

**Отчёт по качеству РСА расшифровки
структуры
Acyl-CoA оксидазы (PDB ID = 1w07)**

Матвейшина Елена

4 курс, 402 группа
Ф-т биоинженерии и биоинформатики
МГУ им. М.В.Ломоносова

2018г.

Аннотация

В данной работе был проведен анализ качества расшифровки структуры фермента Асuл-СоА оксидазы с PDB ID = 1w07, полученной методом рентгеноструктурного анализа. Рассмотрены ключевые параметры качества модели, такие как R-фактор, R-free, карты Рамачандрана и др.

Введение

Асuл-СоА оксидаза катализирует окисление Асuл-СоА в *trans*-2-еноил-СоА, имеет FAD в качестве кофактора и участвует в трех метаболических путях: метаболизм жирных кислот, синтез полиненасыщенных жирных кислот и в PPAR (рецепторы, активируемые пероксисомным пролифератором) сигнальном пути. Семейство пероксисомальных Асuл-СоА оксидаз играет важную роль в метаболизме жирных кислот, а именно в бета-окислении.

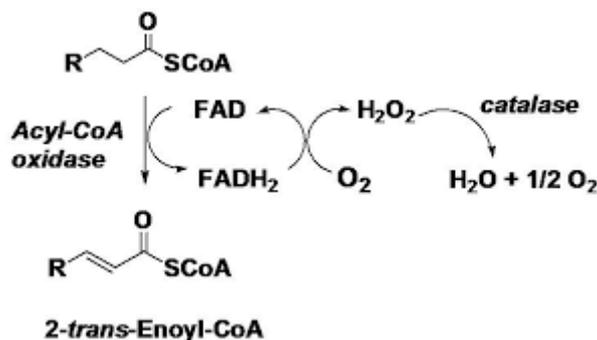


Рис 1. Реакция, катализируемая Асuл-СоА оксидазой.

Данная структура является FAD содержащей Асuл-СоА оксидазой 1 из *Arabidopsis thaliana*, первой структурой Асuл-СоА оксидазы, полученной из растения [1].

A

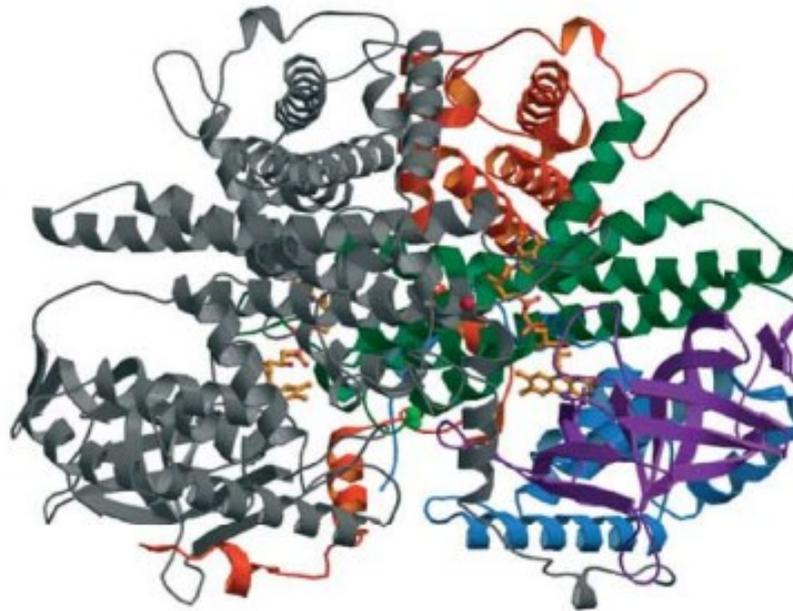


Рис 2. [1] Структура Acyl-CoA оксидазы. Цепь А покрашена серым, а цепь В в разные цвета по доменам: Na домен - синий, Nb домен - фиолетовый, Ca1 домен - зеленый и Ca2 домен - оранжевый. Две FAD молекулы представлены в виде ball-and-stick. Ионы Ca^{2+} представлены в виде красных сфер.

В работе Pedersen L, Henriksen A (2005) было проведено сравнение между полученной ими структурой пероксисомальной Acyl-CoA оксидазы и митохондриальной acyl-CoA дегидрогеназы из животного, чтобы объяснить субстратную специфичность Acyl-CoA оксидазы и предпочтение длинно-цепочечных, моно-ненасыщенных субстратов. Было показано, что пероксисомальная Acyl-CoA оксидаза содержит больший гидрофобный карман, чем митохондриальная Acyl-CoA дегидрогеназа. [1]

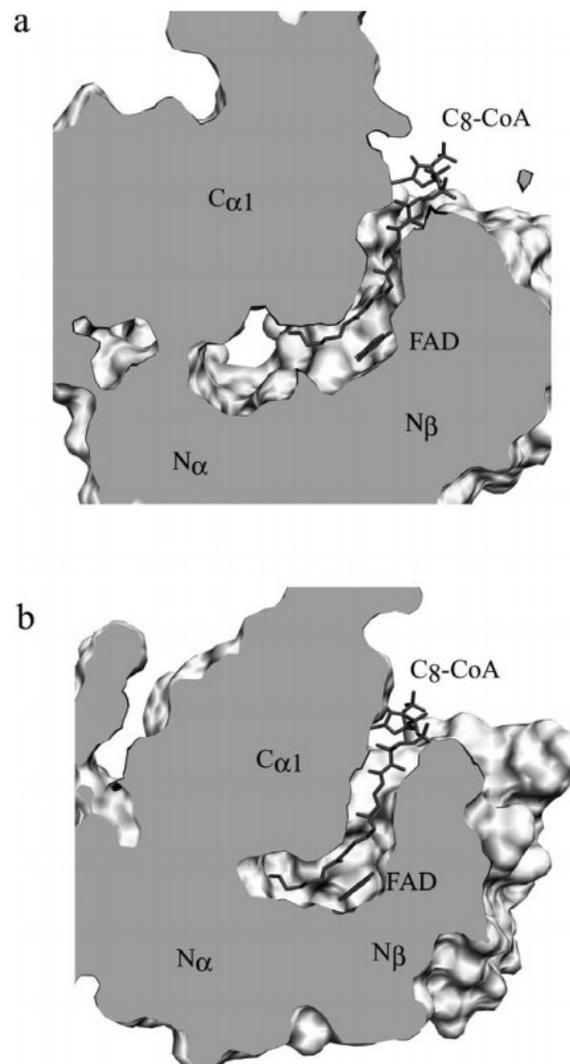


Рис 3. [1] Гидрофобный карман пероксисомальной асул-СоА оксидазы (а) и митохондриальной асул-СоА дегидрогеназы (б).

