

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В.ЛОМОНОСОВА**

**ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И  
БИОИНФОРМАТИКИ**

**Отчет по качеству расшифровки РСА структуры  
тромбина (PDB ID: 2JH5).**

*Выполнила студентка 4-го курса*

**Наумова Юлия Олеговна**

## Оглавление

|   |    |
|---|----|
| Аннотация .....                                   | 2  |
| Введение .....                                    | 3  |
| Результаты .....                                  | 4  |
| Общая информация о модели.....                    | 4  |
| Значения индикаторов качества модели в целом..... | 5  |
| Анализ маргинальных остатков.....                 | 8  |
| Сравнение модели PDB и PDB_redo.....              | 14 |
| Заключение.....                                   | 16 |
| Список литературы .....                           | 17 |

## Аннотация

В данной работе был проведен анализ качества расшифровки структуры тромбина (PDB ID: 2JH5), полученной методом рентгеноструктурного анализа. В результате работы был сделан вывод о качестве расшифровки структуры белка.

## Введение

Тромбин – сериновая протеиназа семейства химоотрипсина, центральный белок свертывания крови человека и животных. Он относится к гидролазам (классу ферментов), катализирует гидролиз пептидных связей, а также обладает ограниченной субстратной специфичностью.

Избыточная активность тромбина может приводить к внутрисосудистому тромбообразованию, так как фермент превращает фибриноген в фибрин, который является основой тромба. Также тромбин может выполнять противосвертывающую функцию, связываясь с тромбомодулином. Фермент, помимо всего этого, принимает участие в процессах воспаления через рецепторы, активируемые протеазами.

Однако результатом тромбообразования могут быть инсульт, инфаркт миокарда, нестабильная стенокардия и другие опасные заболевания. Поэтому вопрос поиска оптимальных ингибиторов этого фермента для снижения риска сердечно-сосудистых заболеваний на данный момент является достаточно важным.

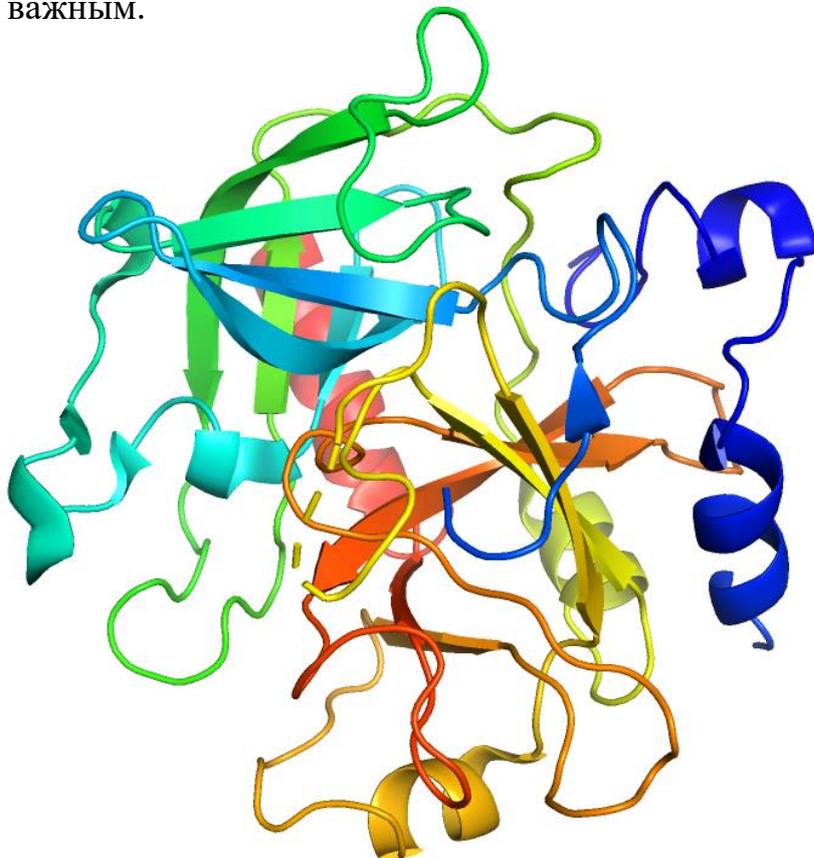


Рисунок 1. Структура тромбина (PDB ID: 2JH5).

