**Отчет о качестве расшифровки структуры белка spsE, участвующего в биосинтезе покровного полисахарида споры в клетках *Bacillus subtilis* (PDB код 1VLI) методом рентгеноструктурного анализа**

Отчет выполнила студентка 4 курса Калинина Марина

Содержание

[Аннотация 3](#_Toc404971693)

[Введение 4](#_Toc404971694)

[Результаты и обсуждение 5](#_Toc404971695)

[1. Общая информация о модели 5](#_Toc404971696)

[2. Значения индикаторов качества модели 6](#_Toc404971697)

[3. Маргинальные остатки 9](#_Toc404971698)

[4. Детальный анализ 5 маргинальных остатков 9](#_Toc404971699)

[i. Глицин-245 и аланин-246, маргиналы по карте Рамачандрана 9](#_Toc404971700)

[ii. Аспартат-247, маргинал по плохому углу ω и большому В-фактору. 10](#_Toc404971701)

[iii. Тирозин-185, маргинал по плохому углу ω и странному окружению 11](#_Toc404971702)

[iv. Глицин 296, маргинал по Zscore 12](#_Toc404971703)

[5. Сравнение модели из PDB с моделью из PDB\_redo 12](#_Toc404971704)

[Заключение 13](#_Toc404971705)

[Список литературы 14](#_Toc404971706)

# Аннотация

В работе был проведен анализ структуры белка, принимающего участие в биосинтезе полисахаридной оболочки споры. Приведены некоторые показатели качества построенной модели (карта Рамачандрана, R-фактор, R-free) и список из 12 маргинальных остатков. Подробно рассмотрены 5 маргинальных остатков структуры белка.

Количество маргинальных остатков в структуре не может служить показателем модели, так как маргинальность не всегда объясняется плохим согласованием электронной плотности с моделью. Часть маргиналов можно объяснить, изучив их окружение. В конкретной структуре большинство маргиналов связано с проблемой расшифровки и плохим соответствием с электронной плотностью, но также есть маргиналы, играющие функционально важную роль в белке, и их маргинальность, на мой взгляд, биологически обоснована. В результате проведенного анализа можно сделать вывод, что структура 1vli разрешена не очень хорошо (сразу по нескольким показателям качества).

# Введение

Белок, участвующий в синтезе полисахаридной оболочки споры Bacillus subtilis (spsE) – гликопируваттрансфераза, принадлежащая к суперсемейству альдолаз. В сам путь биосинтеза полисахаридной споровой оболочки вовлечены еще 10 белков sps, которые предсказаны как «потенциальные партнеры» spsE (однако, прямых доказательств этому пока нет) [[1](#_ENREF_1)]. Белку предсказана активность переноса пирувата с фосфоенолпирувата на сахар. Путь биосинтеза полисахаридной споровой оболочки изучен только в одном организме – Bacillus subtilis, однако кристаллическая структура предполагаемо spsE недавно была получена также для *Chromobacterium violaceum* (структура 3g8r).

К сожалению, не удалось найти статью о расшифровке структуры spsE, хотя обнародована структура была в PDB в 2004 году (последние изменения были внесены в 2011 году). Пришлось взять статью [[2](#_ENREF_2)] о расшифровке структурного гомолога spsE, а именно NeuB. Этот белок синтезирует сиаловую кислоту (N-ацетилнейраминовую кислоту, входящую в состав гликокаликса) во множестве клеток как прокариот, так и эукариот. Авторы заинтересовались белком, так как сиаловая кислота достаточно широко распространена на поверхности бактериальных клеток, а также некоторые патогенные бактерии (кишечная палочка и менингококк, например) синтезируют сиалированные производные, предотвращающие иммунный ответ организма-хозяина. Сиаловая кислота является привлекательной лекарственной мишенью, но для этого важно понимать полный путь ее синтеза. Работа посвящена получению первой структуры NeuB из менингококка в комплексе с малатом и с кофактором Mn2+ , а также в комплексе с N-ацетилнейраминовой кислотой (продуктом реакции).Также в работе доказан механизм реакции путем мечения O18 , демонстрирующий разрыв C-O связи в ходе ферментативной реакции.

