

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ имени М.В.ЛОМОНОСОВА  
ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И  
БИОИНФОРМАТИКИ

Отчет по качеству расшифровки PCA структуры  
пепсина (PDB ID:1PSO)

Выполнила студентка 4-го курса Трифонова Софья

Москва, 2021 год

## Оглавление

Введение ..... 3



..... 4

Результаты..... 4

    Модель ..... 4

Значения индикаторов качества модели в целом..... 5

Сравнение модели PDB и PDB\_redo. .... 13

Заключение ..... 14

Литература ..... 15

## Введение

Пепсин – аспарагиновая протеаза, которая вырабатывается слизистой оболочкой желудка человека в семи различных изоформах зимогена. Они были разделены на два типа: пепсиноген I (пепсиноген A), состоящий из PGA 1-5, и пепсиноген I1 (пепсиноген C или прогастринсин), состоящий из PGC 6 и 7.

Пепсины играют значительную роль в пищеварении у млекопитающих, в том числе у человека, являясь ферментом, выполняющим один из важных этапов в цепочке превращений белков пищи в аминокислоты. Железами

желудка пепсин вырабатывается в неактивном виде, переходит в активную форму при воздействии на него соляной кислоты. Пепсин действует только в кислой среде желудка и при попадании в щелочную среду двенадцатиперстной кишки становится неактивным.

Структурное исследование пепсина человека и других аспарагиновых протеиназ имеет большое применение в разработке лекарственных препаратов. Для обеспечения пероральной биодоступности и специфического действия препаратов, направленных на ингибирование аспарагиновых протеиназ, таких как ренин или протеаза ВИЧ, препарат не должен связываться с пепсином или другими аспарагиновыми протеиназами в организме. С этой целью важно понять происхождение субсайтовых специфичностей и различия между специфичностями ферментов человека.



Рисунок 1. Структура пепсина(1PSO)

## Результаты

Модель

Состав комплекса: Пепсин 3A, цепь E, пепстатин, цепь I.

Год: 1995

Авторы: Fujinaga, M.; Chernaia, M.M.; Tarasova, N.; Mosimann, S.C.; James, M.N.G.

Статья: The crystal structure of human pepsin and its complex with pepstatin

Метод решения фазовой проблемы: NULL

Число измеренных рефлексов: 30584

Разрешение: 2 Å

Полнота набора структурных факторов: 98, 6%

Параметры кристаллографической ячейки:

Кристаллографическая группа: **P 21 21 21**

$a = 72.068\text{Å}$      $\alpha = 90^\circ$

$b = 150.972\text{Å}$      $\beta = 90^\circ$

$c = 40.854\text{Å}$      $\gamma = 90^\circ$

Значения индикаторов качества модели в целом R-factor – мера согласованности между кристаллографической моделью и экспериментальным массивом рентгеновских данных. В представленной модели R-factor равен 0,185, что является хорошим показателем.

Значение R-free 0,166, что также является хорошим показателем.

С помощью сервера MolProbity была получена таблица:

All-Atom Contacts	Clashscore, all atoms:	2.96	99 <sup>th</sup> 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	5	1.81% Goal: <0.3%
	Favored rotamers	266	96.03% Goal: >98%
	Ramachandran outliers	2	0.61% Goal: <0.05%
	Ramachandran favored	323	98.78% Goal: >98%
	MolProbity score <sup>^</sup>	1.28	99 <sup>th</sup> percentile* (N=12522, 2.00Å ± 0.25Å)
	Cβ deviations >0.25Å	4	1.35% Goal: 0
	Bad bonds:	4 / 2499	0.16% Goal: 0%
Bad angles:	15 / 3427	0.44% Goal: <0.1%	
Peptide Omegas	Cis Prolines:	1 / 17	5.88% Expected: ≤1 per chain, or ≤5%
	Twisted Peptides:	1 / 330	0.30% Goal: 0

Расшифровка таблицы:

Контакты всех атомов:

ClashScore — число недопустимых наложений на 1000 атомов; 99 перцентиль означает, что у 99% структур с таким же разрешением ClashScore хуже (то есть наша модель отличная).

Геометрия белка:

Poor rotamers — количество маргинальных остатков по отклонению боковой цепи; в данной модели значение составляет 1.81% (5 остатка).

Ramachandran outliers — маргиналы по карте Рамачандрана, которые лежат вне допустимой области; в идеале таких должно быть > 98%, в представленной модели их почти нет (0,61%)

Ramachandran favored — остатки, лежащие в предпочитаемой области; в данной модели значение почти идеальное (96.76%).