# Результаты

## Общая информация о модели

Данные с сайта RCSB были проанализированы и преобразованы в таблицу:

Таблица 1. Общая информация о структуре 2JH5.

|                                       |  |
|---------------------------------------|--|
| Состав комплекса                      | Две цепи тромбина (LIGHT CHAIN, HEAVY CHAIN) и ингибитор гирудин (HIRUDIN IIIA).   |
| Год                                   | 2007   |
| Авторы                                | Senger, S., Chan, C., Convery, M.A., Hubbard, J.A., Young, R.J., Shah, G.P., Watson, N.S.  |
| Статья                                | Senger, Stefan, Chuen Chan, Máire A. Convery, Julia A. Hubbard, Gita P. Shah, Nigel S. Watson, and Robert J. Young. 2007. "Sulfonamide-Related Conformational Effects and Their Importance in Structure-Based Design." <i>Bioorganic and Medicinal Chemistry Letters</i> 17 (10): 2931–34. doi:10.1016/j.bmcl.2007.02.034. |
| Метод решения фазовой проблемы        | Молекулярное замещение   |
| Число измеренных рефлексов            | 11571  |
| Разрешение                            | 2.50 Å   |
| Полнота набора структурных факторов   | 98.0%  |
| Параметры кристаллографической ячейки | $a = 70.363\text{Å}$ , $b = 71.478\text{Å}$ , $c = 72.106\text{Å}$<br>$\alpha = 90.00^\circ$ , $\beta = 100.62^\circ$ , $\gamma = 90.00^\circ$   |
| Кристаллографическая группа           | C 1 2 1  |

## Значения индикаторов качества модели в целом

Параметры для оценки качества полученной модели представлены ниже.

- **R-Value Free** (тестовая выборка): 0.231
- **R-Value Work** (тренировочная выборка): 0.185
- **R-Value Observed** (весь набор рефлексов): 0.188

В выбранной модели R-фактор, который характеризует соответствие модели ее экспериментальным значениям, равен 0.185, что является хорошим показателем (хорошими считаются значения  $< 0.25$ ).

Фактор R-free является контролем модели (чтобы не терять различные параметры и не судить только по R-фактору), и его значение равно 0.231, что в целом вписывается в допустимые значения (хороший R-free  $< 0.2$ ).

Также наша структура была оценена с помощью сервера MolProbity. Для лучшей оценки были добавлены водороды и произведена оптимизация связей.

Таблица 2. Оценка качества структуры 2JH5 с сервера MolProbity.

|                   |  |                  |                 |  |
|-------------------|--|------------------|-----------------|--|
| All-Atom Contacts | Clashscore, all atoms:   | 5.11             |                 | 99 <sup>th</sup> percentile* (N=271, 2.50Å ± 0.25Å)  |
|                   | Clashscore is the number of serious steric overlaps ( $> 0.4$ Å) per 1000 atoms. |                  |                 |  |
| Protein Geometry  | Poor rotamers  | 3                | 1.19%           | Goal: $< 0.3\%$                                      |
|                   | Favored rotamers   | 243              | 96.05%          | Goal: $> 98\%$                                       |
|                   | Ramachandran outliers  | 0                | 0.00%           | Goal: $< 0.05\%$                                     |
|                   | Ramachandran favored   | 269              | 96.76%          | Goal: $> 98\%$                                       |
|                   | Rama distribution Z-score  | $-1.55 \pm 0.45$ |                 | Goal: $\text{abs}(Z \text{ score}) < 2$              |
|                   | MolProbity score <sup>^</sup>  | 1.53             |                 | 99 <sup>th</sup> percentile* (N=6960, 2.50Å ± 0.25Å) |
|                   | C $\beta$ deviations $> 0.25$ Å  | 0                | 0.00%           | Goal: 0  |
|                   | Bad bonds:   | 3 / 2402         | 0.12%           | Goal: 0%   |
| Bad angles:       | 0 / 3241   | 0.00%            | Goal: $< 0.1\%$ |  |
| Peptide Omegas    | Cis Prolines:  | 1 / 16           | 6.25%           | Expected: $\leq 1$ per chain, or $\leq 5\%$          |

Краткая расшифровка таблицы:

***Контакты всех атомов:***

Здесь 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; данное значение не должно превышать 0.3, однако в нашей модели значение составляет 1.19% (3 остатка). Но я считаю это допустимым.

Ramachandran outliers — маргиналы по карте Рамачандрана, которые лежат вне допустимой области; в данной модели таких остатков нет.

Ramachandran favored — остатки, лежащие в предпочитаемой области; в данной модели значение практически 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.

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

Bad bonds — число связей, существенно отклоняющихся от теории (в данной структуре их 3);

Bad angles — число углов, существенно отклоняющихся от теории (в данной структуре их 0).

По карте Рамачандрана аутлаеров выявлено не было, 96.8% остатков попадают в предпочтительную область и 100% всех остатков находятся в разрешенных областях (рисунок 2).

