

# Отчет по качеству расшифровки структуры родопсина кальмара 2Z73

Выполнен студенткой 4 курса  
Шугаевой Татьяной

2017

# Оглавление

Аннотация.....	3
Введение.....	3
Результаты и обсуждение .....	4
1. Общие сведения о структуре .....	4
Проведение PCA-эксперимента и обработка данных .....	4
Параметры кристаллографической ячейки.....	5
Состав комплекса.....	7
2. Значения индикаторов качества модели в целом.....	8
R-фактор.....	8
R-free .....	9
Карты Рамачандрана .....	9
Real-space R-value (RSR) .....	11
Clashscore.....	13
Торсионные углы боковых цепей.....	14
Дополнительно .....	14
3. Некоторые маргинальные остатки структуры 2Z73.....	15
4. Анализ 5 маргинальных остатков.....	16
Остаток №1: Ala 308, цепь A .....	16
Остаток №2: Pro 90, цепь A .....	19
Остаток №3: Ala 243, цепь A .....	22
Остатки №4 и №5: Met 234 и Asn 70, цепь B.....	25
5. Сравнение с моделью из PDB_redo .....	27
Заключение .....	29
Список литературы .....	30

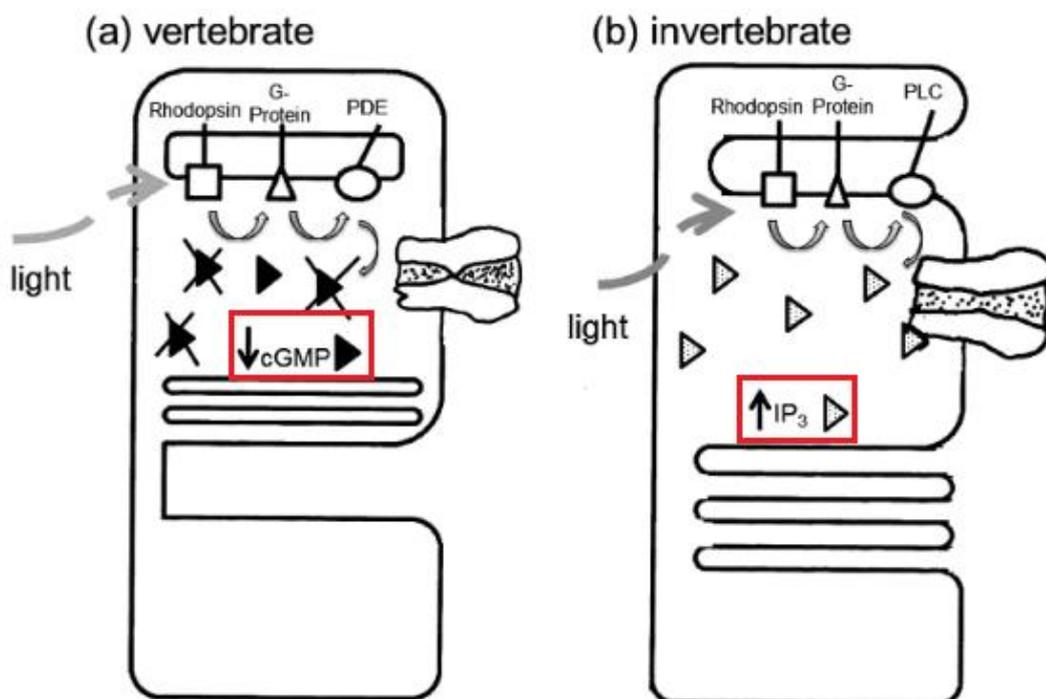
# Аннотация

Данный отчет посвящен оценке качества структуры родопсина кальмара 2Z73, полученной методом рентгеноструктурного анализа. В работе приведены общие сведения о белке и о получении структуры 2Z73, проанализированы значения индикаторов качества модели, рассмотрены имеющиеся в структуре маргинальные остатки. Также сделано сравнение структуры с моделью из PDB-redo и приведены выводы о качестве расшифровки структуры.

# Введение

Белки родопсина являются первичными фоторецепторами в зрительных сигнальных каскадах. Они встречаются как у эукариот, так и у прокариот. В эукариотических организмах используется семейство родопсинов, относящееся к группе рецепторов, сопряженных с G-белком (G-protein-coupled receptors, GPCRs)[1].

В данной работе рассматривается родопсин беспозвоночного – тихоокеанского кальмара *Todarodes pacificus*. В каскаде инвертебрат фотоактивация родопсина приводит к активации  $G_q$  –белка и последующей передаче сигнала через инозитол-1,4,5-трифосфат, что отличает их от позвоночных, у которых вторичным мессенджером является циклическая форма GMP (рисунок 1). Учитывая, что механизм фотоактивации у позвоночных широко изучен, на основании детального рассмотрения данной структуры родопсина кальмара 2Z73 и сравнения её со структурами родопсинов позвоночных можно было бы более детально описать механизм первичной детекции зрительного сигнала у инвертебрат[2].



*Рисунок 1. Сравнение зрительных сигнальных каскадов позвоночных(a) и беспозвоночных(b)[3]. Красной рамкой выделены вторичные мессенджеры: цГМФ и инозитолтрифосфат.*

Родопсин кальмара содержит 448 аминокислот и имеет молекулярную массу приблизительно 50 кДа. Его полипептидная цепь примерно на 100 остатков длиннее, чем у позвоночных за счет богатого пролинами участка на С-конце, необходимого для внутриклеточных перемещений и морфогенеза. Было показано, что данный участок не влияет на способность родопсина активировать G-белки. При получении структуры 2Z73 С-концевой участок удалялся протеазой.

По базе данных SCATH белок содержит один домен класса «mainly alpha», то есть состоящий, в основном, из альфа-спиралей. Доменная архитектура – up-down bundle. Анализ топологии и гомологии же выдал, что белок принадлежит к 7-спиральным трансмембранным родопсином. Семиспиральная трансмембранная структура чрезвычайно консервативна, наложение структуры бычьего родопсина на родопсин кальмара дает практически полное совпадение мембранных участков[2].

## Результаты и обсуждение

### 1. Общие сведения о структуре

РСА эксперимент по получению структуры 2Z73 был проведен **26 мая 2007** года, а готовая модель была добавлена в базу данных PDB 29 августа того же года (то есть на обработку данных ушло около трех месяцев). Работу выполнили **Мидори Мураками** (Midori Murakami) из Нагойского университета и **Тсутому Коуяма** (Tutomu Kouyama) из Института физико-химических исследований Японии.

#### Проведение РСА-эксперимента и обработка данных

Тихоокеанские кальмары заказывались на рыбном рынке Янагибаши города Нагои. Из их глаз выделялись мембраны, содержащие родопсин, которые на следующем шаге обрабатывались протеазой V<sub>8</sub> для удаления участка на С-конце. Далее происходило выделение очищенного родопсина по методу, разработанному для выделения родопсина быка[4], с минимальными изменениями.

Очищенный родопсин кристаллизовали при 277 кельвинах методом паровой диффузии сидячей капли (sitting drop vapor diffusion crystallization) с помощью коммерческого набора компании ICM Biomedicals. На выращивание кристаллов размером 200 микрометров ушло более 6 месяцев.

При получении рентгенограммы образец подвергался рентгеновскому излучению длиной волны 1Å. Обработка пятен на рентгенограмме проводилась программой Mosflm 6.2. Расшифровка структуры проводилась программами CCP4.

Фазовую проблему решили **методом молекулярного замещения**. В качестве базовой структуры использовали структуру мембранной части родопсина быка (1GZM). Кажется, что в условиях крайнего сходства между родопсинами разных групп, этот метод решения фазовой проблемы подходит лучше всего. Уточнение структуры проводилось программами CNS-1.1[5] и Xtal-View-4.0.

В файле структурных факторов приведено **43 511 рефлексов** (эта же цифра указана на странице 2z73 EDS-сервера). Из них для оптимизации использовалось 41 337 рефлексов, что составляет 95%.

На сайте PDB во вкладке Experiment указано, что всего было получено 46 274 рефлекса с разрешением от 2,5 до 57 Å, полнота набора 99,2%. При уточнении использовалось 43 279 рефлексов с разрешением от 2,5 до 15 Å, на расчет r-free был оставлен 2161 рефлекс (5%).

### Параметры кристаллографической ячейки

Кристаллографическая ячейка характеризуется 3 векторами (a, b и c) и тремя углами между ними ( $\alpha$ ,  $\beta$  и  $\gamma$ ). В случае структуры 2Z73  $a=b=122,55$  Å,  $c=158,75$  Å. Значения углов:  $\alpha=\beta=90^\circ$ ,  $\gamma=120^\circ$ .

**CRYST1 122.550 122.550 158.740 90.00 90.00 120.00 P 62 12**

Из того, что  $a=b \neq c$  и  $\alpha=\beta=90^\circ$ ,  $\gamma=120^\circ$  следует, что элементарная ячейка относится к гексагональному типу.  $P6_2$  – обозначение кристаллографической группы в символике Германа-Могена (171 группа). Первый символ – это символ решетки Браве, обозначает наличие дополнительных узлов трансляции внутри элементарной ячейки. Буква P расшифровывается как primitive, что означает, что дополнительного узла трансляции в ячейке нет. Цифры же  $6_2$  означает наличие винтовой оси симметрии, большая цифра означает угол поворота, индекс – то, насколько нужно сдвинуться по оси при повороте. В случае структуры 2Z73 необходимо совершить поворот на  $60^\circ$  и сдвинуться на  $2/6$  трансляции.

Асимметрическая ячейка содержит 2 молекулы родопсина (рисунок 2), элементарная ячейка – 12 молекул. В клеточных мембранах родопсин содержится в виде димеров (рисунок 3).

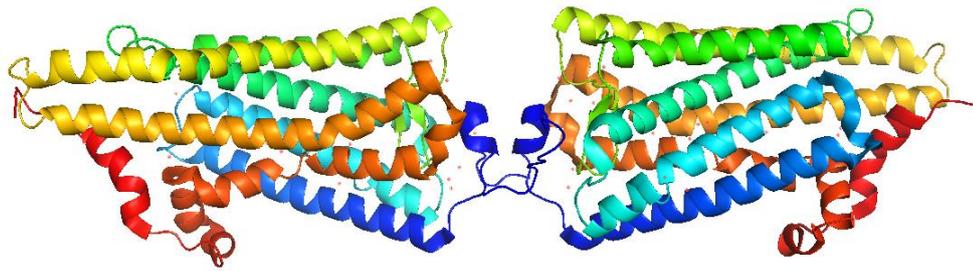


Рисунок 2. Две молекулы родопсина в асимметрической ячейке. Цепи А и В покрашены по типу spectrum (градиент от С-конца до N-конца)