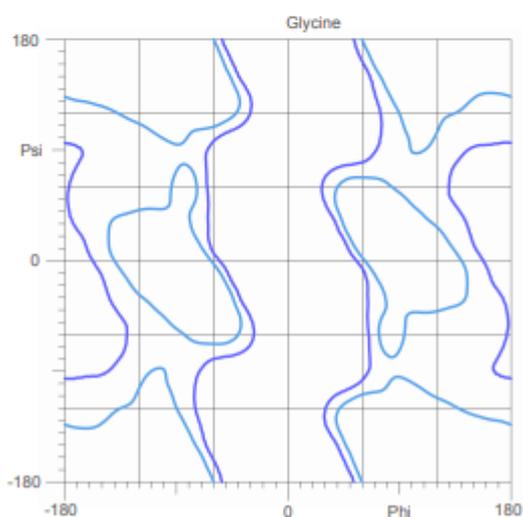
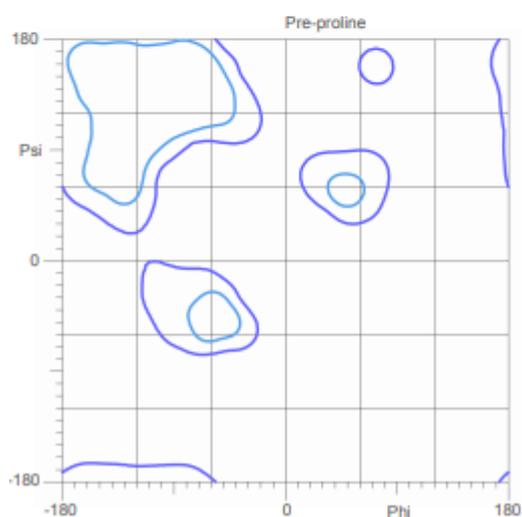
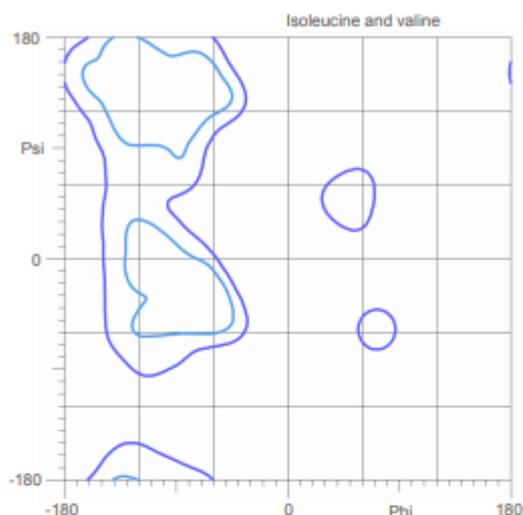
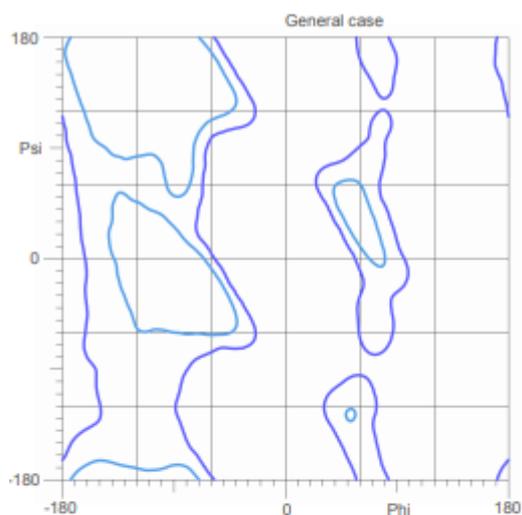
MolProbity score — интегральная оценка структуры по данным этого сервиса; 99 перцентиль означает, что у 99% структур с таким же разрешением MolProbity score хуже.

Cβ deviations > 0.25Å — число Cβ-атомов с неприемлемым отклонением от ожидаемого положения; таких в структуре 0.

Bad bonds — число связей, существенно отклоняющихся от теории (в данной структуре их 4);

Bad angles — число углов, существенно отклоняющихся от теории (в данной структуре их 15).

По карте Рамачандрана аутлаеры составляют 0,61%, 99,1% остатков попадают в предпочтительную область и 99,7% всех остатков находятся в разрешенных областях.



