

Отчет о качестве расшифровки структуры 3BP5 методом PCA

ВЫПОЛНИЛ СТУДЕНТ 4-ГО КУРСА ФАКУЛЬТЕТА
БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ МГУ ИМ. М.В.
ЛОМОНОСОВА
ПРОСВИРОВ КИРИЛЛ АНТОНОВИЧ

Аннотация

В отчёте рассмотрены некоторые индикаторы качества модели структуры 3VP5, а также проведён анализ индикаторов локального качества структуры для ряда аминокислотных остатков. В ходе работы над отчётом удалось познакомиться с возможностями, которые предоставляют серверы для оценки качества структур белковых молекул.

Введение

PD1 – мембранный белок, который участвует в клеточной дифференцировке иммунных клеток. Имеет два лиганда – PD-L1, PD-L2. PD-L1 предотвращает активацию Т-лимфоцитов, повышая аутоотолерантность. PD-L2 участвует в активации дендритных клеток.

Общая инфо о модели

- Две цепи. А и В. А – белок, В – лиганд, рис. 1.
- Расшифрована группой ученых во главе с **Eszter Lazar-Molnar** в **2008** году в работе: "*Crystal structure of the complex between programmed death-1 (PD-1) and its ligand PD-L2*". [2]
- Получено рефлексов **33562**, использовано для оптимизации **33099**
- Фазовая проблема была решена с помощью **молекулярного замещения** гомологичной структурой **PD-L2 N-terminal IgV domain**
- Разрешение данной структуры **1.8 Å**.
- Диапазон разрешений структурных факторов: **24.95 – 1.80 Å**
- Полнота набора структурных факторов **98.9 %**
- Параметры кристаллографической ячейки: **a = 43.93 Å, b = 84.06 Å, c = 52.11 Å, alpha = 90.00, beta = 107.88, gamma = 90.00**

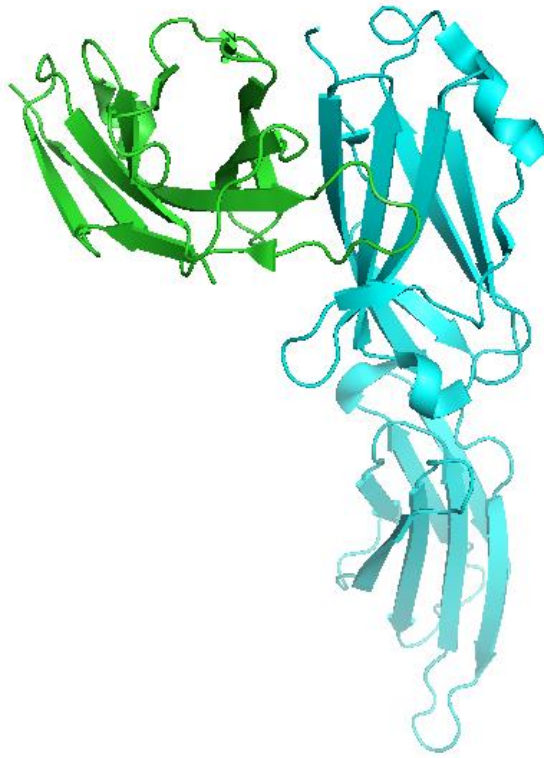


Рис. 1. Пространственная структура. Цепь А (белок) – зеленая, цепь В (лиганд) – бирюзовая.

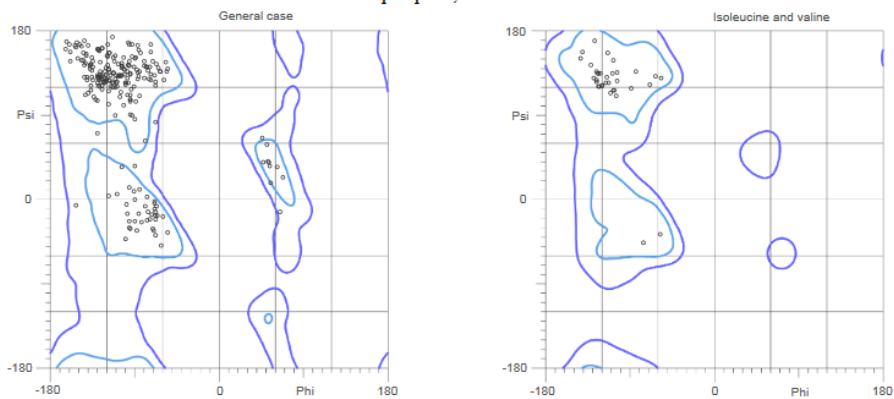
Значение индикаторов качества модели в целом

