

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

---

***ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ***

**Отчёт по качеству RSA расшифровки структуры  
1XTM**

***Выполнил студент 4-го курса  
Попов Ярослав***

**Москва  
2015 г.**

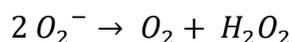
## Аннотация

В данной работе коротко описан анализируемый белок, представлены индикаторы качества структуры 1XTM в целом, приведён список некоторых маргинальных остатков, а также проведён подробный анализ некоторых из них. В отчёте присутствует сравнение моделей структуры 1XTM из PDB и PDB\_redo.

## Введение

Структура 1XTM (см. Рисунок 1) соответствует Cu/Zn супероксид дисмутаза подобному белку из *Bacillus subtilis* с двумя мутациями: Y88H и P104H.

Супероксид дисмутаза (СОД, SOD) катализируют дисмутацию супероксида на кислород и пероксид водорода:



Фермент осуществляет эту реакцию в две стадии [1]:

- 1)  $M^{(n+1)+} - SOD + O_2^- \rightarrow M^{n+} - SOD + O_2$
- 2)  $M^{n+} - SOD + O_2^- \rightarrow M^{(n+1)+} + H_2O_2$

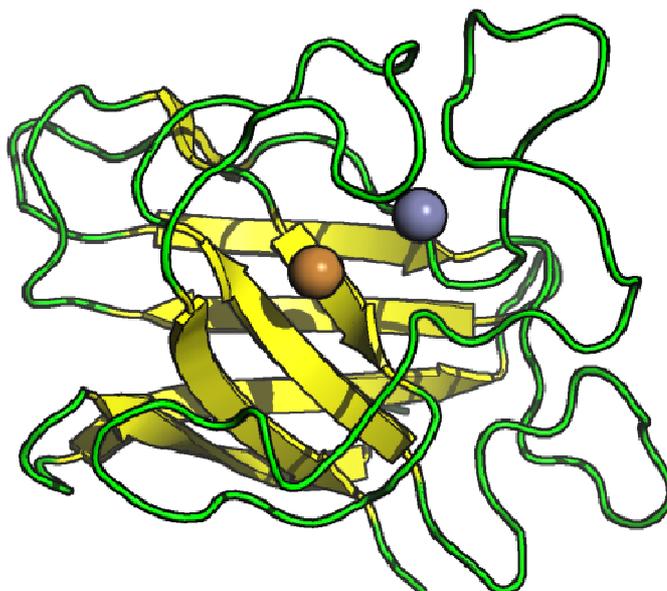


Рисунок 1. Модель структуры 1XTM в представлении cartoon.

Выделяют несколько типов СОД в зависимости от входящих в их активный центр ионов металлов [1]:

1. Cu/Zn-SOD – в большинстве своём эукариотические
2. Fe-SOD, Mn-SOD – встречаются у протистов, прокариот и в митохондриях
3. Ni-SOD – прокариотические

Cu/Zn-SOD встречаются и у некоторых бактерий. Всего среди 138 геномов архей и бактерий найдено 57 предположительных гомологов SOD. Большинство таких бактериальных СОД сохраняют Zn- и Cu-связывающие мотивы и являются полностью функционирующими ферментами. Однако у некоторых бактерий такие СОД, например BsSOD у *Bacillus subtilis*, не связывают 2 лиганда меди и не способны функционировать [2].

Группа Viezzoli, авторы модели структуры 1XTM, создали двойной мутант BsSOD по остаткам Tyr88 и Pro104, заменив их на гистидины, тем самым создав частично активную форму BsSOD. Для структурной характеристики получившегося белка авторы расшифровали его структуру методом рентгеноструктурного анализа (РСА). Интересно, что в растворе данный белок структурно неупорядочен, в то время как в кристалле имеет определённую третичную структуру и образует димеры путём координации на интерфейсе иона цинка [2].

## Результаты и обсуждение

### *Общая информация о модели*

В 2004 году 22 октября в банке данных PDB было размещено 2 структуры 1XTM и 1XTL. Авторами модели 1XTM являются Calderone, V., Mangani, S., Banci, L., Benvenuti, M., Bertini, I., Fantoni, A., Viezzoli, M.S [3]. Фазовая проблема была решена методом молекулярного замещения с помощью структуры 1S4I (нативная форма белка) [2].

