
Оценка качества расшифровки структуры комплекса MnSOD-азид методом РСА (PDB ID 5T30)

Отчет

Марина Гладкова - 12 января 2019 г.

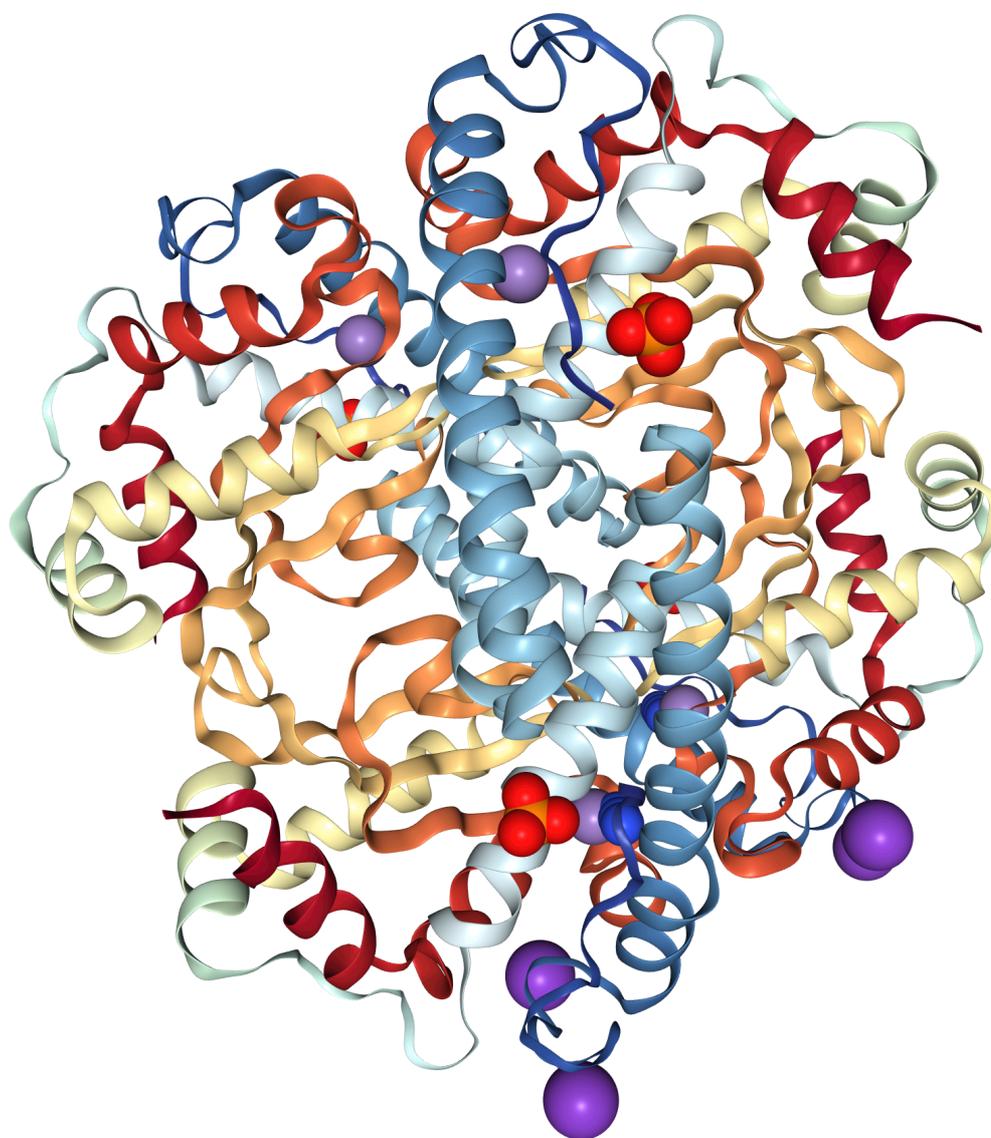


Рис. 1. Структура биологически активной MnSOD в комплексе с азидом.

Аннотация

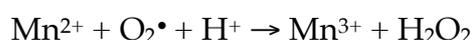
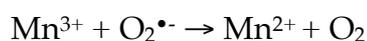
Цель работы - оценка качества расшифровки структуры человеческой магниевой супероксиддисмутазы MnSOD в комплексе с азидом (PDB ID 5T30), полученной методом рентгеноструктурного анализа (РСА) авторами . Были рассмотрены основные показатели качества модели в целом и несколько примеров маргинальных остатков.

Введение

Супероксиддисмутаза (СОД) - фермент, играющий ключевую роль в защите клеток от токсичных метаболитов кислорода. Они способствуют диспропорционированию двух молекул супероксида до молекулярного кислорода и перекиси водорода путем циклического восстановления и окисления в металле активного центра. У эукариот MnSOD обнаруживается в митохондриальном матриксе и устраняет образующиеся в результате функционирования ЭТЦ супероксидные радикалы [1].

Человеческий фермент, представляющий собой гомотетрамер, кодируется геном *SOD2*, расположенным на 6 хромосоме. Мутации в этом гене связаны с идиопатической кардиомиопатией (IDC), преждевременным старением, спорадической болезнью двигательных нейронов и развитием онкологических заболеваний. Альтернативный сплайсинг этого гена приводит к множественным транскриптам. Кроме того, был идентифицирован родственный псевдоген на хромосоме 1 [2, 3].