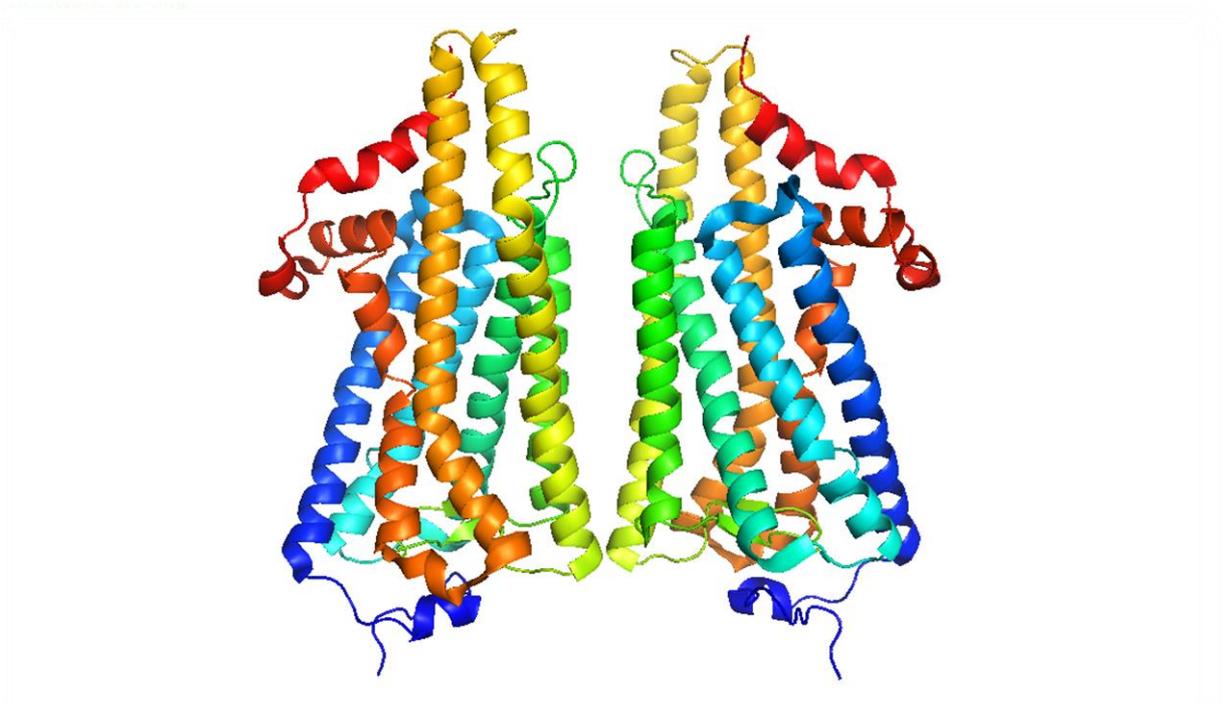


Рисунок 3. Димер родопсина – вероятная биологическая форма. Цепи покрашены по типу spectrum (градиент от С-конца до N-конца)

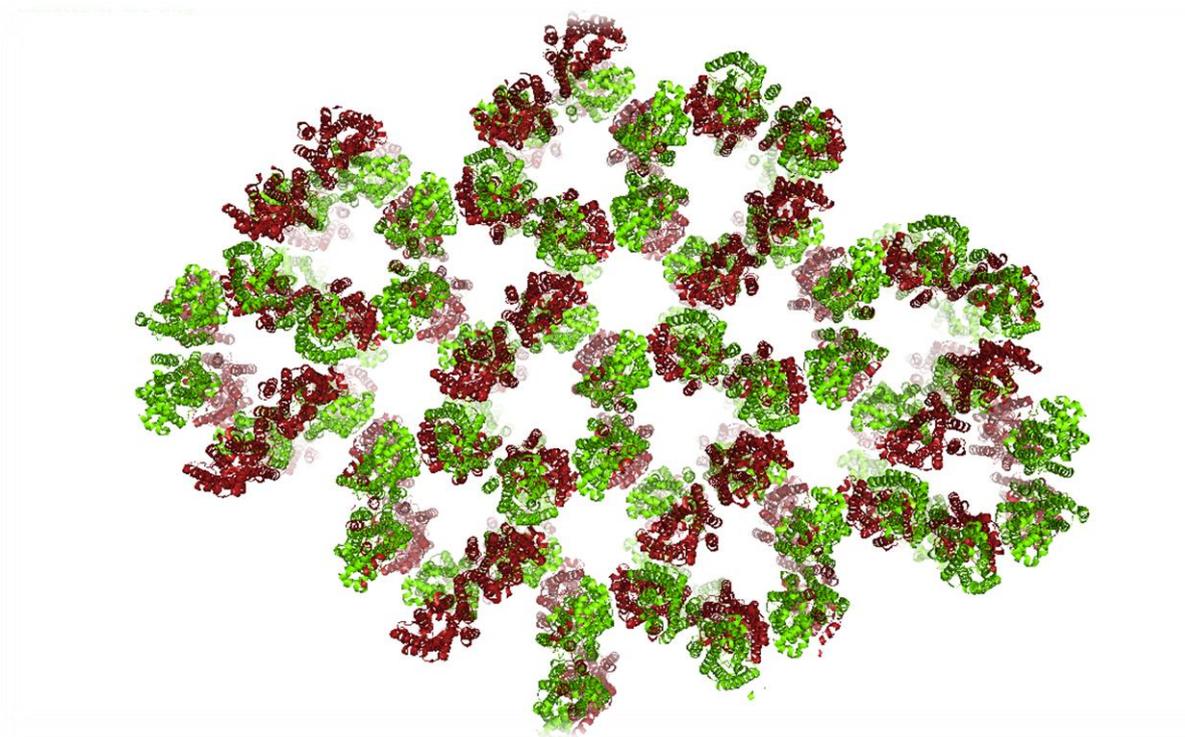


Рисунок 4. Фрагмент кристалла. Красным отмечены субъединицы А, зеленым – субъединицы В.

Поле MTRIX, которое должно содержать информацию о **некристаллографических симметриях**, отсутствует в 2Z73.pdb, поэтому, вероятно, их нет.

### **Состав комплекса**

Помимо двух цепей белка в асимметрической ячейке содержится 6 типов лигандов: 3-SN-фосфатидилхолин, докозан, В-октилглюкозид, ретиналь, сульфат-ионы. Ретиналь является обязательным компонентом фоторецепторов.

## 2. Значения индикаторов качества модели в целом

Значения качества модели приводятся на странице структуры в качестве графика перцентилей распределения значений по всем структурам из базы данных PDB, полученных методом PCA (прямоугольники), и по структурам PDB, близким к 2Z73 по разрешению (белый прямоугольник).

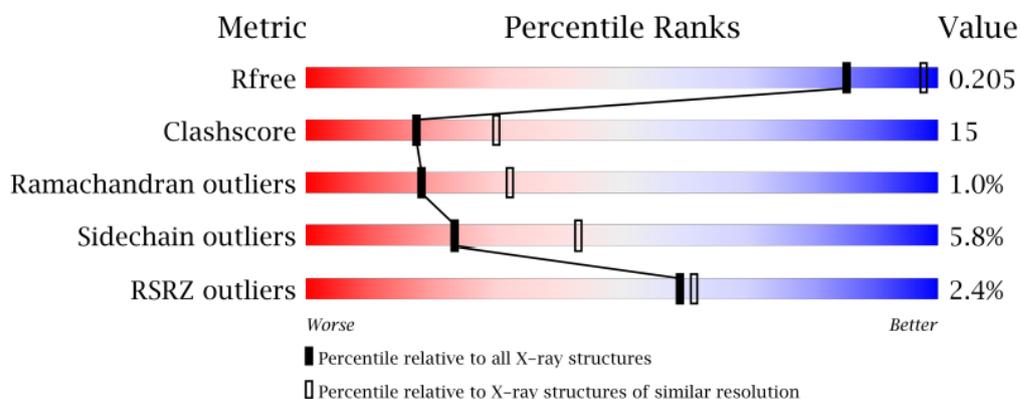


Рисунок 5. График индикаторов качества модели. Синяя область – значения, лучшие, чем в большинстве структур из выборки, красная область – значения, более плохие, чем в большинстве структур из выборки.

Рассмотрим отдельно каждый индикатор качества модели.

### R-фактор

Данный показатель характеризует, насколько полученная модель структуры близка реальной, то есть к той, по которой были получены экспериментальные структурные факторы –  $F^{obs}$ . По полученной после оптимизации модели можно рассчитать, какие структурные факторы она даст в эксперименте –  $F^{calc}$ . Тогда R-фактор вычисляется по формуле:

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

Часто, однако, он приводится не в долях, а в процентах, поэтому полученное после применения формулы число бывает необходимо умножить на 100%. Считается, что хорошее значение R-фактора <25%

Для структуры 2Z73 **R-фактор равен 19,7%** (EDS-сервер), то есть попадает в диапазон хороших значений.

## R-free

Было обнаружено, что если оптимизировать структуру слишком сильно, то можно довести R-фактор до отличных значений, удалившись при этом от реального белка. Для контроля подобной переоптимизации был придуман R-free. 5% структурных факторов не используются в процедуре оптимизации модели. После ее завершения применяют к этой группе формулу для расчета R-фактора. Полученная величина и называется R-free. Хорошо, если она меньше 20% и если R-фактор – R-free < 10%.

Для рассматриваемой в отчете структуры **R-free равен 20,6%**, R-фактор, рассчитанный по остальным рефлексам – 18,8%, то есть разность **R – R-free равна 1,8%** (сгенерированный отчет по качеству структуры со страницы PDB, данные от авторов структуры). Программа DCC выдала чуть другую пару значений: 20,5% и 18,6%, но они не слишком отличаются от первых. Именно значение 20,5% указано на графике, на рисунке 5. Видно, что сам по себе R-free находится на самой границе между хорошими и нормальными значениями, отличие от R-фактора очень хорошее, переоптимизации структуры, вероятно, не было.

Как описано выше, для расчета R-free был использован 2161 рефлекс, что составляет 5% от общего количества.

От показателей, характеризующих только модель в целом, перейдем к рассмотрению показателей качества, которые могут быть рассчитаны для отдельных остатков.

## Карты Рамачандрана

Существуют определенные «комфортные» значения углов  $\psi$  и  $\phi$  остова полипептидной цепи. Для оценки качества структуры можно посмотреть, насколько значения этих углов в структуре близки к наименьшим по энергии. Визуализируются углы с помощью карт Рамачандрана, где по оси абсцисс отложены значения угла  $\phi$ , по оси ординат угол  $\psi$ , а замкнутыми линиями обведены изоэнергетические, выгодные по энергии области.

Карта Рамачандрана по всем остаткам приведена на рисунке 6, карта для остатков изолейцина и валина приведена на рисунке 7. Данные карты приведены на странице структуры PDB, они были получены с помощью сервиса MolProbity[6]. 91,5% остатков находятся в предпочтительной области, 99% остатков находятся в допустимой области. Сервис выдал 7 маргинальных остатков, которые на картах находятся совсем рядом с допустимыми областями. Наиболее отличающимся является 243 аланин цепи А.

Количество маргиналов достаточно мало, хотя по гистограмме на рисунке 5 видно, что среди всех структур PDB значение не очень-то и хорошее.

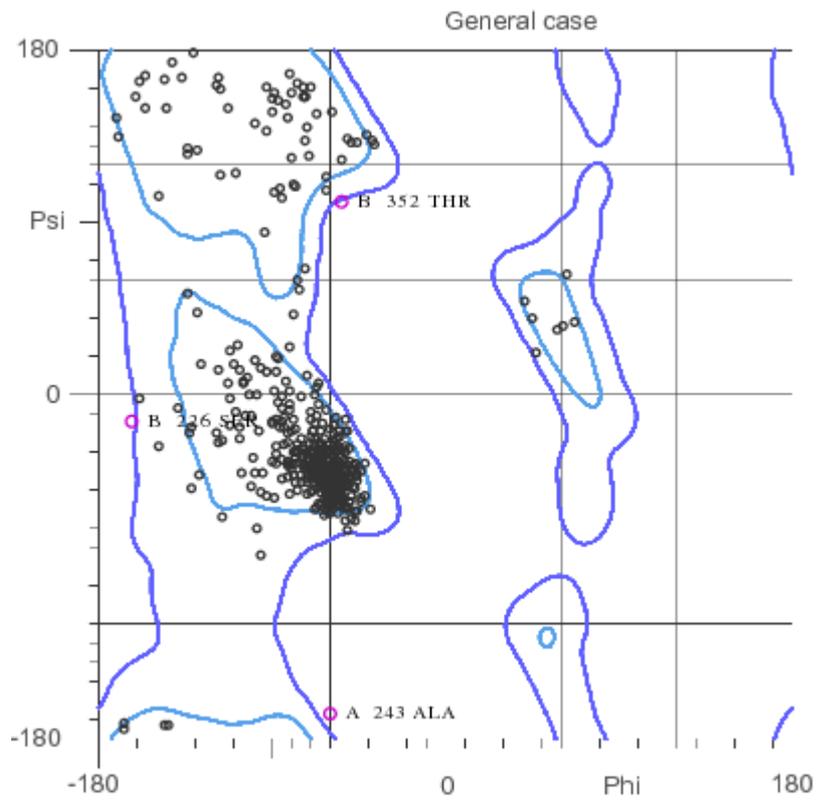


Рисунок 6. Полная карта Рамачандрана для остатков родопсина кальмара, в фиолетовые кружочки обведены маргинальные остатки B352 Thr, A243 Ala и B226 Ser .