Результаты

Общая информация о модели

Таблица 1. Общая информация о модели

Параметр	Значение
состав комплекса	Цепи А, В (659 а.к.); Кофактор 2xFAD; Лиганды 1xCa ²⁺ , 2xCl ⁻ , 1xPt ⁺²
год	15.12.2004 (загрузка структуры)
фамилии авторов	Pedersen L, Henriksen A
метод решения фазовой проблемы	Молекулярное замещение
число измеренных рефлексов	87331
разрешение	2Å
полнота набора структурных	96.9%

факторов	
диапазон разрешений структурных факторов	2Å - 29.59Å
параметры кристаллографической ячейки	a: 85.201Å b: 117.001Å c: 131.047Å α: 90° β: 90° γ: 90°
наличие некристаллографических симметрий в асимметрической ячейке	нет
Тип кристаллографической симметрии	P2 ₁ 2 ₁ 2 ₁

Значения индикаторов качества модели в целом

R-фактор - интегральная мера различия F^{calc} и F^{obs} , где F^{calc} и F^{obs} - модули структурных факторов, построенные по модели и полученные из эксперимента соответственно.

$$R = \frac{\sum_{hkl} |F_{hkl}^{calc} - F_{hkl}^{obs}|}{\sum_{hkl} F_{hkl}^{obs}} * 100\%$$

Однако для оптимизации модели используют составной R-фактор, учитывающий не только R-фактор, представленный выше (R_{X-ray}), но еще и два других R-фактора, которые устроены аналогично R_{X-ray} , однако являются интегральной мерой различия между:

- расстояниями между валентно взаимодействующими атомами, полученными по модели и известными из литературы (R_{dist})
- углами между валентно взаимодействующими атомами, полученными по модели и известными из литературы (R_{angle})

$$R_{mixed} = w_{X-ray} R_{X-ray} + w_{dist} R_{dist} + w_{angle} R_{angle}, \text{ где } w_i - \text{соответствующие веса}$$

R-фактор для данной модели равен 0.204, что меньше 0.25, а значит считается хорошим R-фактором.

R-free - фактор, посчитанный аналогично R-фактору, только не некотором количестве (в данной структуре 5.05%) случайно выбранных рефлексов, которые не использовались при оптимизации модели по R-фактору.

R-free для данной модели равен 0.249, что не меньше 0.2, поэтому не очень хорошо, но и не больше 0.4, так что и не плохо. Кроме того разница между R-фактором и R-free меньше 0.1, значит можно не подозревать переоптимизацию модели.

Карта рамачандра отражает 2 торсионных угла (ψ и ϕ), которые определяют укладку полипептидной цепи. Для аминокислот известны их характерные значения, если же значения этих торсионных углов, полученное по модели, для какой-то аминокислоты отличаются, то у нее

конформации высокая энергия, и она считается выбросом (маргинальным остатком). Это может быть связано как и с особенностями укладки пептидной цепи, так и с неточностью модели. На рис. 4 показана карта Рамачандрана для рассматриваемой структуры, полученная с помощью online-сервиса EDS [2].

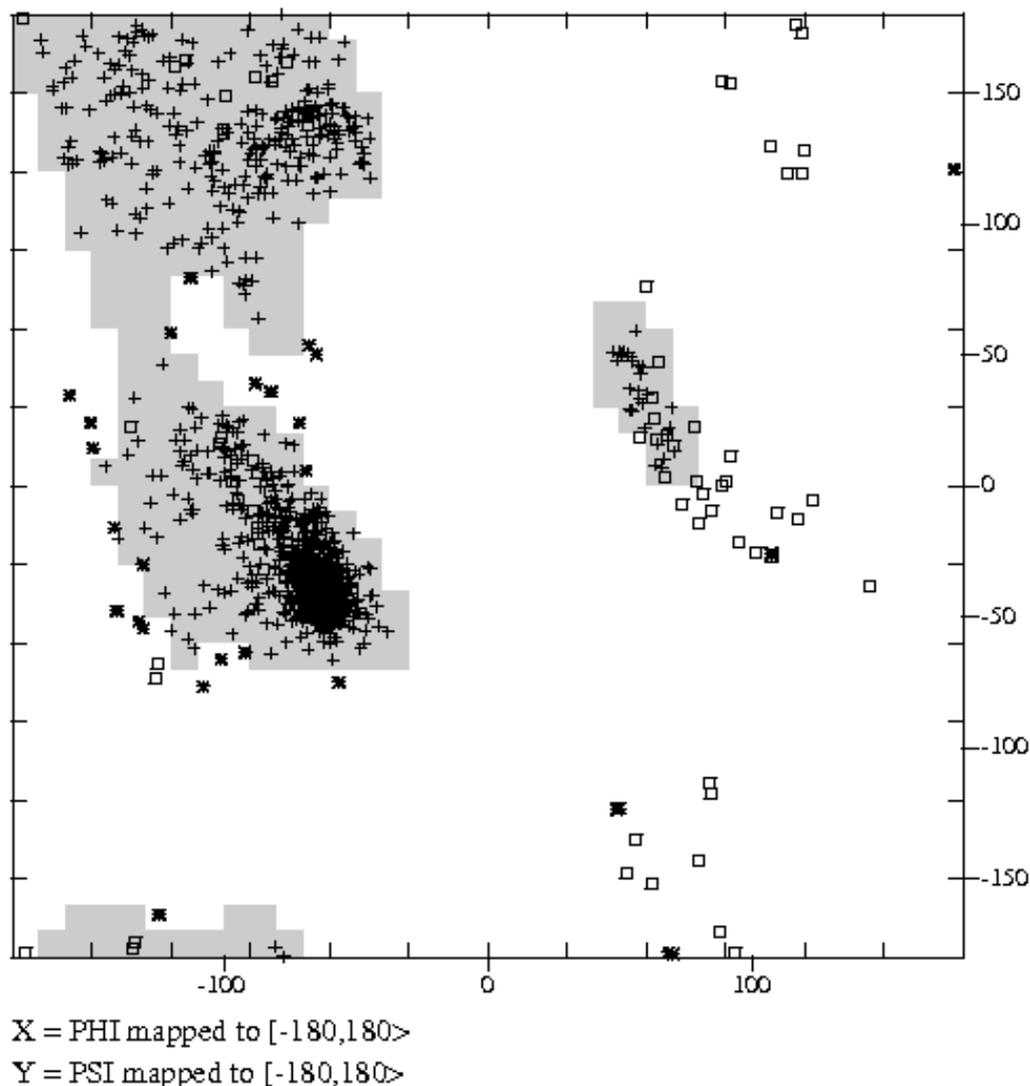


Рис 4. Карта Рамачандрана для цепей А и В. [2]

В данной структуре было найдено с помощью сервиса EDS 28 маргинальных остатков, что составляет 2,3% (ожидаемая цифра 0-5%), притом 12 из них в цепи А (2%) и 16 в цепи В (2,6%).