Азид-анион является мощным конкурентным ингибитором, который связывается непосредственно с металлом и используется в качестве субстратного аналога супероксида в исследованиях СОД. Ниже приведена общая схема ферментативной реакции MnSOD:



В статье [4] описана кристаллическая структура человеческого MnSOD-азидного комплекса. Было определено предполагаемое положение связывания супероксида, что стало основой для создания модели связывания субстрата с активным центром. Азид связывает ион марганца в шестом координационном положении. Были рассчитаны тетрамерные электростатические с учетом точных частичных зарядов для активного центра в трех состояниях, включая комплекс с супероксидом, представлена модель электростатически-опосредованной диффузии и эффективного связывания супероксида для катализа.

Результаты и обсуждение

Общая информация о модели

Состав комплекса: 2 полипептидные цепи (А и В, каждая длиной 199 а.о.) человеческой MnSOD в комплексе с азидом, выступающим в качестве аналога супероксида. Лиганды: K^+ (цепь А), Mn^{2+} (цепи А, В), PO_4^{3-} (цепи А, В), N_3 (цепь А). Число атомов белка — 3162, гетероатомов — 17, атомов растворителя — 354.

Модель была опубликована в PDB 17 мая 2017 года. Разработана группой авторов из института Эппли, США (J. Azadmanesh, S.R. Trickle, G.E.O. Borgstahl, C.H. Kolar, J.J. Lovelace, the Eppley Institute for Research in Cancer and Allied Diseases).

Разрешение модели: 1.77 Å. Число измеренных рефлексов: 40061 (42583 в оригинальном файле CIF). Полнота набора структурных факторов — 94.1 %. Диапазон разрешений: 67.46 — 1.77 Å (согласно депозитору); 38.95 – 1.77 Å (согласно EDS).

Метод решения фазовой проблемы — молекулярное замещение. Нативная MnSOD и MnSOD-азидный комплекс были решены с использованием банка данных белка 1JA8, имеющим одинаковые с исследуемым комплексом параметры элементарной ячейки и пространственной группой [5]. После удаления растворителя и металлов из активного центра было выполнено простое молекулярное замещение посредством rigid body refinement с использованием REFMAC5 (с разрешением 1,82 Å и 1,77 Å для нативной структуры и комплекса соответственно) [6]. С помощью программы Coot были проанализированы карты электронной плотности и построена белковая модель [7].

Параметры кристаллографической ячейки: $a = 77.90 \text{ \AA}$, $b = 77.90 \text{ \AA}$, $c = 238.44 \text{ \AA}$, $\alpha = 90.00^\circ$, $\beta = 90.00^\circ$, $\gamma = 120.00^\circ$. Тип кристаллографической симметрии — $P 6_1 2 2$. Некристаллографических симметрий в асимметрической ячейке нет. Биологическая ячейка включает 2 асимметрические.

Индикаторы качества модели

Для оценки качества модели наиболее часто используются следующие индикаторы:

- 1) R-фактор. Итеративно минимизируется в процессе оптимизации, отражает соответствие модели экспериментальным данным;
- 2) R free (свободный R-фактор). Является несмещенной оценкой, рассчитывается аналогично R-фактору на некотором небольшом количестве

случайно выбранных рефлексов, которые не включались в оптимизацию, позволяет учесть переоптимизацию. Хороший $R_{free} < 25\%$, плохой $> 40\%$. Если $(R_{free} - R) > 10\%$, высока вероятность переподгонки модели.

3) Карта Рамачандрана. Показывает распределение аминокислотных остатков в соответствии с торсионными углами и углами связей остова. Используется для оценки конформации остова, как правило, не зависит от метода оптимизации. Ramachandran outliers — доля маргинальных остатков соответственно карте Рамачандрана;

4) Clashscore — критерий перекрытия Ван-дер-Ваальсовых радиусов (более чем на 0.4 \AA) атомов в расчете на 1000 атомов;

5) Sidechain outliers — доля маргинальных остатков боковой цепи, вычисляется в соответствии с торсионными углами и ротамерами (наиболее выгодными конформациями);

6) RSR (RSRZ) — Real space R. Характеризует, насколько модель группы атомов (или даже отдельного атома) соответствует «экспериментальной» электронной плотности. Z-score - относительная оценка RSR. Показывает, на сколько стандартных отклонений RSR для остатка отличается от среднего RSR для такого же типа остатков в структурах PDB со сходным разрешением.

