

Факультет биоинженерии и биоинформатики
Московского государственного университета имени М.В.Ломоносова

Отчёт по качеству РСА расшифровки структуры
белка 1S4I

Выполнил студент 4 курса
Лозинский Ярослав

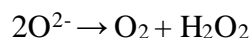
Москва, 2017

Аннотация

Данный отчет представляет собой анализ качества расшифровки структуры супероксид-дисмутаза-подобного белка из *Bacillus subtilis*, PDB-код 1S4I, на основе определенных показателей, таких как параметры кристаллической решетки, R-фактор, R_free, карта Рамачандрана, и других, полученных с помощью различных сервисов и баз данных.

Введение

Целью расшифровки белка 1S4I было расширить представления об устройстве и функционировании прокариотических гомологов Cu/Zn супероксид дисмутазы (СОД). СОД катализирует реакцию дисмутации супероксида на кислород и перекись водорода [1]:



Для этого класса ферментов характерно образование гомодимера, а также наличие в каждом мономере лигандов, связывающих ионы Zn и Cu. Последний обуславливает каталитическую активность СОД.

Среди 138 геномов архей и бактерий найдено 57 предположительных гомологов эукариотической СОД. Большинство из них содержат консервативные сайты связывания ионов Zn и Cu представляют собой нормально функционирующие фермента. Однако некоторые из них, например, BsSOD у *B. subtilis*, не связывают два иона меди и не проявляет СОД-активности [2].

Расшифровке структуры BsSOD была посвящена работа группы учёных во главе с Viezzoli [2]. Полученной модели был присвоен идентификатор 1S4I и именно качеству её определения посвящён мой отчёт.

Примечательно, что BsSOD существует в растворе в форме мономера, но при кристаллизации формирует димеры. В димеризации белков участвует иона цинка, который координируется Asp-His диадой каждого из мономеров (рис. 1) [2].



Рисунок 1. Димер белка BsSOD (1S4I) в кристалле, образуемый за счёт координации иона Zn боковыми радикалами остатков Asp и His.

Результаты и обсуждение

Общая информация о модели

В состав модели входит 4 молекулы белка, которые организованы в два димера, а также 6 ионов цинка: по одному из них связано с каждым из белков, а каждый из последних двух находится на границе между мономерами и участвует в их димеризации [2], [3]. Фазовая проблема была решена методом аномального рассеяния с ионами цинка в роли аномально рассеивающих атомов [2].

Структура 1S4I была загружена в PDB 16 января 2016 года [4]. Её авторами являются Banci, L., Bertini, I., Calderone, V., Cramaro, F., Del Conte, R., Fantoni, A., Mangani, S., Quattrone, A., Viezzoli, M.S. [2].

Информация, полученная из экспериментальных данных [5]:

- Число измеренных рефлексов: 96807;
- Разрешение: 1.80 Å;
- Полнота набора структурных факторов: 94.6 %;
- Диапазон разрешений структурных факторов: 37.07 - 1.80 Å;
- Параметры кристаллографической ячейки представлены в таблице 1.

Таблица 1. Параметры кристаллографической ячейки.

a	38.22 Å	alpha	84.35
b	61.11 Å	beta	76.02
c	64.92 Å	gamma	90.42

Пространственная группа относится к P1 группе с 4 молекулами в асимметрической единице [3]. Асимметрическая единица некристаллографической симметрией не обладает.

Значения индикаторов качества модели в целом

R-фактор: 0.221 [3]

R_{free}: 0.258 [3]

Среднее RSR: 0.120 [4]

С помощью сервиса MolProbity к структуре были добавлены атомы водорода и выполнены предложенные инверсии остатков (таблица 2). Также была построена карта Рамачандрана (рис. 2). Данным сервисом определено, что 87.97% остатков находятся в предпочтительных областях, а 96.48% лежат в разрешённых областях.

MolProbity также предоставляет некоторые характеристики качества структуры, указанные в Таблице 3.