В состав модели входит 2 почти идентичные цепи белка А и В, а также 5 ионов цинка и 2 иона меди, по одному из которых связаны с цепью А, а остальные - с В.

В таблице Таблица 1 приведены параметры кристаллографической ячейки. Пространственная группа относится к группе  $P_{2_1}2_12$  с 2 молекулами в асимметрической единице [3], [2], [4]. Асимметрическая единица некристаллографической симметрией не обладает.

**Таблица 1.** Параметры кристаллографической ячейки 1XTM

Длина (Å)	Угол (°)
a = 52.46	$\alpha = 90.00$
b = 104.35	$\beta = 90.00$
c = 58.76	$\gamma = 90.00$

Ниже приведена информация об экспериментальных данных [4]:

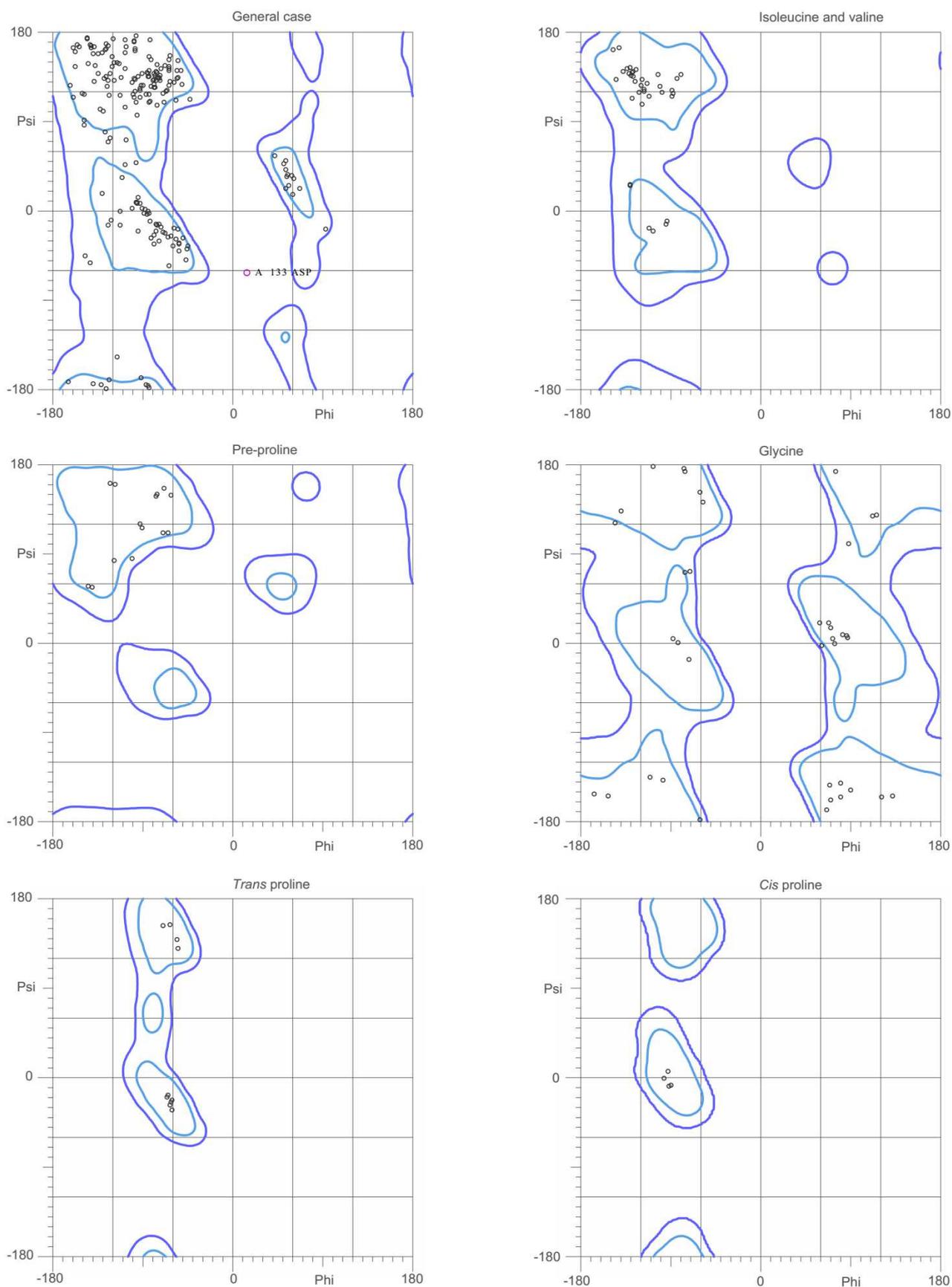
- число измеренных рефлексов: 81899
- разрешение: 1.60 Å
- полнота набора структурных факторов: 99.5 %
- диапазон разрешений структурных факторов: 25.60 - 1.60 Å

