

Практикум 4. Валидация.

Введение.

В данном практикуме рассматривалась структура гомо-4-мера L-аспарагиназы (PDB ID: 1O7J), полученного из грамотрицательной бактерии *Dickeya dadantii* (ранее известной как *Erwinia chrysanthemi*). L-аспарагиназа катализирует гидролиз L-аспарагина до аспарагиновой кислоты. Данный фермент долгое время использовался в качестве терапевтического средства при лечении острого лимфобластного лейкоза. Структура 1O7J была получена с целью изучения механизма ферментативной реакции и субстратной специфичности терапевтической функции белка [1].

Задание 1.

На странице PDB записи приведены следующие данные о качестве структуры:

- Разрешение: 1.0 Å
- R-value: 0.110
- R-free: 0.128
- График с общей информацией о качестве структуры в сравнении с ранее депонированными в PDB записями (рис.1.)

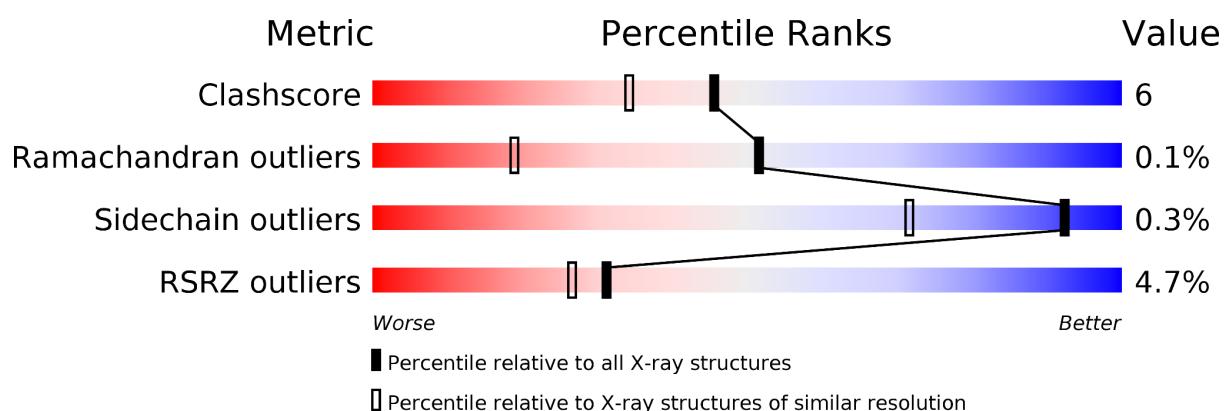


Рис.1. Общая информация о качестве структуры 1O7J в сравнении с ранее депонированными в PDB записями.

Значение R-value меньше 0.2, что означает, что качество у модели достаточно хорошее. Значения R-free и R-value достаточно близки. Это говорит о том, что переобучения в ходе получения модели не произошло.

На графике (рис.1.) черным прямоугольником обозначена доля записей, которые равны или хуже этой структуры по показателю качества среди всех PDB записей, полученных с помощью X-ray. Белым прямоугольником - доля записей, у которых параметры качества равны или хуже данной структуры, среди X-ray структур с примерно таким же разрешением.

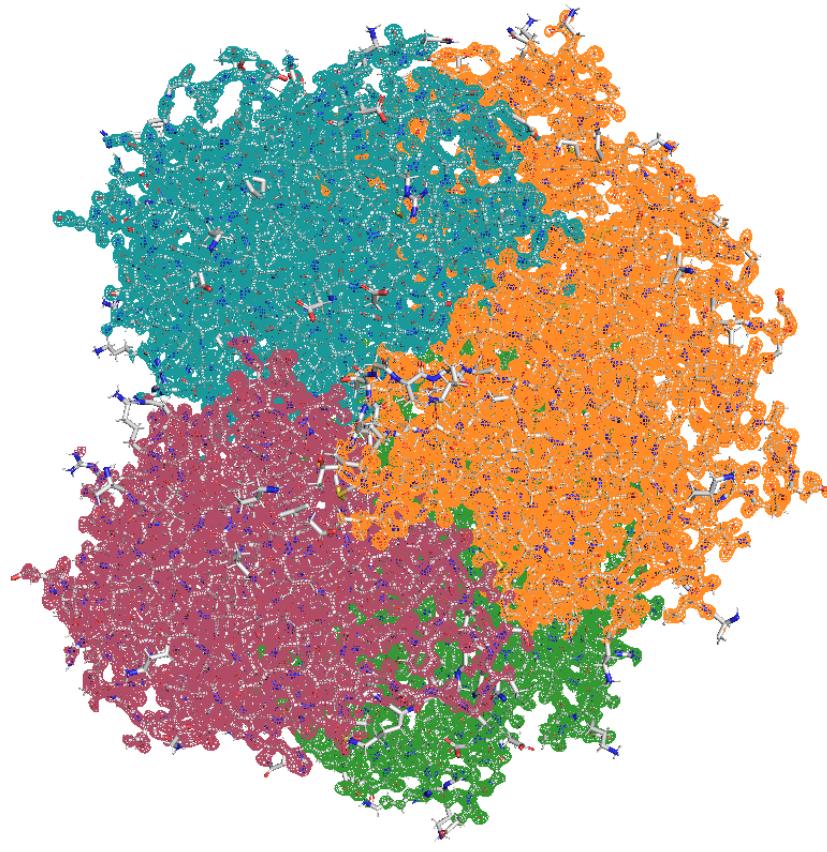
Видно, что значение clashscore (оценка, основанная на числе пар атомов в модели, которые нетипично близки друг к другу) у рассматриваемой структуры хуже.