- R-фактор – **0.191**, R-free – **0.228**. R-фактор меньше **0.25**, что является показателем хорошего качества модели, но R-free больше **0.2**
- RSR – **0.113** со стандартным отклонением в **0.068**. Это является также хорошим показателем.

Анализ в MolProbity

MolProbity Ramachandran analysis

3bp5.pdb, model 1



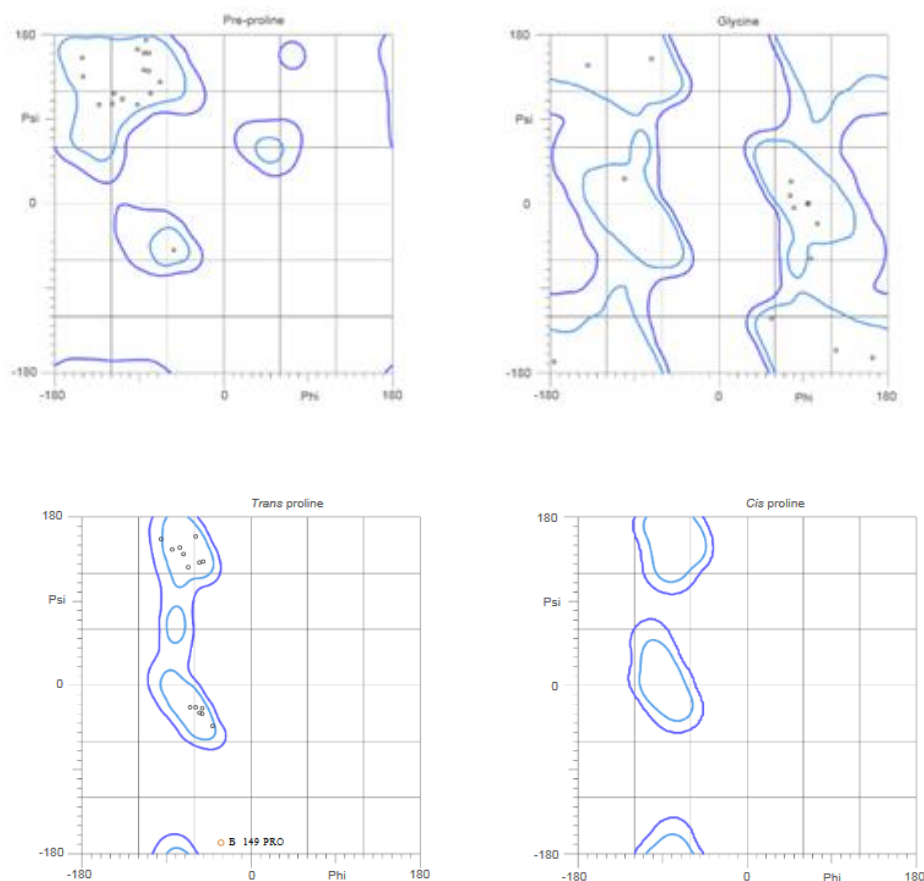


Рис. 2. Карты Рамачандрана для 6 типов аминокислотных остатков. Виден один маргинал – Pro149:B.

Таблица 1. Значение критериев качества модели, предоставленное сервисом MolProbity. Красные значения не удовлетворяют понятию хорошей модели.

Protein geometry	Poor rotamers	13	4.63%	Goal: < 0.3%
	Favored rotamers	249	88.61%	Goal: > 98%
	Ramachandran outliers	1	0.33%	Goal: < 0.05%
	Ramachandran favored	292	97.01%	Goal: > 98%
	Cb deviations	0	0.00%	Goal: 0
	Bad bonds	0 / 2539	0.00%	Goal: 0%
	Bad angles	0 / 3467	0.00%	Goal: < 0.1%
Peptide omegas	Cis prolines	0 / 16	0.00%	Goal: < 5%
	Twisted peptides	2 / 304	0.66%	Goal: 0

Из таблицы 1 видно, что, например, в белке 1 маргинальный остаток, есть плохие ротамеры (аминокислоты повернуты не так как нужно) и есть перекрученные пептиды. Во всем остальное модель хорошая, нет ошибок в связях и углах, все пролины в транс-конформации.