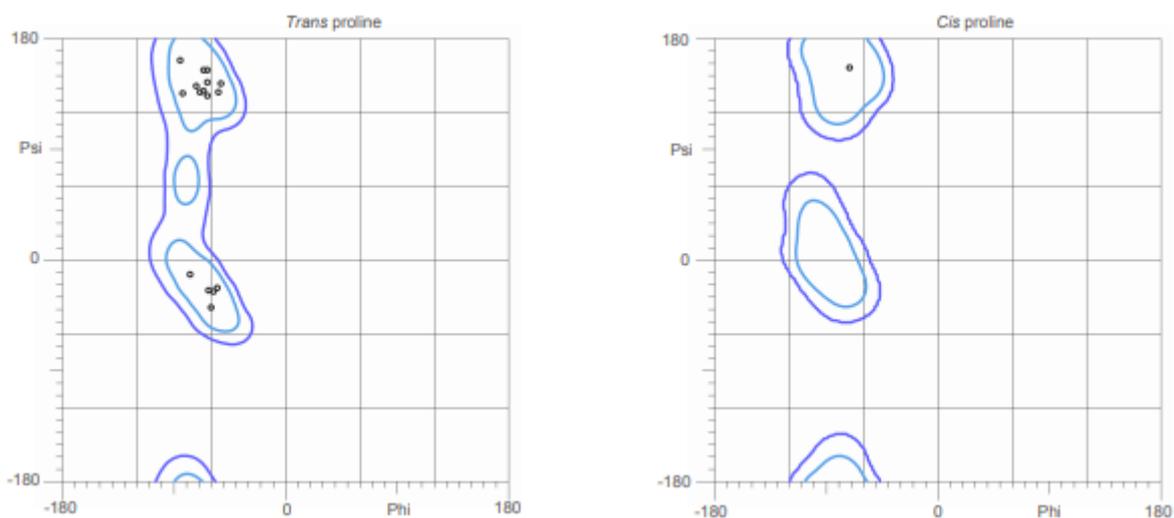


Рисунок 2. Карты Рамачандрана для структуры 1PSO.

## Анализ маргинальных остатков

Остаток	Показатель
E:247:VAL	Перекрывается с E:275:ILE- 0.71
E:222:THR	Перекрывается с E:300:ILE- 0.58
E 142 ASN	Неротамер
E 266 GLN	Неротамер
E:296:GLY	RSRZ=2.1
D:186(B):GLU	Z=6.04

Рассмотрение пяти маргиналов.

### Chain E – VAL-247

Этот остаток является маргинальным по пересечению с ILE-275 по атому HD12. Перекрывание атомов составляет 0.71 Å. Боковая цепь у данного остатка- ротамер.