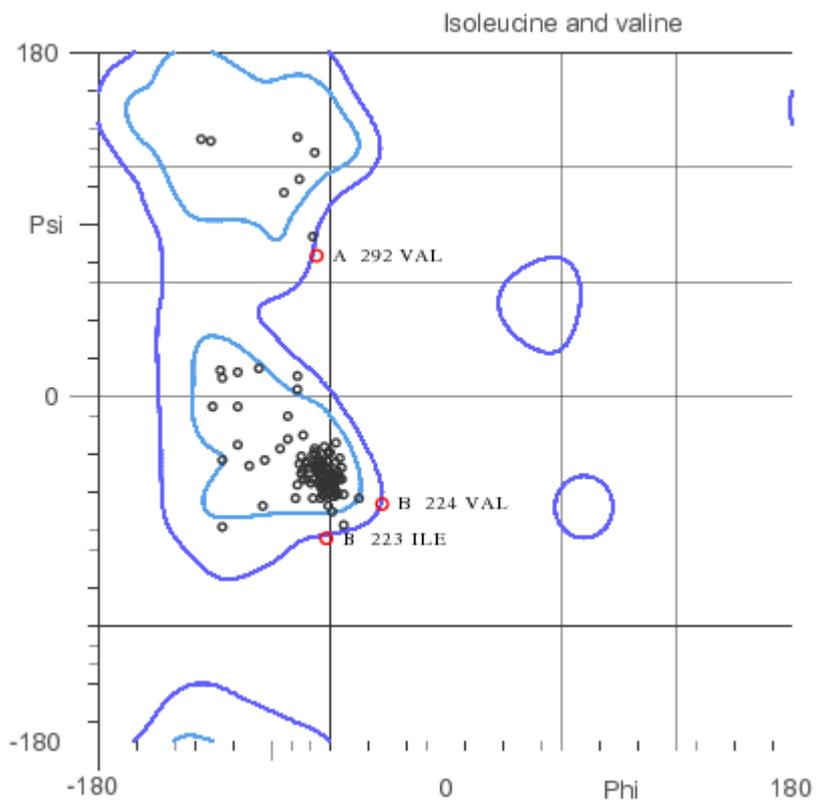


Рисунок 7. Карта Рамачандрана для остатков валина и изолейцина. В красные круги обведены маргинальные остатки A292 Val, B224 Val и B223 Ile.

## Real-space R-value (RSR)

RSR, или пространственный R-фактор, является характеристикой того, насколько полученная структура вписывается в “экспериментальную” электронную плотность. На сервере EDS[7] приводятся гистограммы пространственных R-факторов для отдельных остатков структуры 2Z73 (рисунки 8 и 9).

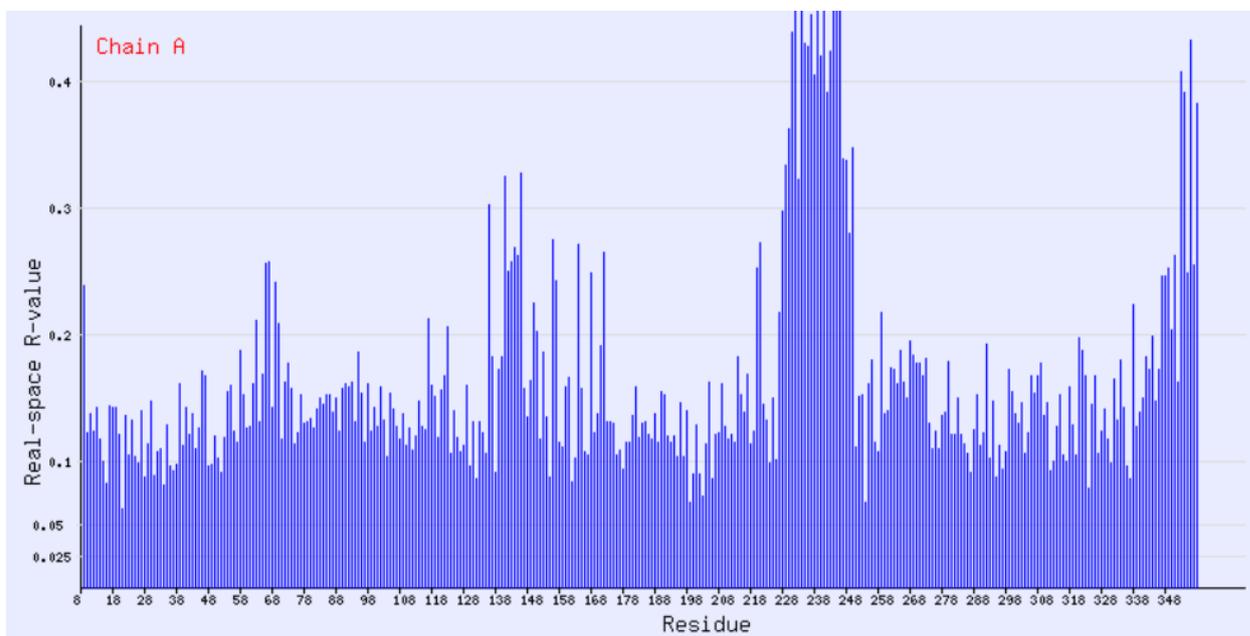


Рисунок 8. Гистограмма значений RSR для остатков цепи A структуры 2Z73.

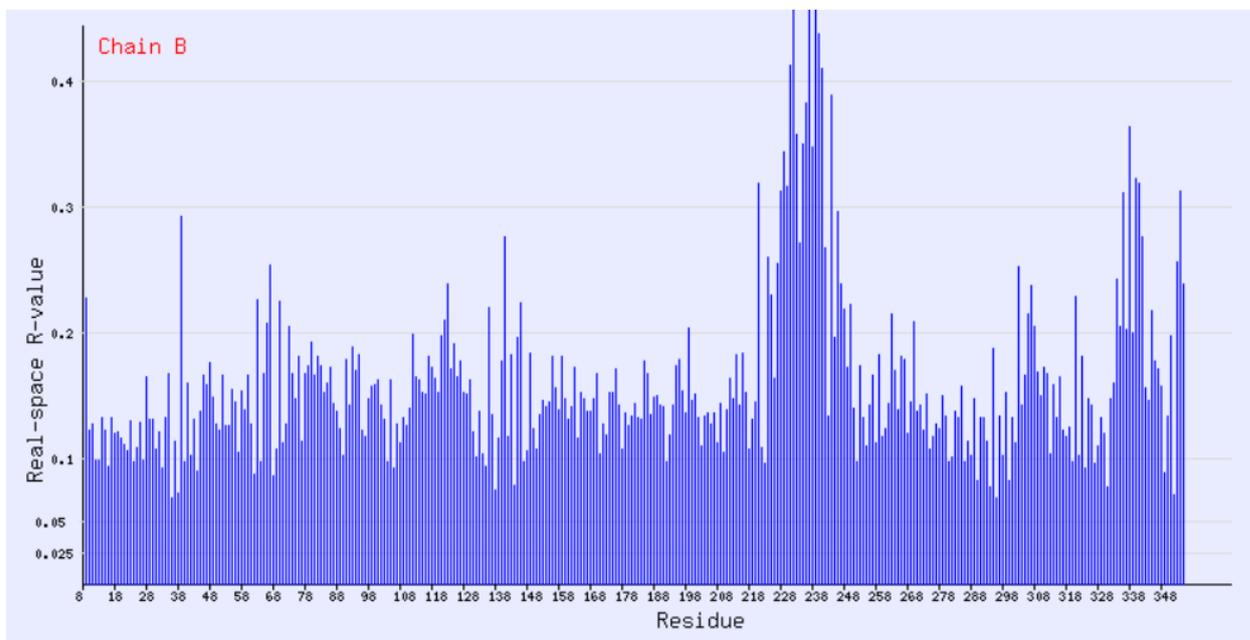
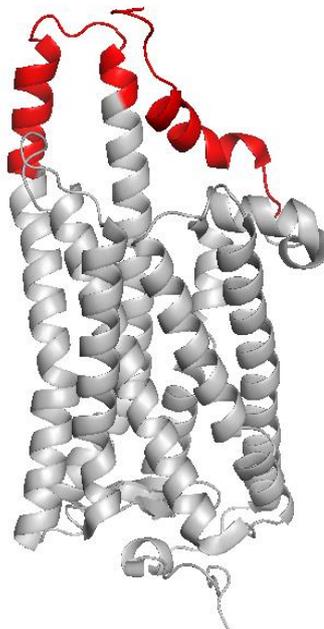


Рисунок 9. Гистограмма значений RSR для остатков цепи B структуры 2Z73.

Из графиков видно, что с RSR происходит какая-то беда в районе 228-248 остатков и около 338 остатка и дальше. Эти остатки соответствуют элементам структуры, выдающимся из мембраны в цитоплазму (рисунок 10).



*Рисунок 10. Участки цепи с большим RSR (228-248, 238-258 остатки) выделены красным цветом, цепь А.*

Должно быть, эти петли и участки альфа-спиралей достаточно подвижны, и в итоге электронная плотность от них получилась “размазанной”, отсюда и несовпадение с моделью.

Хорошие значения RSR <10%, плохие >20%. Для данной структуры среднее значение RSR по всем остаткам равно 16,8% (стандартное отклонение 9,3%), что можно оценить как удовлетворительно.

Часто используется не сам RSR, а Z-score RSR. Для вычисления Z-score берется RSR данного остатка и сравнивается со средним RSR для данного типа остатков. На сервере EDS приведены диаграммы Z-score для структуры 2z73 (рисунки 11 и 12). Тенденции все те же, что и на рисунках 8 и 9. Остатки в регионах 228 – 248 и 338-358 определились плохо и по Z-score, то есть причина высокого RSR лежит не в аминокислотной последовательности.

Остатки считаются маргинальными, если Z-score больше 2. Процент маргинальных остатков по данному критерию в исследуемой структуре составляет 7,17%. Распределены маргинальные остатки неравномерно, можно сказать, что мембранная часть белка определилась очень хорошо, а вот цитоплазматическая была расшифрована плохо.

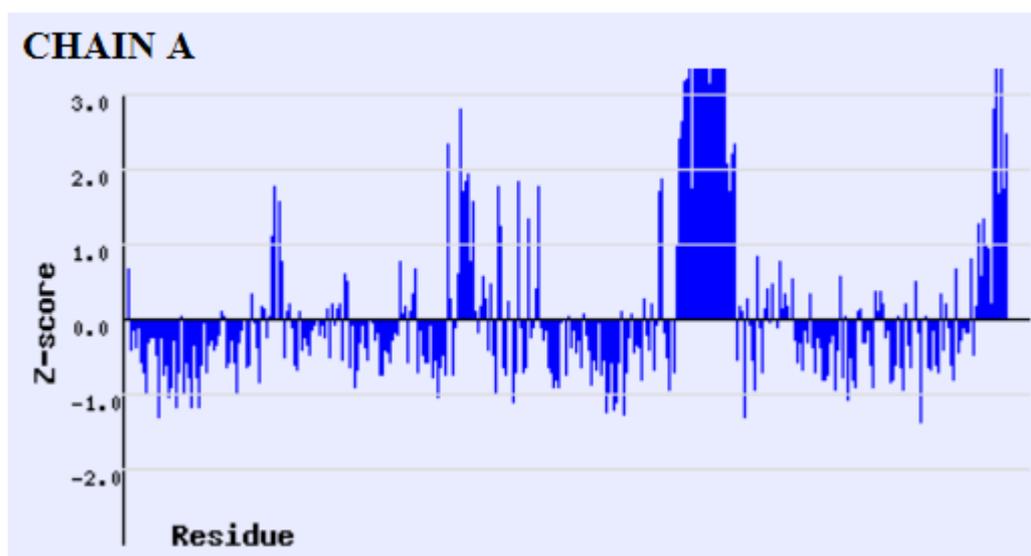


Рисунок 11. Гистограмма Z-score для остатков цепи А.