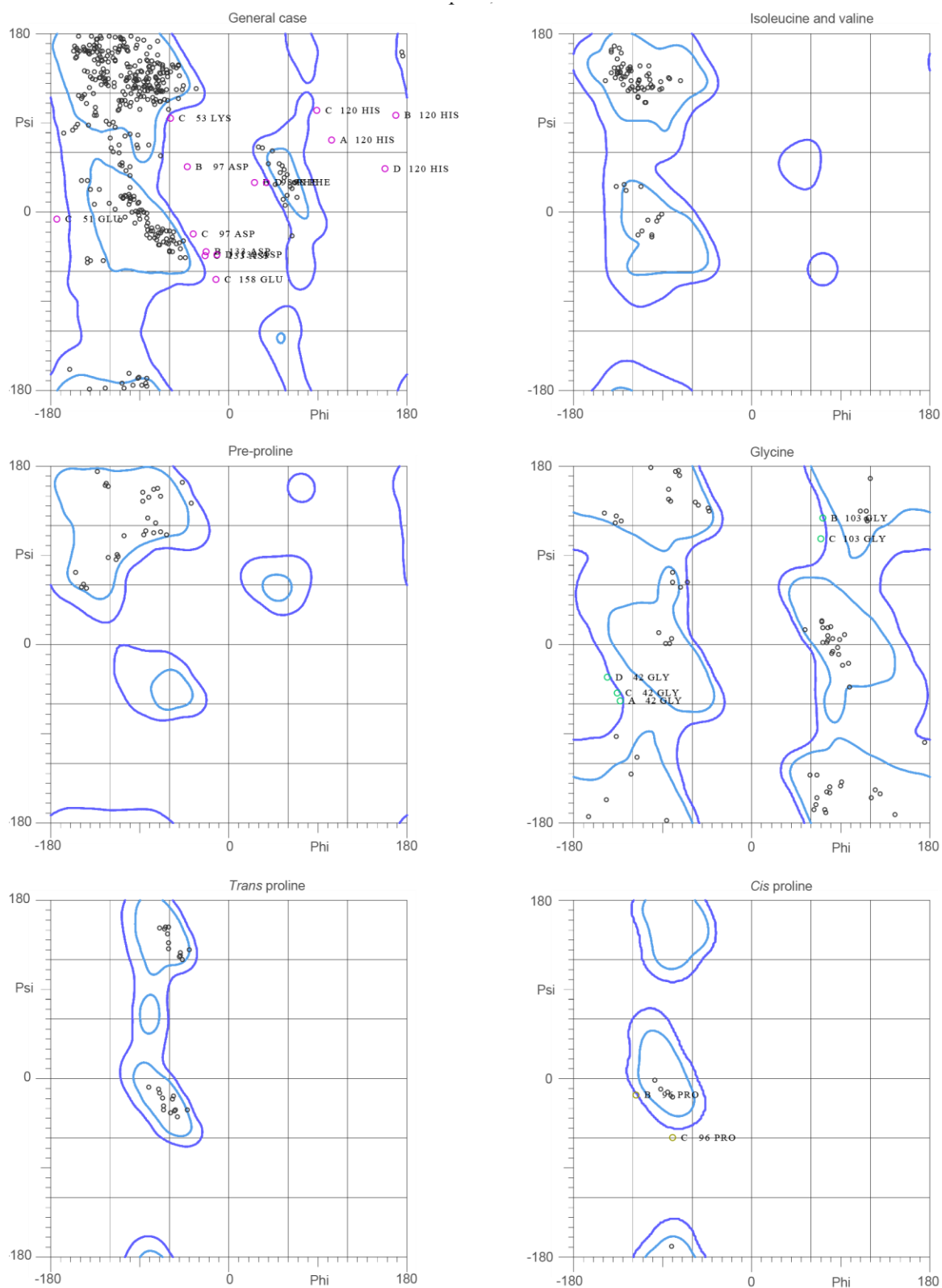


Рисунок 2. Карты Рамачандрана структуры 1S4I, полученные MolProbity. Голубым выделены

предпочтительные области значений углов ϕ и ψ , синим - разрешённые. Розовым и зелёным выделены маргинальные остатки.

Таблица 2. Остатки структуры 1XTM, инвертированные для анализа качества структуры

Цепь	Номер остатка	Тип остатка
A	75	ASN
A	127	ASN
B	46	GLN
B	75	ASN
B	127	ASN
C	75	ASN
C	127	ASN
C	175	ASN
D	75	ASN
D	115	ASN
D	127	ASN

Таблица 3. Характеристики модели, полученные с помощью MolProbity

All-Atom Contacts	Clashscore, all atoms:	26.72	9 th percentile* (N=837, 1.80Å ± 0.25Å)	
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	68	14.17%	Goal: <0.3%
	Favored rotamers	363	75.62%	Goal: >98%
	Ramachandran outliers	21	3.52%	Goal: <0.05%
	Ramachandran favored	525	87.94%	Goal: >98%
	MolProbity score [^]	3.39	1 st percentile* (N=11444, 1.80Å ± 0.25Å)	
	C β deviations >0.25Å	25	4.73%	Goal: 0
	Bad bonds:	33 / 4620	0.71%	Goal: 0%
	Bad angles:	128 / 6245	2.05%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	7 / 32	21.88%	Expected: ≤1 per chain, or ≤5%
	Cis nonProlines:	13 / 569	2.28%	Goal: <0.05%
	Twisted Peptides:	21 / 601	3.49%	Goal: 0

Маргинальные остатки

Для определения всех маргинальных остатков была использована информация из отчёта о валидации структуры 2S4I [3]. Имеющиеся в нем схемы (не приведены в целях экономии места) показывают, какие остатки обладают отклонениями в своей геометрии по различным показателям и позволяют оценить маргинальность каждого остатка по совокупности показателей. Зеленый цвет остатка в такой схеме означает отсутствие маргинальности по всем показателям, желтый – маргинальность по одному, оранжевый – по двум, красный – по трем и более критериям. Серыми выделены остатки, не представленные в структуре. Также на

схеме отмечены остатки в показателем $RSR-Z > 2$. Информация о маргинальных остатках и критериях их маргинальности представлена в таблице 4.

Таблица 4. Некоторые маргинальные остатки и показатели маргинальности. Жирным выделены остатки, которые были рассмотрены подробнее.