**Целью** данной работы был анализ и оценка качества белка spsE из организма Bacillus subtilis.

# Результаты и обсуждение

## Общая информация о модели

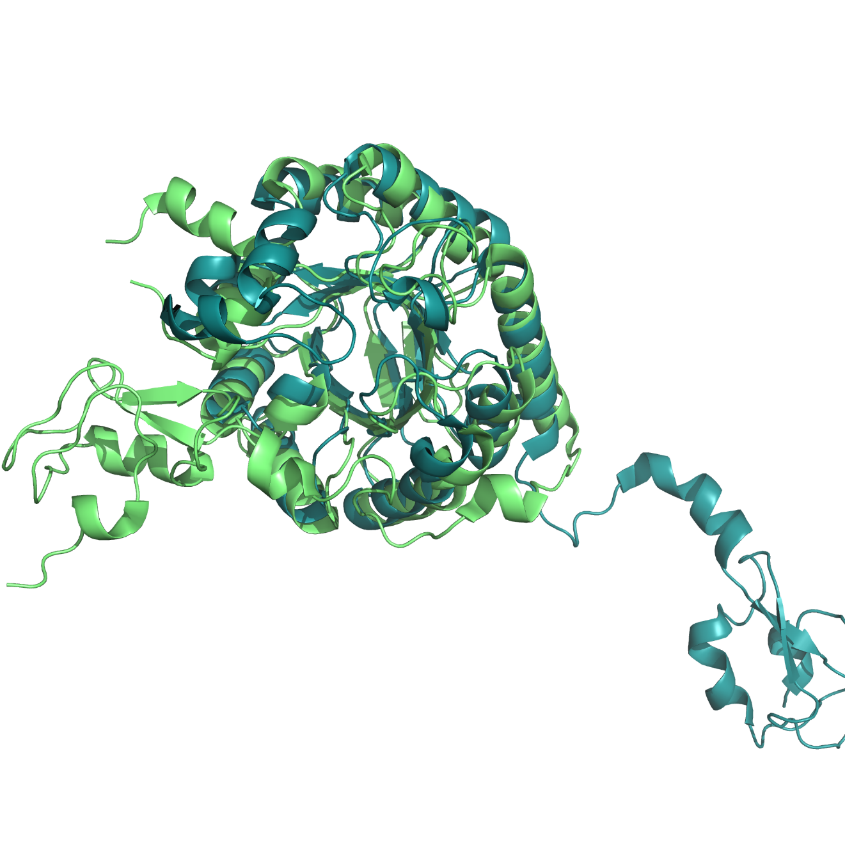
***Состав комплекса****:* Белок биосинтеза полисахаридной споровой оболочки, имеющий одну цепь А (содержит 373 аминокислотных остатка). Биологическая единица также является мономером, то есть совпадает с ассимметрической. Элементарная ячейка содержит 6 молекул. Пространственная группа - P 32 2 1.

***Дата размещения в PDB-банке***: 2004-07-27. Последнее обновление структуры в банке - 2011-07-13.

***Авторы структуры***: Joint Center for Structural Genomics.

***Метод решения фазовой проблемы***(на основе статьи [[2](#_ENREF_2)])

Гомологичность структур 1vli и 1xuz (NeuB) показана на изображении ниже (Рисунок 1). Для решения фазовой проблемы в случае расшифровки структуры 1xuz использовался многоволновой метод аномального рассеяния. Для этого в молекулу белка были введены селеновые производные метионинов, и были идентифицированы 8 из 10 мест вставок. Карты электронной плотности были улучшены посредством сглаживания растворителя. Изначальная модель была построена вручную (с использованием XTALVIEW), а финальная модель, содержащая 349 остатков в термодинамически предпочтительных позициях была получена после нескольких раундов улучшения до разрешения 1,9 Å. Качество модели было проанализировано PROCHECK, обнаружено, что 93% остатков располагаются в предпочитаемых областях карты Рамачандрана и ни одного остатка не располагается в запрещенных областях.



**Рисунок 1 Наложение структур 1vli (зеленая) и 1xuz (голубая). Видно, что основные каталитические домены совпадают**

*Далее будет рассматриваться только белок spsE*

***Число измеренных рефлексов (в файле структурных факторов):*** 24643

***Число и процент рефлексов с силой сигнала >3.0:*** 23455 (из 24643) = 95,2%

***Разрешение, в том числе минимальное и максимальное разрешение для использованных рефлексов:*** 29.29 Å – 2,38 Å

***Число использованных рефлексов и полнота рефлексов в процентах:*** 24620 (из 24643) – полнота данных составляет 99,9%.

***Наличие некристаллографических симметрий в асимметрической ячейке***– присутствует вращательная симметрия - поворот на 120 градусов и перемещение на 2/3 вектора а.