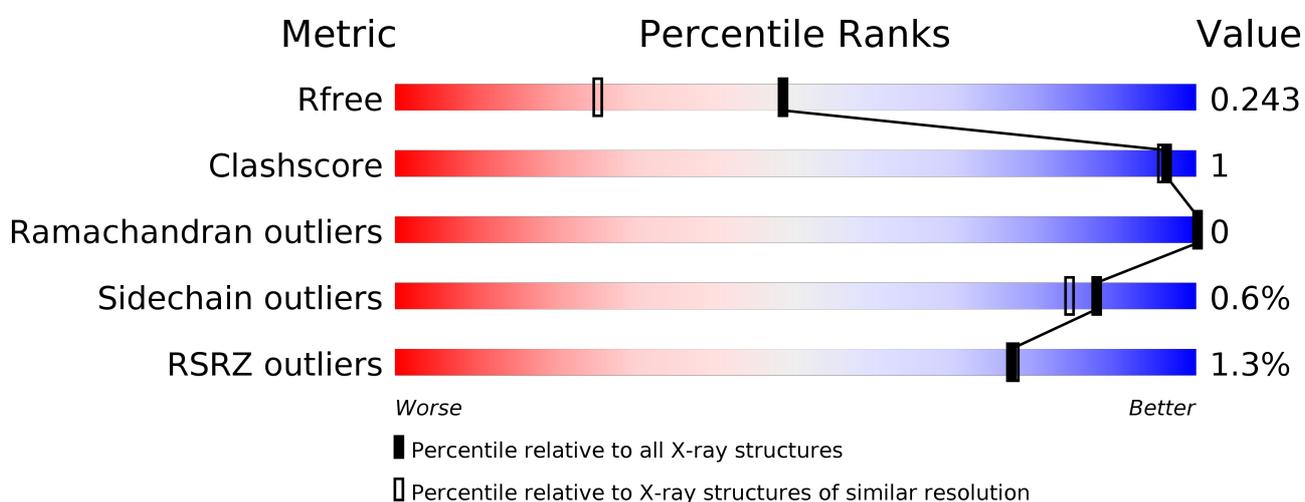


Рис. 2. Основные показатели качества модели из PDB Validation Report [8]. Градиентом показано, насколько параметры данной структуры лучше (синий) или хуже (красный) средних значений (белый). Закрашенные прямоугольники отражают качество относительно всех структур, прозрачные — относительно структур со схожим разрешением.

Табл. 1. Значения R-факторов для 5T30.

Параметр	Значение
R-фактор	0.209 (согласно депозитору), 0.216 (согласно DCC)
R free	0.239 (согласно депозитору), 0.243 (согласно DCC)
Доля структурных факторов, по которым вычисляли R free	1964 (4.9%)

Согласно табл. 1, оба R-фактора попадают в предельные значения из допустимой границы. Их разница составляет 0.03 (0.027), что составляет 12.5% (11.1%) от значения R free и говорит о приемлемой обобщающей способности модели при отсутствии переобучения. Однако стоит отметить, что показатель R free данной модели в сравнении с факторами других структур значительно уступает.

Как видно из рис. 2, выбранная модель является достаточно хорошей по основным критериям качества. Согласно отчету PDB [8] в параметрах стандартной геометрии процент маргинальных находок ~ 0%.

Для более подробного анализа был использован сервис MolProbity [9]. При валидации структуры он добавляет атомы водорода, оптимизируя водородные связи, и производит при необходимости инверсии боковых цепей аспарагина, глутамина и гистидина.

Неправильно ориентированных а.о. в структуре 5T30 не нашлось, но было добавлено 3096 атомов H, позиции 73 из них были оптимизированы для создания наиболее устойчивой сети водородных связей.

All-Atom Contacts	Clashscore, all atoms:	1.12	100 th percentile* (N=841, 1.77Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	0	0.00% Goal: <0.3%
	Favored rotamers	314	96.02% Goal: >98%
	Ramachandran outliers	0	0.00% Goal: <0.05%
	Ramachandran favored	384	97.46% Goal: >98%
	MolProbity score [^]	0.93	100 th percentile* (N=11232, 1.77Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.55% Goal: 0
	Bad bonds:	2 / 3275	0.06% Goal: 0%
Bad angles:	0 / 4448	0.00% Goal: <0.1%	
Peptide Omegas	Cis Prolines:	2 / 20	10.00% Expected: ≤1 per chain, or ≤5%

In the two column results, the left column gives the raw count, right column gives the percentage.