Дополнительный анализ параметров

Z-score для RSR – параметр, позволяющий сравнить RSR остатка модели со средним RSR в выборке моделей PDB с таким же разрешением. Если Z-score остатка выше 2, значит остаток вписывается в плотность значительно хуже остальных в структурах с таким же разрешением. Исходя из этих показателей можно утверждать, что остаток - маргинал. Распределения Z-score для двух цепей показаны на рис. 3.

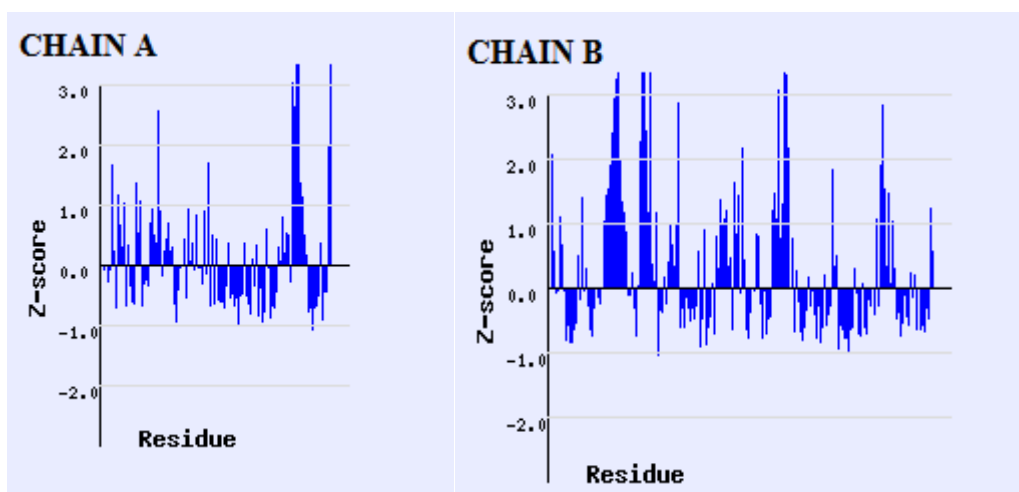


Рис. 3. Значения Z-score для разных цепей – А и В.

Итого можно привести примеры различных маргиналов, найденных разными способами. Некоторые из них приведены в таблице 2.

Таблица 2. Примеры маргинальных остатков в структуре.

Критерий	Остаток
Z-score > 2	Pro97:A – 5.14800
	Leu52:B – 4.7001
	Asn64:B – 5.665
Ramachandran outlier	Pro149:B
Большой температурный фактор	Glu51:B – 82.49 против среднего 32.54
Нехарактерные углы ротамеров	Ile93:A
	Thr43:A
	Leu52:B
	Thr50:B

Примеры маргиналов

Ниже будут рассмотрены 5 примеров различных маргинальных остатков.

Asn64:B – Z-score > 2

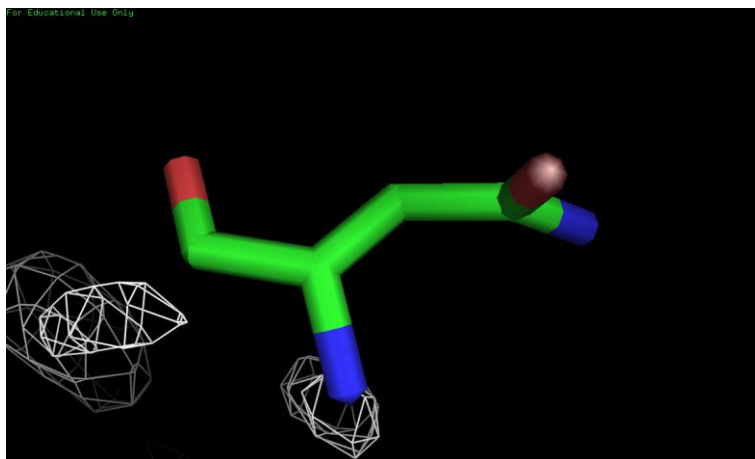


Рис. 4. Аспарагин 64 в цепи В в виде lines, подрезка плотности 1.5.