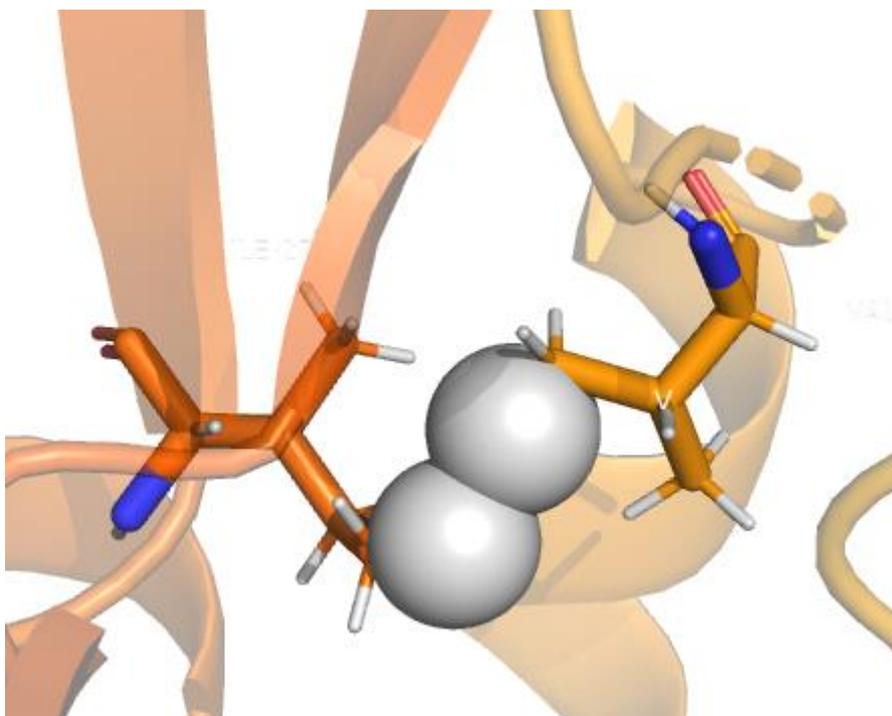


Рисунок 3. Перекрывание атомов HG-11 VAL-247 и HD-12 ILE-275

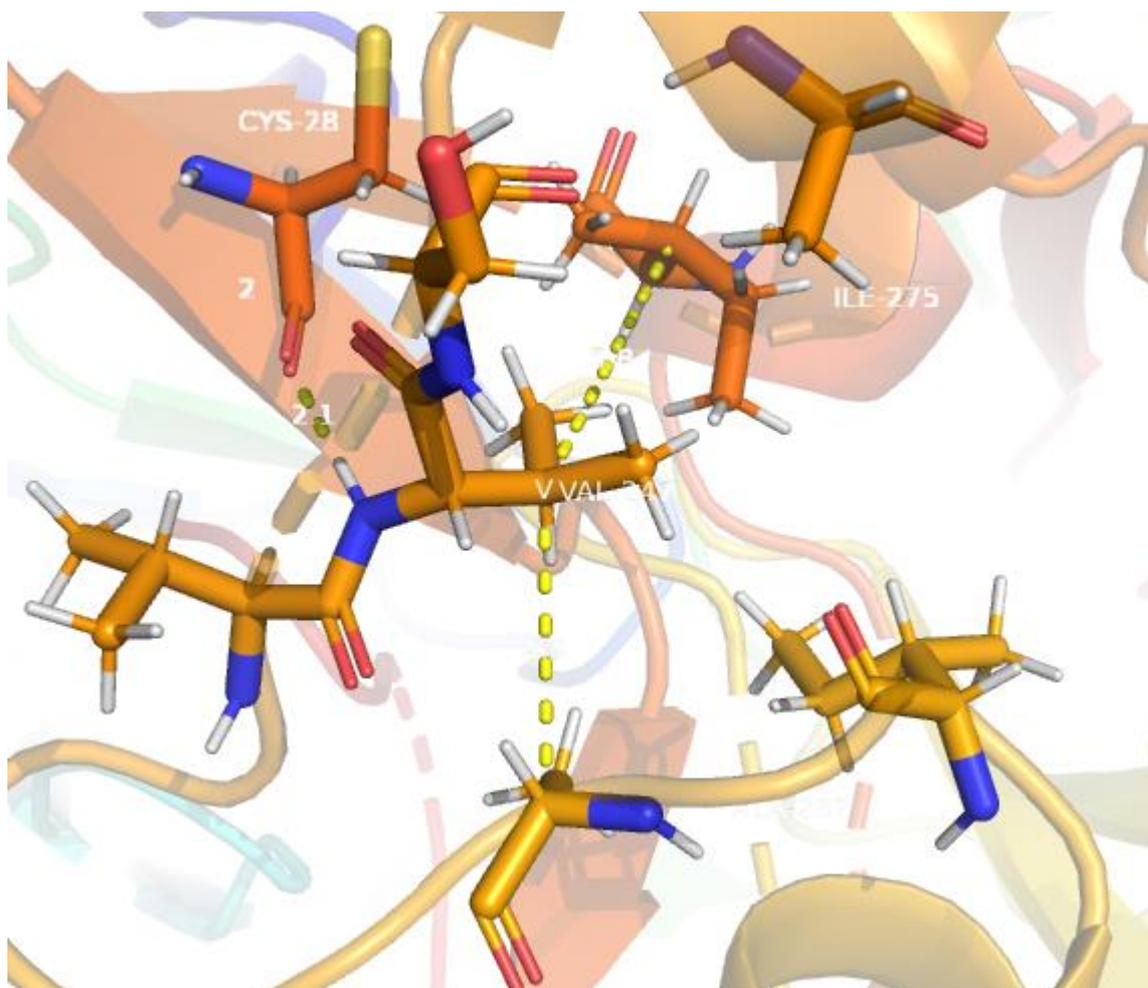


Рисунок 4. Взаимодействия, стабилизирующие VAL-247.

**Chain E: THR-222**

В этой ситуации атом НВ THR-222 перекрывается с атомом НВ ILE-300. Перекрывание - 0.58 Å. Значение перекрывания не является критичным.