### *Индикаторы качества модели в целом*

Значения индикаторов качества модели в целом структуры 1XTM следующие [4]:

- R-фактор: 0.254
- R-free: 0.266
- Среднее RSR: 0.138
- Среднее температурного фактора, взвешенного по заселённости: 26.0 Å<sup>2</sup>

С помощью сервиса MolProbity к структуре были добавлены атомы водорода и выполнены предложенные инверсии остатков (см. таблицу Таблица 2). Также была построена карта Рамачандрана (см. Рисунок 2). Данным сервисом определено, что 92% (277/301) остатков находятся в предпочтительных областях, а 99.7% (300/301) лежат в разрешённых областях. MolProbity также предоставляет некоторые характеристики качества структуры, указанные в таблице Таблица 3.



**Рисунок 2.** Карты Рамачандрана структуры 1XTM, полученные MolProbity. Голубым выделены предпочтительные области значений углов  $\phi$  и  $\psi$ , синим - разрешённые. Фиолетовым выделен маргинальный аспарат 133.

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

Цепь	Номер остатка	Тип остатка
B	49	ASN
B	75	ASN
B	127	ASN
A	75	ASN
A	127	ASN
A	179	ASN

**Таблица 3.** Характеристики модели, полученные с помощью MolProbity.

All-Atom Contacts	Clashscore, all atoms:	13.3		42 <sup>nd</sup> percentile* (N=718, 1.60Å ± 0.25Å)
Protein Geometry	Poor rotamers	21	8.71%	Goal: <0.3%
	Favored rotamers	195	80.91%	Goal: >98%
	Ramachandran outliers	1	0.33%	Goal: <0.05%
	Ramachandran favored	277	92.03%	Goal: >98%
	MolProbity score <sup>^</sup>	2.83		4 <sup>th</sup> percentile* (N=7200, 1.60Å ± 0.25Å)
	Cβ deviations >0.25Å	4	1.51%	Goal: 0
	Bad bonds:	4 / 2324	0.17%	Goal: 0%
	Bad angles:	17 / 3138	0.54%	Goal: <0.1%
Peptide Omega	Cis Prolines:	4 / 14	28.57%	Expected: ≤1 per chain, or ≤5%
	Cis nonProlines:	8 / 289	2.77%	Goal: <0.05%
	Twisted Peptides:	5 / 303	1.65%	Goal: 0

### Маргинальные остатки

На основе данных MolProbity и Full wwPDB X-ray Structure Validation Report [5] была составлена таблица нескольких маргинальных остатков структуры 1XTM (см. Таблица 4).

**Таблица 4.** Некоторые маргинальные остатки структуры 1XTM. Жирным шрифтом отмечены остатки, подвергнутые дальнейшему анализу.

