

Отчет по качеству расшифровки структуры белка 2DOB, полученной методом рентгеноструктурного анализа

Выполнил студент 4 курса

Чаплий Константин

## Оглавление

Аннотация	3
Введение	3
Результаты	
1. Общая информация о модели	3
2. Значение индексов качества модели в целом	4
3. Анализ маргинальных остатков	6
4. Анализ пяти маргинальных остатков	6
5. Сравнение моделей из PDB и PDB-REDO	9
Выводы	9
Список литературы	10

## Аннотация

В данной работе представлен анализ качества расшифровки структуры 2DOB, рассмотрены основные индикаторы качества модели, а также несколько маргинальных остатков.

## Введение

Для анализа был выбран белок сапозин А человека. Сапозины – небольшие лизосомные белки, которые служат активаторами разнообразных ферментов, осуществляющих деградацию липидов [1]. Конкретно сапозин А осуществляет деградацию галактозилцерамида [2]. Обычно сапозины действуют за счет изоляции липидного субстрата от мембранного окружения, таким образом делая их более доступными для растворенных деградирующих ферментов. Сапозины образуются в результате разрезания молекулы-предшественника – просапозина (Рис.1) [3].

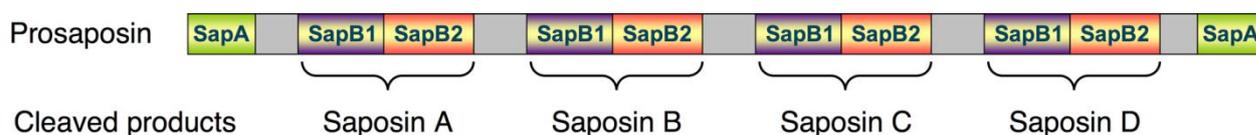


Рис. 1. Просапозин – молекула-предшественник сапозинов.

## Результаты

### 1. Общая информация о модели

Таблица 1. Параметры структуры

Параметр	Значение
Статья	Ahn, V. E., Leyko, P., Alattia, J.-R., Chen, L., & Privé, G. G. (2006). Crystal structures of saposins A and C. <i>Protein Science</i>
Состав комплекса	Сам белок и ион кальция
Разрешение	2 Å
Диапазон разрешений	33.82 - 2.00 Å

Полнота данных	98.1 %
Количество рефлексов	5477
Тип симметрии	P 21 21
Параметры ячейки	a: 45.68, b: 50.04, c: 33.82 Å вектора ортогональны
Метод решения фазовой проблемы	Многоволновое аномальное рассеяние

## 2. Значение индексов качества модели

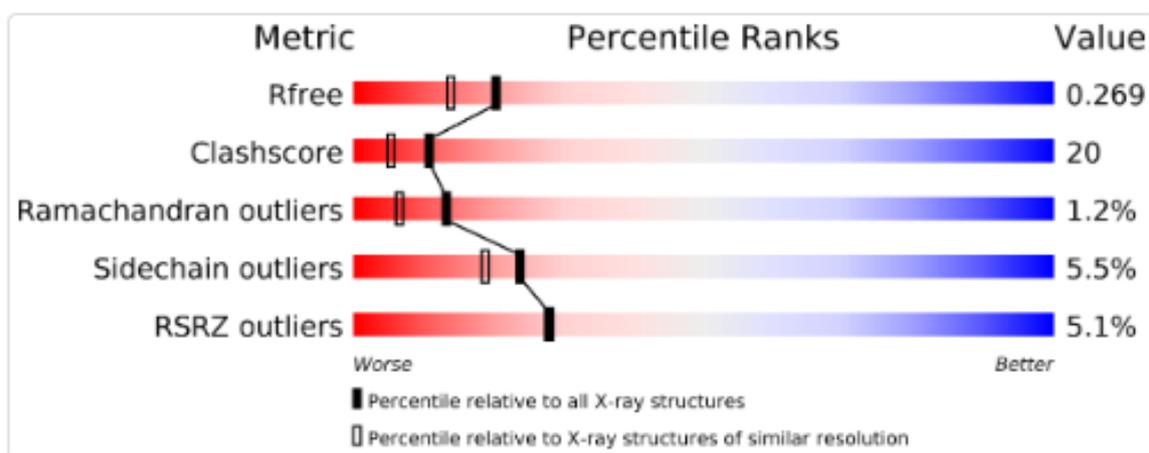


Рис. 2. Значение различных индексов качества модели.