Значение Ramachandran outliers (оценка, основанная на количестве аутлаеров по картам Рамачандрана) хуже.

Значение sidechain outliers (оценка, основанная на количестве боковых радикалов, расположенных в нетипичном положении) достаточно хорошее.

Значение RSRZ outliers (оценка, основанная на количестве аминокислотных остатков с плохим значением RSRZ) хуже.

Далее посмотрим на расположение электронной плотности вокруг рассматриваемой модели структуры L-аспарагиназы (рис.2.).



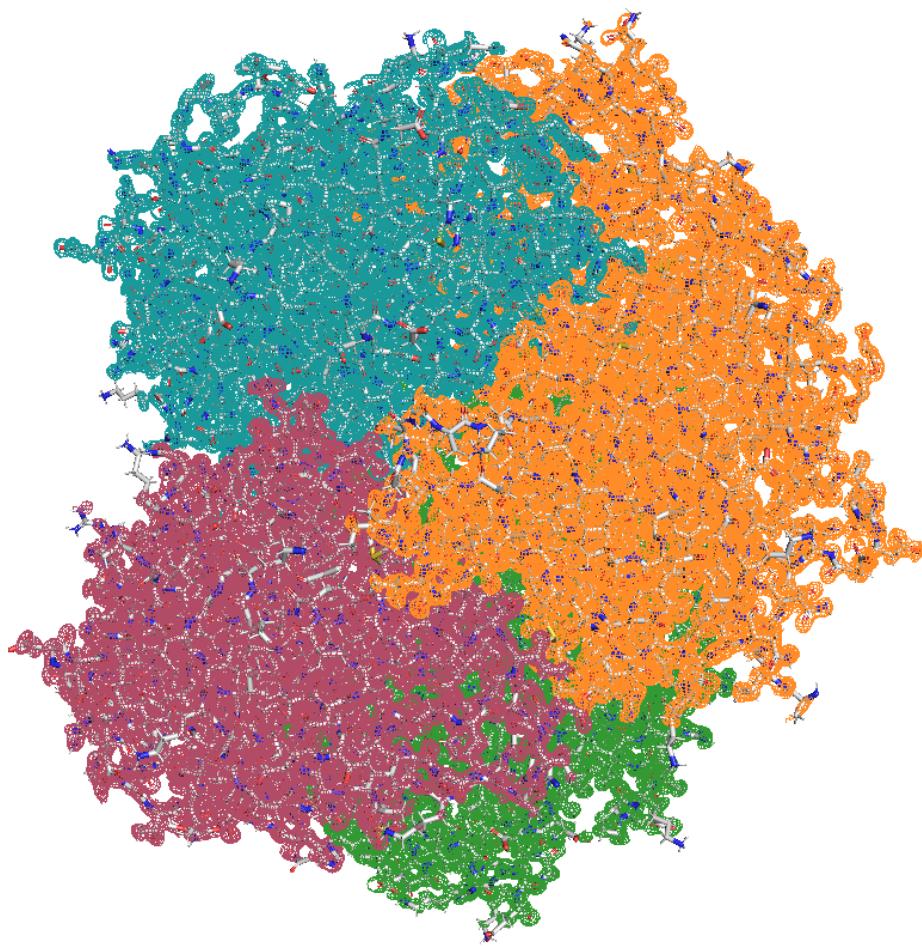


Рис.2. Электронная плотность на разных уровнях подрезки вокруг структуры 1O7J. ЭП покрашена по цепям. А) Уровень подрезки 1.5; Б) Уровень подрезки 1.

При отображении электронной плотности вокруг структуры видно, что некоторые боковые аминокислотные остатки практически не имеют поддержки. Рассмотрим некоторые такие участки структуры.