Остаток	Причины маргинальности
В:133:Asp	Из запрещённой области карты Раманчандрана. Нехарактерный ротамер боковой цепи. Нехарактерный угол СВ-CG-OD2 (124.28° вместо 180.30° в идеале). Нехарактерный угол СВ-CA-C (99.70° вместо 110.40° в идеале).
A:118:Met	Нехарактерный ротамер боковой цепи.
A:42:Gly	Из запрещённой области карты Раманчандрана. Аутлайер по планарности (нарушение плоскости sp_2 -гибридного атома). Нехарактерный угол C-N-CA
В:97:Asp	Значение RSR 0.318 (больше 0.2). Длина связи C-N (1.12 Å вместо 1.34 Å в идеале). Длина связи C-O (1.38 Å вместо 1.23 Å в идеале). Нехарактерный угол O-C-N (93.42° вместо 122.70° в идеале). Нехарактерный угол CA-C-N (135.27° вместо 117.20° в идеале). Нехарактерный угол CA-C-O (103.3° вместо 120.10° в идеале). Аутлайер по планарности.
D:100:Ser	Значение RSR 0.286 (больше 0.2). Нехарактерный угол O-C-N (113.14° вместо 122.70° в идеале). Слишком близкий контакт внутри одной асимметрической единицы (перекрывание D:94:Val:HB с D:100:Ser:HB2 на 0.72 Å).
D:101:Ala	Значение RSR 0.266 (больше 0.2). Нехарактерный угол СВ-CA-C (122.17° вместо 110.10° в идеале). Нехарактерный угол C-N-CA (107.48° вместо 122.30° в идеале). Нехарактерный угол N-CA-CB (118.81° вместо 110.10° в идеале). Нехарактерный угол O-C-N (113.03° вместо 123.20° в идеале). Аутлайер по хиральности – отклонение хирального центра от ожидаемой конфигурации. Аутлайер по планарности. Слишком близкий контакт внутри одной асимметрической единицы (перекрывание D:92:SER:HB3 с D:101:ALA:HB3 на 0.76 Å, перекрывание D:90:LYS:HB2 с D:101:ALA:HB1 на 0.43 Å).
C:99:Glu	Значение RSR 0.335 (больше 0.2). Слишком близкий контакт внутри одной асимметрической единицы (перекрывание C:99:GLU:O с C:101:ALA:HA на 0.53 Å)
C:62:Glu	Нехарактерный ротамер боковой цепи. Слишком близкий контакт внутри одной асимметрической единицы (перекрывание C:62:GLU:HB2 с C:681:HOH:O на 0.99 Å, перекрывание C:62:GLU:CA с C:681:HOH:O на 0.56 Å, перекрывание C:62:GLU:HA с C:681:HOH:O на 0.52 Å, перекрывание C:62:GLU:CB с C:681:HOH:O на 0.5 Å).
C:668:HOH	Не фиксирована водородными связями.
A:75:Asn	Подозрение на необходимость инверсии боковой цепи.

Анализ маргинальных остатков

1. Аспаргат 97 цепи В.

Параметры маргинальности указаны в таблице 4.

На рис. 2 изображён остаток аспартата 97 и визуализация его электронной плотности при уровне подрезки 2. Видно, что атом кислорода пептидной связи образует ковалентную связь с атомом азота, что противоречит законам химии. Это делает понятным причину нарушения планарности sp_2 -гибридизованного атома, наличие нехарактерных длин связей C-N и C-O, а также нехарактерных углов O-C-N и CA-C-O. Также стоит отметить плохое попадание карбоксильной группы в облако электронной плотности, соответствующее ей (рис. 3). Наблюдаемые аномалии следует отнести к ошибкам расшифровки электронной плотности и, вероятно, являются следствием недостаточного разрешения электронной плотности на данном участке.

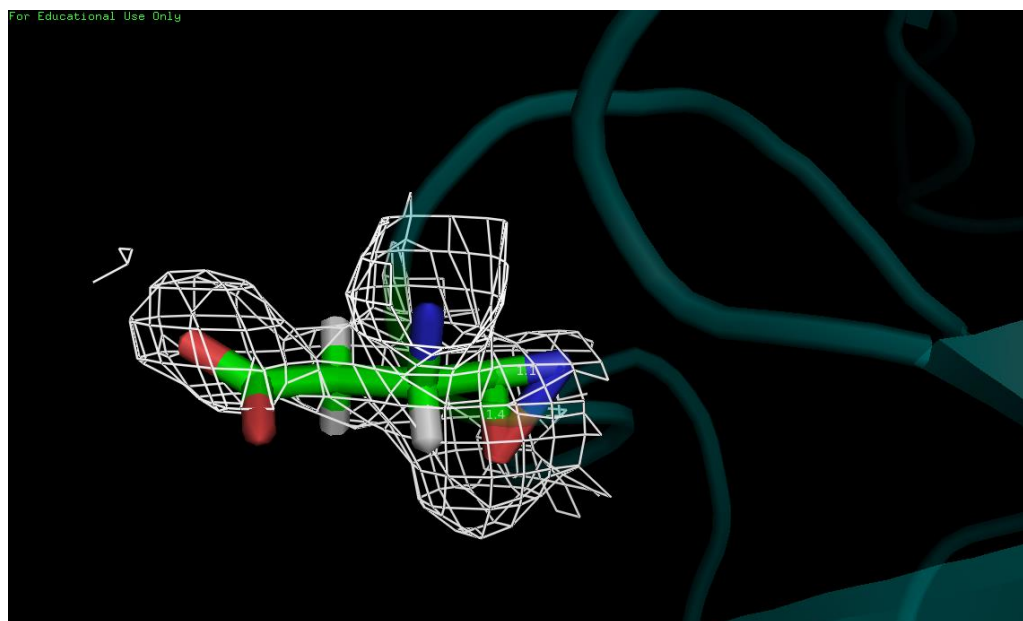


Рисунок 2. Изображение электронной плотности вокруг остатка аспартата 97 цепи В. Уровень подрезки 1. Обозначены расстояния C-O (1.4 Å) и C-N (1.1 Å).