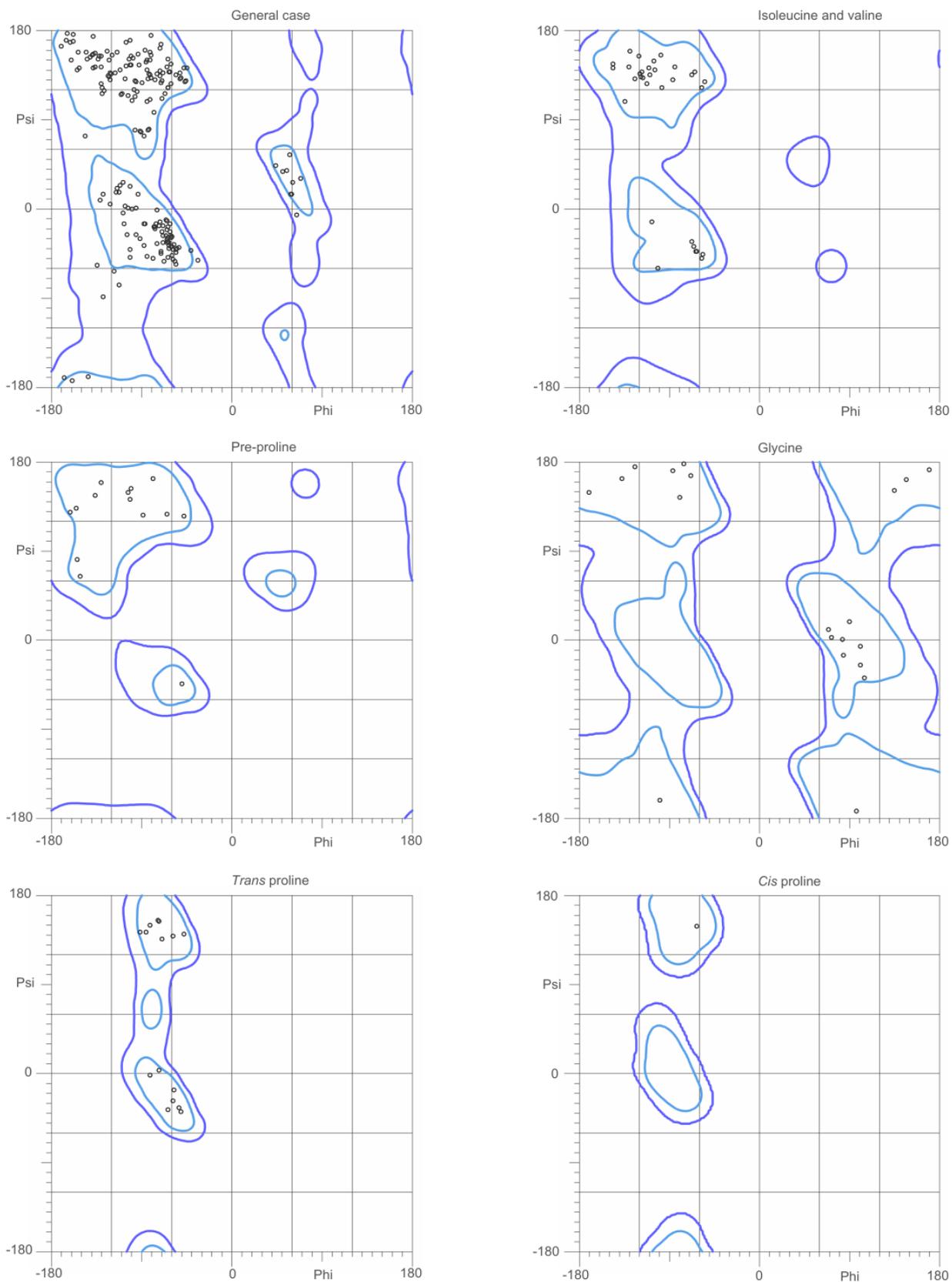


Рисунок 2. Карты Рамачандрана для структуры 2JH5.

## Анализ маргинальных остатков.

Таблица 3. Список маргинальных остатков по различным параметрам.

| Остаток   | Показатель  |
|-----------|---|
| C:CYS-1   | Перекрывается с D:CYS-122<br>1:CYS:C - 122:CYS:SG - 0.407   |
| C:LYS-9   | Перекрывается с D:PRO-5<br>5:PRO:CA – 9:LYS:HG3 - 0.421<br>5:PRO:HA – 9:LYS:HG3 - 0.662<br>Неротамер                                    |
| D:ARG-75  | Перекрывается с D:GLU-77<br>75:ARG:HD2 - 77:GLU:HG2 - 0.987<br>75:ARG:CD - 77:GLU:HG2 - 0.969<br>75:ARG:HD2 – 77:GLU:CG - 0.946         |
| D:GLN-30  | Неротамер   |
| D:GLN-38  | Неротамер   |
| D:GLN-244 | Неротамер   |
| C:ASP-15  | RSRZ = 3.4  |
| D:PRO-37A | Пересекается с D:SER-37<br>37:SER:HA - 37A:PRO:C - 0.614<br>* Также является единственным пролином с цис конформацией во всей структуре |

\* При плохом перекрывании атомов значения должны быть  $\geq 0.4$  ангстрем.

Далее 5 маргиналов были рассмотрены более подробно.