Остаток	Причина маргинальности
<b>A:Asp65</b>	Очень высокое значение RSRZ (7.3), нехарактерный угол CB-CG-OD2 (124° вместо 118°)
<b>A:Asn75</b>	Нехарактерный ротамер
<b>A:Gly119</b>	Короткая связь C-N (0.76 вместо 1.34Å), нехарактерные углы O-C-N (103 вместо 122°), C-N-CA (147 вместо 121°) и CA-C-N (136 вместо 117°), аутлайер по плоскости пептидной группы, высокое значение RSRZ (2.8)
<b>A:His120</b>	Длинная связь C-N (1.78 вместо 1.34Å), нехарактерные углы

	CA-C-N (93 вместо 117°) и CA-C-O (138 вместо 120°), высокое значение RSRZ (3.3)
A:Asn127	Нехарактерный ротамер
<b>A:Asp133</b>	Запрещённые торсионные углы ( $\phi=14.4$ , $\psi=-62.1$ )
B:Pro143	Длинная связь C-N (1.66 вместо 1.34Å)
B:Asp144	Длинная связь C-N (1.50 вместо 1.34Å), нехарактерный угол CB-CA-C (121 вместо 110°), высокое значение RSRZ (2.0)

### Анализ маргинальных остатков

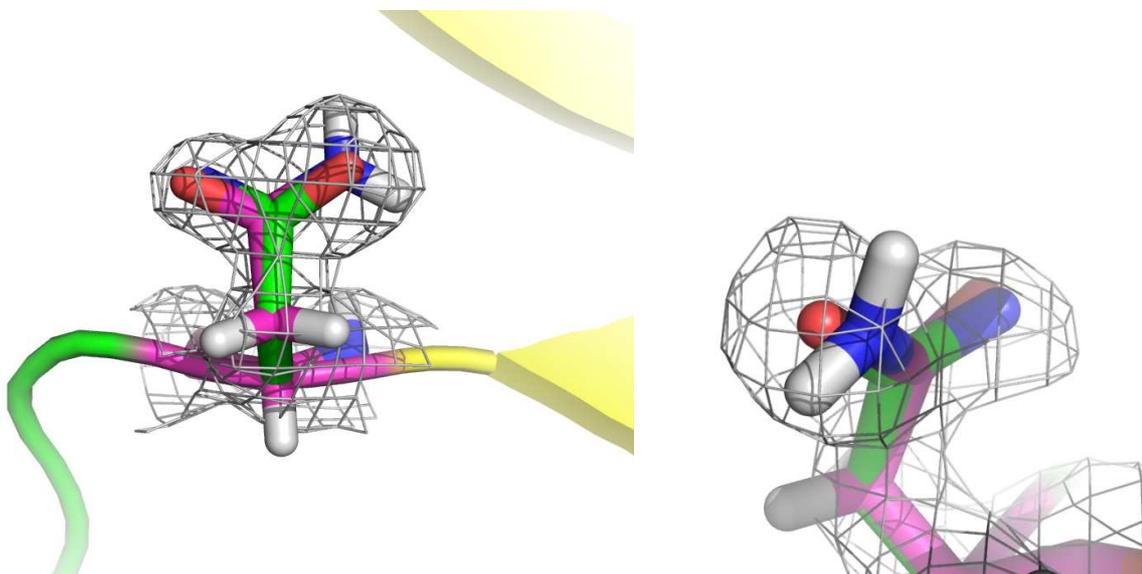
Для анализа были выбраны остатки, выделенные жирным шрифтом в таблице Таблица 4.

#### 1. Аспарат 65 цепи А

Данный остаток имеет очень плохое значение  $RSR=0.475$  (не больше 0.2 в норме), т.е. он плохо соответствует экспериментальной электронной плотности. Значение температурного фактора для этого остатка - 50.8, что довольно много. Данный случай подробно рассмотрен в практикуме: [http://kodomofbb.msu.ru/~popov/term7/ED\\_show/index.html](http://kodomofbb.msu.ru/~popov/term7/ED_show/index.html). Этот остаток явно является слабым местом экспериментальных данных, поэтому сложно говорить о качестве расшифровки.

#### 2. Аспарагин 75 цепи А

По версии MolProbity необходимо инвертировать боковую цепь данного остатка.