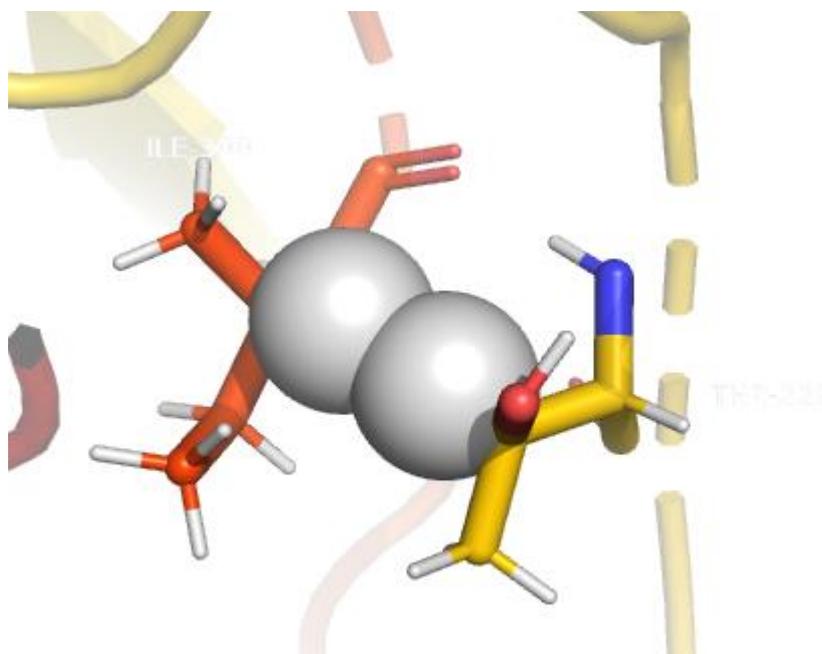


Рисунок 5. Перекрывание атомов НВ ILE-300 и THR-222

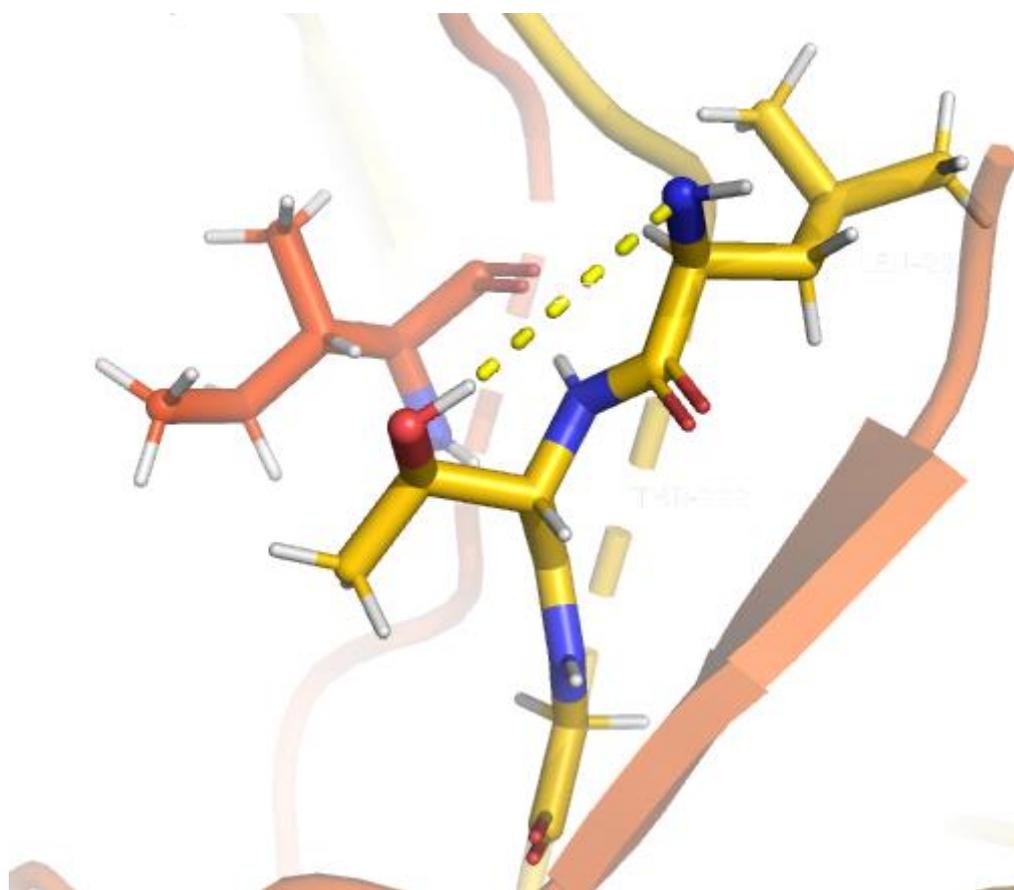


Рисунок 6. Стабилизация THR-222.

## Chain E: ASN-142

Данный остаток является неротамером. Расположен в альфа-спирали, не взаимодействует ни с кем из окружения.