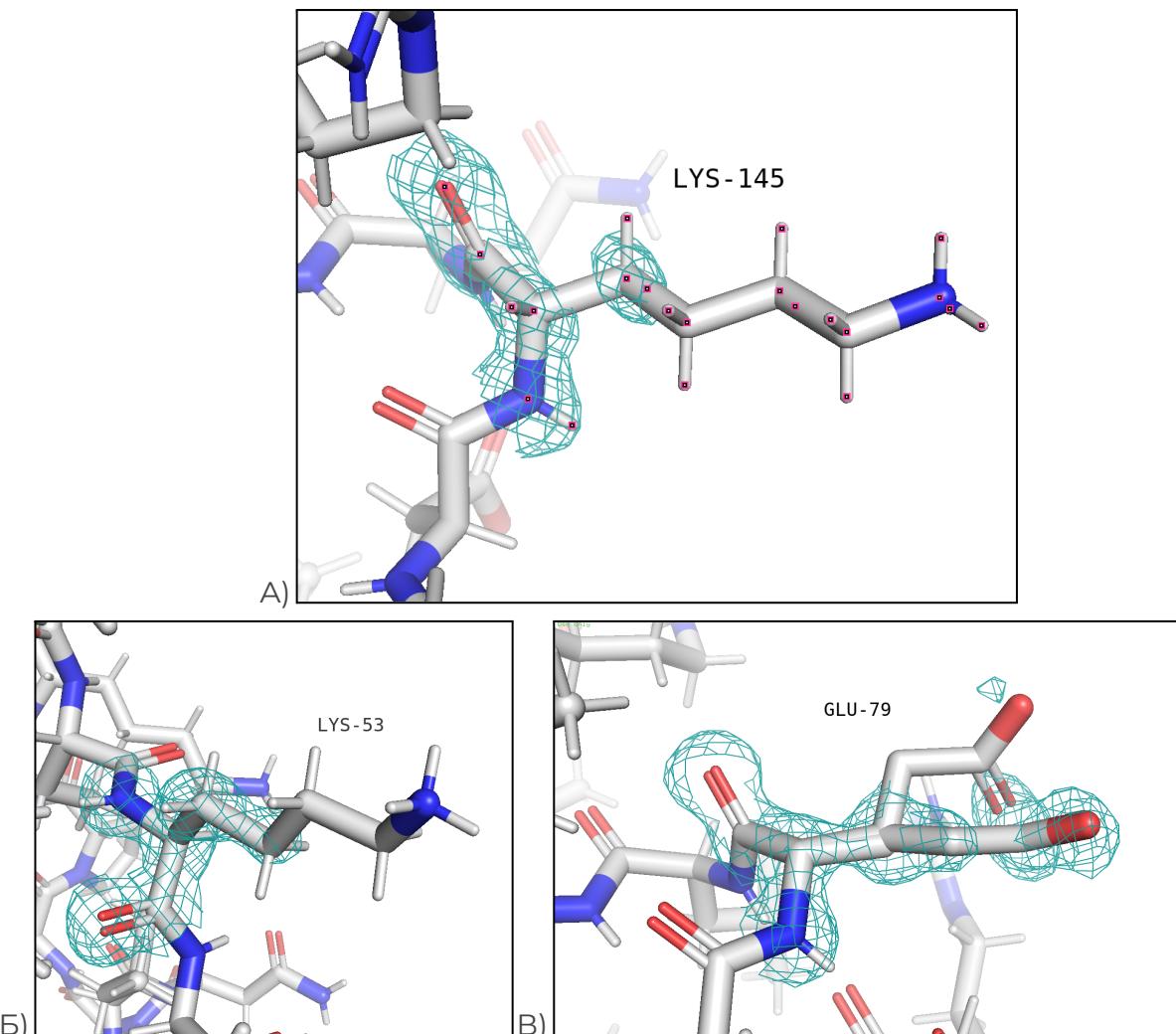


Рис. 3. Электронная плотность на уровне подрезки 1 вокруг боковых остатков, не имеющих достоверной поддержки ЭП. А) Лизин 145; Б) Лизин 53; В) Глутамат 79.

На рисунке 3 представлены три боковых радикала, электронная плотность которых как мне показалось плохо подходит под их структуру.

На рисунках 3А и 3Б видно, что терминальная часть лизина в обоих случаях вообще не покрыта ЭП. На рисунке 3В видно, что авторы модели предположили альтернативное положение глутамата. “На глаз” кажется, что одно из положений глутамата (верхнее) не совсем обосновано тк остаток практически не покрыт ЭП.

У рассматриваемой структуры 1O7J существует 15 альтернативных структур. Одна из них получена с помощью AlphaFold - AF_AFP06608F1, остальные с помощью рентгеноструктурного анализа. Альтернативная структура [5HWO](#) имеет разрешение чуть меньше (1.70 Å), но при этом параметры качества модели имеет более

внушительные (на графике сравнения с другими структурами из PDB большинство параметров находятся в большом процентильном ранге).

Для исследования L-аспарагиназы я бы выбрала модель **5HWO**. При этом ЭП, на основе которой строилась модель 1O7J, на мой взгляд достаточно перспективная (низкое разрешение). Возможно, в PDB redo найдется хорошая перестроенная по этой электронной плотности модель.