По значению параметров видно, что структура очень плохая. R-фактор = 0,218; R-free = 0,267. Поскольку  $R < 25\%$ , значение среднее,  $R\text{-free} > 25\%$  => значение не очень хорошее.

Таблица 2. Критерии качества модели, выданные сервисом MolProbity. Зеленым цветом выделены значения, удовлетворяющее условиям хорошей модели, желтым – значения, отклоняющиеся от условий хорошей модели и красным – значения, не удовлетворяющие условиям хорошей модели

All-Atom Contacts	Clashscore, all atoms:	19.5	36 <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	6.85%
	Favored rotamers	66	90.41%
	Ramachandran outliers	1	1.25%
	Ramachandran favored	78	97.50%
	MolProbity score <sup>^</sup>	2.52	32 <sup>nd</sup> percentile* (N=12522, 2.00Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%
	Bad bonds:	0 / 628	0.00%
Bad angles:	0 / 854	0.00%	
Peptide Omegas	Cis Prolines:	0 / 5	0.00%
			Expected: ≤1 per chain, or ≤5%

По таблице видно, что структура содержит много плохих ротамеров, мало предпочтительных ротамеров, и имеет очень низкий score.

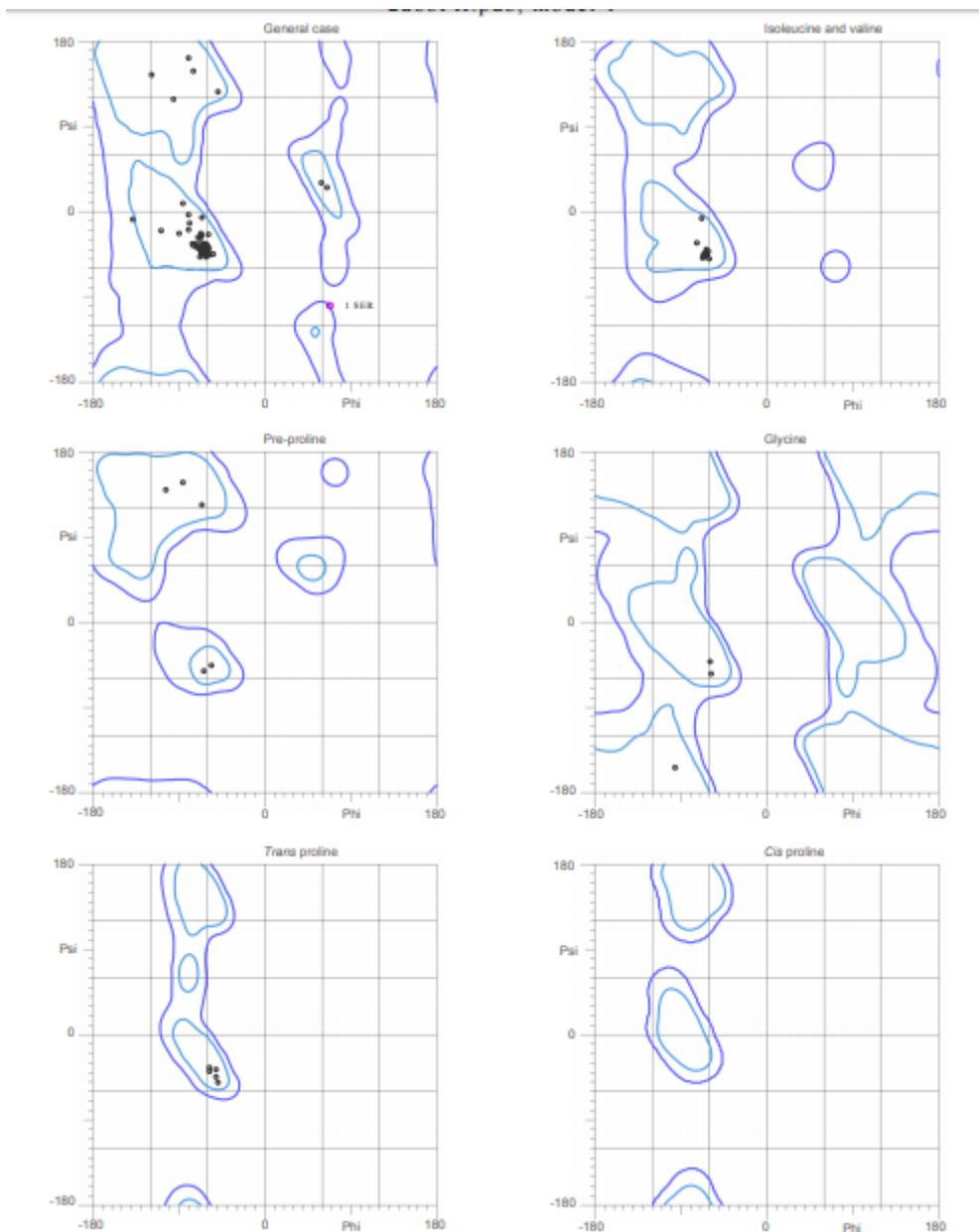


Рис. 3. Карта Рамачандрана, построенная сервисом MolProbity  
Выявлен один остаток, не попавший в допустимую зону – Ser1

### 3. Анализ маргинальных остатков

Поиск маргинальных остатков осуществлялся при помощи результатов работы сервиса MolProbity.

Таблица 2. Примеры маргинальных остатков

Остаток	Критерий маргинальности
Ser1	Ramachandran outlier
Leu2	Poor rotamer
Asp20	Poor rotamer
Tyr30	Poor rotamer
Lys40	Poor rotamer
Glu49	Poor rotamer

### 4. Анализ пяти маргинальных остатков

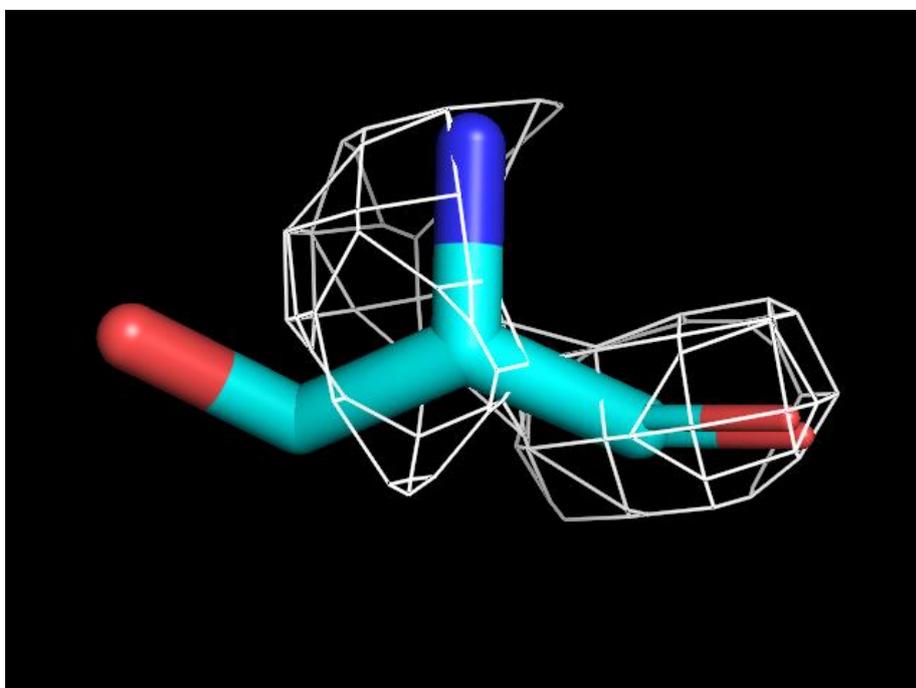


Рис.5. Ser1, уровень подрезки здесь и далее 1. Электронная плотность не полностью покрывает остаток, значит он верно был определен как