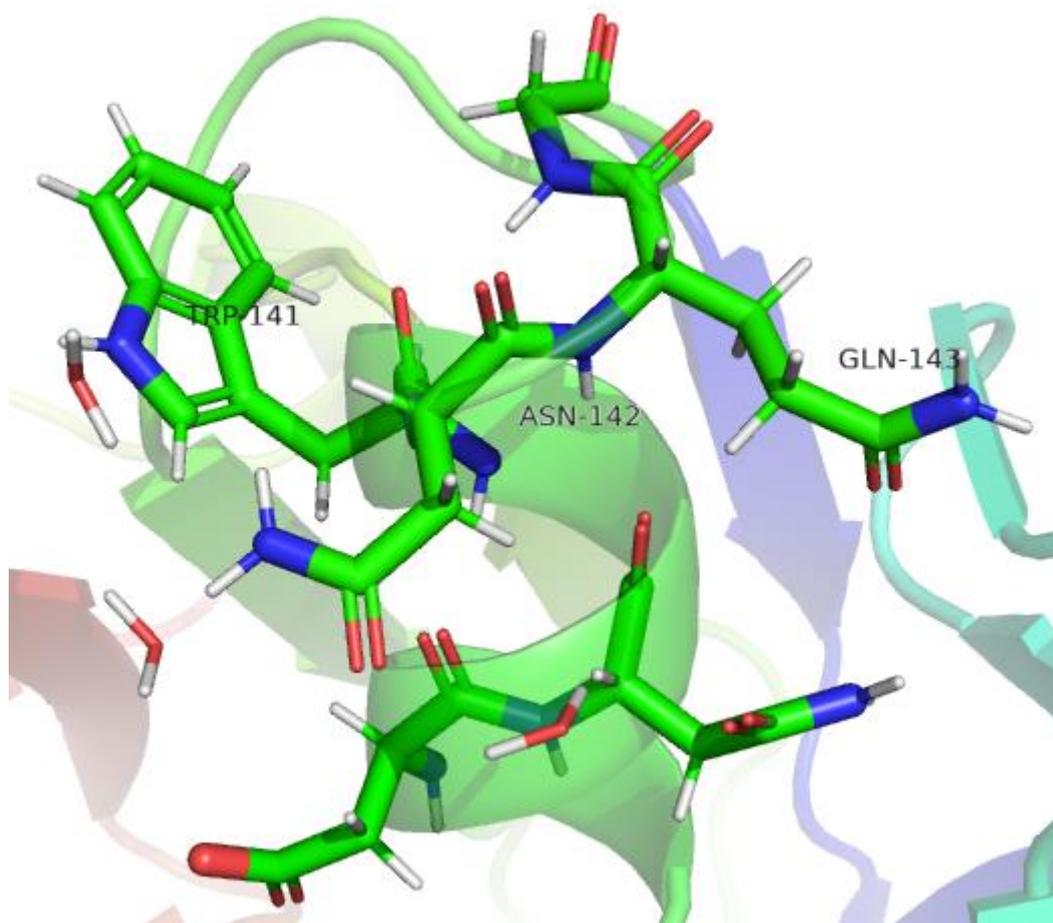


Рисунок 7. ASN-142 и его окружение в диапазоне 3 ангстрем.

## Chain E: GLY – 296

Для данного остатка параметром выбора было значение RSRZ = 2,1. Оно не слишком большое, а значит остаток должен неплохо вписаться в электронную плотность.

При уровне подрезки 0.5 и  $\sigma_{\text{cut}}=2$ , GLY-296 хорошо вписывается в ЭП.