Однако, если использовать сервис MolProbity [3], ситуация выглядит по-другому.

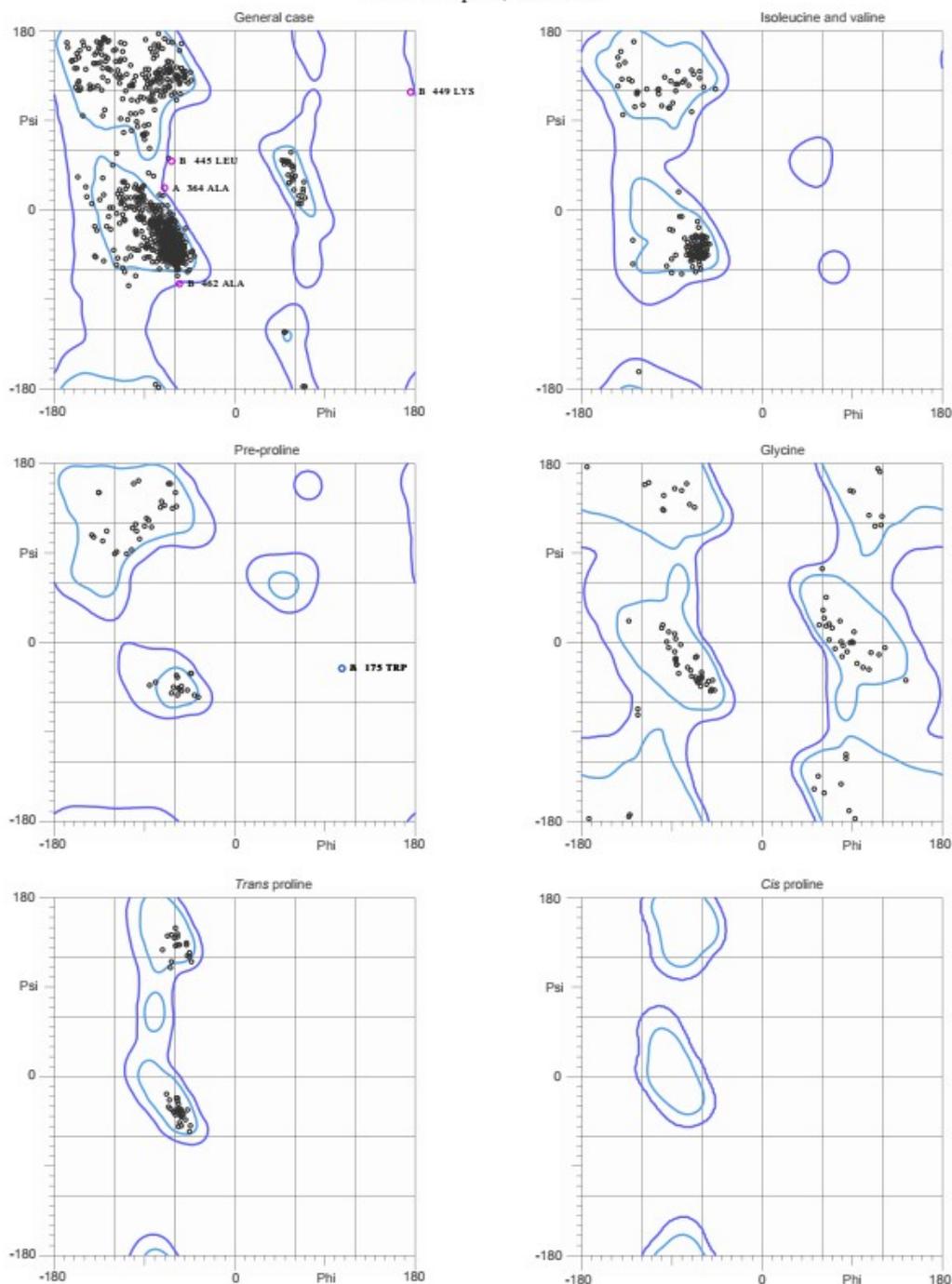


Рис 5. Карта Рамачандрана. [3]

Найдено было всего 6 выбросов, при том 2 из цепи А и 4 из цепи В, что составляет 0.46%. Правда в рамках требований этого сервиса это значение недостаточно хорошо, т.к. больше 0.05%. Еще было посчитано количество остатков, находящихся в областях с очень хорошей энергией. Их оказалось 1250, что составляет 95.64%, правда желательная цифра - больше 98%.

Также в данной структуре 7.9% выбросов по параметру RSR, который отражает согласованность экспериментальной электронной плотности и электронной плотности, построенной по модели.

$$RSR = \frac{\sum_{A \in L} |\rho_{\text{эксп}} - \rho_{\text{модель}}|}{\sum_{A \in L} \rho_{\text{эксп}}} [\cdot 100\%]$$

Сумма берется по узлам пространственной решетки в окружении группы атомов. Экспериментальная плотность считается, используя фазы из модели и модули структурных факторов из эксперимента, даже вернее $2F_{\text{эксп}} - F_{\text{модель}}$.