Задание 2.

Для поиска маргинальных остатков воспользуемся инструментом оценки на сайте [SwissModel](https://swissmodel.expasy.org/assess) (<https://swissmodel.expasy.org/assess>).

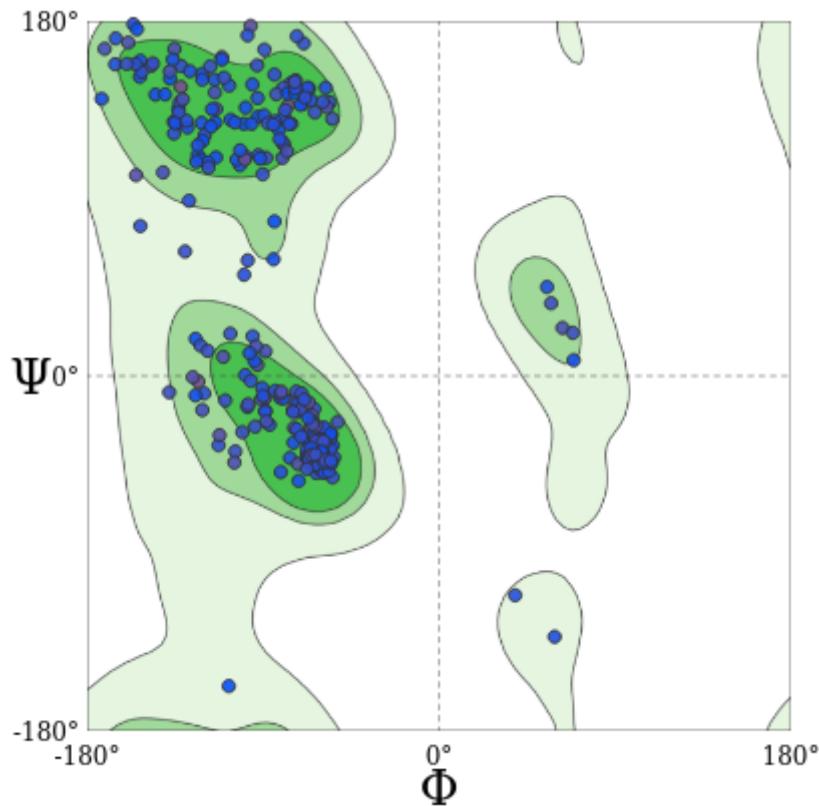


Рис. 4. Карта Рамачандрана для всей структуры 1O7J.

В выдаче для структуры 1O7J есть несколько карт Рамачандрана. Общая карта для всех остатков изображена на рисунка 4. Видно, что нет остатков, которые бы лежали в белой области (что хорошо), хотя некоторые все-таки лежат в не самых типичных положениях.

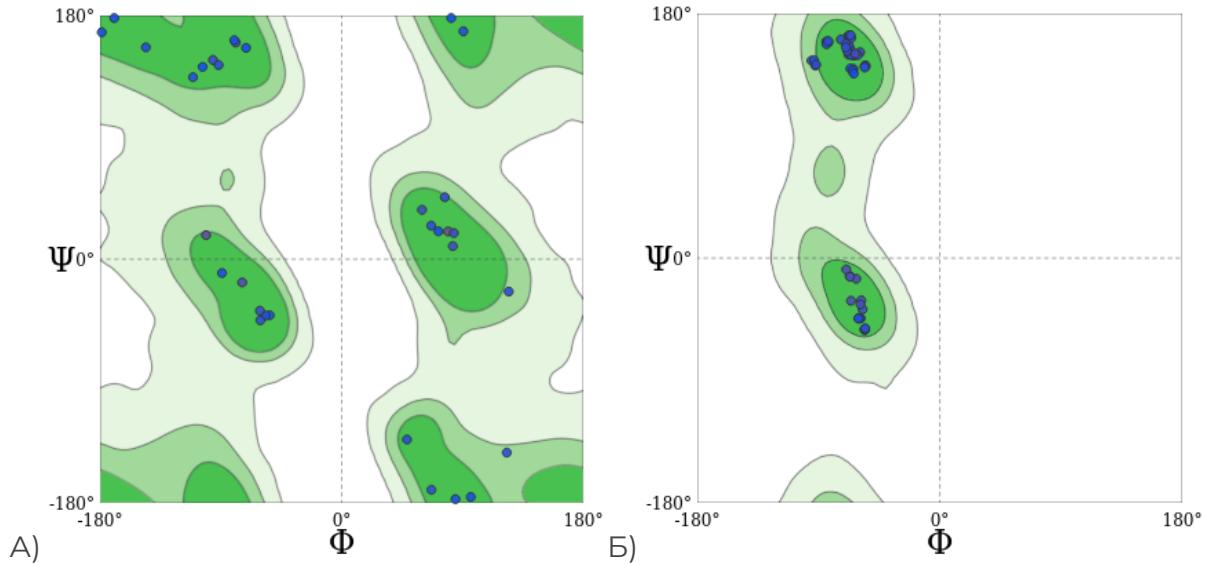


Рис. 5. Карта Рамачандрана структуры 1O7J для глицина А); и пролина Б).