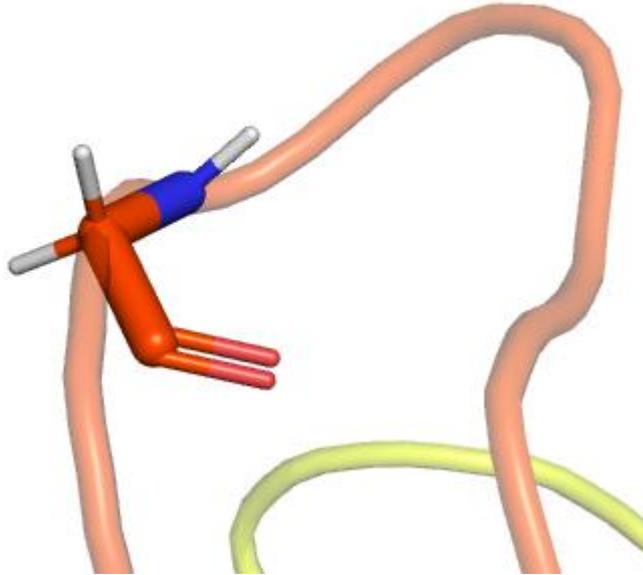


Рисунок 8. GLY-296 в составе белка

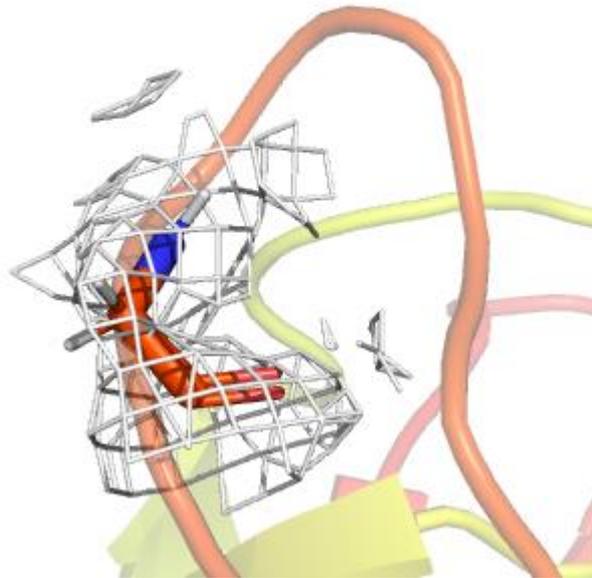


Рисунок 9. Остаток GLY-296 и его ЭП, уровень подрезки 0.5, carve=2.

### **Chain E: GLN-266**

Неротамер, не взаимодействует ни с кем из окружения.

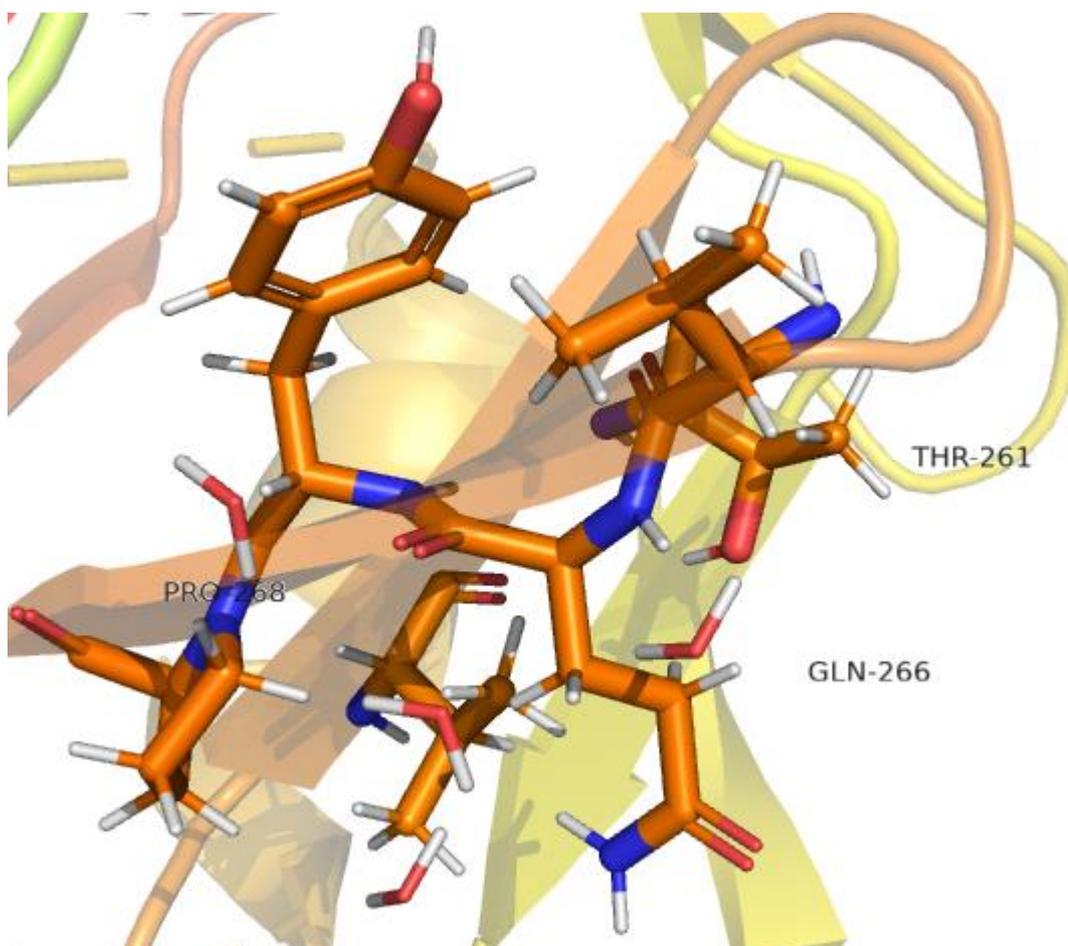


Рисунок 10. Остаток GLN-266 и его окружение в диапазоне 3 ангстрем.