маргинальный. Вероятно, структура плохо расшифрована из-за высокой подвижности концевой остатка.

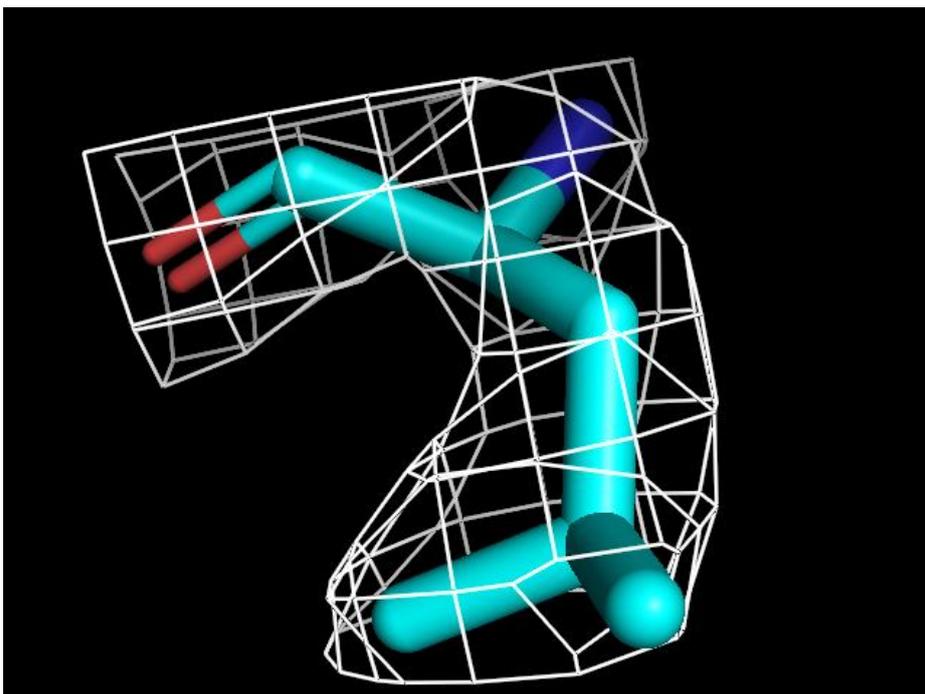


Рис.6. Leu2. Электронная плотность полностью покрывает остаток, скорее всего, он не является маргинальным.