2. Серин 100 цепи D.

На рисунке 3 изображён остаток серина цепи D. Параметры маргинальности указаны в таблице 4. На рисунке указан угол O-C-N с нехарактерным значением (113.14° , атом азота принадлежит остатку 101), а также расстояние между остатком серина 100 и остатком валина 94 цепи D (1.7 Å), которое свидетельствует о перекрытии атомов этих двух остатков. Вероятно, причиной таких отклонений от «нормы» является неправильная расшифровка координат атомов серина. Разрешения в этой области довольно плохое, что хорошо видно на рисунке 4, где остаток серина находится в окружении двух соседних остатков. Остаток глутмина 99 и остаток аланина 101 также отмечены в [3] как маргинальные и плохо вписываются в свою электронную плотность, что в свою очередь также может осложнять определение координат атомов серина 100 между ними. Кажется, что, если немного «подвинуть» боковой радикал серина 100 и лучше вписать его в соответствующую электронную плотность, то удастся уменьшить степень перекрывания с остатком валина 94.

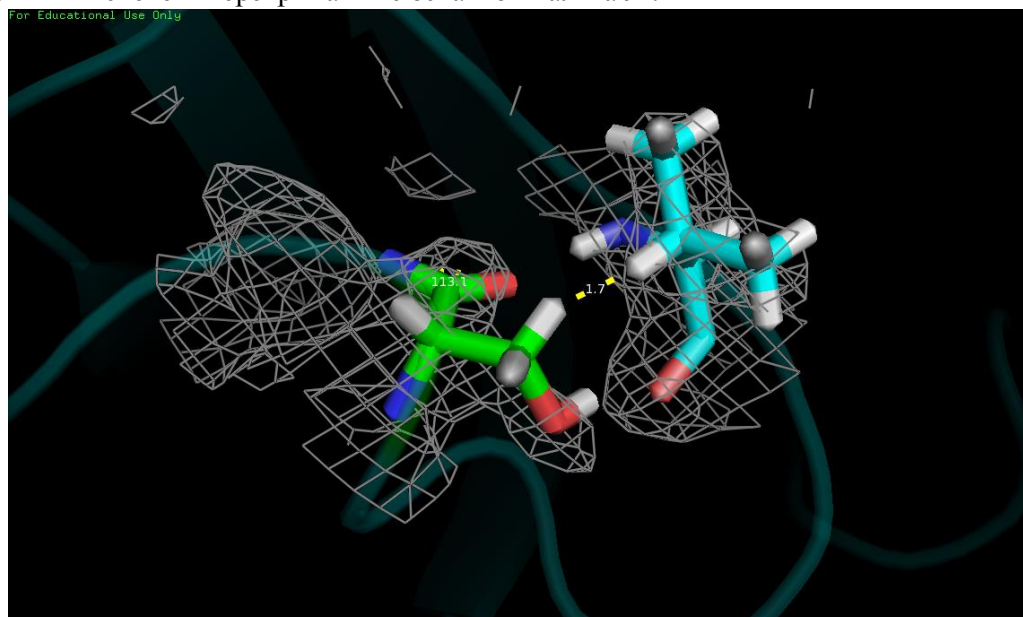


Рисунок 3. Изображение электронной плотности вокруг остатка серина 100 цепи D. Уровень подрезки 1. Обозначено расстояние между D:94:Val:HB с D:100:Ser:HB2 (1.7 Å). Обозначен угол O-C-N (113.1°).

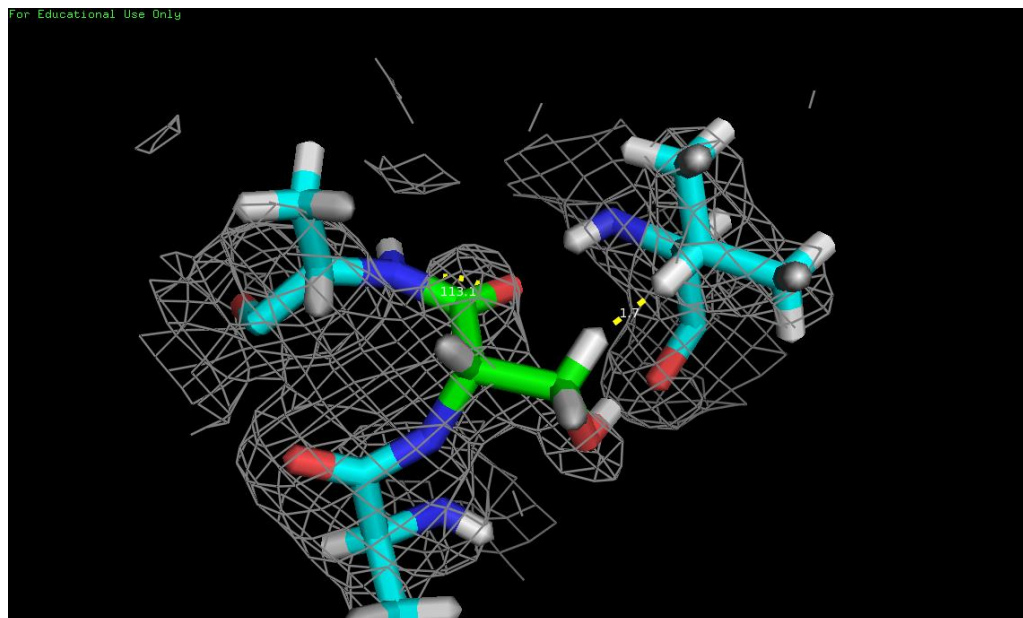


Рисунок 4. Изображение электронной плотности вокруг остатка серина 100 цепи D соседних с ним остатков: глутамин 99 и аланин 101. Уровень подрезки 1. Обозначено расстояние между D:94:Val:HB с D:100:Ser:HB2 (1.7 Å). Обозначен угол O-C-N (113.1°).