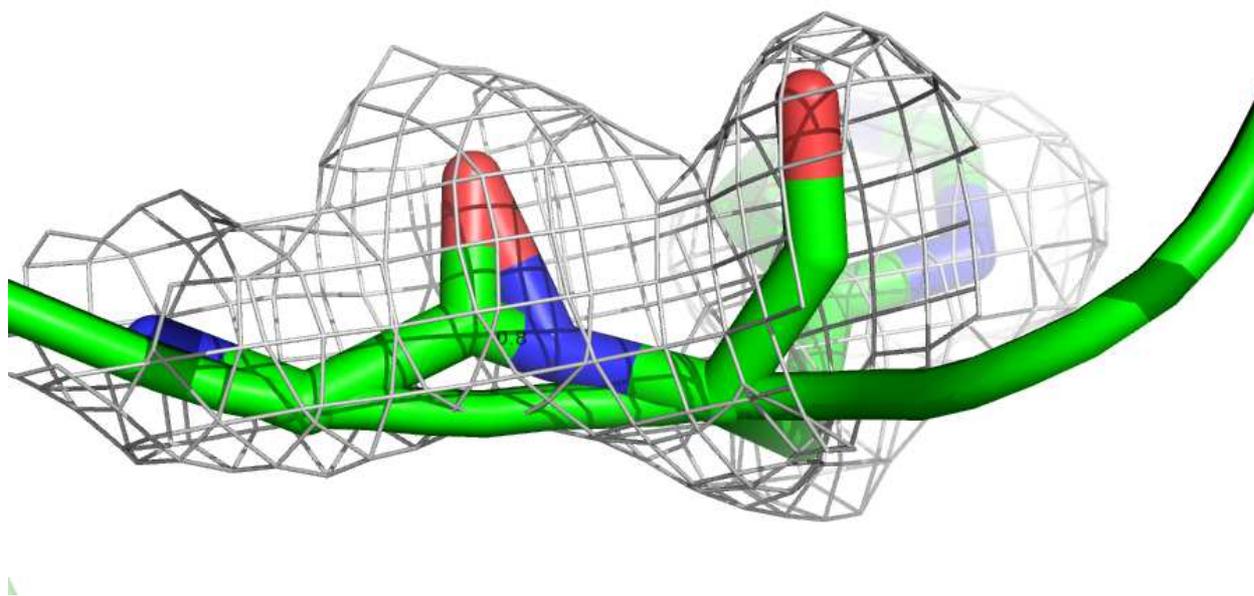
**Рисунок 3.** Соответствие аспарагина 75 электронной плотности. Слева общий вид остатка, справа - его боковая цепь. Зелёным изображена исходная модель, фиолетовым - инвертированная модель с добавленными водородами.

На рисунке Рисунок 3 видно, что электронная плотность (ЭП) боковой цепи не симметрична, и для одного из атомов O и N она больше, чем для другого. С одной стороны, атом кислорода больше атома азота, поэтому логично предположить, что "большая" ЭП принадлежит кислороду, а "маленькая" - азоту. С другой стороны, у азота есть 2 атома водорода, у которых также должна быть своя ЭП. Однако на рисунке Рисунок 3 справа видно,

что добавленные MolProbity атомы водорода не вписываются в экспериментальную ЭП. Таким образом, нельзя утверждать, что данный остаток неправильно расшифрован, т.к. нужна электронная плотность с бóльшим разрешением.