## Значения индикаторов качества модели

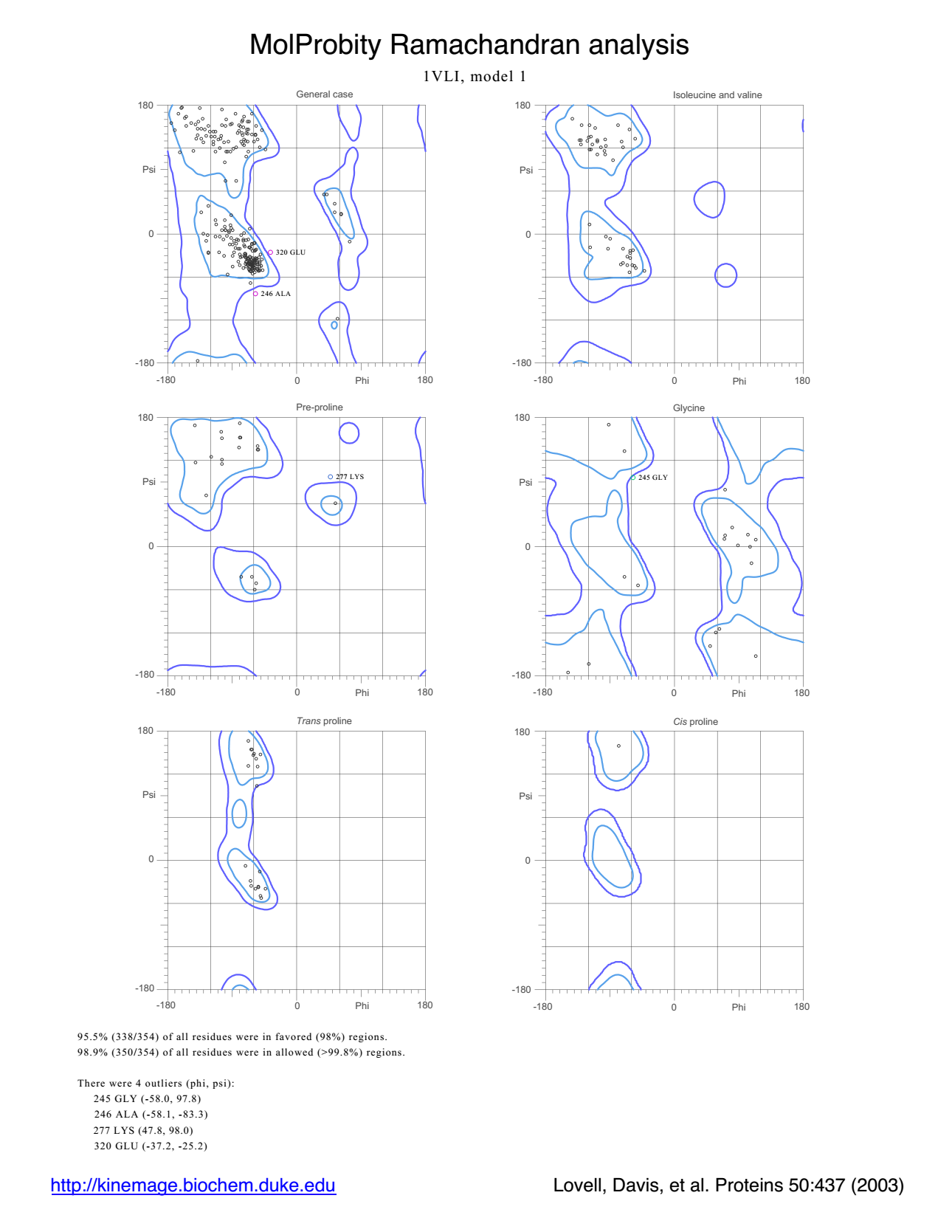
***R-фактор:***

Rx-ray=19,3%, Rfree=23,8% ( Rfree посчитан по 1255 рефлексам - 5,10% от общего числа), разница между ними меньше условного порога в 10% и считается индикатором хорошей оптимизации (точное значение равно 4,8%).

***Карта Рамачандрана:***

Для построения карты был использован сервер MolProbity [[3](#_ENREF_3)]. Результат анализа торсионных углов этим сервером представлен на Рисунок 2 и содержит 6 карт (множество пар двугранных углов φ и ψ для всех типов аминокислотных остатков разбито для удобства на 6 групп). На карте находится 4 остатка, находящихся в запрещенной области – глицин-245, аланин-246, лизин-277 и глутаминовая кислота-320.