3. Остаток глутамата 62 цепи С.

Параметры маргинальности представлены в таблице 4. Для остатка глутамата 62 цепи С были обнаружены нехарактерный ротамер боковой цепи с углами χ_1 , χ_2 , χ_3 равными 164.2°, 120.6°, 7.4°, соответственно [6]. Также для остатка глутамата 62 наблюдается перекрывание с молекулой воды 681 по четырём атомам: CA, HA, CB, HB2 (рис.5).

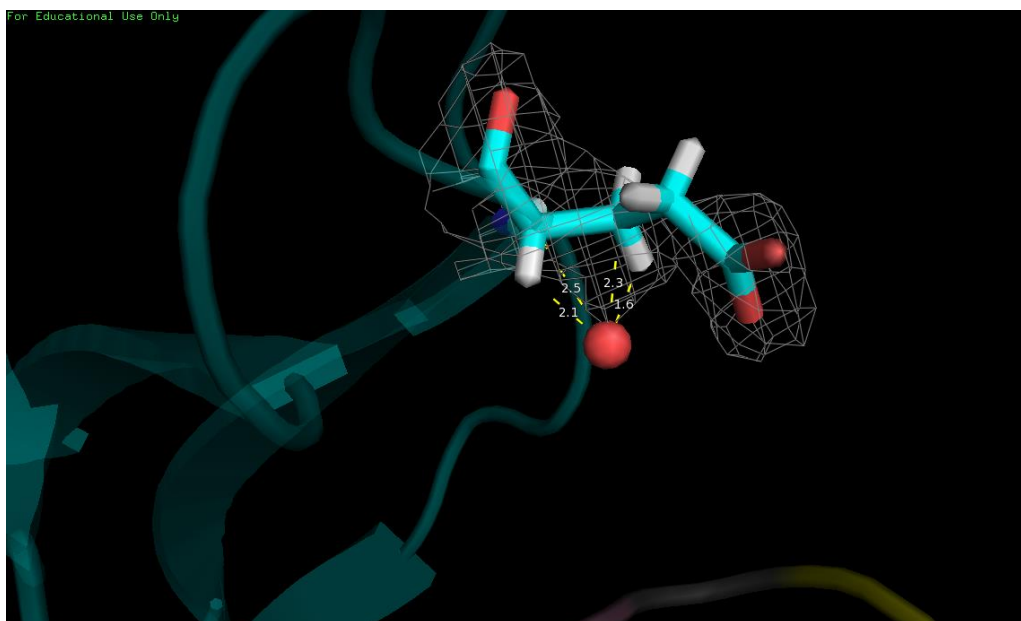


Рисунок 5. Остаток глутамата 62 и его электронная плотность. Показаны расстояния от атома кислорода молекулы воды 681 до атомов, с которыми возникает перекрытие. Уровень подрезки 1.

Остаток глутамата 62 не слишком хорошо вписывается в своё облако электронной плотности (рис. 5 и рис. 6): отсутствует электронная плотность для двух атомов водорода при CG атоме, нечёткие границы электронной плотности карбоксильной группы. Однако корректность определения координат для молекулы воды 681 также внушает сомнения.

Для молекулы воды 681 были найдены полярные контакты. По-видимому, молекула воды 681 является донором водородной связи для атома OE1 остатка глутамата (рис. 7). Полярные взаимодействия были детектированы также между молекулой воды 681 и молекулой 634. Молекула воды 634 в свою очередь образует ещё три полярных контакта, в который дважды выступает акцепторов водородной связи и один раз донором (лизин 148, лейцин 147 – акцептор, серин 146 – донор) (рис. 7), а также ещё один контакт с остатком глутамата, в котором может являться донором водородной связи. Нельзя точно сказать, какой из двух атомов – O воды 681 или OE1 остатка глутамата 62 является акцепторов второй водородной связи для молекулы воды 634: в обоих случаях угол между двумя предполагаемыми водородными связями (угол между связью с атомом OG серина и контактом с атомом OE1 остатка глутамата 62 или угол между связью с атомом OG серина и контактом с атомом кислорода молекулы воды 634) составляет примерно 146° (результаты измерения не представлены), что, однако, отличается от тетраэдрического. При всём этом образование водородной связи через молекулу воды 634 с другими остатками белка могло бы стабилизировать положение ротамера глутама и делать его более выгодным, что могло бы объяснить нехарактерные значения углов χ .