### 3. Глицин 119 цепи А

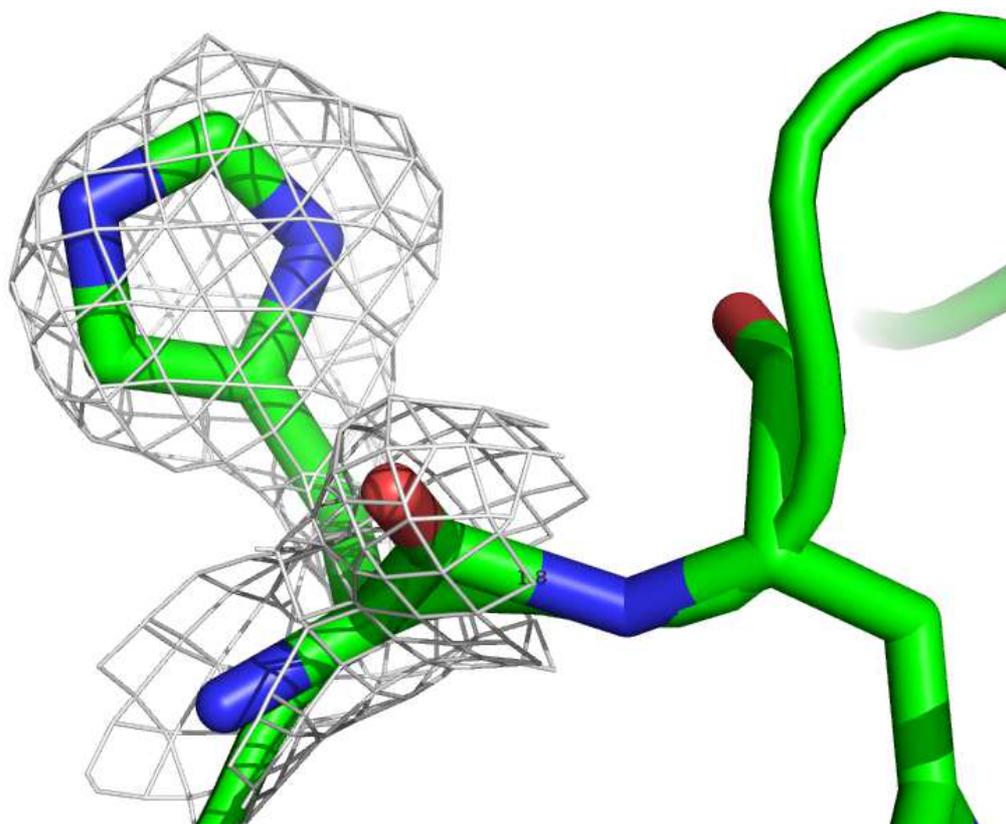
Данный остаток является маргинальным по целому ряду параметров: короткая связь C-N (0.76 вместо 1.34Å), нехарактерные углы O-C-N (103 вместо 122°), C-N-CA (147 вместо 121°) и CA-C-N (136 вместо 117°), аутлайер по плоскости пептидной группы, высокое значение RSRZ (2.8). Уже по рисунку Рисунок 4 видно, что причиной этому является неправильна расшифровка координат атомов глицина: атомы пептидной группы настолько сближены, что в отрисовке РуMol между атомами O и N появляется ковалентная связь. Для улучшения качества нужно хотя бы "сдвинуть" правее остаток His120, тем самым удлинив связь C-N и улучшив соответствие модели гистидина его ЭП (см. ниже).



**Рисунок 4.** Соответствие глицина 119 и гистидина 120 электронной плотности. Подписана длина связи C-N между этими остатками.