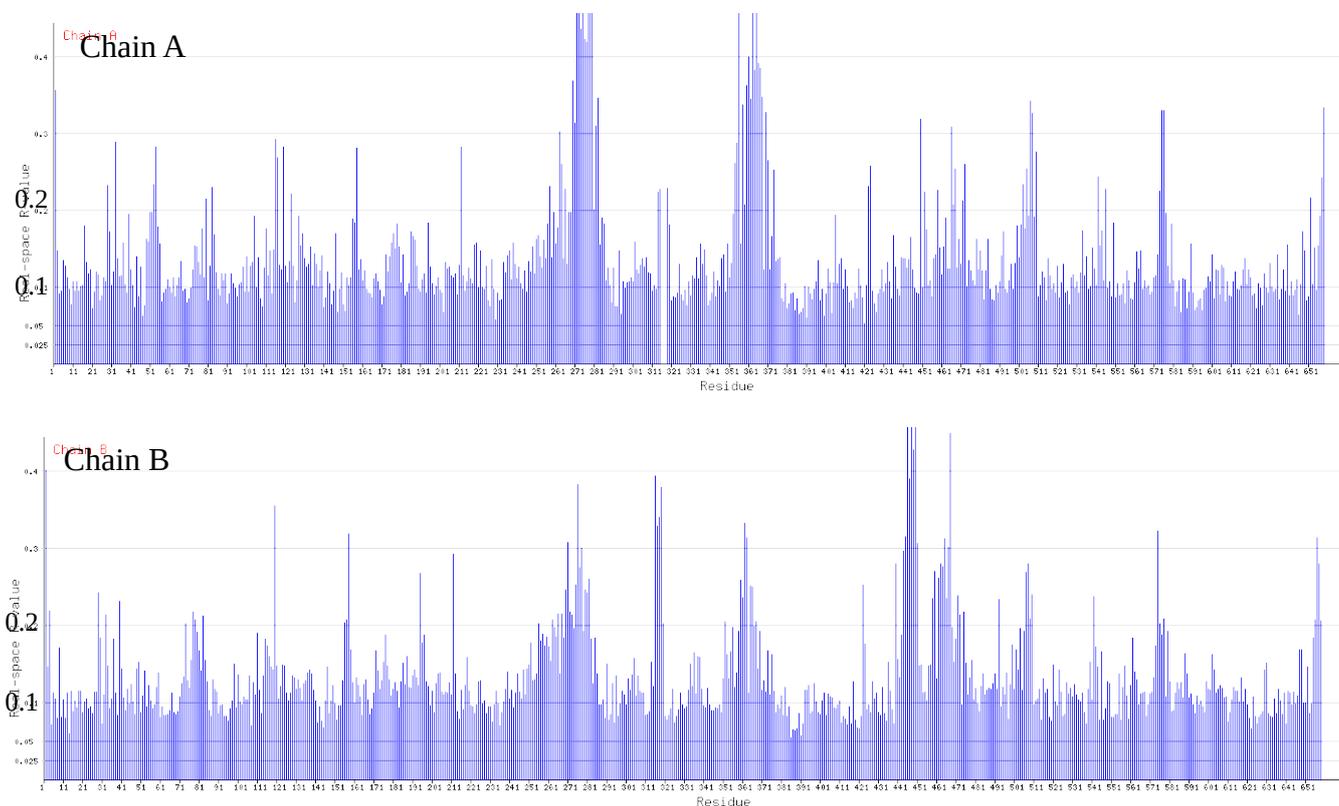


Рис 6. RSR для аминокислот цепей A и B.

На рис.5 изображено распределение значений RSR для всех аминокислот рассматриваемой структуры. Как можно видеть, в основном RSR колеблется в районе 0.1, что принято считать границей для хороших значений. Однако аминокислот с RSR больше чем 0.1 довольно много, и даже немало аминокислот с RSR больше чем 0.2, что считается уже плохими значениями пространственной R фактора. Среднее значение RSR составляет 0.136 со стандартным отклонением 0.071.

Можно посмотреть на Z-score, который определяется насколько плохо данный остаток вписан в экспериментальную электронную плотность по сравнению с остатками такого типа в структурах с таким же разрешением (были рассмотрены структуры с разрешениями от 1.8 до 2Å).

Таблица 2. Z (для RSR)

Chain	Nres	<Z>	Sigma	%<-1	%<0	%>1	%>2
A	655	0.46	1.54	0.46	48.09	16.34	8.85
B	658	0.39	1.25	0.30	45.29	16.87	8.05
All	1313	0.43	1.40	0.38	46.69	16.60	8.45

Исходя из таблицы 2 получается, что в целом атому данной структуры вписаны в электронную плотность чуть хуже, чем в среднем. Немного меньше половины атомов вписаны в электронную плотность лучше чем в среднем по PDB, и более 8% остатков могут считаться маргиналами, т.к. имеют $Z > 2\%$.

Еще можно рассмотреть температурный фактор (B-фактор), который отражает динамическую и статистическую неупорядоченность (ширину гауссианы электронной плотности).



Рис 7. Температурный (B) фактор для цепей A и B. [2]

Как можно видеть по рис. 7 довольно много аминокислотных остатков с B-фактором больше 20, которые нельзя отнести к хорошим, однако аминокислот с плохим значением B-фактора больше 60 в данной структуре нет. Средние значения B-фактора, взвешенное по коэффициенту заполнения составило 26.5\AA^2 со стандартным отклонением 11.3\AA^2 .

Ниже, в таблице 3 представлена общая сводка параметров модели, полученная сервисом MolProbity [3]. Как можно видеть, больше всего проблем вызвали длины связей и углы, которые не соответствуют литературным данным, их нашлось 8 и 19 соответственно. В случае неправильных длин связей MolProbity считает, что их не должно быть совсем. Однако clashscore, отражающий число стерических перекрываний, у данной структуры весьма неплохой. Кроме того найден 9 плохих ротамеров – аминокислотных остатков, у боковых цепей которых торсионные углы не совпадают с типичными для боковых цепей данной аминокислоты торсионными углами.

Таблица 3. Общие параметры модели по MolProbity [3]

All-Atom Contacts	Clashscore, all atoms:	7.99	90 th percentile* (N=715, 2.00Å ± 0.25Å)	
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	9	0.83%	Goal: <0.3%
	Favored rotamers	1038	96.02%	Goal: >98%
	Ramachandran outliers	6	0.46%	Goal: <0.05%
	Ramachandran favored	1250	95.64%	Goal: >98%
	MolProbity score [^]	1.74	89 th percentile* (N=12522, 2.00Å ± 0.25Å)	
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	8 / 10642	0.08%	Goal: 0%
Bad angles:	19 / 14465	0.13%	Goal: <0.1%	
Peptide Omegas	Cis Prolines:	0 / 46	0.00%	Expected: ≤1 per chain, or ≤5%

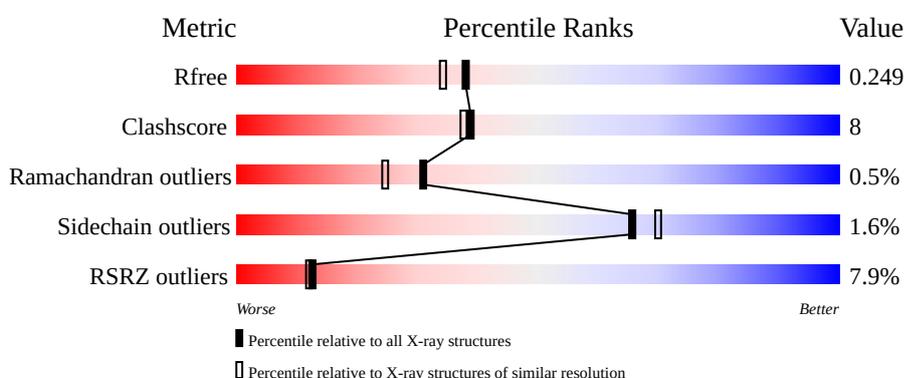


Рис 8. Параметры модели относительно других моделей PDB.