### Chain C – LYS-9.

Данный остаток является маргинальным по пересечению с PRO-5 по нескольким атомам (атомы CA - HG3 и HA - HG3). Также по выдаче PDB report у данного остатка боковая цепь – неротамер.

В целом, перекрытие атомов не сильное (для CA – HG3 значение 0.421, а для HA – HG3 значение 0.662).

Скорее всего перекрытие атомов и неротамерность получаются, так как наш LYS-9 стабилизирован двумя водородными связями и такое его положение является наиболее оптимальным в данной структуре (рисунок 5).

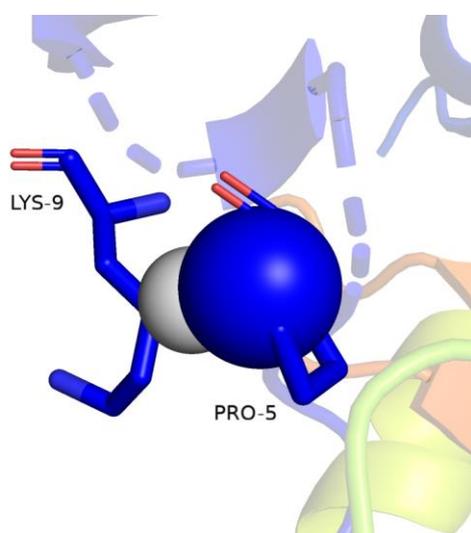


Рисунок 3. Перекрытие атомов CA PRO-5 и HG3 LYS-9.

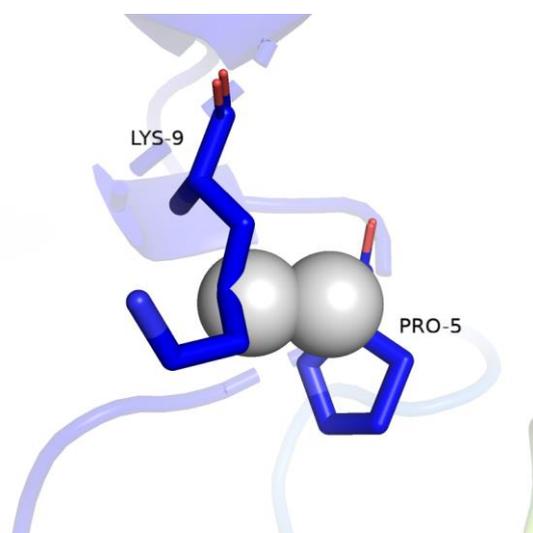


Рисунок 4. Перекрытие атомов HA PRO-5 и HG3 LYS-9.

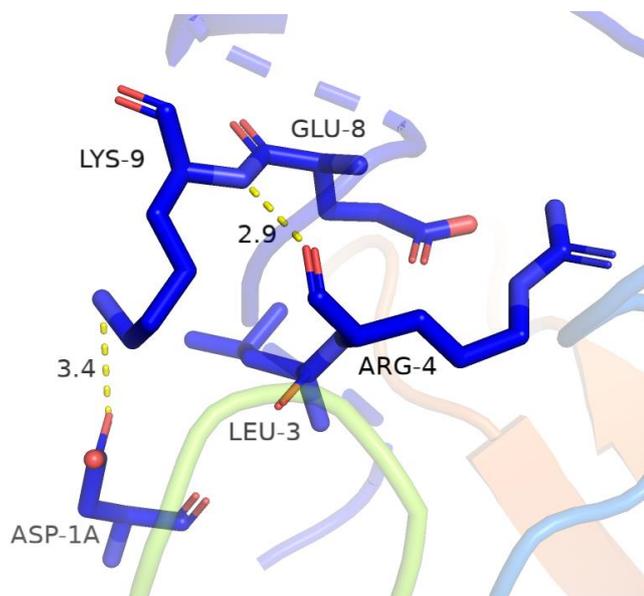


Рисунок 5. Взаимодействия, стабилизирующие LYS-9.

### Chain D - ARG-75.

В данном случае, как и в предыдущем главная проблема – перекрытие атомов ARG-75 с атомами GLU-77 той же цепи. Однако, здесь уже значения довольно большие, что является критичным (для HD2 - HG2 значение 0.987, для CD - HG2 значение 0.969, для HD2 - CG значение 0.946).

Хотя, судя по всему, это происходит все по той же причине – самое оптимальное положение, стабилизирующееся 4 водородными связями.

На рисунке 8 показаны взаимодействия, стабилизирующие данный остаток.

И, можно сказать, что связь с GLU-77 является основной (из-за чего и происходит перекрытие атомов), так как у других ротамеров данного взаимодействия просто нет.