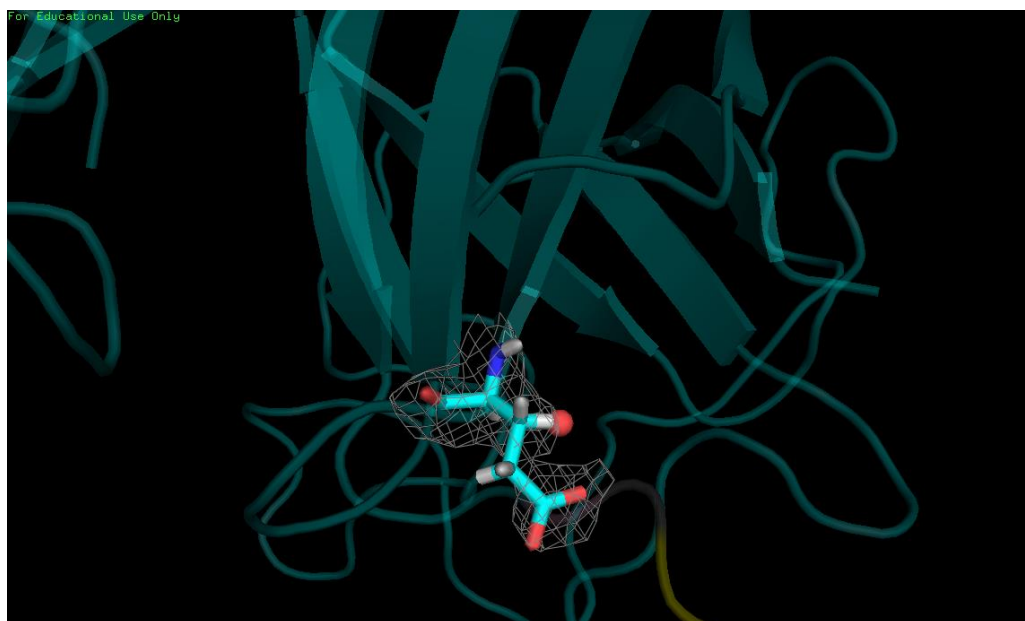


Рисунок 6. Остаток глутамата 62 и его электронная плотность. На заднем плане молекула воды 681. Уровень подрезки 1.

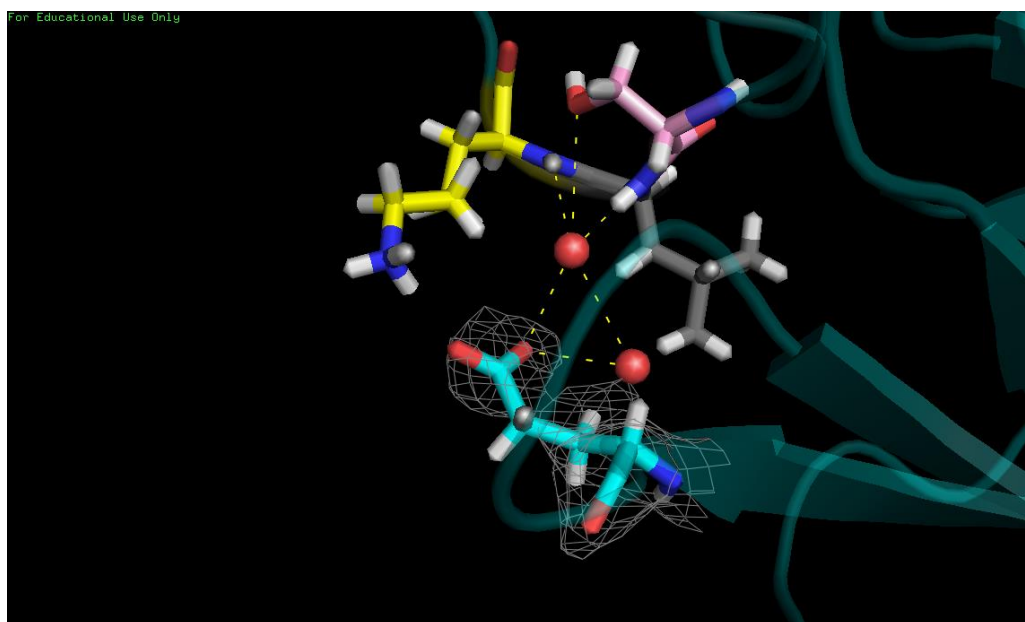


Рисунок 7. Остаток глутамата 62 и его электронная плотность. Показаны полярные контакты молекулы воды 634 (в центре) с OE1 атомом остатка глутамата 62, молекулой воды 681, атомом N остатка лизина 148 (жёлтый), атомом N остатка лейцина (серый), атомом OG сериан 146 (розовый). Также показаны полярные контакты молекулы воды 681 с OE1 атомом остатка глутамата 62 и молекулой воды 634. Уровень подрезки 1.

Итог. Проблема с перекрытием атомом глутамата 62 и молекулы воды 681 могла быть вызвана неточностью расшифровки и присвоением молекуле воды 681 неверных координат (температурный фактор составляет 56.13, что выше среднего 46.4 и выше значения для молекулы воды 634 - 43.3). Проблема, связанная с нехарактерными углами ротамеров могла быть вызвана описанным выше эффектом и может отражать особенности структуры белка.

4. Молекула воды 878, относящаяся к цепи В.