На рис. 8 показано, как соотносятся значения параметра качества рассматриваемой модели с другими модели во всем PDB и среди моделей со сходным разрешением. И как мы можем видеть, в целом данная модель немного хуже, чем типичная модель и сильно хуже по параметру RSR.

В таблице 4 приведены примеры маргинальных остатков.

Таблица 4. Примеры маргинальных остатков. Красным помечены маргинальные значения параметров. “-” - нет данных.

Остаток	ψ и φ [3]	RSR [2]	Z [2]	Ротамер [3]	В-фактор[2]
A 175 TRP	107.6, -26.1	0.140	0.997	Favored 49.1%	13.470
A 364 ALA	-71.2, 23.8	0.382	5.793	-	54.100
B 175 TRP	107.7, -26.3	0.129	0.722	Favored 51.9%	15.840
B 445 LEU	-64.8, 50.0	0.315	4.595	Favored 95.7%	45.310

B 449 LYS	176.8, 120.9	0.428	4.120	-	44.310
B 462 ALA	-56.3, -75.0	0.261	-	-	50.740
B 450 VAL	-	0.545	9.175	Favored 13.5%	45.910
A 278 TYR	Favored 3.79%	0.573	10.644	-	59.610

Анализ маргинальных остатков

Рассмотрим подробнее некоторые маргинальные остатки, представленные в таблице 4.

A 364 ALA

Он маргинал по карте Рамачандрана (-71.2, 23.8), RSR (0.382), Z (5.793) и B-фактору (54.100).

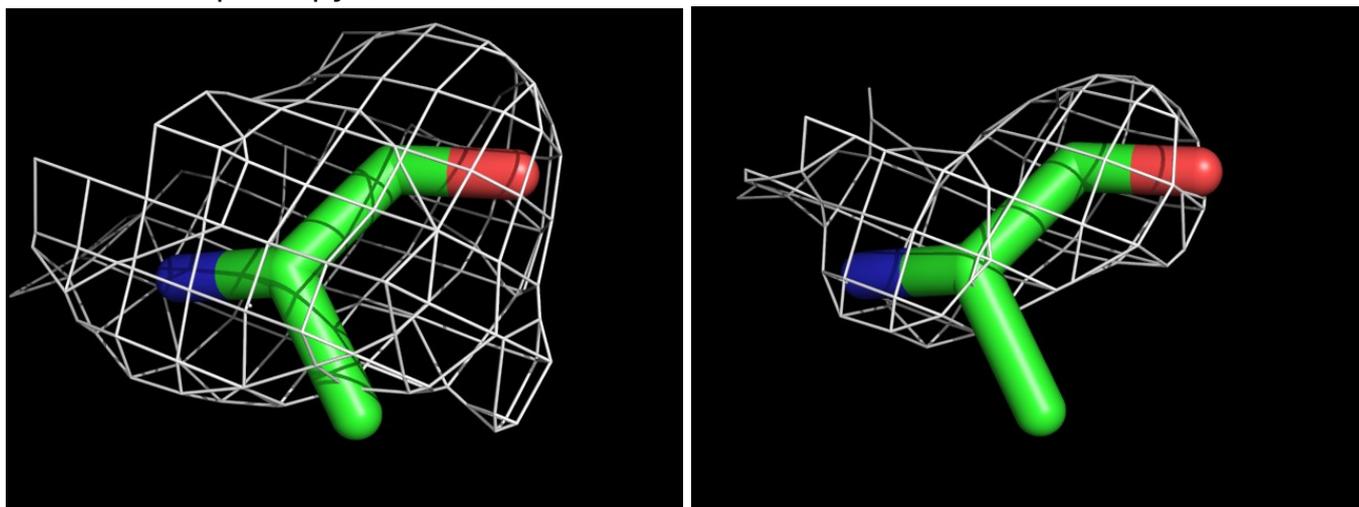


Рис 9. 364ALA цепи A с электронной плотностью на уровне подрезки $\sigma=0.5$ и $\sigma=1$

Как можно видеть на рис.9 электронная плотность не отражает структуру данной аминокислоты, CH_3 группа не покрыта электронной плотностью, и возможно, что данный остаток является ошибкой расшифровки.

A278TYR

Он маргинал по RSR (0.545), Z(9.175) и B-фактору (45.910).

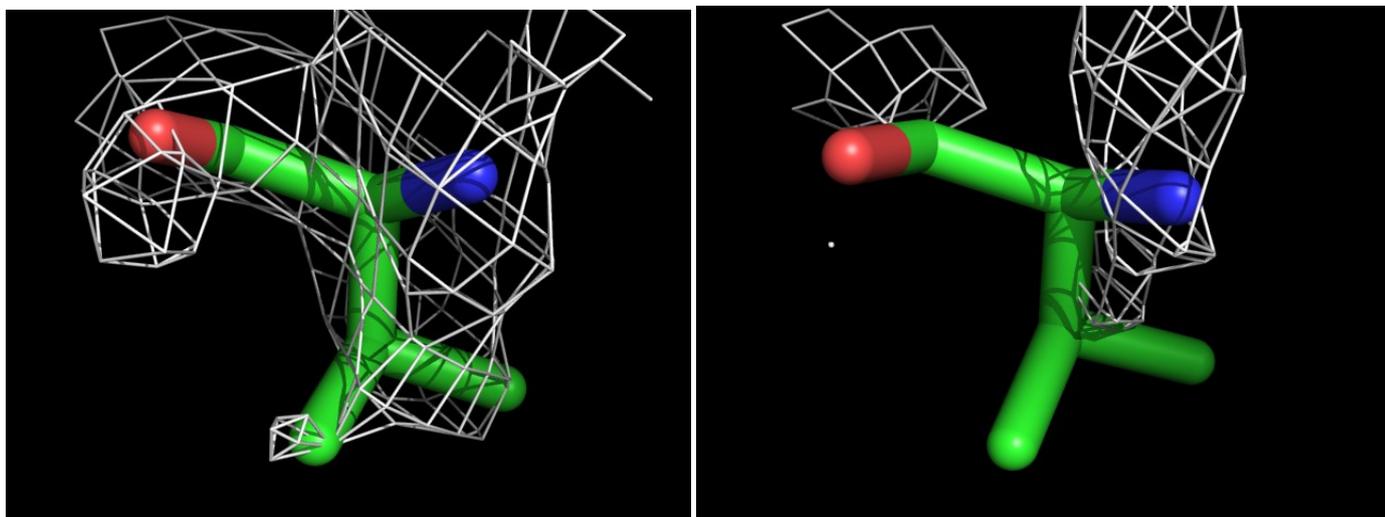


Рис 10. 450 VAL цепи В с электронной плотностью на уровне подрезки $\sigma=0.5$ и $\sigma=1$.