В картах для глицина и пролина (рис.5.) особых аутлаеров не наблюдается, что говорит о хорошем качестве модели с точки зрения углов Ψ и Φ .

Помимо карт Рамачандрана сайт выдает таблицу об аутлаерах по разным параметрам (рис.6.).

MolProbity Score	1.49	
Clash Score	6.44	(D72 LYS-D1331 EDO), (D72 LYS-D1333 EDO), (B19 SER-B30 LYS), (D72 LYS-D1334 EDO), (A243 LYS-A1335 EDO), (A49 LEU-A136 BLEU), (D231 AGLU-D265 LYS), (B232 BTYR-D221 ASP), (B30 LYS-B30 LYS), (A208 BVAL-A313 BARG), (C298 BLEU-C298 LEU), (C164 ARG-C298 BLEU), (B22 THR-B25 GLN), (B232 BTYR-D219 BLYS), (B27 THR-B27 THR)
Ramachandran Favoured	97.29%	
Ramachandran Outliers	0.08%	B28 GLY
Rotamer Outliers	0.62%	A268 GLU, C110 LYS, D30 LYS, B30 LYS, B25 GLN, A71 LEU, C233 LEU
C-Beta Deviations	11	B260 ILE, B71 LEU, B71 LEU, D71 LEU, C199 ILE, A199 ILE, B308 MET, B221 ASP, B43 VAL, B199 ILE, D232 TYR
Bad Bonds	16 / 10348	A1334 EDO-A1335 EDO, D308 MET, B1330 EDO-B1331 EDO, A1333 EDO-A1334 EDO, D302 BHIS, D1333 EDO-D1334 EDO, B1329 SO4, C308 MET, B302 BHIS, B302 AHIS, D1329 SO4, A240 HIS, A1331 EDO-A1332 EDO, D302 AHIS, D1332 EDO-D1333 EDO
Bad Angles	143 / 14032	A198 ARG, A93 HIS, (B28 GLY-B29 TYR), C51 AASN, A41 ASN, A178 ARG, C206 BARG, C51 BASN, D235 ASP, B93 HIS, C206 AARG, D93 HIS, (D212 AARG-D213 AGLY), A51 ASN, C93 HIS, B316 ASP, C308 MET, B148 ARG, A84 ASP, D178 ARG, D296 ASP, A296 ASP, B64 ASN, C178 AARG, D240 HIS, B209 PHE, (B32 GLY-B33 ALA), A146 GLN, B71 LEU, C221 ASP, D68 ASP, A316 ASP, D210 ASP, A1334 EDO, A29 TYR, B192 ARG, C198 BARG, D29 TYR, A158 ASP, C296 ASP, B229 ASP, B240 HIS, A199 ILE, B296 ASP, C199 ILE, D71 LEU, B199 ILE, C51 ASN, C240 HIS, (A31 ALA-A32 GLY), (B31 ALA-B32 GLY), C302 HIS, D199 ILE, B260 ILE, A302 HIS, B308 MET, C72 LYS, R75 GLN, A64 ASN, D299 ASN, A192 ARG, A19 ASFR, (A19 ASFR-A20 ALA), (D213 RG1 Y-D214 F11), A239 RG1 N, R221 ASP, D46 VAI

Рис. 6. Таблица об аутлаерах для структуры 1O7J.