На MolProbity [[3](#_ENREF_3)] приведены критерии качества по некоторым характеристикам. Количество остатков, находящихся в разрешенных и предпочитаемых областях карты Рамачандрана, ниже стандартных порогов (пороги – должно быть менее 0,05% остатков в запрещенной области (у данной структуры это 1,13% остатков) и более 98% остатков в предпочитаемой области – в данной структуре эта цифра равняется 95,48%).



**Рисунок 2 Карта Рамачандрана для структуры белка 1vli. Карта построена сервисом MolProbity. На карте видны 4 маргинала, находящиеся в запрещенной зоне углов – 245 Gly, 246 Ala, 277 Lys и 320 Glu**

***Ротамеры:***

MolProbity [[3](#_ENREF_3)] нашел 10 остатков с конформацией боковой цепи, не соответствующей ни одному из ротамеров. Это число (3,53%) также меньше допустимого порога (менее 1%).

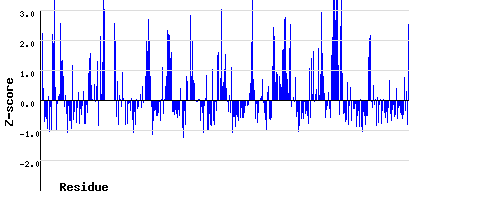
***Clashscore***

Число недопустимых наложений атомов составляет 4.42 атома на 1000. Это соответствует 99-й перцентили, что отмечается как хорошее значение.

Все значения статистик, выданных сервисом MolProbity [[3](#_ENREF_3)], суммируются в Таблица 1.

***Z-score***

Считается, что z-score меньше 2 – допустимое значение для остатка. В структуре белка 1vli по мнению сервера EDS [[4](#_ENREF_4)] содержится 33 остатка (9,22%), z-score которых превышает значение 2 (Рисунок 3).



**Рисунок 3 Z-score для всех остатков структуры [**[**4**](#_ENREF_4)**]**

***Геометрия полипептидного остова.***

В данном случае плохим остатком считается такой, у которого Z больше 5. Есть два маргинала по длине связей - 270 Glu по трем связям (Z = 23.64, 15.27 и 6.60) и 316 Glu (Z = 5.34). По величине углов остова находятся 5 остатков, самое большое значение Z у 363 Asp (Z = 6.99) и 370 Glu (Z = 6.04). Маргиналов по хиральности и планарности данная структура не имеет.

**Таблица 1 Суммарная статистика по качеству структуры 1vli из сервиса Molprobity [**[**3**](#_ENREF_3)**]**

|  |  |  |
| --- | --- | --- |
| Clashscore для всех атомов | 99-я перцентиль | 100 процентов |
| Плохие ротамеры | 3,53% | <1% |
| Аутлайеры в карте Рамачандрана | 1,13% | <0.05% |
| Остатки в предпочитаемой области в карте Рамачандрана | 95,48% | >98% |
| Сbeta отклонения> 0.25 А | 1,18% | 0 |
| Плохие связи остова | 0,29% | 0 |
| Плохие углы остова | 0,05% | <0,1% |

## Маргинальные остатки

В Таблица 2 представлен список из 12 маргинальных остатков.

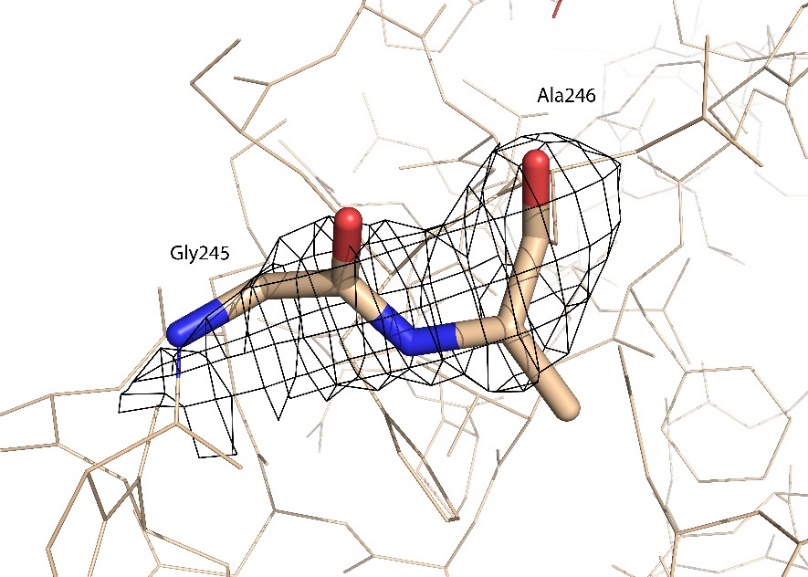
**Таблица 2 Список 12 маргинальных остатков с указателем критерия, по которому были отобраны**