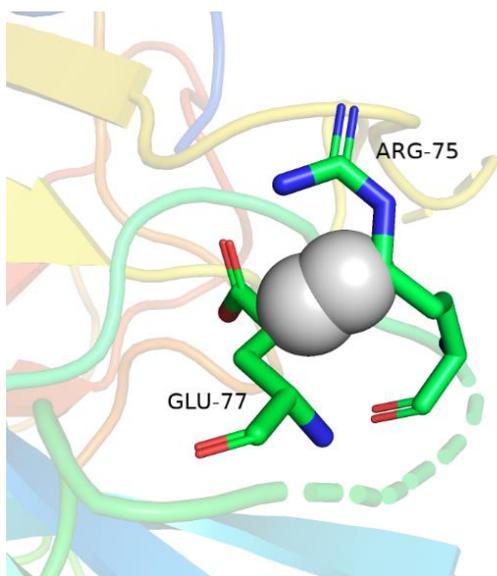


Рисунок 6. Перекрытие атомов HD2 ARG-75 и HG2 GLU-77.

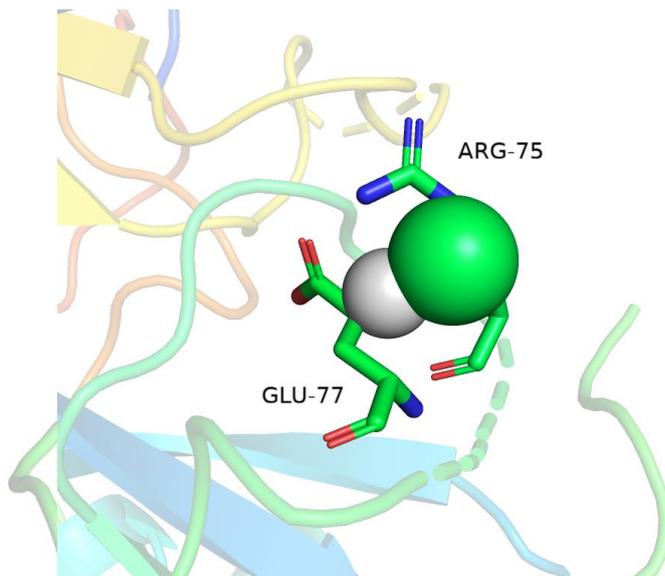


Рисунок 7. Перекрытие атомов CD ARG-75 и HG2 GLU-77.

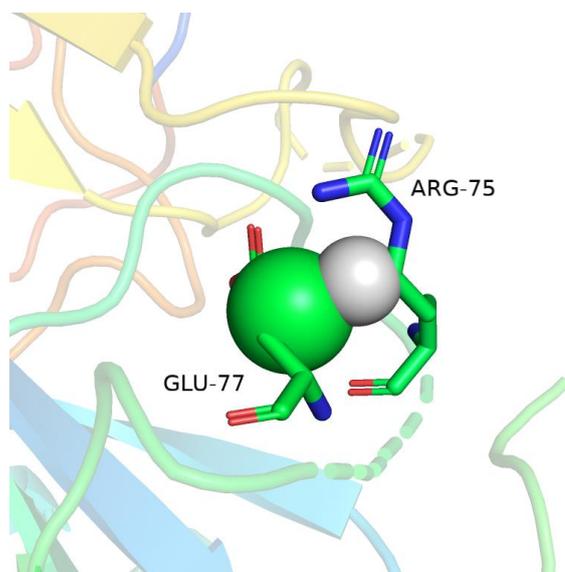


Рисунок 8. Перекрывание атомов HD2 ARG-75 и CG GLU-77.

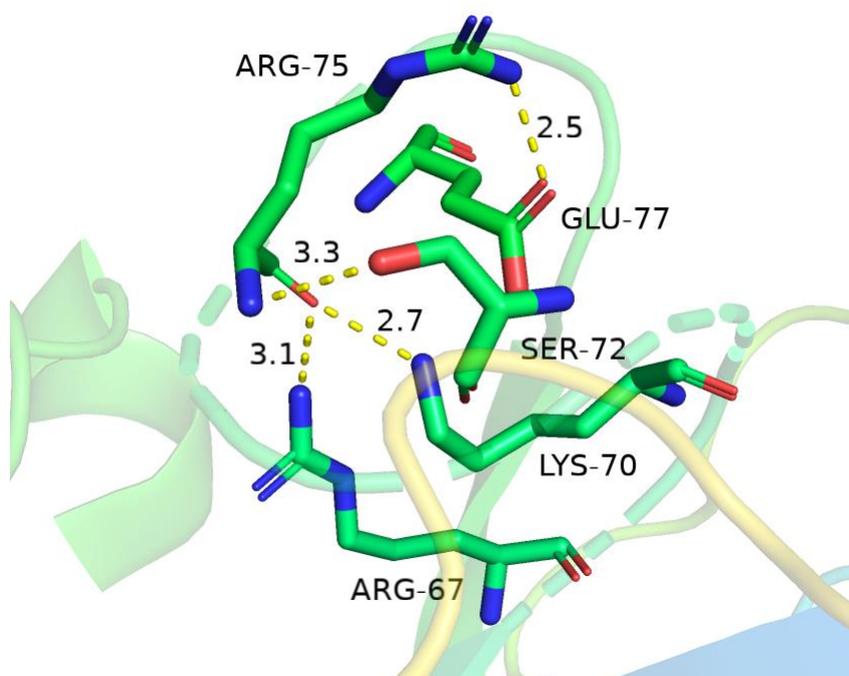


Рисунок 9. Взаимодействия, стабилизирующие ARG-75.