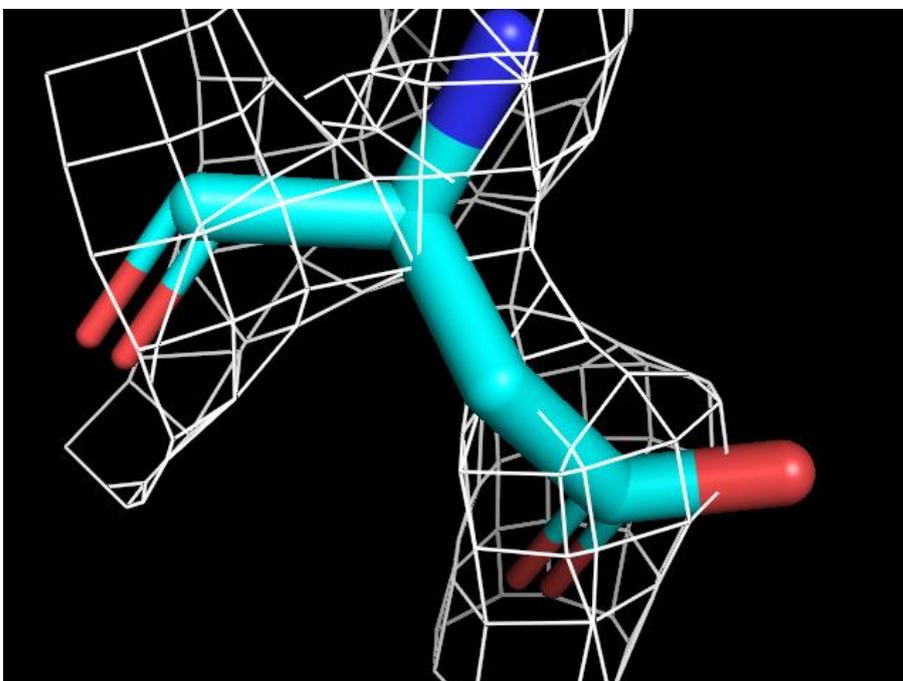


Рис.7. Asp20. В некоторых местах электронная плотность не полностью покрывает остаток, скорее всего, он действительно маргинальный. Возможно, это вызвано ошибками в расшифровке данных.

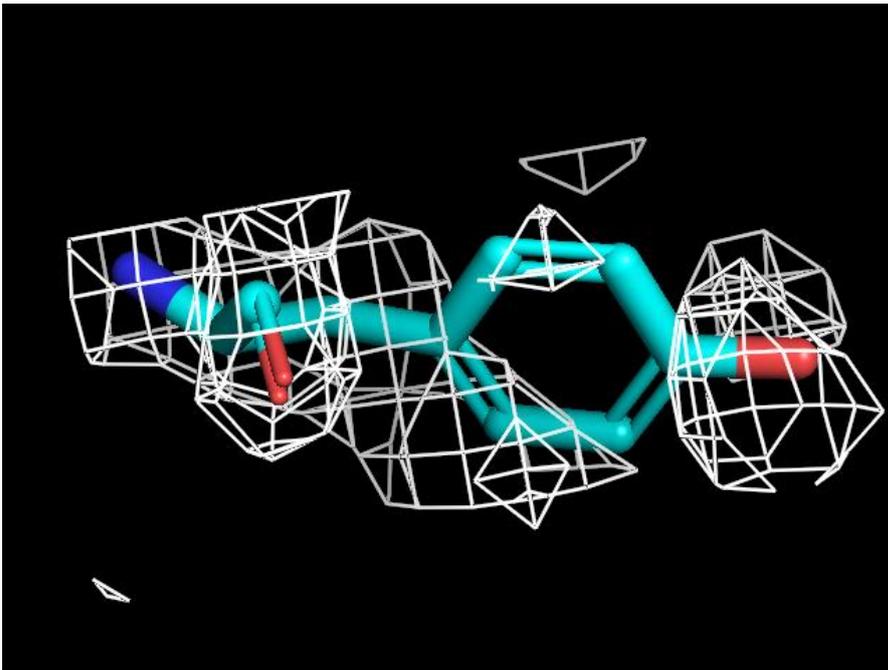


Рис.8. Tyr30. Электронная плотность не полностью покрывает бензольное кольцо тирозина, следовательно остаток маргинальный. Вероятно, это вызвано особенностями структуры остатка.