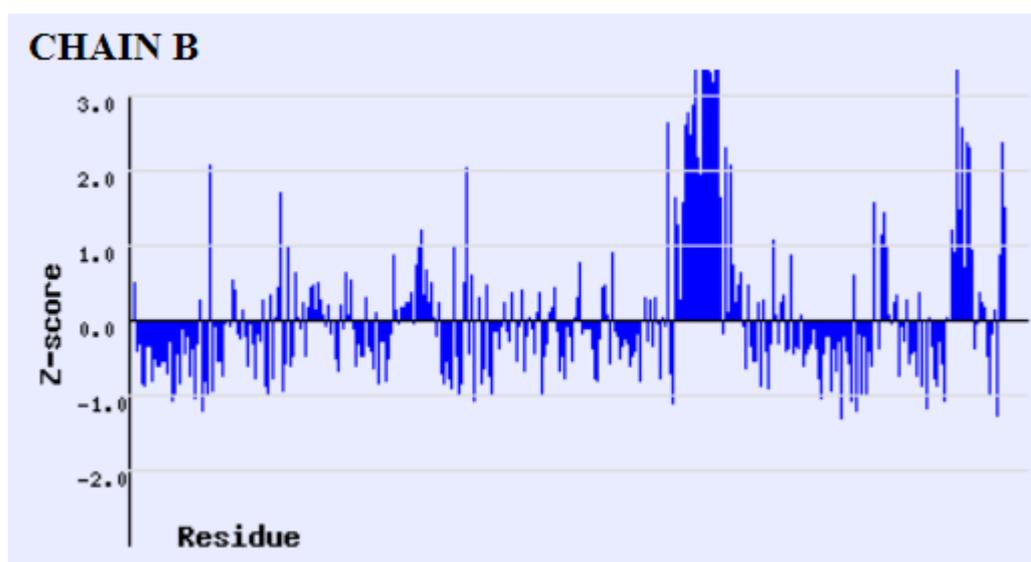


Рисунок 12. Гистограмма Z-score для остатков цепи В.

## Clashscore

Судя по диаграмме параметров структуры, один из худших показателей для 2Z73 по сравнению с другими структурами в PDB является clashscore – количество “слишком близких” контактов несвязанных атомов, неблагоприятных энергетически перекрываний ван-дер-Ваальсовых радиусов, на 1000 аминокислотных остатков (включая водороды). Рассчитывается данный параметр с помощью сервиса MolProbity, при этом в структуру искусственно вписываются атомы водорода.

Для рассматриваемой структуры clashscore равен 15. Контакт, при котором произошло перекрытие от 0,4 до 0,96 Å, в данной структуре 180. “Топовые” перекрытия >0,9Å произошли между атомами водорода 150 аргинина цепи В (0,96 Å), между атомами

водорода 354 изолейцина цепи В (0,94 Å) и между водородом 21 пролина цепи А и водородом 24 аргинина цепи А (0,92 Å).

### **Торсионные углы боковых цепей**

Программа MolProbity сравнивает торсионные углы в боковых цепях аминокислот изучаемой структуры с известными для данных типов остатков ротамерами (наборами “комфортных” торсионных углов) и оценивает, маргинален остаток или нет. Структура 2Z73 содержит 34 маргинальных остатка, что составляет 6% от общего числа проанализированных остатков. По диаграмме на рисунке 5 видно, что этот показатель ниже среднего по всем записям PDB, полученным методом PCA.

Также в структуре найдено 11 остатков, боковые цепи которых можно было бы инвертировать для лучшего образования водородных связей и меньшей поверхности перекрытия ван-дер-Ваальсовых радиусов. Их 11 остатков 7 составляют аспарагины, 4 глутамины.

### **Дополнительно**

С помощью сервиса WHAT\_CHECK[8] были проверены длины связей. Средний Z-score по структуре составил 0,306, программа не выявила явных систематических ошибок. В том же пакете была использована программа Fine Packing Quality Control, чтобы проверить комфортность окружения атома. Средний Z-score по структуре равен -0,31, что по комментарию авторов программы говорит, что структура очень хорошая. Также указано, что остатки со значениями ниже -2,5 являются маргинальными: они могут связывать лиганд или ион, находиться в активном центре или быть ошибочно расшифрованными. Таких в структуре было найдено 3 (Ala 308:A, Ala 308:B, Val 333:B). С помощью еще одной программы из WHAT\_CHECK были проверены конформации пролинов (Proline puckering или складчатость пролина). У двух пролинов значения параметра были равны правой границе интервала нормальных значений – 0,45, у одного пролина (90:A) значение было превышено на 0,01 и равнялось 0,46. Этот же пролин фигурирует в списке остатков с аномальными торсионными углами боковой цепи.

### 3. Некоторые маргинальные остатки структуры 2Z73

Таблица 1. Маргинальные остатки в структуре 2Z73		
№	Остаток	Критерий
1	<u>Ala 243:A</u>	Положение на карте Рамачандрана, $Z > 2$ , слишком близкий контакт атома водорода с атомом водорода Glu 245:A
2	Val 292:A	Положение на карте Рамачандрана, слишком близкий контакт атомов водорода с атомами водорода Ala 284:A
3	Ile 39:B	$Z > 2$ (единственный в цепи B), слишком близкий контакт с атомом кислорода 35 Tyr
4	<u>Ala 308:A</u>	Имеет некомфортное окружение
5	Val 333:B	Имеет некомфортное окружение
6	Gln 328:A	Может быть необходима инверсия боковой цепи, некомфортные контакты атомов водорода с водородами 325 Ala, некомфортные внутримолекулярные контакты атомов водорода
7	Ser 226:B	Положение на карте Рамачандрана
8	<u>Asn 70:B</u>	Может быть необходима инверсия боковой цепи
9	<u>Met 234:B</u>	Экстремальное значение $Z = 5.6$ , аномальные торсионные углы боковой цепи
10	<u>Pro 90:A</u>	Аномальные торсионные углы боковой цепи, критическое значение параметра 'складчатости' (puckering), некомфортные атомов водорода с водородами соседнего остатка

Судя по Z-score, вполне возможно, что цитоплазматические участки белка в целом определены не очень хорошо, и там находится много маргиналов по нескольким показателям. Интересно было бы найти маргинальные остатки в мембранной части белка, так как они с большей вероятностью будут нести какой-то биологический смысл.

## 4. Анализ 5 маргинальных остатков

### Остаток №1: Ala 308, цепь A

Я начала рассмотрение с 308 аланина, так как это практически единственный маргинальный остаток, который мне удалось найти в глубине мембранной части белка (рисунок 13). Более того, этот остаток оказался маргинальным в обеих цепях асимметрической ячейки. Цепи практически одинаковые, но конформации у них немного отличаются, и список маргиналов в них разный.

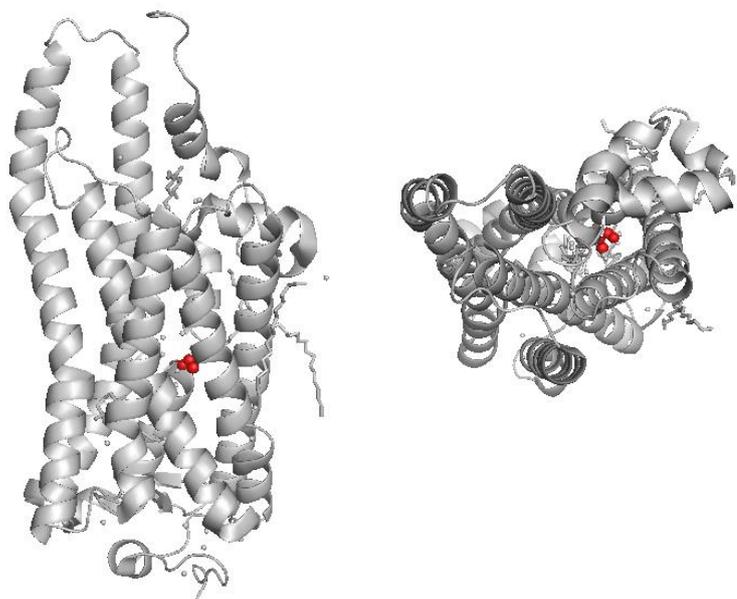


Рисунок 13. Положение аланина 308:А в структуре родопсина кальмара. Показана только субъединица А, весь белок выделен серым цветом, аланин 308 показан красными шариками.

Мембранная часть явно расшифрована достаточно хорошо, но сервис WHAT\_CHECK (алгоритм Fine Packing Quality Control) показал, что у аланина 308 **quality control value** равняется **-2,78**, в то время как авторами отмечено критическое минимальное значение равное -2,5. Это гидрофобный остаток, он не участвует в связывании лигандов. Если взглянуть на структуру, становится понятно, почему программа выдала такой плохой показатель – боковая цепь аланина 308 направлена внутрь “тоннеля” из альфа-спиралей, который заполнен молекулами воды (рисунок 14).

Минимальное расстояние между молекулами воды и углеродом боковой цепи равно 3,4 Å. Надо сказать, что температурный фактор у трех ближайших молекул воды, которые находятся на расстоянии менее 4 Å относительно низкий – 30-40, плюс “нижняя” молекула воды образует водородные связи с парой кислородов остова, а “верхняя” может быть стабилизирована за счет серина 122 и аспартата 80, которые лежат “выше”

обсуждаемого аланина и также развернуты боковыми цепями внутрь кармана, образуя водородную связь между собой (рисунок 14).

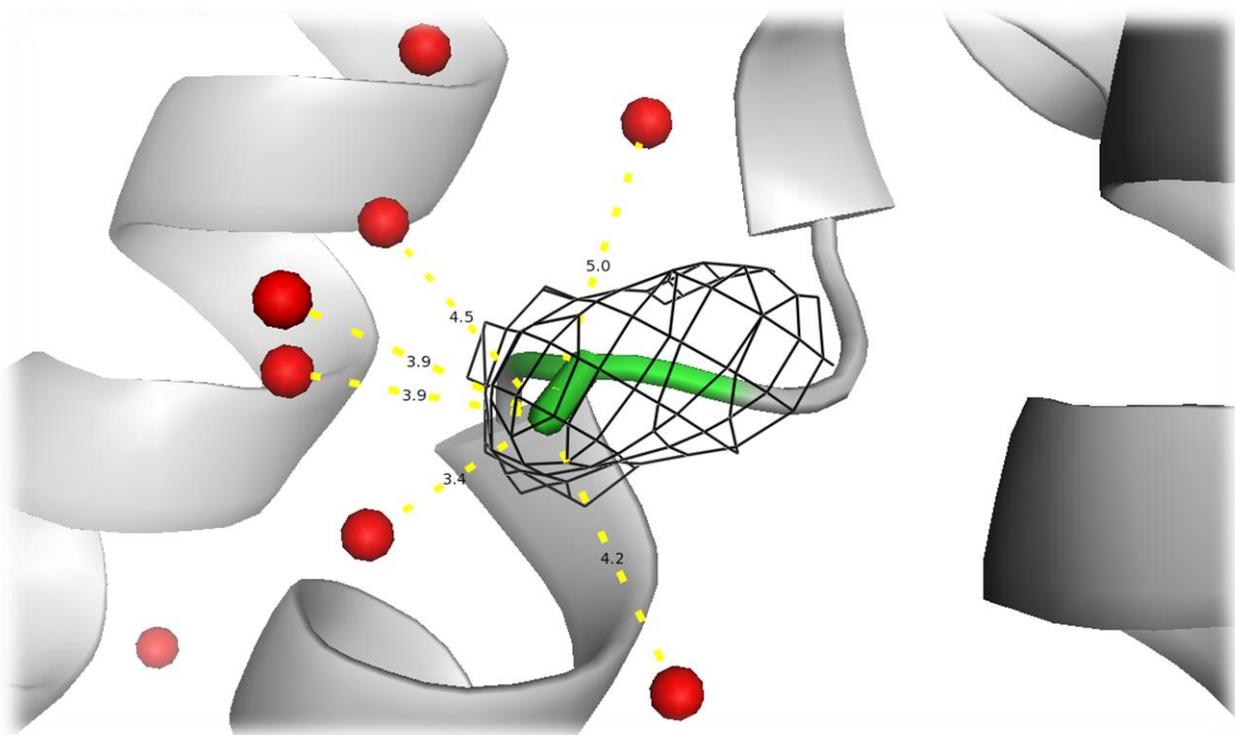


Рисунок 14. Аланин 308:А в окружении молекул воды. Остаток выделен зеленым цветом, боковая цепь показана в виде "sticks", вода показана красными шариками, остальной белок выделен серым. Черной сеткой показана электронная плотность, уровень подрезки 1.0.