### Chain D - GLN-38.

Данный остаток был выбран, так как его боковая цепь – неротамер. GLN-38 находится в петле, смотрящей «наружу» от центра белка.

Также этот остаток не взаимодействует ни с кем из его окружения. Отсюда я могу предположить, что такая его конформация может быть нужна для белок-белкового взаимодействия.

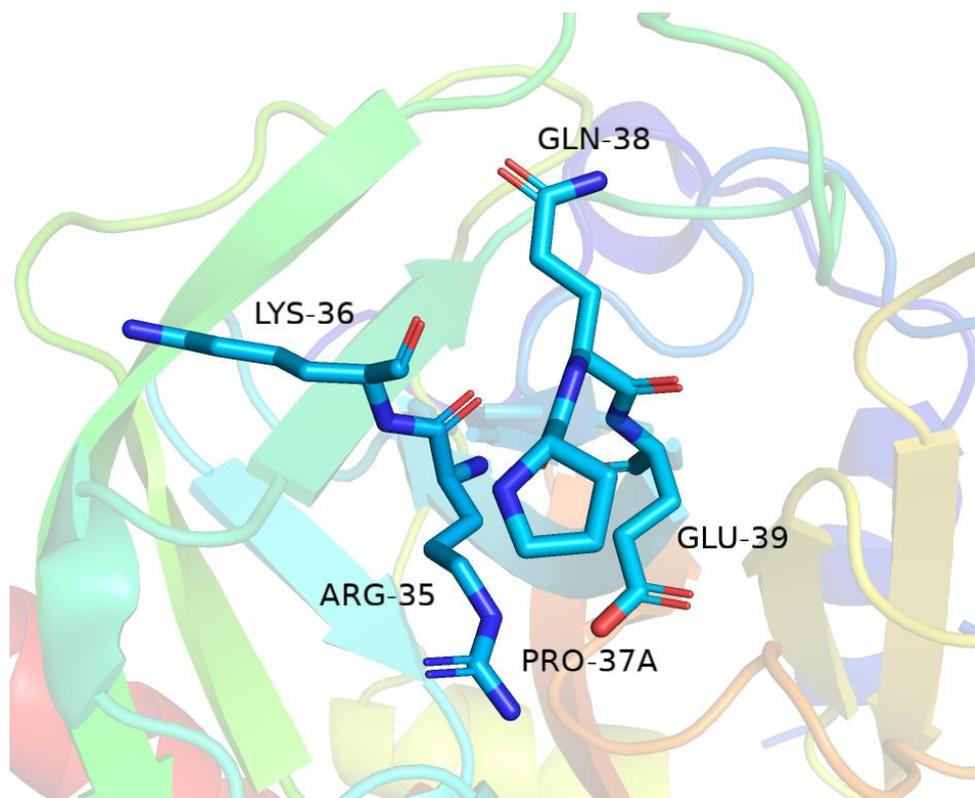


Рисунок 10. Маргинальный остаток GLN-38 и его окружение в диапазоне до 3.5 ангстрем.

### Chain C - ASP-15

Для данного остатка параметром выбора было значение  $RSRZ = 3.4$ . Оно слишком большое, а значит остаток должен плохо вписаться в электронную плотность.

Хотя, при уровне подрезки 0.5 и  $carve=2$ , ASP-15 вполне неплохо вписывается в ЭП (кроме одного атома).

На мой взгляд, большое значение RSRZ связано с тем, что данный остаток является последним остатком цепи.