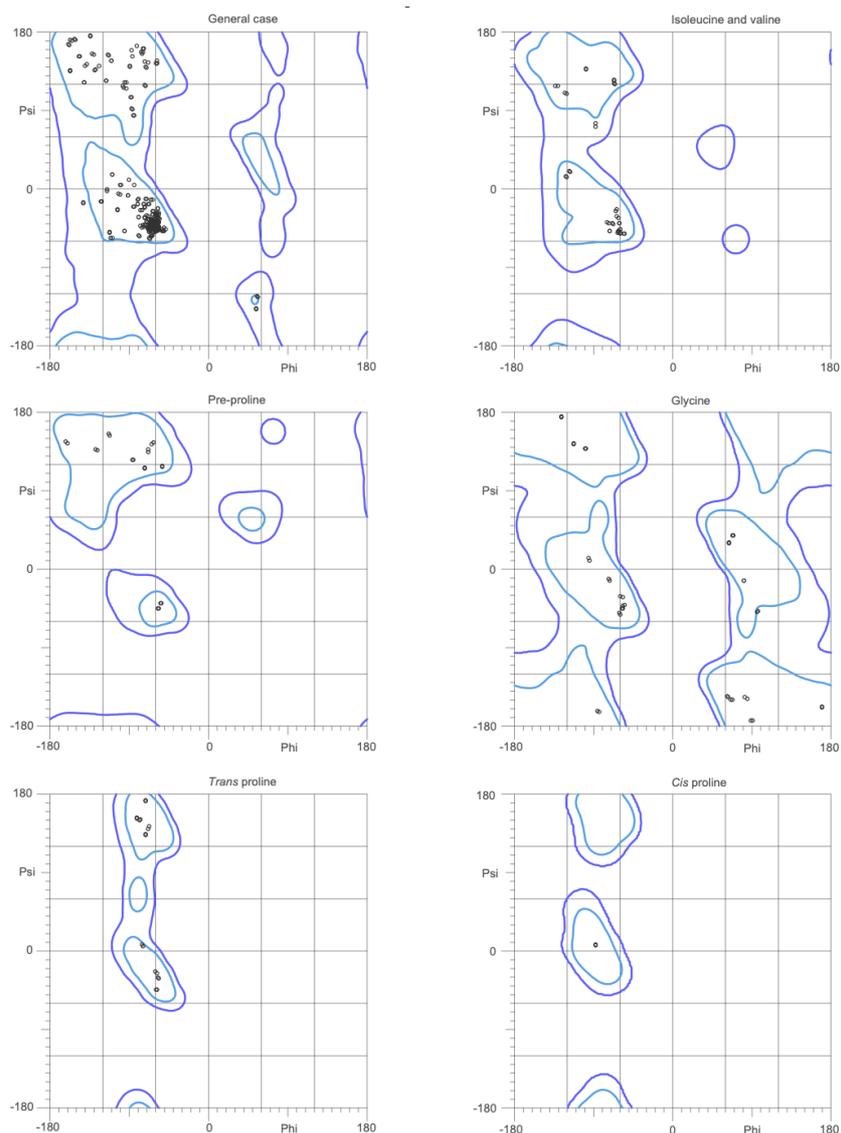
* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

[^] MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Рис. 3. Выдача сервиса MolProbity для 5T30.

Как можно видеть по рис. 3, clashscore составляет 1.12 и является 100-ым перцентилем, что соответствует наилучшему качеству по данному параметру среди

моделей со сходным разрешением (1.75 ± 0.25 Å). Большинство ротамеров (96.02%) находятся в предпочитаемых для них конформациях, что немного меньше порога (>98%). Модель содержит две неоптимальные по длине связи, две цис-пептидные связи с участием пролина, 2 C β -атома с неприемлемым отклонением от ожидаемого положения. По торсионным углам (рис. 4) маргинальных остатков обнаружено не было. MolProbity score — суммарная оценка модели этим сервисом. Модель 5T30 является 100-ым перцентилем, и по данному интегральному критерию.



97.5% (385/395) of all residues were in favored (98%) regions.

100.0% (395/395) of all residues were in allowed (>99.8%) regions.

There were no outliers.

Рис. 4. Построенные MolProbity карты Рамачандрана для 5T30. Голубые линии — границы предпочитаемых областей, синие — допустимые.

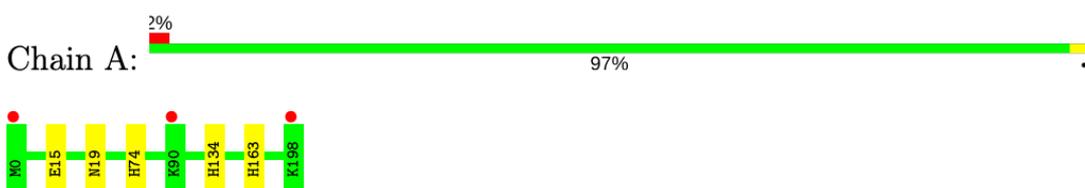
Маргинальные остатки

В табл. 2 показаны некоторые маргинальные остатки, которые были отмечены в отчете PDB [8] или сервисом MolProbity [9].

Табл. 2. Примеры маргинальных остатков в 5T30.

Аминокислотный остаток	Цепь	Обоснование маргинальности
HIS134	A, B	Неблагоприятная конформация боковых цепей ($\chi_1 = 65.4$, $\chi_2 = 210.1$)
HIS163	A	Clashscore = 0.55 Å, перекрывание CE1 с A 205 AZI N3
VAL116	B	Clashscore = 0.58 Å, перекрывание HG23 с B 177 LEU CD1; неблагоприятная конформация боковых цепей ($\chi = 154.4^\circ$)
ARG99	B	RSRZ = 2.6; выброс по углу NE-CZ-NH1 (112.92°)
VAL172	A, B	Из разрешенной области Рамачандрана ($\phi = -88.4$, $\psi = 76.6$); неблагоприятная конформация боковых цепей ($\chi = 198^\circ$)
ASN19	A*, B	Clashscore* = 0.46 Å, перекрывание HA with A 15 GLU HG2; C β deviation = 0.27 Å

- Molecule 1: Superoxide dismutase [Mn], mitochondrial



- Molecule 1: Superoxide dismutase [Mn], mitochondrial

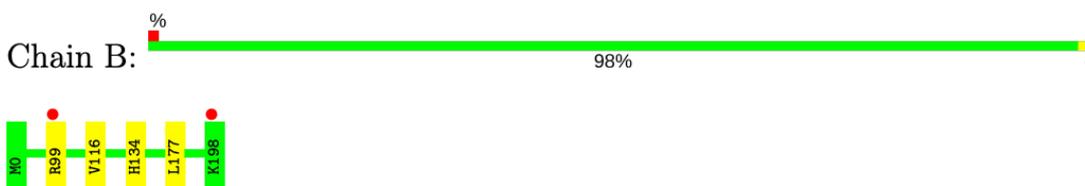


Рис. 5. График качества остатков из отчета PDB [8]. Последовательность остатков отражает их относительное положение в полипептидной цепи, цветом показаны критерии качества. Зеленый — хорошие остатки, желтый — маргинальные по 1 признаку и т.д. Красная точка — признак плохого показателя RSRZ (>2).