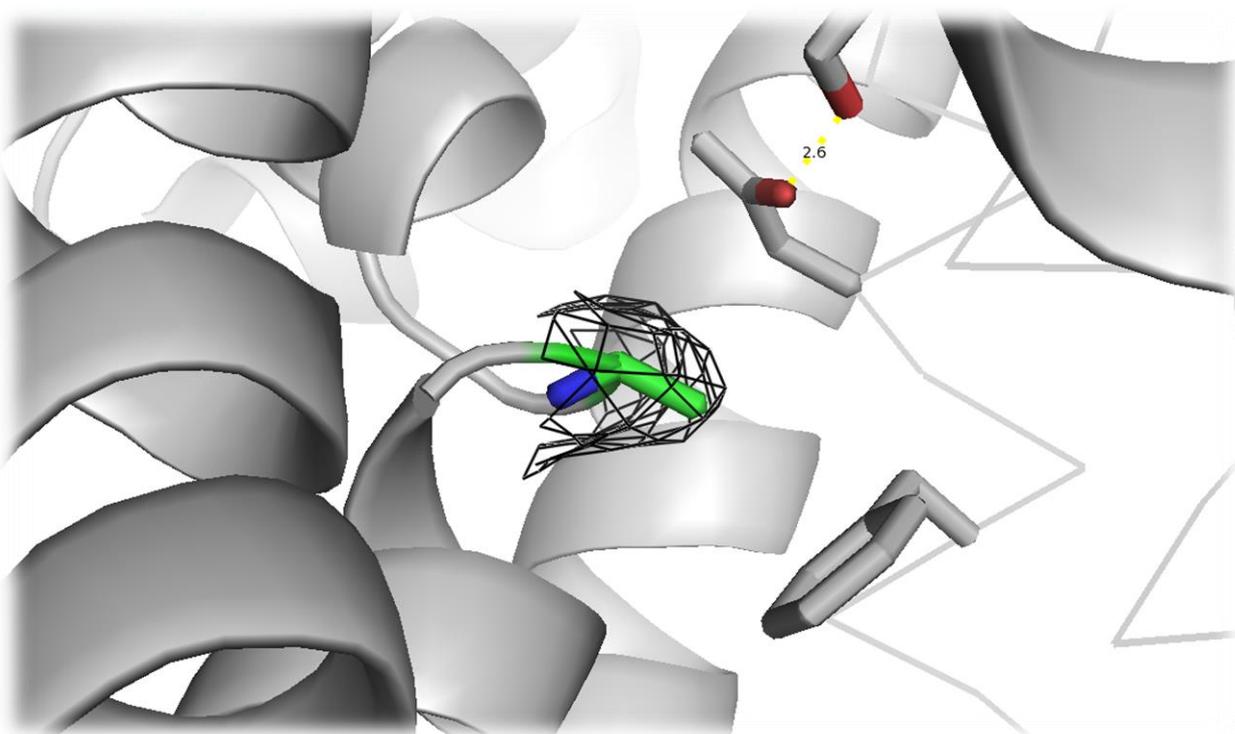
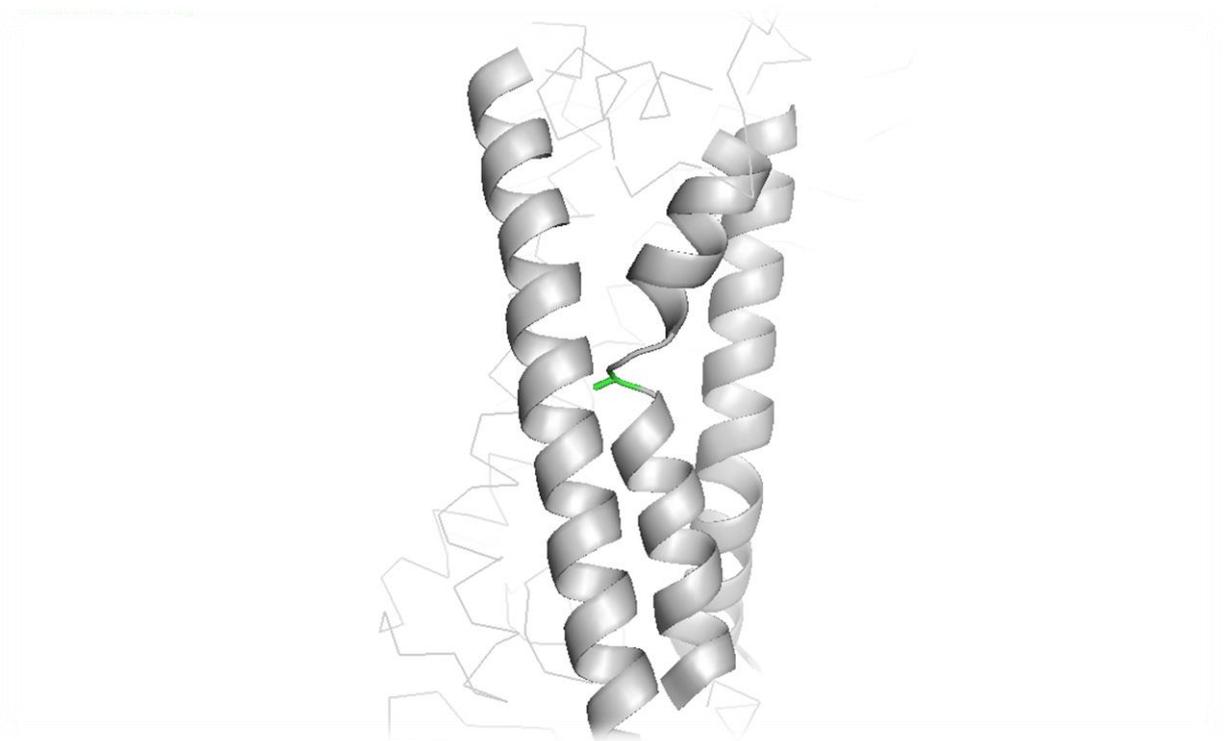


Рисунок 14. Аналогично рисунку 13 показан аланин 308:А и три близких остатка: аспарагин 80, серин 122 и фенилаланин 83.

Но вода все-таки не очень достоверный элемент структур, поэтому было решено посмотреть на ближайшее аминокислотное окружение остатка. Действительно, выше по тоннелю серин 122 и аспартат 80 образуют полярный контакт, но напротив самого аланина находится фенилаланин 83 (см также рисунок 14). Фенилаланин – исключительно гидрофобная аминокислота. Вполне возможно, что вместе с интересующим нас аланином 308 он образует своего рода “гидрофобный контакт”, аланину выгоднее быть повернутым в эту сторону, чем в какую-либо другую.

Привлекает внимание еще одна интересная деталь: наш аланин 308 находится в центре небольшого (3 аминокислоты длиной) петлевого участка между двумя альфа-спиралями. Посмотрев на структуру с чуть большего расстояния, мы можем заметить, что петлевой участок будто бы ломает одну цельную альфа-спираль, позволяя ей изогнуться и “прижаться” центром к срединному тоннелю, а краями образовать расширение (рисунок 15).



*Рисунок 15. Аланин 308 и изгиб альфа-спирали. Картонками показаны альфа-спирали, окружающие петлю с аланином 308, бантиком весь остальной белок. Аланин 308 зеленого цвета.*

Мне кажется, эта яркая структурная особенность в совокупности с “контактом” с фенилаланином 83 и хорошей поддержкой электронной плотности указывает на важность подобной конформации аланина 308.

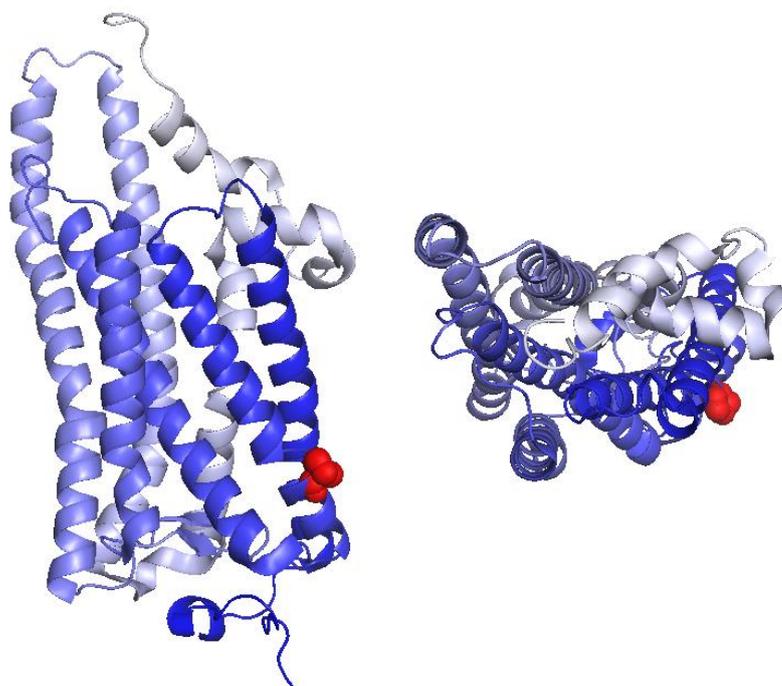
**Вывод:** не ошибка расшифровки, а особенность.

## Остаток №2: Pro 90, цепь A

Теперь рассмотрим пролин 90:A. Во-первых, остаток был причислен к маргиналам двумя разными алгоритмами, но по одной и той же причине: нарушение конформации боковой цепи (MolProbity), неестественная складчатость боковой цепи (WHAT\_CHECK). Как я указывала ранее, **амплитуда складчатости** данного пролина, рассчитанная методом Кремера и Попла (Cremer and Pople), **равнялась 0,46**, при норме 0,2 – 0,45. Отличие от нормы минимальное, всего 0,01, но MolProbity тоже выделила данный остаток, поэтому было решено изучить его пристальнее.

Также было обнаружено перекрывание Ван-дер-Ваальсовых радиусов атома водорода с атомом водорода соседнего фенилаланина 89 (**interatomic distance = 1.89 Å, clash overlap = 0.40 Å**).

Находится пролин 90 в альфа спирали мембранной части белка, близко от поверхности мембраны, выходящей в межклеточную среду(рисунок 16).