Для данного остатка наблюдается достаточно большой Z-score относительно средней величины. Посмотрев, можно понять, что остаток ужасно вписан в электронную плотность, это, вероятно, и так сильно искажает Z-score.

Glu51:B – большой температурный фактор (в 2.5 раз больше среднего)

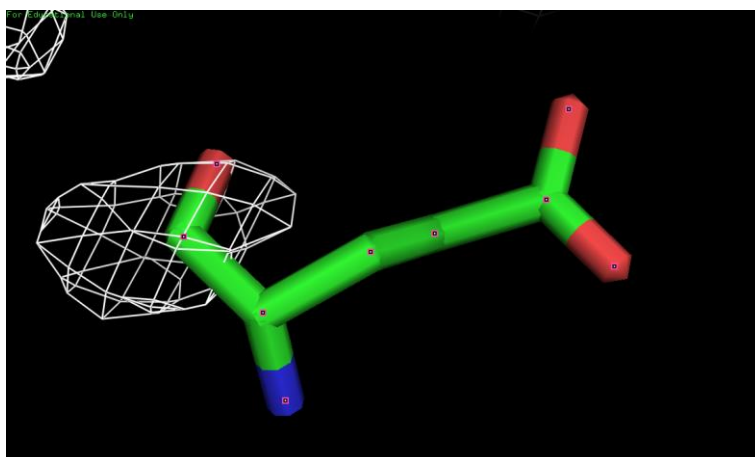


Рис. 5. Глутамин 51 в цепи В в виде lines, подрезка плотности 1.5.

Маргинальность данного остатка вызвана большим температурным фактором. Все остальные параметры находятся в зоне допустимых значений согласно MolProbity. Поэтому маргинальность, в конечном итоге, выражена недостаточной точностью экспериментальных данных.

Pro97:A – Z-score > 2

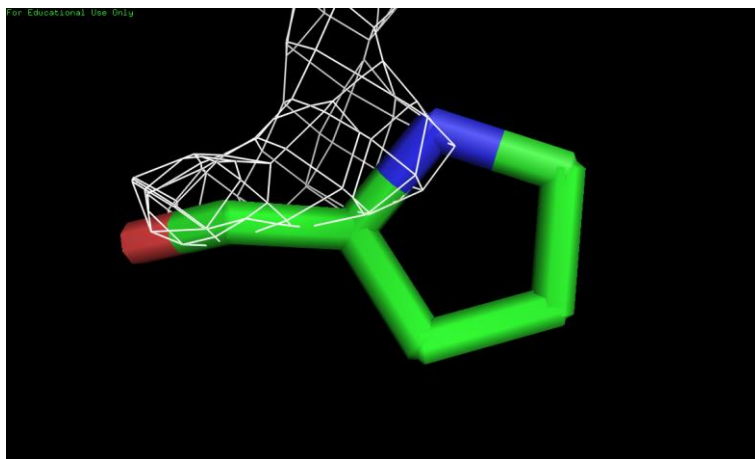


Рис. 6. Пролин 97 в цепи А в виде lines, подрезка плотности 1.5.

Здесь ситуация аналогичная аспарагину 64 из цепи В. Тут Z-score в несколько раз выше среднего. Температурный фактор находится в районе среднего. Поэтому ошибкой в данном случае является – недостаточная точность экспериментальных данных.