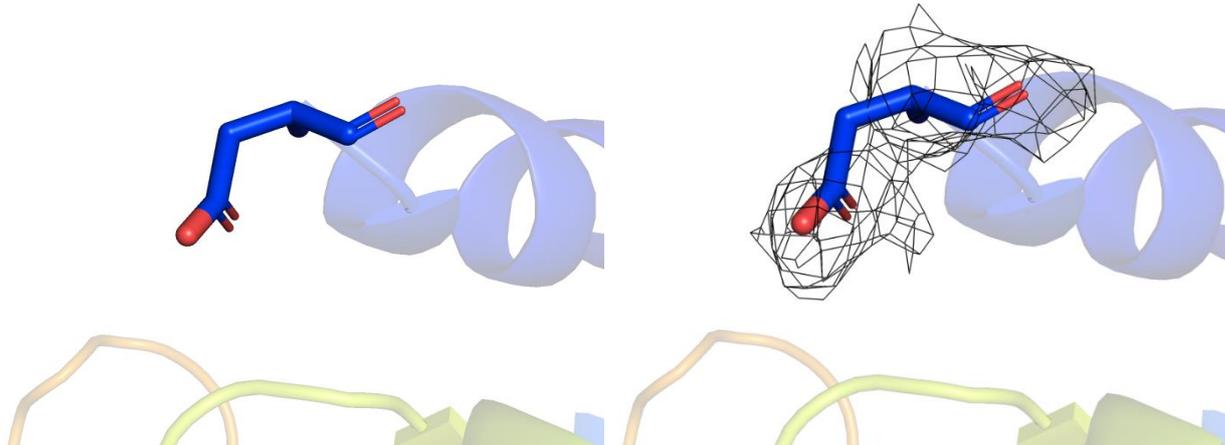


Рисунок 11. Остаток ASP-15 в структуре белка.

Рисунок 12. Остаток ASP-15 и его ЭП, уровень подрезки 0.5, carve=2.

### Chain D - PRO-37A

Данный остаток также как и несколько предыдущих имеет проблемы с перекрыванием атомов, за что и был выбран как маргинальный. Также, этот пролин находится в цис положении (что в целом для данного остатка нормально), однако все остальные пролины в данной структуре находятся в транс положении.

В целом, данное перекрывание атомов является не сильно критичным (HA – C с значением 0.614).

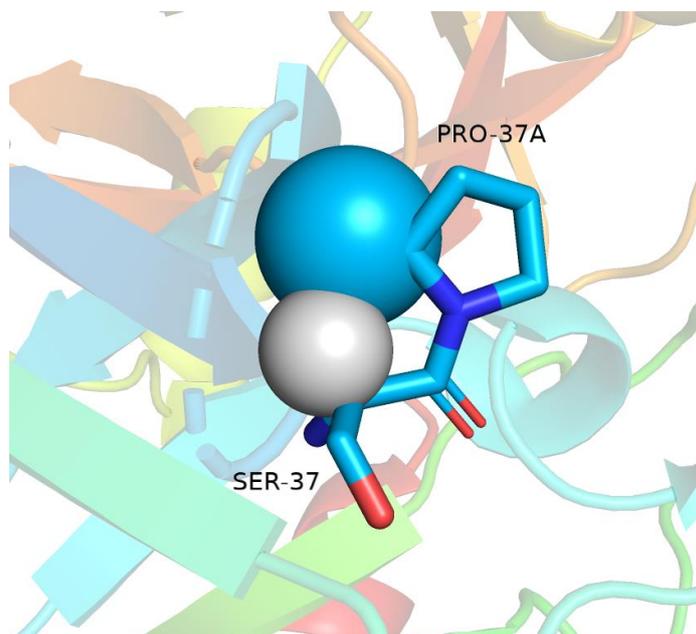


Рисунок 13. Перекрывание атомов C PRO-37A и HA SER-37.

## Сравнение модели PDB и PDB\_redo

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

После оптимизации нашей структуры значение R-фактора почти не изменилось (было 0.185, а стало 0.188). R-free уменьшился (был 0.231, а стал 0.218). Также были изменены 2 ротамера, удалено 7 молекул воды.