### 4. Гистидин 120 цепи А

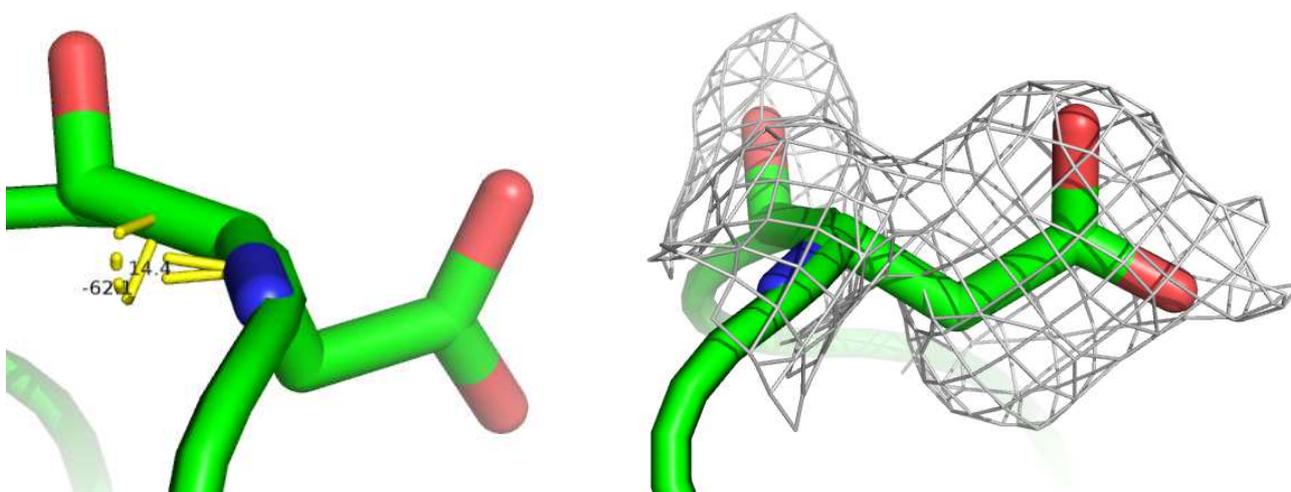
Данный остаток имеет ряд несоответствий: длинная связь C-N (1.78 вместо 1.34Å), нехарактерные углы CA-C-N (93 вместо 117°) и CA-C-O (138 вместо 120°), высокое значение RSRZ (3.3). На рисунке Рисунок 5 видно, что боковая цепь гистидина не совсем точно вписана в ЭП, а пептидная связь выглядит действительно очень длинной. Устранить данные несоответствия можно было бы, как описано в предыдущем пункте, "сдвинув" гистидин правее.



**Рисунок 5.** Соответствие гистидина 120 электронной плотности. Подписана длина связи C-N между ним и гистицином 121.

### 5. Аспаргат 133 цепи А

Данный остаток находится в запрещённой области на карте Рамачандрана и имеет торсионные углы  $\phi=14.4$ ,  $\psi=-62.1$  (см. Рисунок 6 слева).



**Рисунок 6.** Торсионные углы  $\phi$ ,  $\psi$  аспартата 133 (слева) и соответствие аспартата 133 ЭП (справа).

Данный остаток находится в петле на поверхности белка, и поэтому, скорее всего, совершает колебания, о чём может говорить его размытая ЭП (см. Рисунок 6 справа). Вероятно,

данное несоответствие торсионных углов предпочтительным значением не является биологически значимым, а является порождённой экспериментальными данными ошибкой.

### *Сравнение с моделью из PDB\_redo*

На сервере PDB\_REDO размещена улучшенная модель структуры 1XTM [6], сравнение параметров качества которой с исходной моделью представлено в таблице Таблица 5.

**Таблица 5.** Сравнение индикаторов качества моделей из PDB и PDB\_redo. Зелёным выделены улучшенные параметры, жёлтым - незначительно лучшие, красным - худшие.

Характеристика	Модель из PDB	Полностью оптимизированная модель
R	0.2530	0.2295
R-free	0.2660	0.2571
Показатель карты Рамачандрана	-1.660	-0.859
Нормальность $\chi$ -1/ $\chi$ -2 ротамеров	-3.084	-1.366
Конформация остова	-1.796	-1.648
RMS Z-score длин связей	0.725	0.777
RMS Z-score углов связей	0.999	0.998
Ненасыщенные доноры/акцепторы водородных связей	15	14
Улучшенные остатки	0	104
Ухудшенные остатки	0	2

В целом, оптимизированная модель обладает улучшенными индикаторами качества.

На рисунке Рисунок 7 представлено сравнение моделей структуры 1XTM из PDB и PDB\_redo. В оптимизированной модели присутствуют более длинные  $\beta$ -тяжи, а некоторые петли лежат немного по-другому.

Глицин 119 цепи А, маргинальный по ряду параметров в исходной модели, в оптимизированной модели выглядит куда более приличней (см. Рисунок 8). Теперь, длина связи C-N составляет, как и положено, 1.3 Å, что говорит о действительном улучшении модели структуры 1XTM.