#### Сравнение модели PDB и PDB\_redo.

PDB\_redo – это база данных, где хранятся оптимизированные модели структур из PDB.

После оптимизации структуры значение R-фактора стало 0,1402 . R-free уменьшился до 0,1764. Также были изменены 5 ротамеров, удалено 68 молекул воды.

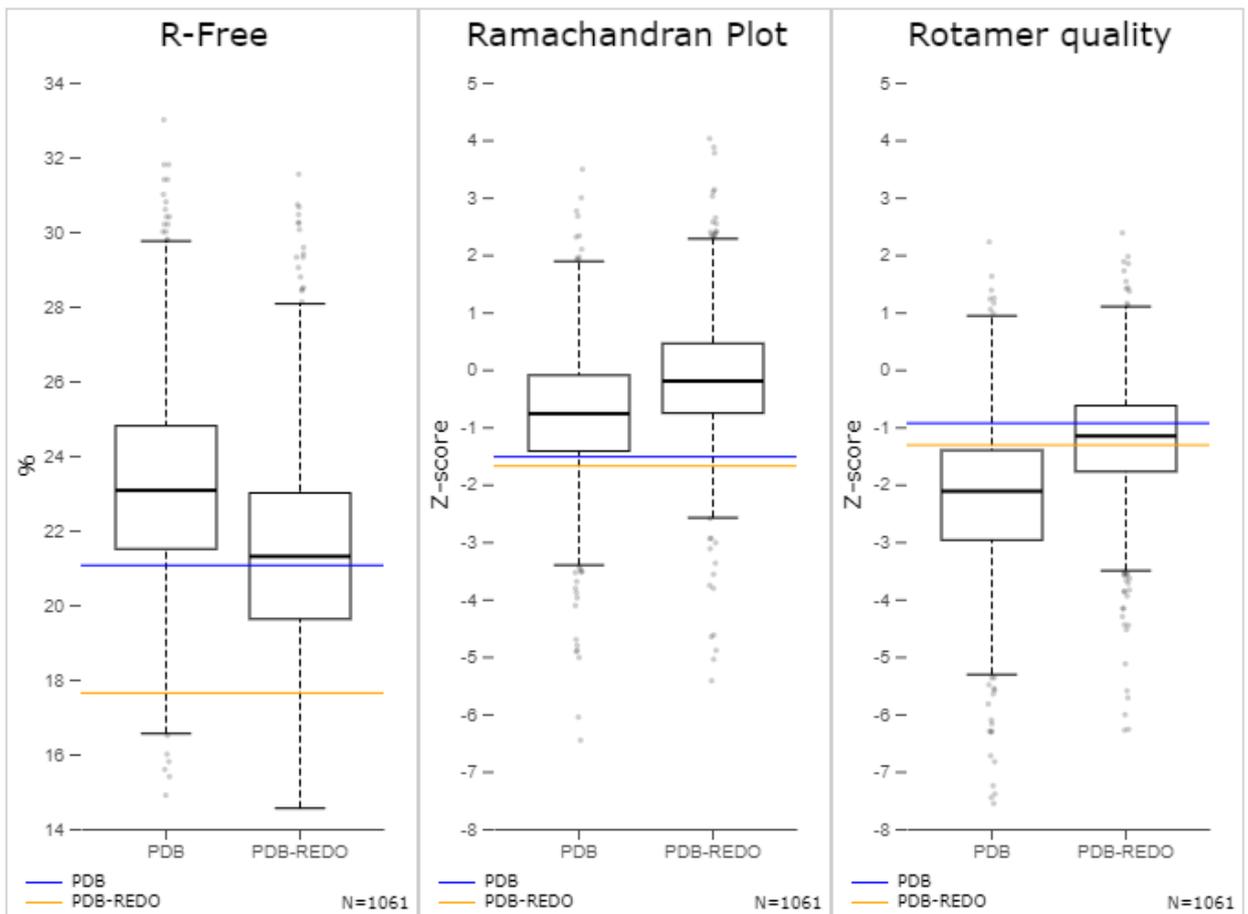


Рисунок 11. Сравнение качества показателей модели для pdb и pdb\_redo.

## Заключение

В данной работе был рассмотрен пепсин из базы данных PDB (1PSO) и проведен анализ качества его структуры.

Структура относительно неплохая, количество маргиналов небольшое.

## Литература

1. <https://www.rcsb.org/structure/1PSO>
2. <http://molprobitry.manchester.ac.uk/index.php>
3. <https://pdb-redo.eu/db/1pso>