Рассмотрим выделенные в таблице аминокислоты подробнее. Поскольку при выполнении практикума 1 для структуры 5Т30 оптимальным для отображения электронной плотности уровнем подрезки был определен $Z = 2$, при дальнейшей визуализации и анализе будет использоваться это значение. На рис. 6а, 6б показан ротамер His134.

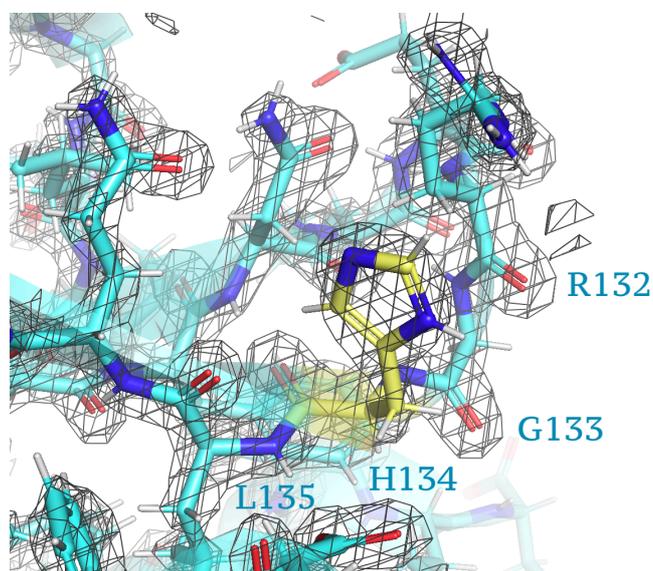


Рис. 6а. Изображение His134 и его окружения, ЭП при $Z = 2$.

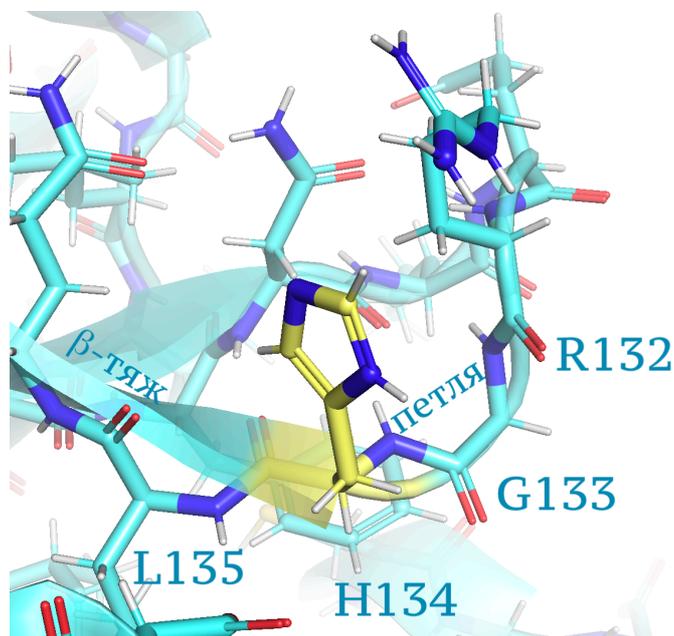


Рис. 6б. Изображение His134 и его окружения, отмечены элементы вторичной структуры.

Он идеально покрыт электронной плотностью и визуально не выглядит изуродованным или насильно вписанным в модель. Видно, что этот остаток находится на границе β -тяжа, который сменяется петлей. Вероятно, отмеченные выше значения углов χ являются особенностью данного гистидина в условиях формирования структуры петли и могут объясняться стерическими затруднениями в тесном окружении.

На рис. 7 изображен His163. Поскольку данный остаток формирует активный центр фермента, нет ничего удивительного в том, что его clashscore превышает стандартное значение.

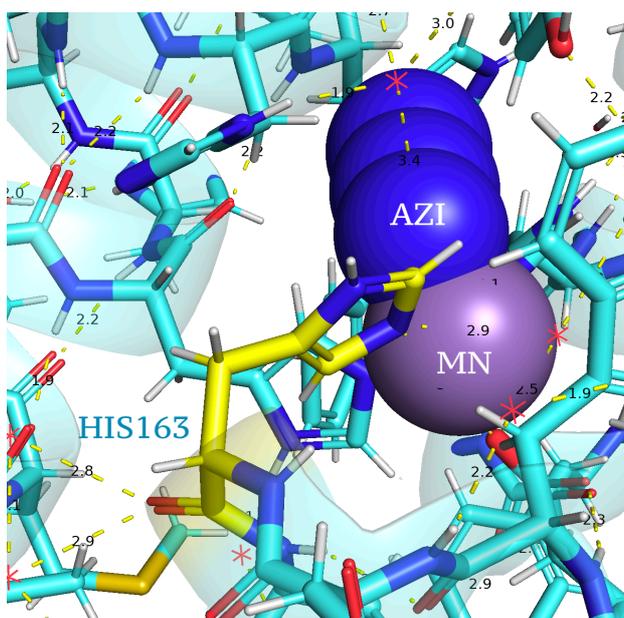


Рис. 7а. Изображение His163, формирующего активный центр.