*Рисунок 16. Положение пролина 90:A в структуре родопсина кальмара. Красными шариками выделен остаток Pro 90:A, картонками с градиентом от синего цвета к белому покрашен остальной белок (синий – начало нумерации остатков, белый – конец нумерации).*

Есть ли какие-то причины, вызывающие повышенную амплитуду складчатости? Посмотрим на ближайшее окружение пролина 90:A (рисунок 17). Во-первых, видно, что рядом с пролином связалась вода ( $B = 36.67$ ). Остатки вокруг гидрофобные, однако в витке альфа-спирали выше пролина 90 образовался карман, в который выставлено несколько кислородов остова, по-видимому, стабилизирующих молекулу.

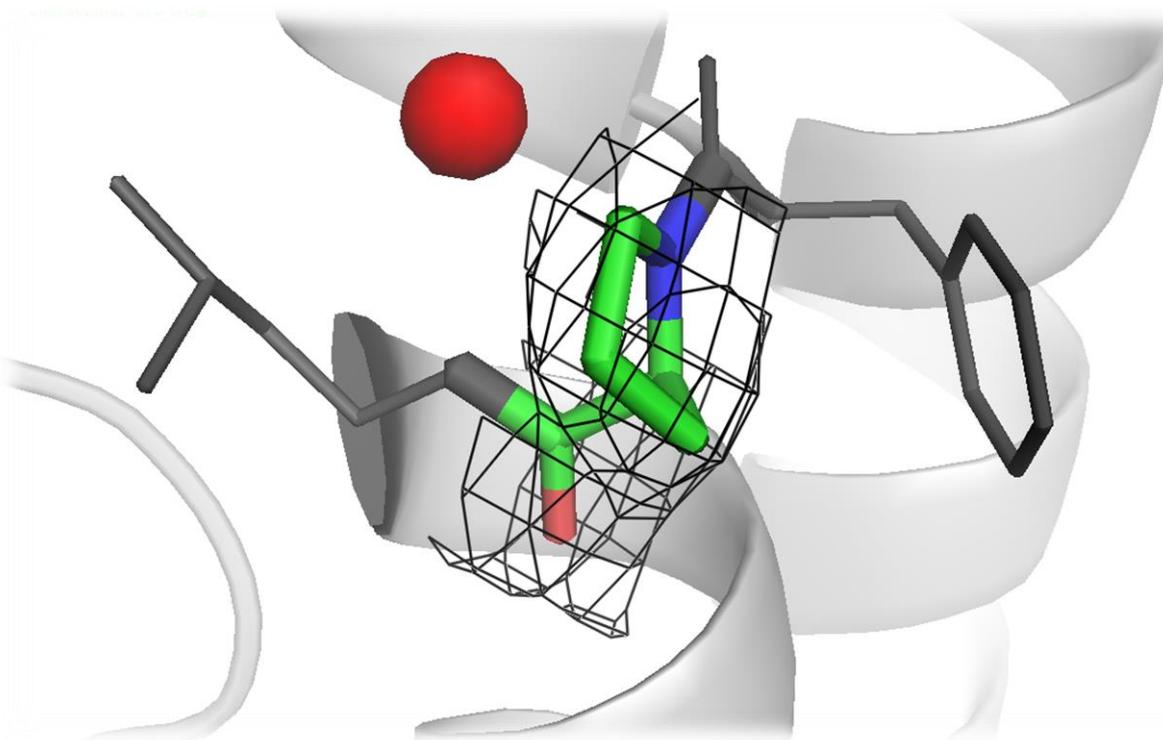


Рисунок 17. Пролин 90:А и соседние остатки (Phe 89 и Leu 91). Пролин раскрашен поатомно, красный шарик – атом кислорода воды, соседние остатки раскрашены поатомно с углеродами темно-серого цвета, остальной белок светло-серый. Вокруг пролина построена электронная плотность с уровнем подрезки 1.0.

Правда, расстояние от складки пролина до воды все-таки значительное, не похоже, что она может как-то влиять на конформацию боковой цепи.

Если посмотреть на электронную плотность пролина при разных уровнях подрезки, то заметно, что он определился четко и вписан очень хорошо. Закрадывается подозрение, что даже слишком хорошо (рисунок 18).

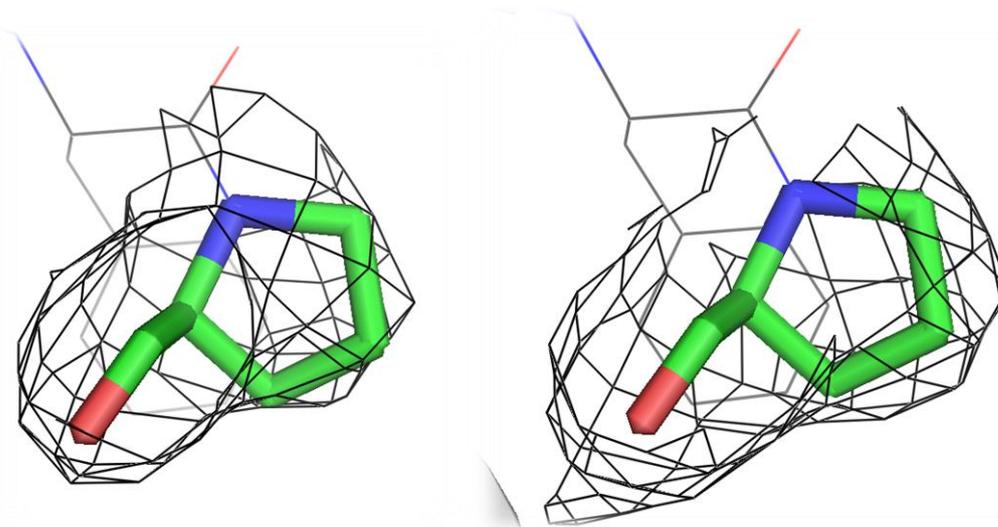
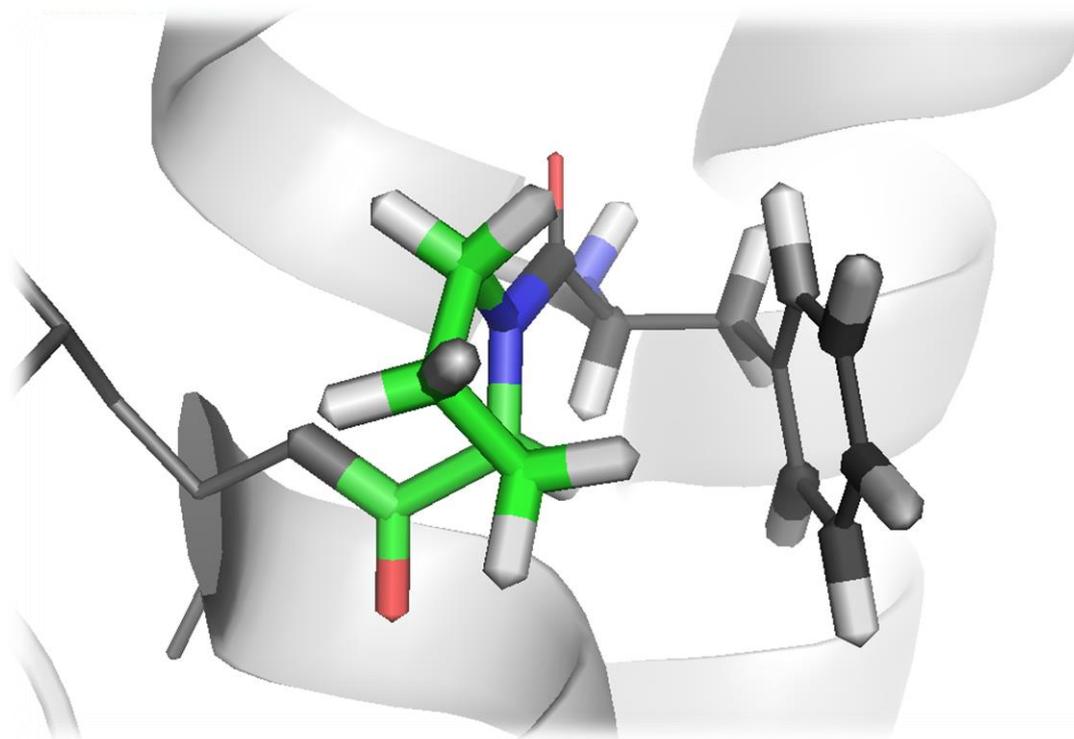


Рисунок 18. Пролин 90:А. Слева: уровень подрезки = 2.0, справа: уровень подрезки = 1.0

На рисунке 18 (слева) видно, что атомы углерода находятся на самой границе электронной плотности при уровне подрезки равном 2.0. Возможно, программа при подборе конформации боковой цепи слишком сильно хотела вписать пролин в эту электронную плотность, поэтому и сжала его до критического значения амплитуды складчатости.

Что касается слишком близкого расположения к фенилаланину 89, кажется, что отнесение данного пролина к маргиналам по этому критерию не совсем обосновано. В списке перекрывающихся атомов пара Pro 90 – Phe 89 записана последней, так как площадь перекрытия очень мала – 0.4 Å. Визуально ничего необычного не обнаруживается (рисунок 19).



*Рисунок 19. Атомы водорода пролина 90 и фенилаланина 89 (светло-серые толстые палочки).  
Рассчитаны программой PyMOL.*

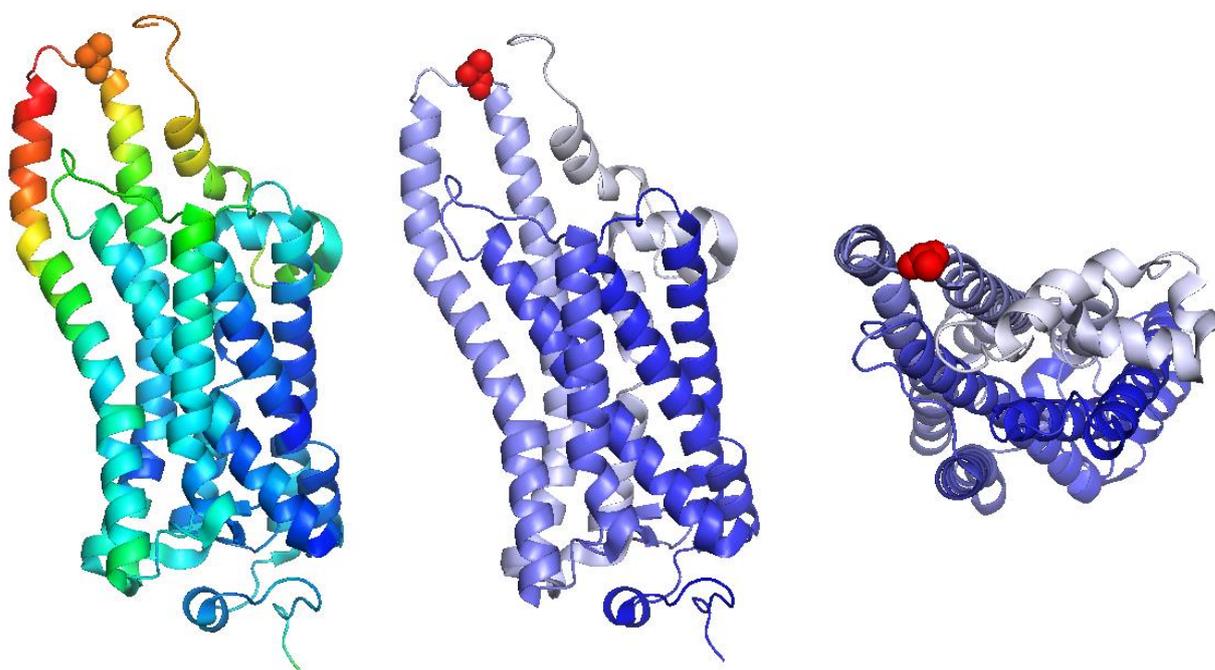
**Вывод:** показатель складчатости повышен из-за ошибки расшифровки, не особенность.