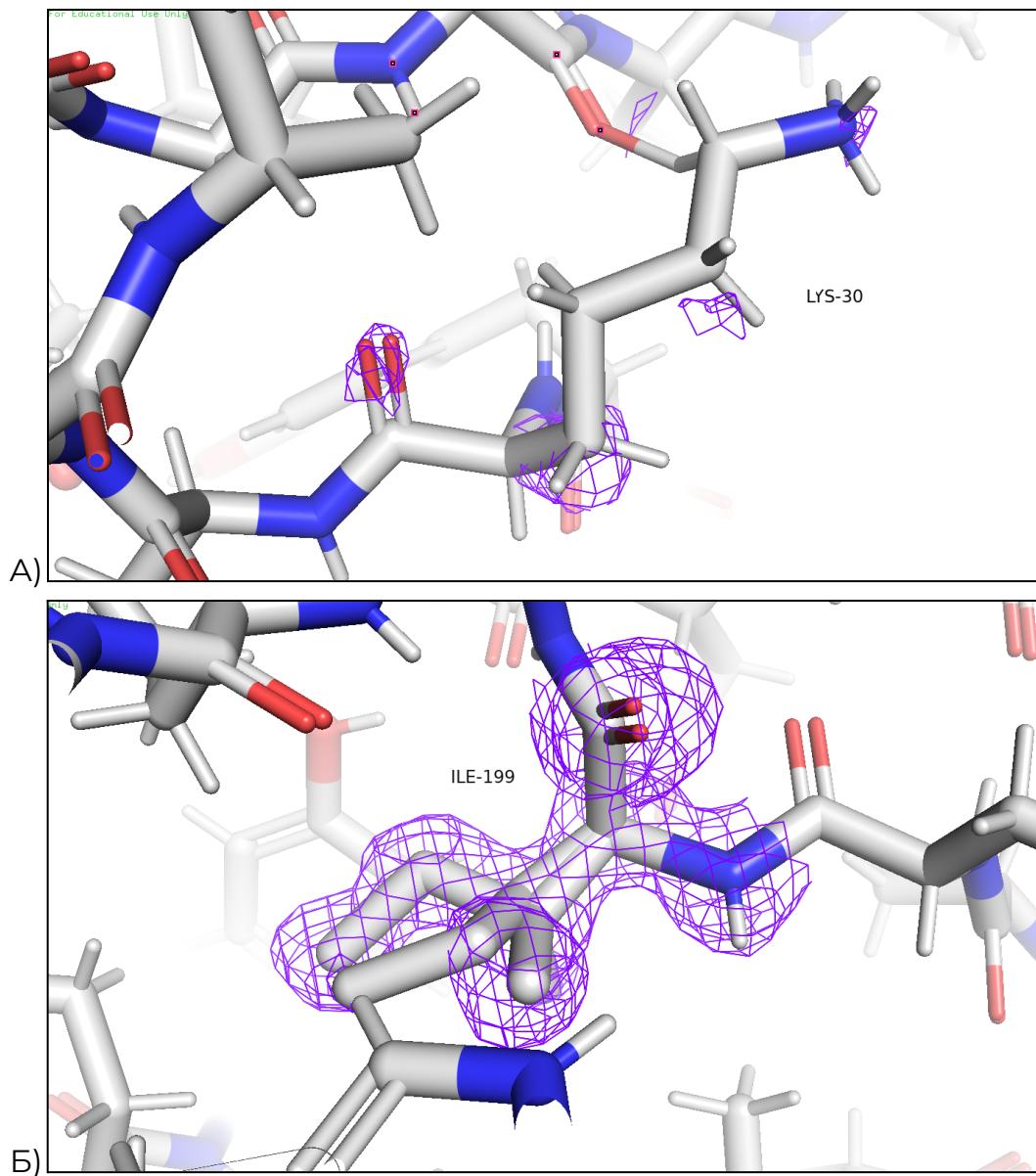
Так как структура была снята в виде гомо-4-мера, в сводной таблице об аутлаерах указаны остатки всех 4 цепей. Можно заметить, что не много остатков являются аутлаерами для всех цепей. Это связано с тем, что в целом все остатки, указанные в таблице находятся примерно на периферии белка. При этом в гомо-4-мере некоторые остатки,

которые были терминальными в одной цепи, в другой попадают внутрь структуры (из-за разной ориентации мономеров).

Выберем несколько остатков на основе таблицы с аутлаерами, которые можно назвать маргинальными.

30 LYS (аутлаер в В и D), 199 ILE(аутлаер во всех цепях), 308 MET (аутлаер в В, D, C).

Посмотрим, соответствуют ли они своей ЭП.