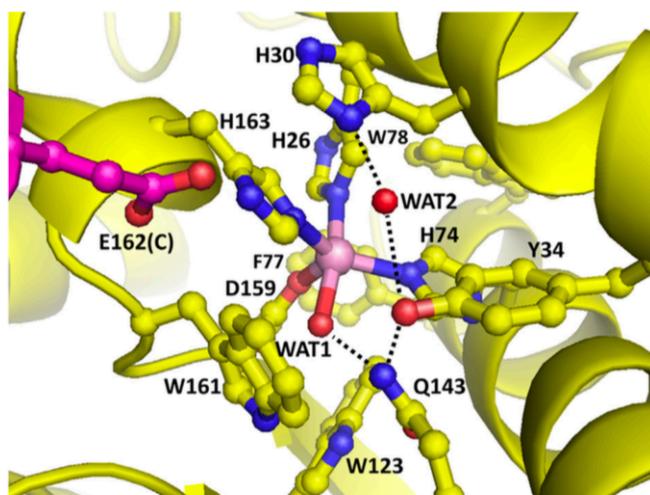


Рис. 7б. Иллюстрация активного центра в статье [4].

На рис. 8a, 8b показаны маргинальный остаток Val116 и перекрывающийся с ним остаток Leu177. На уровне подрезки $Z = 2$, видно, что боковая цепь валина практически «голая» в плане электронной плотности, а лейцин также покрыт не полностью. Возможно, это и является причиной неправильной взаимной ориентации, приводящей к перекрыванию ван-дер-ваальсовых радиусов, а также проблемами с χ -углами в самом валине. Скорее всего, данный маргинал — следствие ошибки расшифровки.

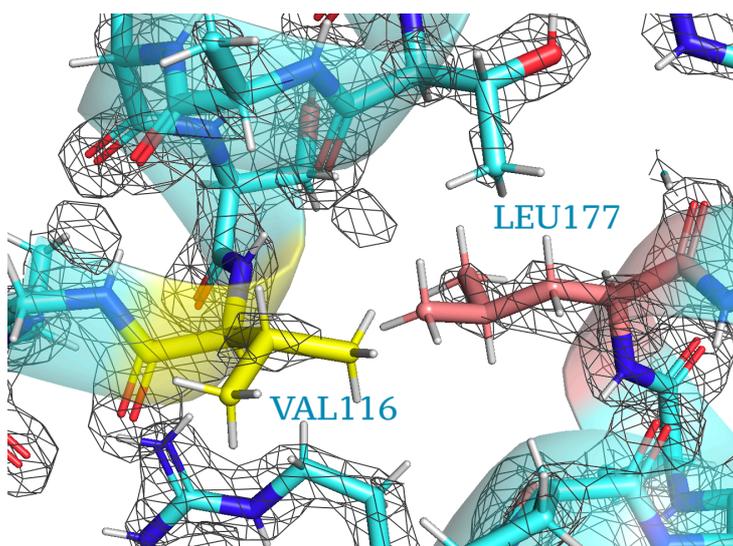


Рис. 8a. Изображение Val116, Leu177, ЭП при $Z = 2$.

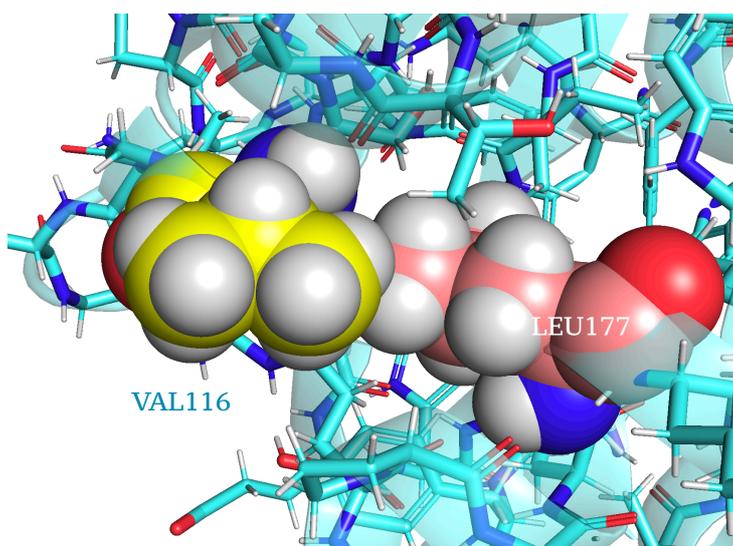


Рис. 8b. Изображение Val116, Leu177 и их ван-дер-ваальсовых радиусов.