Pro149:B – Ramachandran outlier

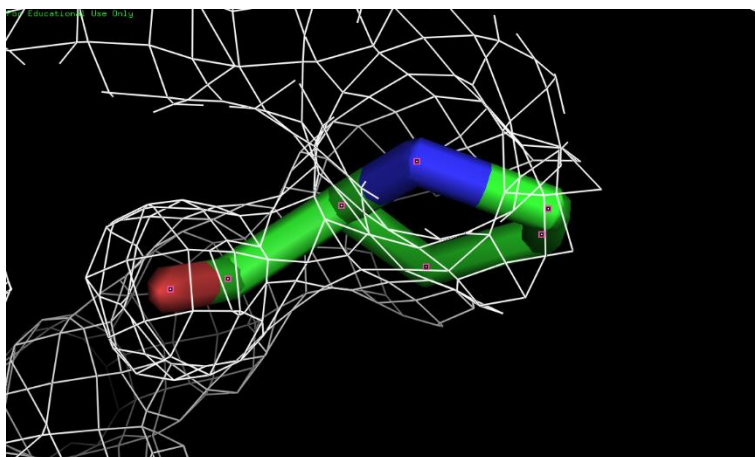


Рис. 7. Пролин 149 в цепи В в виде lines, подрезка плотности 1.5.

Здесь остаток пролин лежит в запрещенной зоне на карте Рамачандрана. Уже по рис. 7. видно, что остаток сильно перекошен и идет как будто бы змейкой. Маргинальность остатка – вероятно, не следствие ошибки расшифровки, т.к. он хорошо вписан в электронную плотность, за исключением азота, там отсутствует ЭП.

Thr43:A – не характерный углы ротамеров

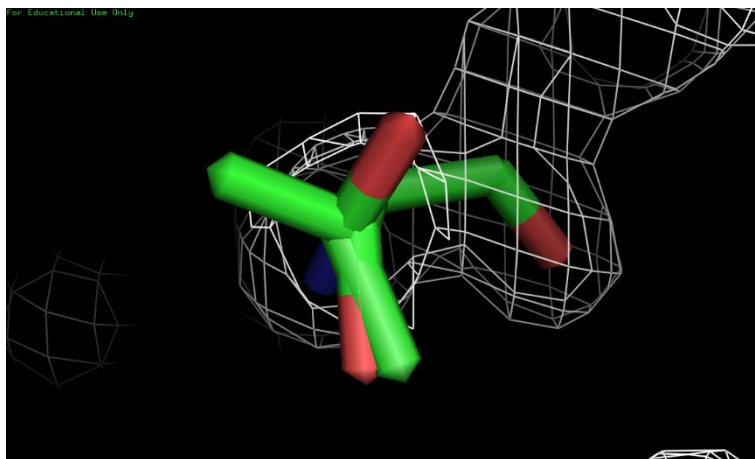


Рис. 8. Треонин 43 в цепи А в виде lines, подрезка плотности 1.5.

Здесь очевидна ошибка, все наложилось друг на друга. Это ошибка расшифровки, т.к. здесь остатки физически «друг в друге», так они ещё и не влезают в плотность.

Сравнение с PDB_redo

Таблица 2. Примеры маргинальных остатков в структуре.

Параметр	3bp5	3bp5_final
R-factor	0.191	0.1754
R-free	0.228	0.2033
R-free – R-factor	3.7%	2.79%
Z-score	0.89	0.84
Ramachandran outliers	1 (0.33%)	2 (0.66%)
Ramachandran favored	292 (97.01%)	292 (97.01%)
Poor rotamers	13 (4.64%)	10 (3.56%)
Good rotamers	249 (88.61%)	260 (92.54%)
Cb derivation	0.00%	0 (0.00%)
Bad bond	0.00%	0 (0.00%)
Bad angles	0.00%	0 (0.00%)
Cis prolines	0 / 16 (0.00%)	2 / 16 (12.50%)

Как видно, почти все характеристики, относящие к параметрам структуры, улучшились. Единственное, что смущает, это появляется цис-пролинов, аж 12.50%, а это в два раза выше порога в 5%. Но в общем, можно заключить, что улучшенная структура имеет большее отношение к реальности, чем старая.

Заключение

Модель 3bp5 является моделью качества чуть выше среднего, хотя все равно присутствуют недостатки: достаточно много остатков в необычной конформации, есть остаток, выпавший в запрещенную область карты Рамачандрана. А в целом структура имеет приемлемое качество.

Отчасти качество связано с недостаточностью экспериментальных данных, т.к. при помощи PDB_redo удалось улучшить основные характеристики структуры, пожертвовав двумя пролинами и ещё одной аминокислотой, попавшей в запрещенную область на карте Рамачандрана.