Внутри кристаллографической ячейки были найдены полярные взаимодействия молекул воды с другими атомами ячейки. При этом для молекулы воды 878 никакого полярного взаимодействия найдено не было. Был восстановлен фрагмент кристалла и определены полярные взаимодействия молекул воды с любыми атомами выбранного фрагмента, содержащего молекулу воды 878. Причём эта молекула воды была окружена белками со всех сторон, что исключало возможность пропустить какие-то связи с атомами соседней ячейки. Молекулы воды 878 в окружении ближайших белков и других молекул воды изображена на рис. 8. Ближайший атом был найден на расстоянии 5 Å, и он оказался атомом кислорода молекулы воды из соседней ячейки. Так как это расстояние никак нельзя считать длиной водородной связи, можно заключить, что молекула воды 878 не вступает ни в какие взаимодействия с другими. Это является ошибкой построения модели. Температурный фактор для неё составляет 58.43, что выше среднего значения 46.4.

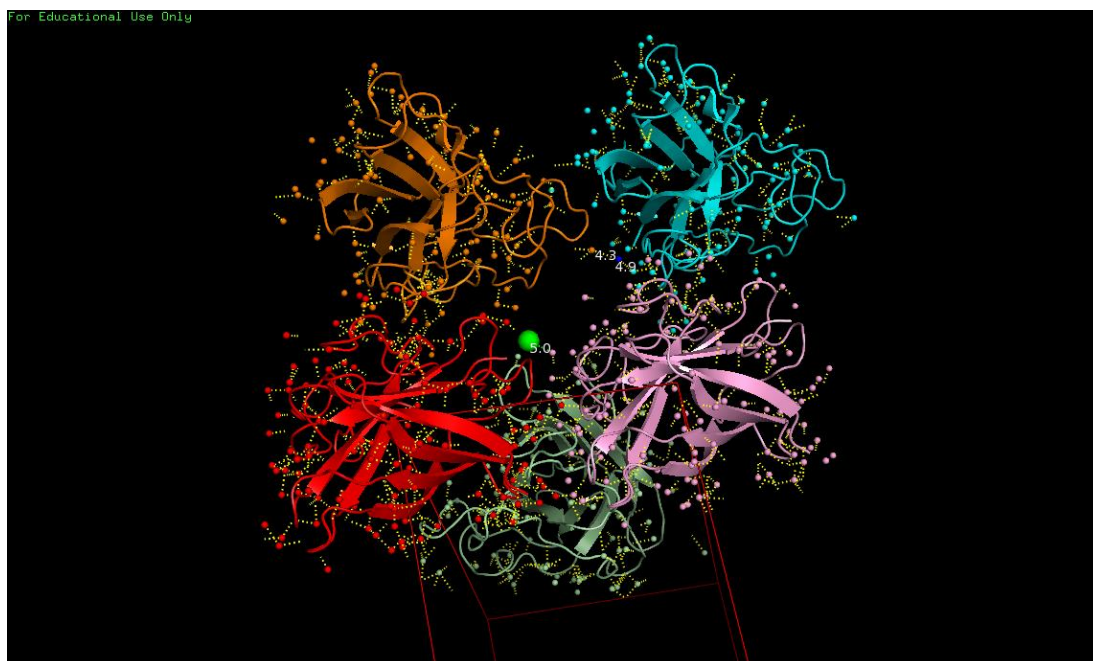


Рисунок 8. Молекулы воды 878 в окружении ближайших белков и других молекул воды (обозначена зелёным). Показано расстояние молекулы воды 978 до ближайшей другой молекулы (5 Å). Синим показана ещё одна «сомнительная» молекулы воды, которая была обнаружена в процессе подготовки картинки. Показаны расстояния от неё до ближайших молекул воды (4.4 и 4.9 Å).

5. Аспарагин 75 цепи А.

Критерий маргинальности для аспарагина 75 цепи А – возможность инверсии. Аспарагин 75 до инверсии изображён на рис. 9. Так как азот связан с двумя водородами, суммарная электронная плотность азота и водородов должна быть больше, чем электронная плотность кислорода. Как видно рис. 10, после инверсии азот стал занимать большую из двух частей электронной плотности в раздвоении на конце бокового радикала. Следовательно, аспарагин 75 является маргинальным и его маргинальность является следствием ошибочной расшифровки структуры.



Рисунок 9. Остаток аспарагин 75 цепи А до инверсии. Изображена электронная плотность остатка. Уровень подрезки 1.



Рисунок 10. Остаток аспарагин 75 цепи А после инверсии. Изображена электронная плотность остатка. Уровень подрезки 1.

Заключение

Не смотря на высокое разрешение структуры 1S4I в 1.8 Å, её качество является далеко не идеальным, а в её модели присутствуют маргинальные остатки, причиной маргинальности которых может быть не один отклоняющийся показатель, а целый ряд нехарактерных параметров. Скорее всего, это является следствием не точно определённой ЭП.

С помощью сервиса PDB_redo [6] данная модель была значительно улучшена, что привело к более точному определению вторичной структуры белка и оптимизации конформации некоторых аминокислотных остатков. При совмещении изначальной и оптимизированной структуры видно, что конформация боковых цепей некоторых остатков значительно изменилась (рис. 11).