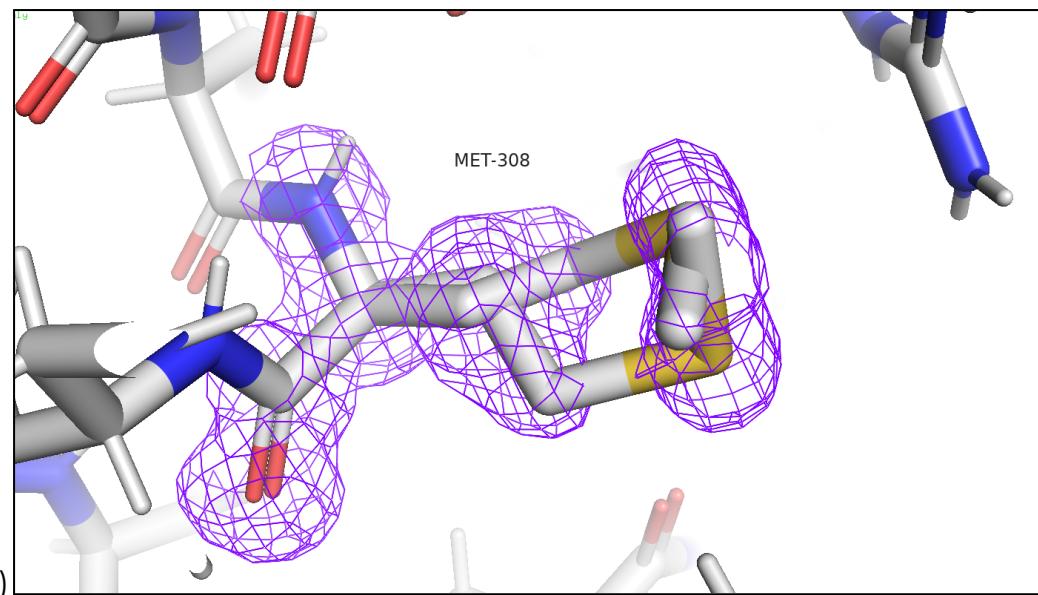


Рис. 7. ЭП (уровень подрезки 1) вокруг маргинальных остатков структуры 1O7J, цепь В. А) 30 LYS; Б) 199 ILE; В) 308 MET.

Вокруг 30 LYS вообще нет ЭП. Думаю его маргинальность связана с особенностью структуры.

Вокруг 199 ILE ЭП хорошо видна, при этом она плохо описывает предложенные альтернативные положения изолейцина. В данном случае ошибка в расшифровке ЭП.

С 308 MET ситуация аналогична - ЭП видна, но плохо описывает предложенную расшифровку. Возможно отдельное облако в конце бокового радикала вообще относится к другому остатку, а оставшаяся ЭП напоминает валин.

Задание 3.

Чтобы подтвердить или опровергнуть предположения об источнике маргинальности воспользуемся моделью из PDB redo, которая была построена на основе той же ЭП, что и 1O7J.

Сначала посмотрим параметры качества перестроенной модели (рис. 8.). Видно, что большая часть параметров изменилась в лучшую сторону. Некоторые показатели (bump severity) изменились достаточно сильно.

В целом новая структура выигрывает первоначальную по параметрам качества.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
Crystallographic refinement		
R	0.1117	0.1062
R-free	⚠ 0.1314	0.1181
Bond length RMS Z-score	0.705	0.599
Bond angle RMS Z-score	1.177	0.958
Model quality (raw scores percentiles)		
Ramachandran plot appearance	55	58
Rotamer normality	93	95
Coarse packing	58	55
Fine packing	91	92
Bump severity	28	74
Hydrogen bond satisfaction	58	60

Рис. 8. Сравнение параметров качества моделей 1O7J и перестроенной структуры из PDB redo.

Что с ЭП вокруг маргинальных остатков?

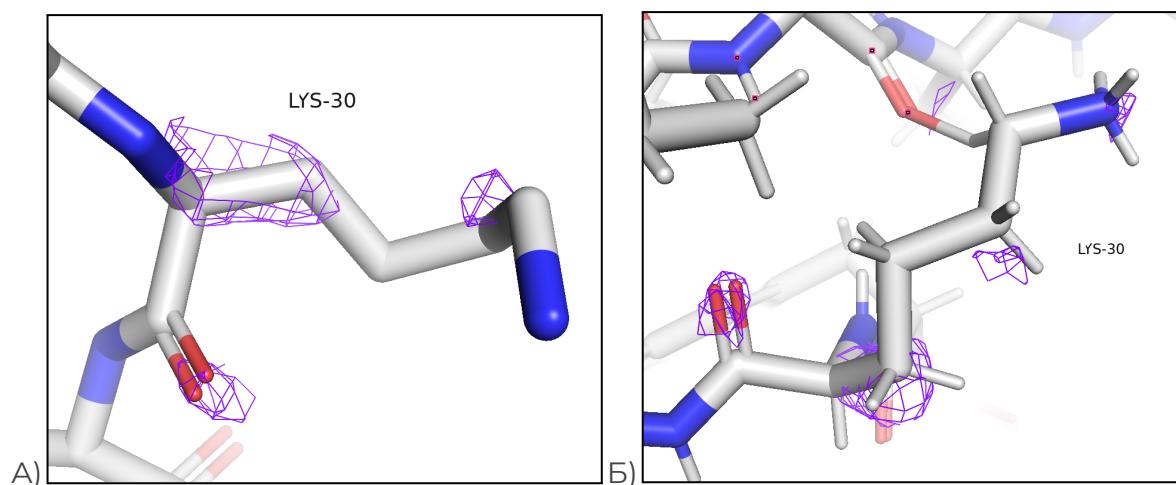


Рис. 9. ЭП (уровень подрезки 1) вокруг 30 LYS в структуре 1O7J, цепь B (А) и перестроенной структуры из PDB redo, цепь B (Б).

ЭП вокруг 30 LYZ в новой структуре все так же не особо покрывает остаток. Хотя расположение остатка (угля между связями) изменилось.