## Заключение

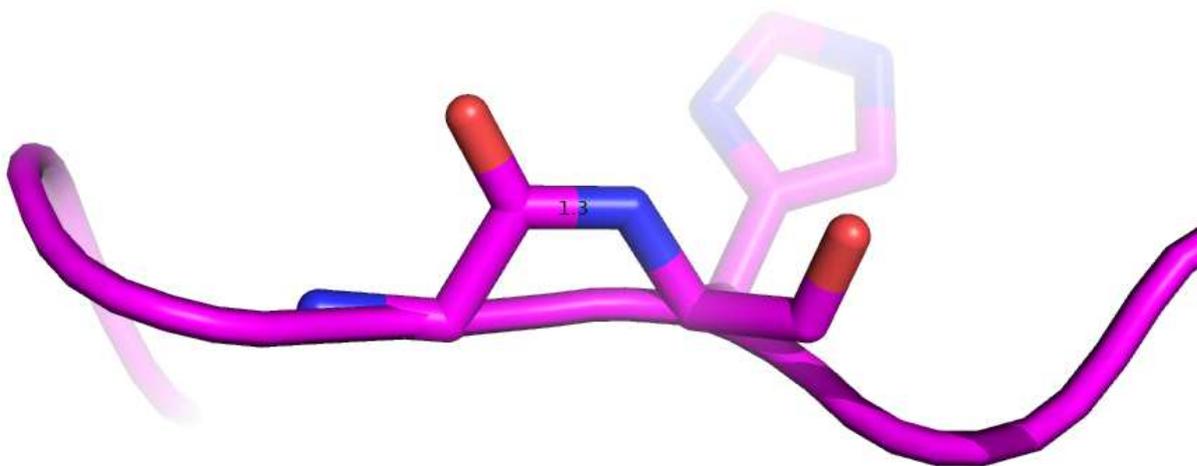
Не смотря на высокое разрешение структуры 1XTM в 1.6 Å, её качество является далеко не идеальным, а в её модели присутствуют маргинальные остатки, причиной маргинальности которых может быть не один отклоняющийся показатель, а целый ряд нехарактерных параметров. Скорее всего, это является следствием не точно определённой ЭП.

С помощью сервиса PDB\_redo данная модель была значительно улучшена, что привело к более точному определению вторичной структуры белка и оптимизации конформации некоторых аминокислотных остатков.

Нельзя забывать, что целью авторов структуры 1XTM не была лишь расшифровка структуры белка. К тому же в статье, в которой представлена данная модель, представлена ещё и структура 1XTL. Поэтому авторы не могли уделить всё своё внимание лишь уточнению модели 1XTM, качество которой и так можно описать как более чем приемлемое.



**Рисунок 7.** Наложение исходной (оранжевый) и оптимизированной (серый) структуры 1XTM.



**Рисунок 8.** Оптимизированные глицин 119 и гистидин 120.

## Список литературы

- [1] “Superoxide dismutase.” [Online]. Available: [https://en.wikipedia.org/wiki/Superoxide\\_dismutase](https://en.wikipedia.org/wiki/Superoxide_dismutase).
- [2] L. Banci, M. Benvenuti, I. Bertini, D. E. Cabelli, V. Calderone, A. Fantoni, S. Mangani, M. Migliardi, and M. S. Viezzoli, “From an inactive prokaryotic SOD homologue to an active protein through site-directed mutagenesis.,” *Journal of the American Chemical Society*, vol. 127, no. 38, pp. 13287–13292, Sep. 2005.
- [3] “RCSB PDB for 1XTM.” [Online]. Available: <http://www.rcsb.org/pdb/explore/explore.do?structureId=1xtm>.
- [4] “EDS: 1xtm.” [Online]. Available: <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=1xtm>.
- [5] “Full wwPDB X-ray Structure Validation Report.” [Online]. Available: [http://ftp.wwpdb.org/pub/pdb/validation\\_reports/xt/1xtm/1xtm\\_full\\_validation.pdf](http://ftp.wwpdb.org/pub/pdb/validation_reports/xt/1xtm/1xtm_full_validation.pdf).
- [6] “1xtm - PDB\_REDO optimisation results.” [Online]. Available: [http://www.cmbi.ru.nl/pdb\\_redo/xt/1xtm/index.html](http://www.cmbi.ru.nl/pdb_redo/xt/1xtm/index.html).