Как можно видеть на рис. 10 этот валин почти совсем не совпадает со своей электронной плотностью на уровне подрезки всего лишь 1 кроме амидной группы. А на уровне подрезки 0.5 электронная плотность не соответствует аминокислоте: одна метильная группа боковой цепи не покрыта, у карбоксидной группы какой-то дополнительный выступ электронной плотности. Скорее всего, данный остаток является ошибкой расшифровки.

A630PHE

Маргинал по торсионным углам боковой цепи (плохой ротамер) (χ angles: 228.8,122.7).

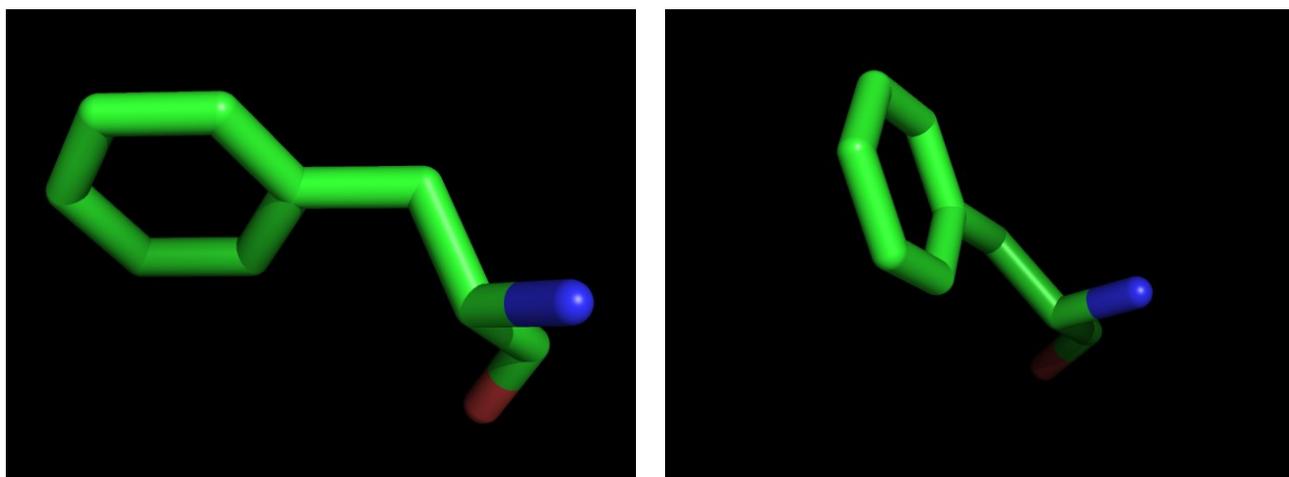


Рис 11. Цепь А PHE 630 (слева) и 607 (справа)

Как можно видеть на рис. 11 данный фенилаланин очень сильно отличается от обычного Favored (26.7%) фенилаланина по своим торсионным углам боковой цепи.

Однако, по карте Рамачандрана проблем нет, RSR довольно хороший (0.093), B-фактор неплохой (20.150). И как можно видеть на рис. 12 электронная плотность на уровне подрезки 1 довольно хорошо соответствует этому аминокислотному остатку.

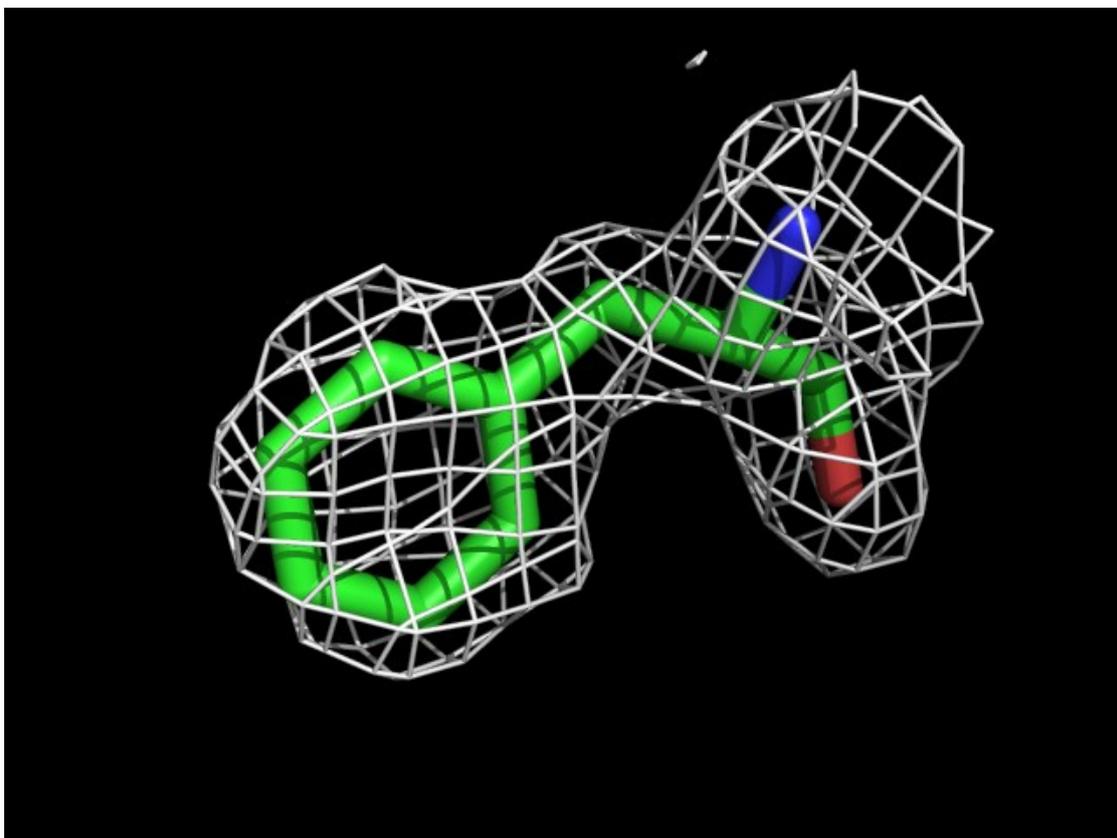


Рис 12. А PHE630 электронная плотность на уровне подрезки $\sigma=1$.

В итоге, я бы сделала вывод, что этот остаток – особенность данной структуры, а не ошибка расшифровки.

B360GLU

Этот остаток является маргиналом из-за серьезного перекрывания (0.77\AA) одного из своих водородов с кислородом воды. Соответственно вода тоже попадает под подозрение в маргинальности.