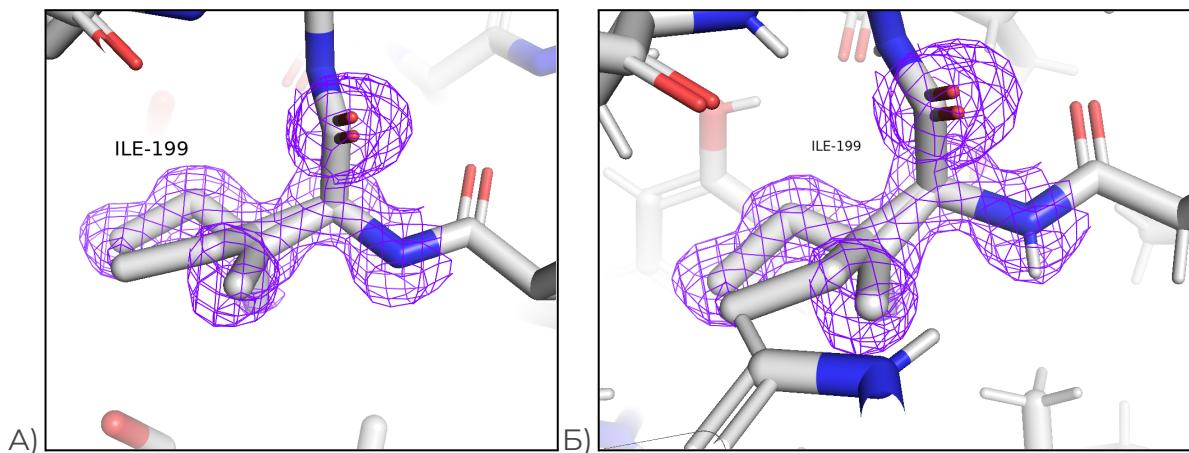


Рис. 10. ЭП (уровень подрезки 1) вокруг 199 ILE в структуре 1O7J, цепь В (А) и перестроенной структуры из PDB redo, цепь В (Б).

Расположение и углы между связями у 199 ILE в старой и новой моделях не поменялись.

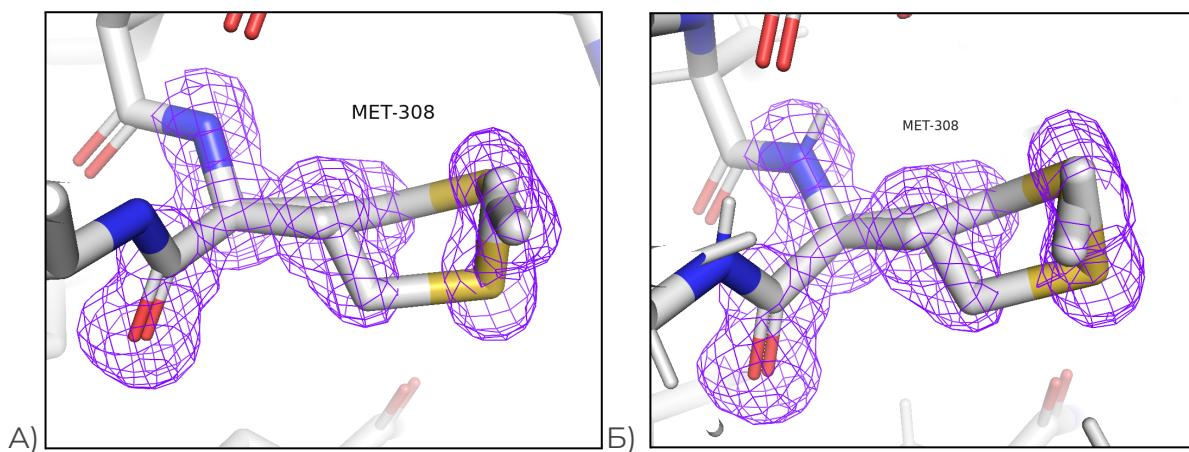


Рис. 11. ЭП (уровень подрезки 1) вокруг 308 MET в структуре 1O7J, цепь В (А) и перестроенной структуры из PDB redo, цепь В (Б).

Альтернативные положения 308 MET немного поменяли положение у углы.

Выводы:

- Структура в целом “нормальная”. Но для анализа данного фермента я бы выбрала либо альтернативную структуру, либо построенную модель по той же ЭП.
- Маргинальные остатки не сильно поменялись в старой и новой структурах. Т.к. 30 LYS маргинален из-за плохой ЭП в том участке (терминальный а.к.о.). 199 ILE и 308 MET немного поменяли положение и углы между атомами боковых радикалов.

Список литературы.

- 1) Lubkowski J, Dauter M, Aghaiypour K, Wlodawer A, Dauter Z. Atomic resolution structure of *Erwinia chrysanthemi* L-asparaginase. *Acta Crystallogr D Biol Crystallogr*. 2003 Jan;59(Pt 1):84-92. doi: 10.1107/s0907444902019443. Epub 2002 Dec 19. PMID: 12499544.

Ссылки на PyMol сессии:

- [1O7J.pse](#)
- [1O7J_redo.pse](#)