Рис. 9а, 9б показывают еще один маргинальный остаток — Arg99 при разных уровнях подрезки. Как можно видеть, при $Z = 2$, электронная плотность одинаково плохо покрывает как остовные атомы, так и боковую цепь. При $Z = 1$ ситуация немного лучше, однако все еще остаются «оголенные» области. Чем может быть вызвано нестандартное значение угла NE-CZ-NH1 понять сложно, поскольку данный остаток находится в боковом положении относительно α -спирали и не имеет явных стерически затрудненных положений. Я склонна думать, что этот маргинал, тоже является артефактом интерпретации структуры при создании модели.

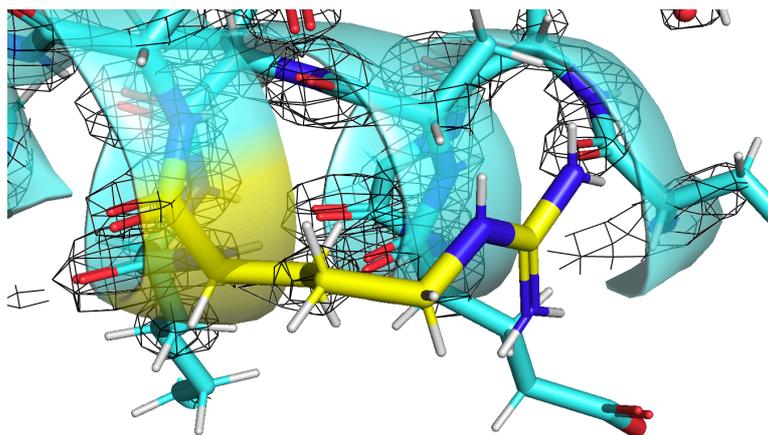


Рис. 9а. Изображение Arg99, ЭП при $Z = 2$.

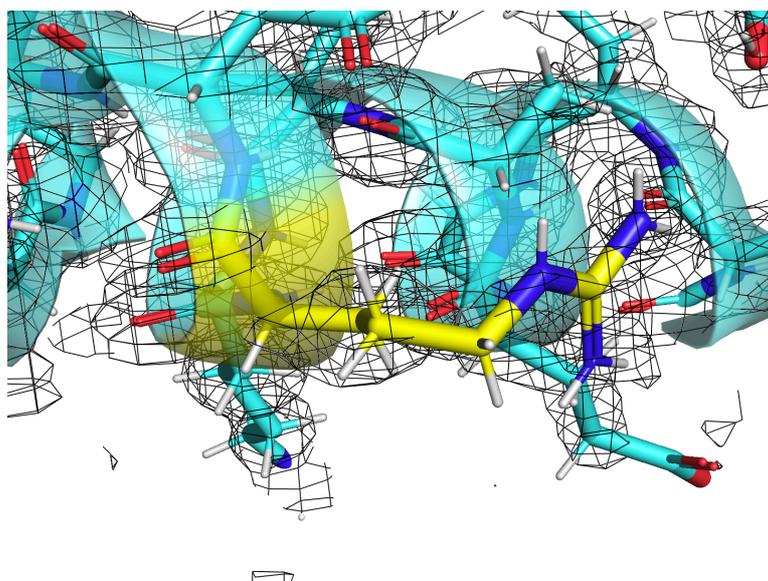


Рис. 9б. Изображение Arg99, ЭП при $Z = 1$.

На рис. 10а, 10b можно видеть Asn19, маргинальный по двум параметрам. Согласно изображению его электронной плотности остов аспарагина хорошо вписан в ЭП, однако боковая цепь вписана не полностью. Полаю, что нестандартные параметры расположения для этого остатка могут являться следствием его положения в полипептидной цепи — как и в первом примере здесь наблюдается чередование элементов вторичной структуры. Здесь уже α -спираль переходит в участок петли. Дополнительная стерическая нагрузка может возникать из-за близко расположенного сайт связывания иона калия.

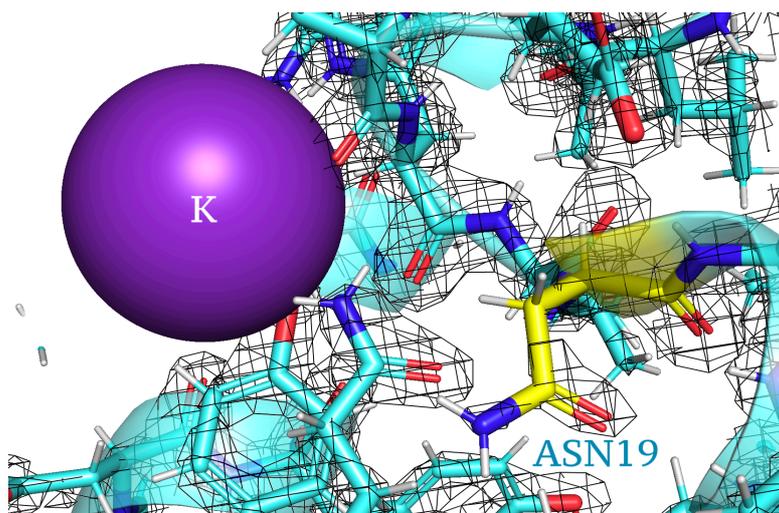


Рис. 10а. Изображение Asn19, ЭП при $Z = 2$.