### Остаток №3: Ala 243, цепь А

Следующий герой повествования – аланин 243, который маргинален по трем показателям:

- Положение на карте Рамачандрана (визуально дольше всех ушедший от приемлемой области, но не слишком далеко)
- **Z-score = 4.5**
- Некомфортный контакт с атомом водорода Glu 245:A

В структуре этот остаток располагается на самой верхушке выходящей в цитоплазму петли (рисунок 20).



*Рисунок 20. Расположение аланина 243:A. Слева: раскраска по градиенту температурного фактора – чем больше, тем краснее, чем меньше, тем синее. По центру и справа: раскраска по градиенту нумерации атомов, синий- начало отсчета, белый – конец цепи. Красными шариками показан аланин 243.*

Помимо положения остатка в структуре на рисунке 19 приведен градиент значений В-фактора (слева). Иллюстрация демонстрирует, что цитоплазмальная часть белка (зелено-желто-красная) гораздо более подвижна, чем мембранная (оттенки синего), что коррелирует со значениями Z-score (рисунок 11, где мембранной части соответствует широкий пик).

Поэтому к остаткам внеклеточной петли будем относиться снисходительно. Рассмотрим торсионные углы 243 остатка (рисунок 21).

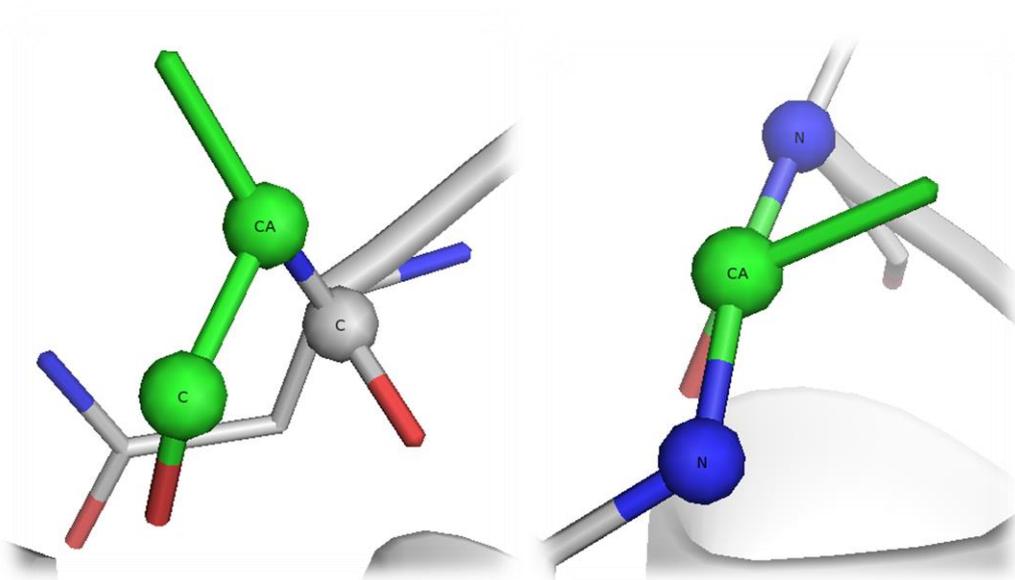


Рисунок 21. Слева: угол  $\phi$ , справа: угол  $\psi$ . Шариками отмечены атомы, между которыми рассчитывался угол, зеленым цветом – исследуемый остаток.

Угол  $\phi = -167^\circ$ , угол  $\psi = -60^\circ$ .

Посмотрим на электронную плотность. С электронной плотностью беда: на уровне подрезки = 1.0 цепь просматривается, а боковые группы нет (рисунок 22).

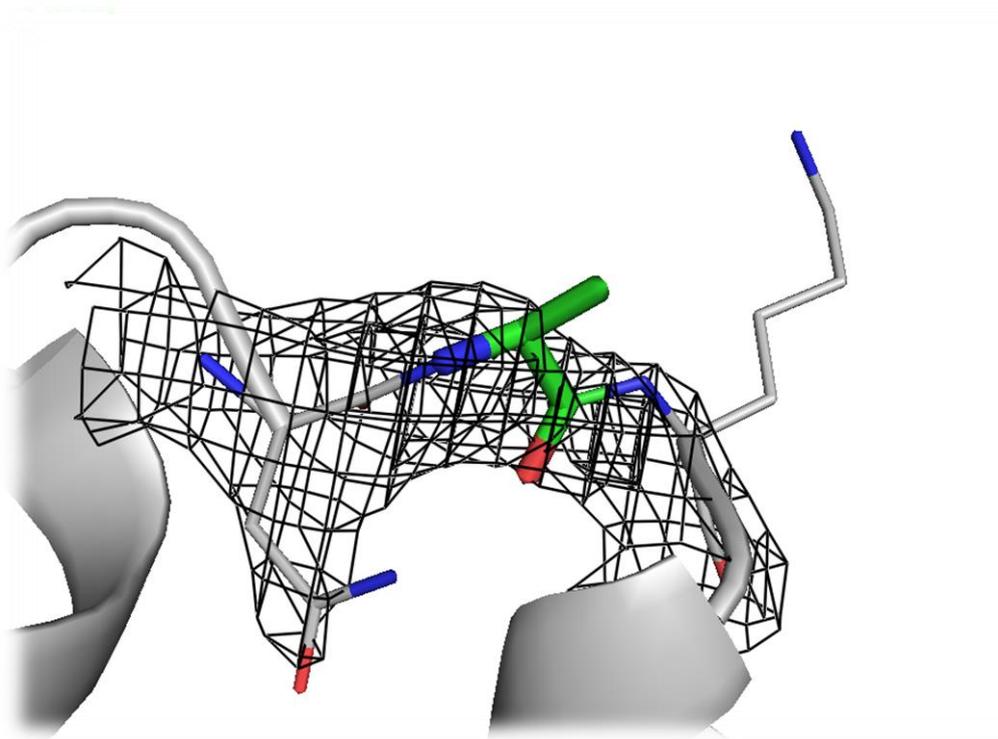


Рисунок 22. Электронная плотность вокруг аланина 243 и двух соседних остатков, уровень подрезки = 1.0.

По этому уровню подрезки судить о правильности конформации сложно. Если же взглянуть на электронную плотность с уровнем подрезки =0.0 со стороны примерно как для расчета угла  $\psi$ , то видно, что скорее всего боковая цепь аланина повернута неправильно (рисунок 23).

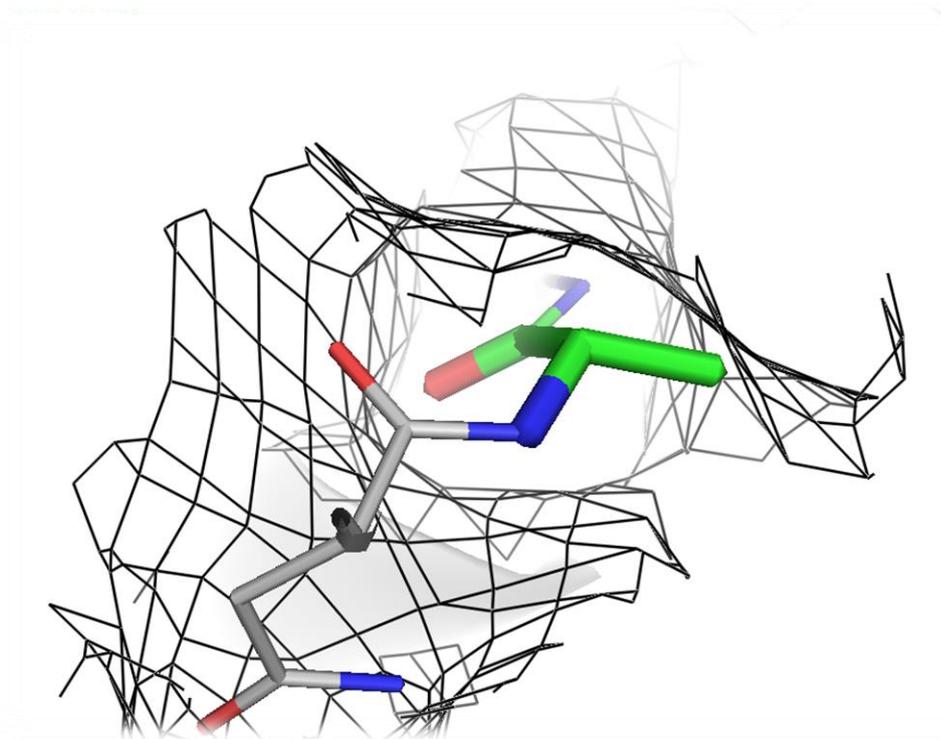


Рисунок 23. Электронная плотность вокруг аланина 243, вид сбоку. Уровень подрезки = 0.0, *carve* = 2.4. Аланин 243 выделен зеленым цветом.

Боковая группа “упирается” в стенку электронной плотности, и кажется, что она должна быть развернута чуть дальше по часовой стрелке. Соседние остатки вписываются в электронную плотность лучше.

Учитывая плохое качество расшифровки этого остатка, вероятно, он неправильно расшифрован.

**Вывод:** ошибка расшифровки, не особенность.

## Остатки №4 и №5: Met 234 и Asn 70, цепь В

Анализ боковых цепей остатков показал, что аспарагину 70 необходим поворот боковой цепи для лучшего образования водородных связей. Метионин похож на предыдущий аланин 243: очень большое значение  $Z=5.6$ , аномальные торсионные углы боковой цепи.

Видно, что в структуре метионин расположен недалеко от аланина 243, но уже не в петле, а в альфа-спирали, направленной в цитоплазму (рисунок 24). Эта альфа-спираль была самой красной на рисунке градиента температурного фактора (рисунок 20). Аспарагин расположен на границе между мембранной и цитоплазматической частью, в начале резкого изгиба альфа-спирали (рисунок 24).

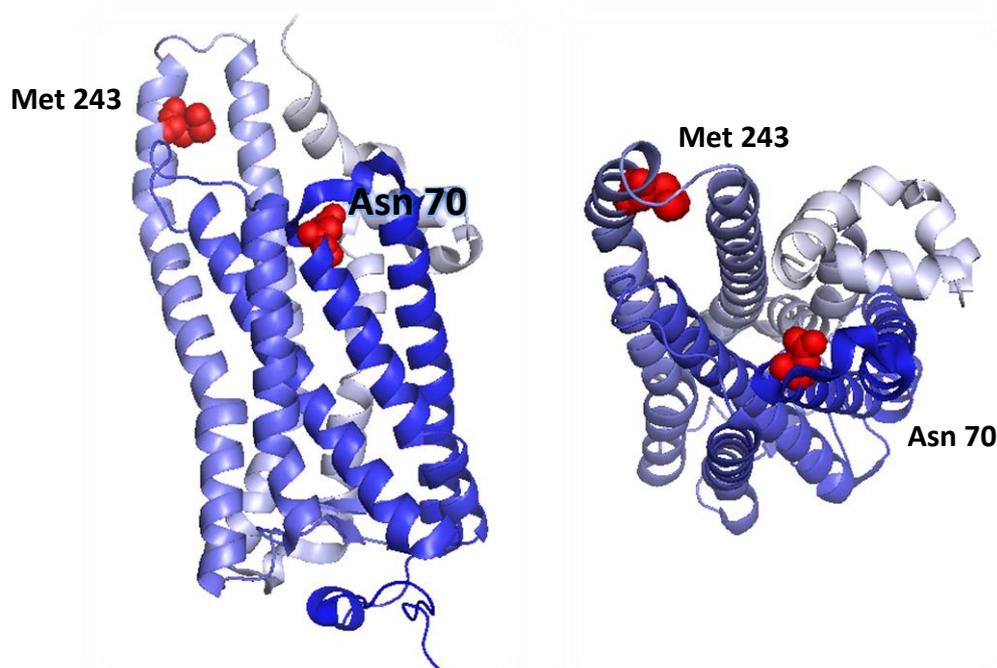


Рисунок 24. Положение остатков метионина 243 и аспарагина 70 в структуре родопсина кальмара.