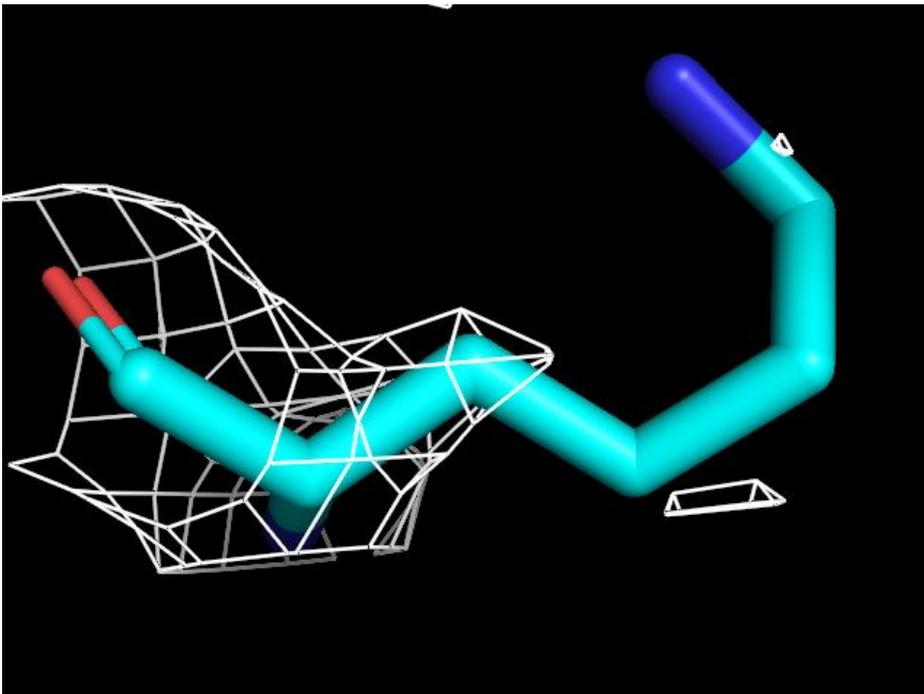


Рис.9. Lys40. Электронная плотность не покрывает боковую цепь, остаток действительно маргинальный. Вероятно, это вызвано особенностями структуры остатка.

## 5. Сравнение моделей из PDB и PDB-REDO

Далее была проведена оптимизация структуры в базе PDB\_REDO (Рис.10). Действительно, в результате оптимизации качество (R, R\_free) значительно улучшилось, а число остатков, лучше вписывающихся в электронную плотность, увеличилось на 17.

Validation metrics from PDB-REDO		
	PDB	PDB-REDO
<b>Crystallographic refinement</b>		
<i>R</i>	0,2105	0,1797
<i>R-free</i>	0,2533	0,2395
<i>Bond length RMS Z-score</i>	0,394	0,660
<i>Bond angle RMS Z-score</i>	0,611	0,780
<b>Model quality (raw scores   percentiles)</b>		
<i>Ramachandran plot appearance</i>	97	95
<i>Rotamer normality</i>	24	20
<i>Coarse packing</i>	N/A	N/A
<i>Fine packing</i>	73	73
<i>Bump severity</i>	70	79
<i>Hydrogen bond satisfaction</i>	43	88
<i>WHAT_CHECK</i>	<a href="#">Report</a>	<a href="#">Report</a>
<b>Significant model changes</b>		
Description	Count	
<i>Rotamers changed</i>	1	
<i>Side chains flipped</i>	0	
<i>Waters removed</i>	7	
<i>Peptides flipped</i>	0	
<i>Chiralities fixed</i>	0	
<i>Residues fitting density better</i>	17	
<i>Residues fitting density worse</i>	0	

Рис.10. Оптимизация PDB\_REDO

## Выводы

Структура довольно низкого качества, есть несколько маргинальных остатков. Сервис PDB\_REDO значительно улучшил качество модели, но все проблемы не решил.

## Список литературы

1. Munford RS, Sheppard PO, O'Hara PJ (August 1995). Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *Journal of Lipid Research*. **36** (8): 1653–63
2. Ahn, V. E., Leyko, P., Alattia, J.-R., Chen, L., & Privé, G. G. (2006). Crystal structures of saposins A and C. *Protein Science*, *15*, 1849–1857.
3. Ponting, C. P. (1994). Acid sphingomyelinase possesses a domain homologous to its activator proteins: Saposins B and D. *Protein Science*, *3*, 359–361.