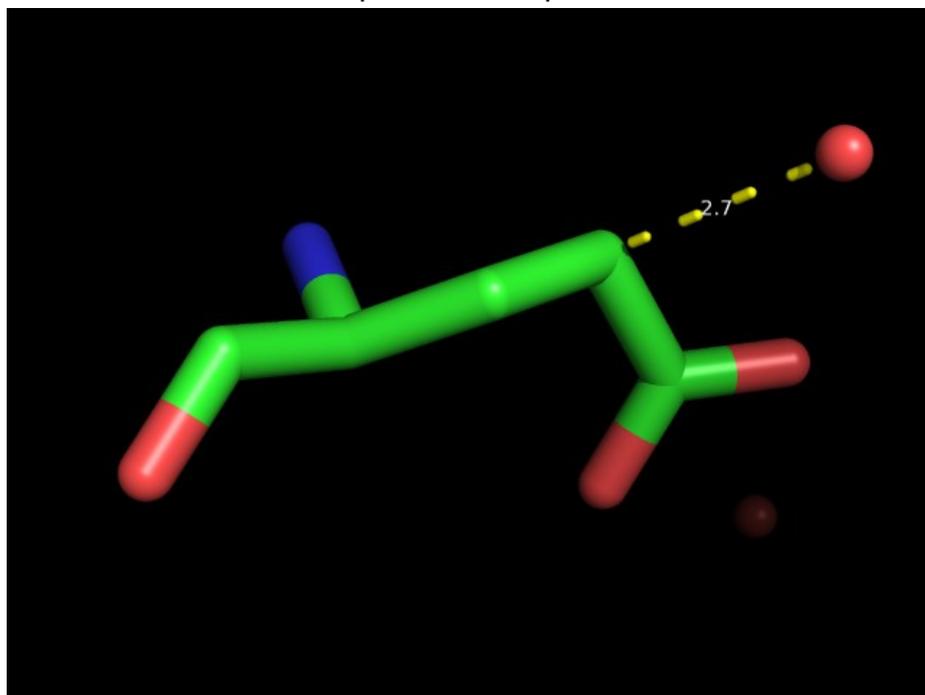


Рис. 13. Расстояния между O воды и C GLU 2.7\AA .

Однако, как мы можем видеть на рис. 14, электронная плотность покрывает и GLU и воду единым облаком на уровне подрезки 1, и пропадает раньше с боковой цепи GLU чем с пространства между водой и GLU (на уровне подрезки 1.5).

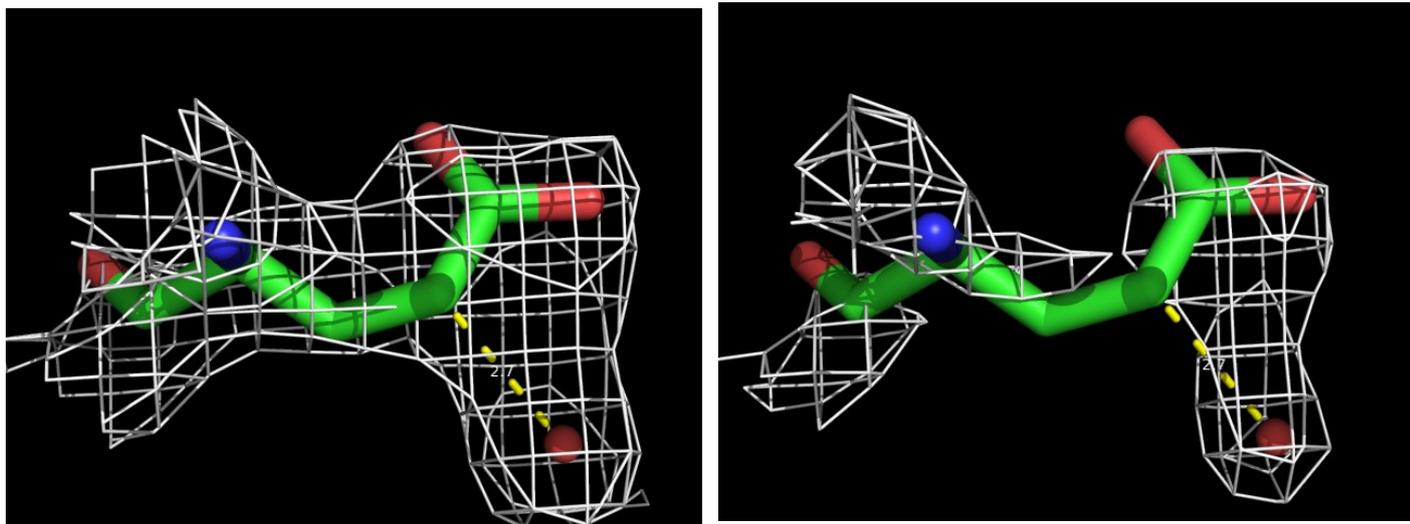


Рис 14. 360 GLU цепи В и вода с электронной плотностью на уровне подрезки $\sigma=1$ и $\sigma=1.5$.

Кроме того это глутамат по карте Рамачандрана является всего лишь допустим, имеет В-фактор 43.520, $Z = 1.688$ и $RSR = 0.256$. Так что в итоге, я бы сделала вывод, что данная вода является ошибкой расшифровки, т.к. ей еще и не с кем рядом взаимодействовать, а может и сам глутамат заодно.

A2227HON

Аналогично можно рассмотреть воду под номером 2227 из цепи А, у которой перекрытие (на 0.64\AA) с 406 TRP из цепи А.

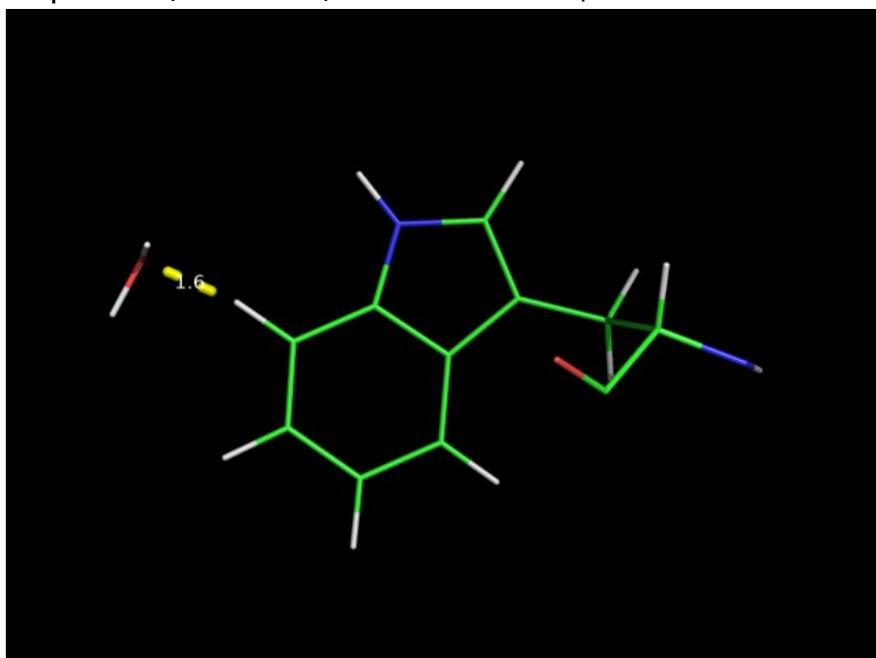


Рис 15. Расстояние между A2227HON и A406TRP.