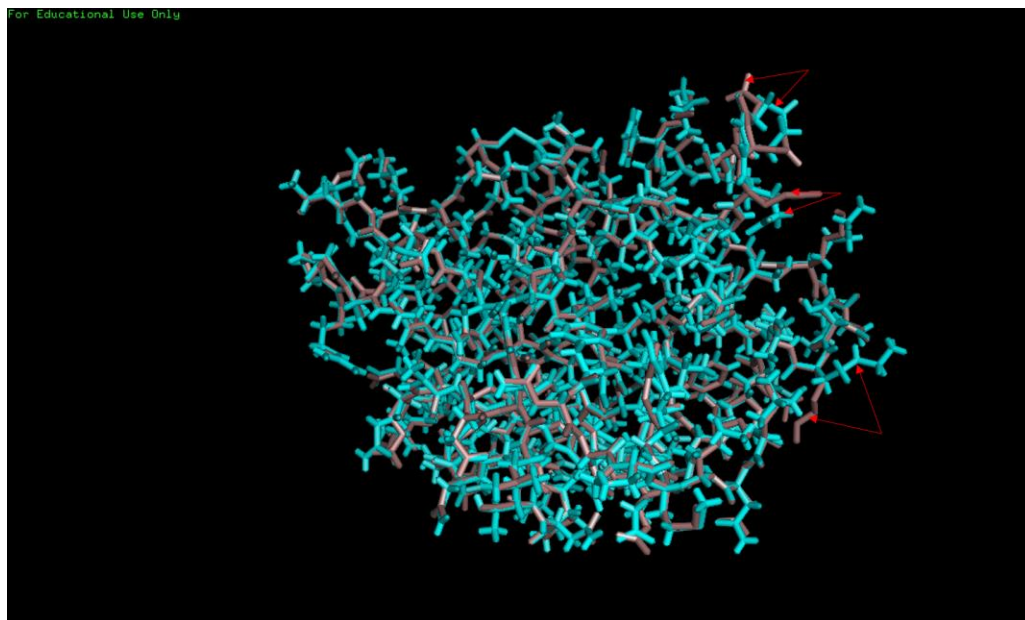


Рисунок 11. Остаток аспарагин 75 цепи А после инверсии. Изображена электронная плотность остатка. Уровень подрезки 1.

При детальном рассмотрении одного из обсуждаемых ранее маргинальных остатков – аспартата 97 цепи D – выяснилось, что в оптимизированной структуре исчезла аномальная связь между кислородом и азотом пептидной связи, что говорит в пользу улучшения качества структуры (рис. 12).

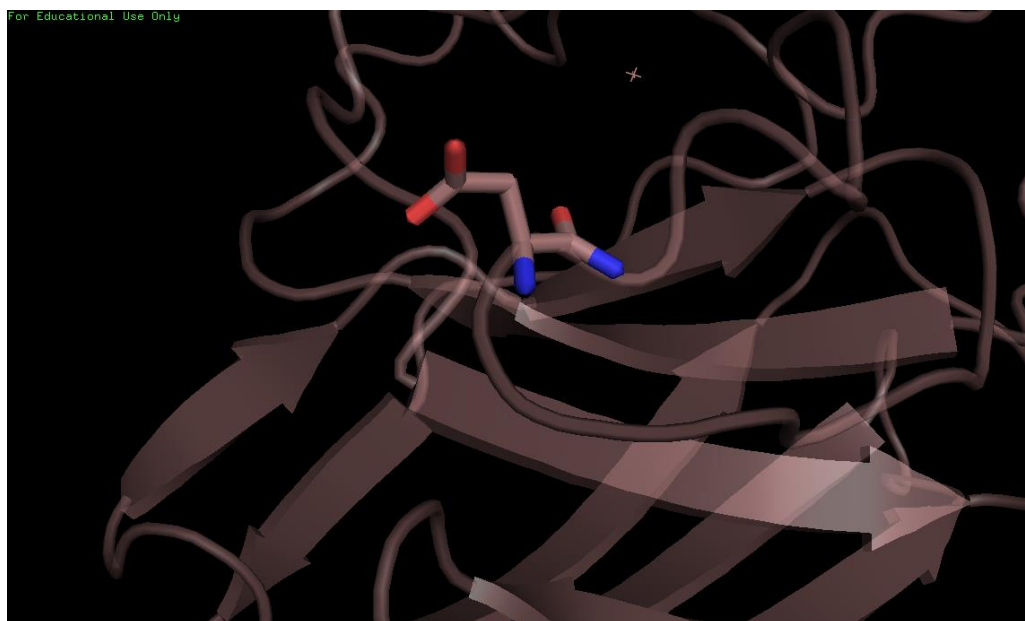


Рисунок 12. Остаток аспарагин 75 цепи А после инверсии. Изображена электронная плотность остатка. Уровень подрезки 1.

Список литературы

[1] "Superoxide dismutase." [Online]. Available: https://en.wikipedia.org/wiki/Superoxide_dismutase.

[2] Banci, Lucia, et al. "A prokaryotic superoxide dismutase paralog lacking two Cu ligands: from largely unstructured in solution to ordered in the crystal." *Proceedings of the National Academy of Sciences of the United States of America* 102.21 (2005): 7541-7546.

[3] "Full wwPDB X-ray Structure Validation Report." [Online]. Available: http://ftp.wwpdb.org/pub/pdb/validation_reports/s4/1s4i/1s4i_full_validation.pdf.

[4] "RCSB PDB for 1S4I." [Online]. Available: <http://www.rcsb.org/pdb/explore/explore.do?structureId=1S4I>.

[5] "EDS: 1s4i." [Online]. Available <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1s4i>.

[6] "1s4i - PDB_REDO optimisation results." [Online]. Available: http://www.cmbi.ru.nl/pdb_redo/s4/1s4i/index.html.