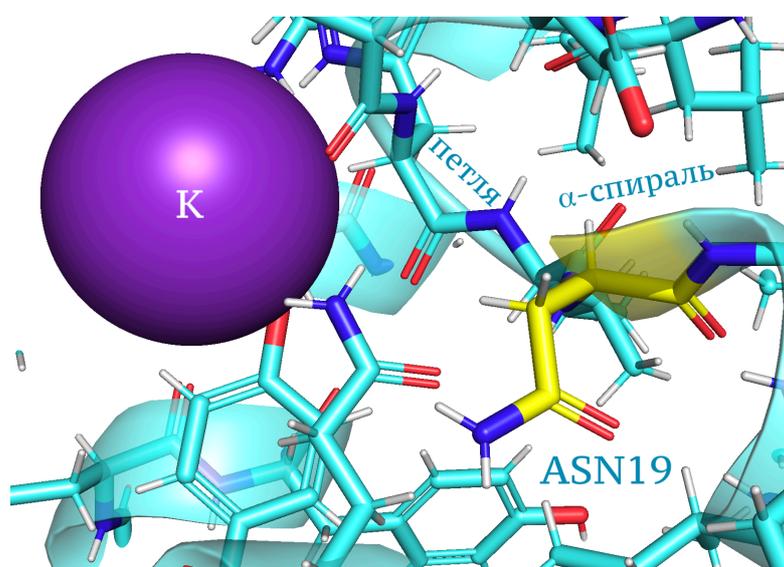


Рис. 10b. Изображение Asn19, отмечены элементы вторичной структуры.

PDB-REDO Databank

Из базы данных PDB-REDO [10] была взята исправленная структура 5T30, с которой были проведены те же манипуляции в сервисе MolProbity, что и с исходной структурой.

Рис. 11. Выдача сервиса MolProbity для исправленной модели 5T30.

All-Atom Contacts	Clashscore, all atoms:	1.75		100 th percentile* (N=841, 1.77Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	2	0.61%	Goal: <0.3%
	Favored rotamers	310	94.80%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	384	97.46%	Goal: >98%
	MolProbity score [^]	1.04		100 th percentile* (N=11232, 1.77Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.55%	Goal: 0
	Bad bonds:	3 / 3275	0.09%	Goal: 0%
Peptide Omegas	Bad angles:	1 / 4448	0.02%	Goal: <0.1%
	Cis Prolines:	2 / 20	10.00%	Expected: ≤1 per chain, or ≤5%

Рис. 11 иллюстрирует параметры, рассчитанные для исправленной модели 5T30. Видно, что незначительно улучшилась лишь интегральная оценка MolProbity (которая и в исходной структуре соответствовала 100-ому перцентилю). Остальные параметры либо не изменились, либо вообще ухудшились. Так, стало меньше разрешенных ротамеров (94.8% против 96.02% в исходной модели) и обнаружилось 2 вообще не разрешенных, нашелся еще один «плохой» угол (здесь их 3 вместо 2 в исходной модели).

Дальнейший анализ этой структуры, считаю, не имеет смысла. Возможно, PDB-REDO помогает хорошо модифицировать структуру в случае изначально плохой модели. Исходная структура 5T30 уже достаточно высокого качества, поэтому PDB-REDO сумел лишь подогнать значения углов, при этом негативно повлияв на характеристики других частей модели.

Выводы

Структура комплекса MnSOD с азидом 5T30 имеет высокое разрешение и хороший уровень качества. Процент маргинальных остатков ничтожно мал при нормировке на длину полипептидной цепи, при этом наличие части маргиналов объясняется с точки зрения биологии фермента. Нет оснований использовать исправленную PDB-REDO структуру.

ИСТОЧНИКИ

1. J.M. McCord, 2002. Superoxide dismutase in aging and disease: an overview. *Methods Enzymol.* 349, 331–341.
2. <https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=ShowDetailView&TermToSearch=6648>
3. P. Becuwe, M. Ennen, R. Klotz, C. Barbieux, S. Grandemange, 2014. Manganese superoxide dismutase in breast cancer: From molecular mechanisms of gene regulation to biological and clinical significance. *Free Radical Biology and Medicine.* 77, 139-151.
4. J. Azadmanesh, S.R. Trickel, G.E.O. Borgstahl, 2017. Substrate-analog binding and electrostatic surfaces of human manganese superoxide dismutase. *Journal of Structural Biology.* 199(1), 68-75.
5. A.S. Hearn et al., 2001. Kinetic Analysis of Product Inhibition in Human Manganese Superoxide Dismutase. *Biochemistry.* 40(40), 12051–12058.
6. G.N. Murshudov et al., 2011. *REFMAC5* for the refinement of macromolecular crystal structures. *Acta Cryst. D67*, 355-367.
7. P. Emsley and K. Cowtan, 2004. Coot: model-building tools for molecular graphics. *Acta Cryst. D60*, 2126-2132.
8. Full wwPDB X-ray Structure Validation Report, 2018: https://files.rcsb.org/pub/pdb/validation_reports/t3/5t30/5t30_full_validation.pdf
9. <http://molprobit.biochem.duke.edu/>
10. <https://pdb-redo.eu/>