
2. <http://eds.bmc.uu.se/eds/>.
3. <http://molprobity.biochem.duke.edu/>.
4. <http://xray.bmc.uu.se/gerard/supmat/chi.html>.