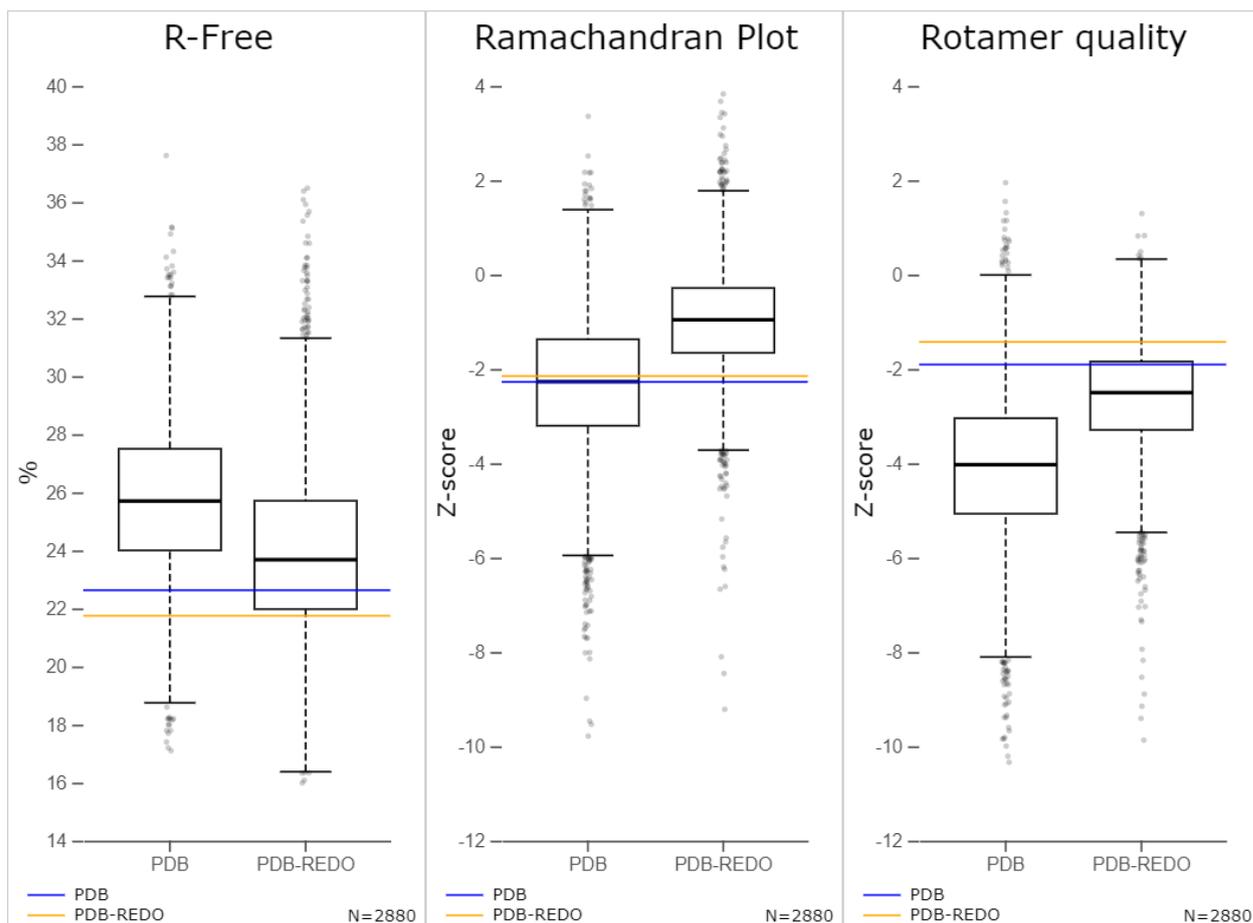
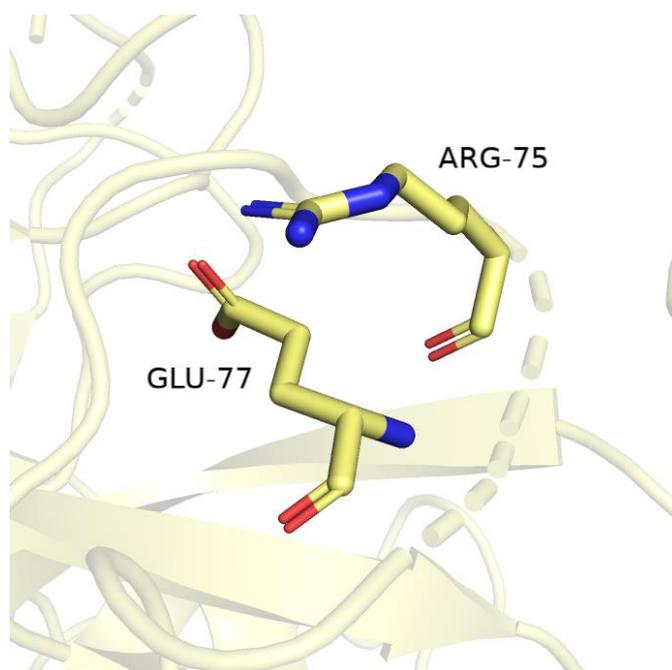


Рисунок 14. Сравнение показателей качества моделей из PDB и PDB-redo.

Из рассмотренных ранее маргиналов изменения произошли только у ARG-75. Наш остаток немного «выгнулся» и атомы перестали перекрываться.



*Рисунок 15. Изменивший свою конфигурацию ARG-75.*

## **Заключение.**

В данной работе была рассмотрена структура тромбина из базы данных PDB (2JH5) и проведен анализ ее качества.

В целом, структура изначально была хорошей (очень малое количество маргиналов по всем параметрам).

И после анализа я только утвердилась во мнении, что структура отличная (хотя PDB\_redo и смог ее еще больше улучшить).

## Список литературы.

1. Senger, Stefan, Chuen Chan, Máire A. Convery, Julia A. Hubbard, Gita P. Shah, Nigel S. Watson, and Robert J. Young. 2007. "Sulfonamide-Related Conformational Effects and Their Importance in Structure-Based Design." *Bioorganic and Medicinal Chemistry Letters* 17 (10): 2931–34. doi:10.1016/j.bmcl.2007.02.034.
2. Запись PDB: <https://www.rcsb.org/structure/2JH5>
3. Сервис MolProbity: <http://molprobity.biochem.duke.edu/>
4. PDB-redo: <https://pdb-redo.eu/db/2jh5>