Аспарагин хорошо вписывается в электронную плотность, образует водородные связи с двумя молекулами воды (рисунок 25). Молекулы воды расположены почти симметрично, по ним сложно судить о необходимости перестановки. Однако видно, что вокруг азота электронная плотность чуть шире, чем вокруг атома кислорода, поэтому, вероятно, перестановка их местами имеет смысл.

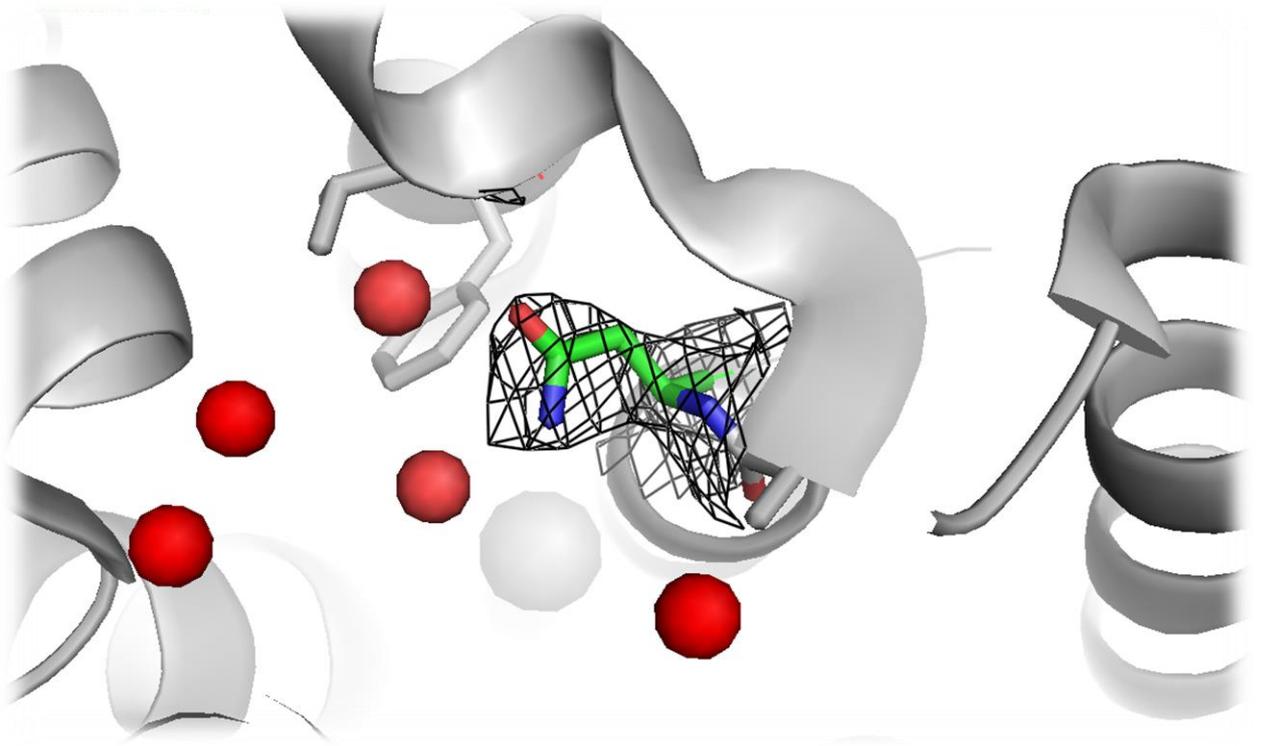


Рисунок 25. Электронная плотность и молекулы воды рядом с аспарагином 70. Уровень подрезки = 1.0,  $\sigma$  = 2.6.

Метионин 234 явно плохо расшифрован (рисунок 26). Хорошо вписывается атом серы, а вот соседствующие с ним углероды выходят за границы электронной плотности даже при меньших уровнях подрезки.

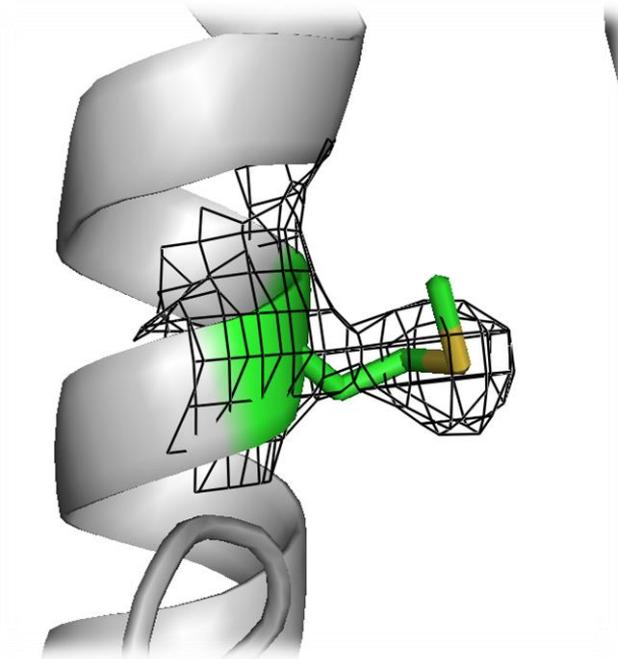


Рисунок 26. Метионин 234 и электронная плотность вокруг него, уровень подрезки = 1.0.

**Вывод:** атомы расшифрованы с ошибками, особенностей не обнаружено.

## 5. Сравнение с моделью из PDB\_redo

Сервис PDB-redo[9] проводит повторную оптимизацию модели, что часто позволяет заметно улучшить структуры из PDB (таблица 2).

Таблица 2. Сравнение параметров структуры 2Z73 с её улучшенной версией		
Параметр	PDB	PDB-redo
R-фактор	0,1974	0,1978
R-free	0,2509	0,2153
Bond length RMS Z-score	0,306	0,418
Bond angle RMS Z-score	0,564	0,606

Из таблицы заметно, что улучшение, проведенное PDB-redo весьма сомнительно. На 3,5 % уменьшился показатель R-free, но странность в том, что в самой структуре со страницы PDB R-free указан 20,5%, что еще меньше, чем посчитанный PDB-redo.

Заметить значительные различия между структурами при визуальном сравнении затруднительно, но некоторые можно (рисунок 27).

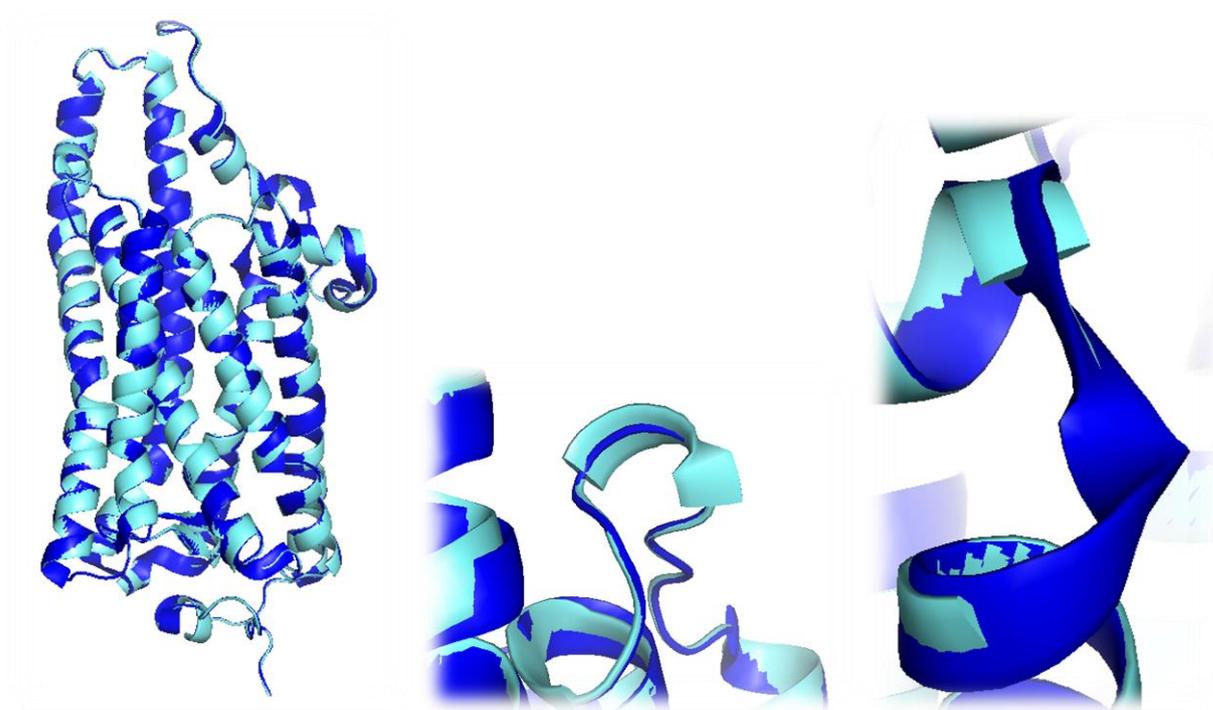


Рисунок 27. Сравнение структур 2z73 и 2z73\_final, полученной сервисом PDB-redo. Слева: наложение двух структур, темно-синяя – исходная версия, светло-синяя – повторная. По

*центру: в 2z73\_final присутствует альфа-спираль, в исходной структуре нет. Спарва: обратный случай, в исходной есть альфа-спирал, в повторной нет.*

На рисунке 27 показаны различия в определении вторичных структур. В одном месте в PDB есть альфа-спираль, а в PDB-redo нет, в другом месте – наоборот. Мне кажется, что и в том и в другом случае логичнее было бы поставить петли. Зачаток альфа-спирали, добавленный PDB-redo, выглядит совершенно излишним.

В целом, даже цитоплазматическая часть легла практически одинаково.

## Заключение

Из рассмотренных в разделе “Результаты и обсуждение” показателей можно сделать вывод, что структура белка в целом хорошая, но качество расшифровки неоднородно. В цитоплазматических участках (228 - 248, 348 – 358 остатки) гораздо больше маргинальных остатков по нескольким показателям, чем в мембранной части, и практически все они имеют большой Z-score.

Интересно было бы больше узнать о роли специфического излома альфа-спирали 294 – 319, в котором находится аланин 308.

## Список литературы

- [1] O. P. Ernst, D. T. Lodowski, M. Elstner, P. Hegemann, and L. S. Brown, "Microbial and Animal Rhodopsins : Structures , Functions , and Molecular Mechanisms," *Chem. Rev.*, 2014.
- [2] M. Murakami and T. Kouyama, "Crystal structure of squid rhodopsin," *Nature*, vol. 453, no. May, pp. 363–368, 2008.
- [3] Y. Tohru and S. Manabu, "Physical aspects of sensory transduction on seeing, hearing and smelling," *Biophysics (Oxf).*, no. December, 2013.
- [4] T. Okada, K. Takeda, and T. Kouyama, "Highly Selective Separation of Rhodopsin from Bovine Rod Outer Segment Membranes Using Combination of Divalent Cation and Alkyl(thio)glucoside," *Photochem. Photobiol.*, vol. 67, no. 5, pp. 495–499, 1998.
- [5] A. T. Brünger *et al.*, "Crystallography & NMR system: A new software suite for macromolecular structure determination.," *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 54, no. Pt 5, pp. 905–21, Sep. 1998.
- [6] "MolProbity." [Online]. Available: <http://molprobity.biochem.duke.edu/>.
- [7] "EDS server, 2Z73." [Online]. Available: <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=2z73>.
- [8] "WHAT\_CHECK." [Online]. Available: <http://swift.cmbi.ru.nl/gv/whatcheck/>.
- [9] "PDB\_redo." [Online]. Available: [http://www.cmbi.ru.nl/pdb\\_redo/](http://www.cmbi.ru.nl/pdb_redo/).