Как можно видеть на рис. 15 расстояние между кислородом воды и водородом триптофана составило 1.6Å. А согласно рис. 16 эти остатки не очень хорошо согласуются с электронной плотностью. Они не образуют единое облако, как в предыдущем случае, но форма облака электронной плотности слишком не явная и вокруг слишком много “ошметков” электронной плотности, чтобы однозначно выделять воду рядом с гидрофобным триптофаном.

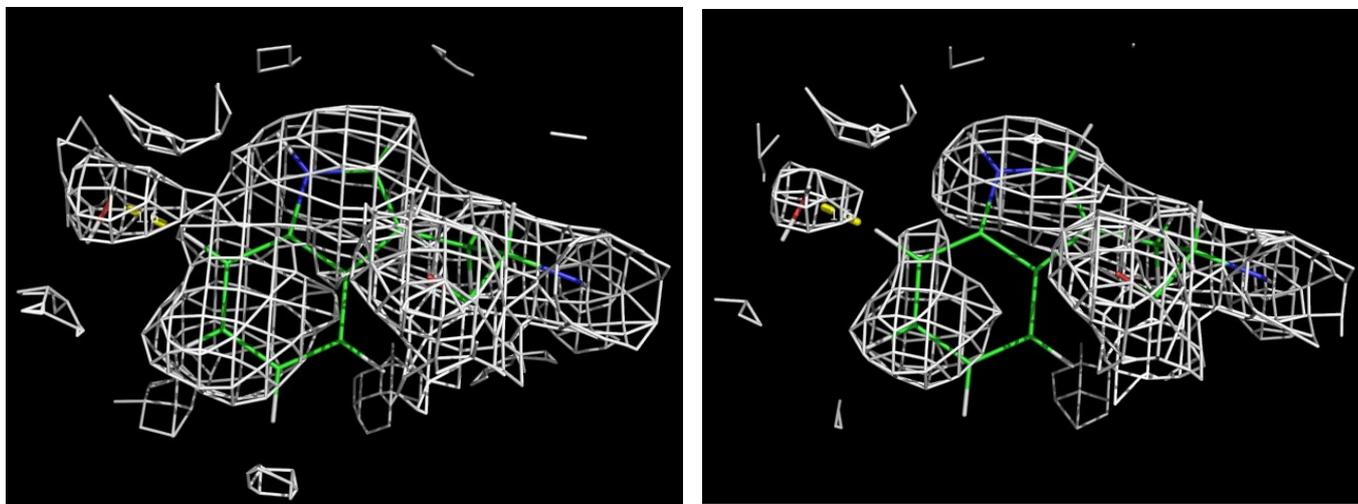


Рис 16. 406 TRP цепи А и вода с электронной плотностью на уровне подрезки $\sigma=0.5$ и $\sigma=1$.

При том температурный фактор у данной воды 29.22, а у триптофана 36.39. И вполне вероятно, что данная вода – просто артефакт построения модели.

Сравнение модели из PDB с моделью из PDB_redo

PDB_REDO [4] – база данных, содержащая оптимизированные версии моделей из PDB. Различия рассматриваемой структуры в PDB и PDB_REDO представлены в таблице 5.

Таблица 5. Серьезные изменения в модели

Описание	Число
Измененные ротамеры	17
Повернутые боковые цепи	21
Удаленная вода	52
Повернутые пептиды	0
Исправленная хиральность	0
Остатки, лучше вписанные в электронную плотность	245
Остатки, хуже вписанные в электронную плотность	4

Для дальнейшего сравнения структура из PDB_REDO была проанализирована сервисом MolProbity.

Таблица 6. Общие параметры модели из PDB_REDO по MolProbity [3]

All-Atom Contacts	Clashscore, all atoms:	2.54		99 th percentile* (N=714, 1.98Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	19	1.72%	Goal: <0.3%
	Favored rotamers	1030	93.47%	Goal: >98%
	Ramachandran outliers	4	0.31%	Goal: <0.05%
	Ramachandran favored	1284	98.24%	Goal: >98%
	MolProbity score [^]	1.22		99 th percentile* (N=12332, 1.98Å ± 0.25Å)
	Cβ deviations >0.25Å	3	0.24%	Goal: 0
	Bad bonds:	20 / 10723	0.19%	Goal: 0%
Bad angles:	14 / 14562	0.10%	Goal: <0.1%	
Peptide Omegas	Cis Prolines:	0 / 46	0.00%	Expected: ≤1 per chain, or ≤5%

Как можно видеть из таблицы 6 оптимизированная модель имеет меньше перекрытий, лучше clashscore, меньше плохих углов, меньше выбросов по карте Рамачандрана и больше остатков с хорошим значением энергии на карте Рамачандрана. Это те параметры, которые были улучшены. Однако, увеличилось количество плохих связей, плохих ротамеров и уменьшилось количество хороших ротамеров. В целом можно предположить, что улучшено состояние главной цепи и ухудшилось состояние боковых. Кроме того, улучшено значение R-фактора с 0.204 на 0.172 и R-free с 0.249 до 0.206.

В целом особых улучшений в модели не наблюдается: что-то улучшается немного, что-то улучшается, что ухудшается. Можно предположить, что сами экспериментальные данные не очень хороши, и не позволяют построить модель лучше.

Заключение

В целом структура 1w07 Асул-СоА оксидазы имеет качество несколько хуже среднего, но и не прям совсем плохое. Самым проблемными параметрами для нее являются RSR и Z. Я бы пользовалась этой структурой со значительной опаской.

Список литературы

1. Pedersen L, Henriksen A. (2005) Acyl-CoA oxidase 1 from *Arabidopsis thaliana*. Structure of a key enzyme in plant lipid metabolism. *J Mol Biol.* 2005 Jan 21;345(3):487-500.
2. <http://eds.bmc.uu.se/eds/>
3. <http://molprobit.biochem.duke.edu/>
4. <https://pdb-redo.eu/>