|  |  |  |
| --- | --- | --- |
| **№** | **Остаток** | **Критерии** |
| **1** | 245 GLY | В запрещенной области на карте Рамачандрана |
| **2** | 246 ALA |
| **3** | 277 LYS |
| **4** | 320 GLU |
| **5** | 270 GLU | Странная геометрия остова (длина связей, углы) |
| **6-7** | 185 TYR + 186 PRO | Странное окружение, плохой угол ω |
| **8** | 125 ALA | Находится не в предпочитаемой зоне на карте Рамачандрана, азот не образует пептидной связи |
| **9** | 296 GLY | ZRSR>5 |
| **10** | 294 ILE | ZRSR >5 |
| **11** | 301 PHE | ZRSR >5 |
| **12** | 247 ASP | Плохой угол ω, высокий В-фактор |

## Детальный анализ 5 маргинальных остатков

### Глицин-245 и аланин-246, маргиналы по карте Рамачандрана

Как показал анализ, глицин-245 и аланин-246 попадают в запрещенную область на карте Рамачандрана (φ =-58.0, ψ = 97.8; φ = -58.1, ψ = -83.3 ). Эти остатки находятся на поверхности белка в участке без периодичной вторичной структуры (вариабельной петле). Так как это два остатка из запрещенной области, идущих подряд, можно предположить инверсию пептидной цепи. На Рисунок 3 изображены два остатка и срез электронной плотности на уровне 1.5σ. Видно, что электронная плотность плохо описывает остаток глицина, поэтому, вероятно, действительно нужно провести инверсию пептидной цепи вокруг пептидной связи на 180°. Помимо нахождения в запрещенной области карты Рамачандрана, остаток глицина-245 также попадает в маргиналы по Zscore=2.75 (больше установленного порога 2).

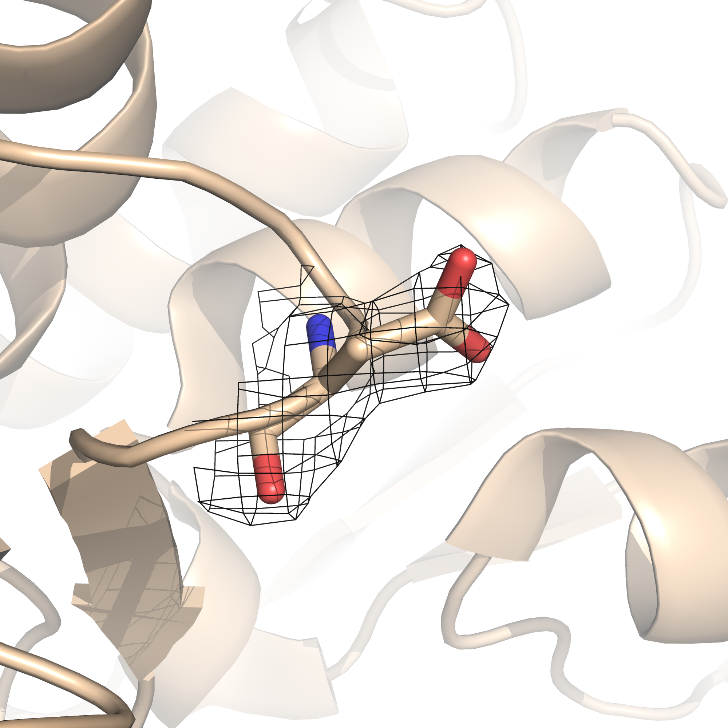


**Рисунок 3 Изображение глицина-245 и аланина-246. Черным показан срез электронной плотности на уровне 1,5 σ**

### Аспартат-247, маргинал по плохому углу ω и большому В-фактору.

В белке spsE неизвестен активный сайт, поэтому опять обратилась к белку NeuB. В статье [[2](#_ENREF_2)] сказано, что есть три высоко-консервативных остатка в активном центре. В белке spsE эти остатки тоже есть, и находятся они в NeuB домене, и если предположить, что белки катализируют похожие реакции, то эти остатки находятся как раз в активном центре. Среди этих остатков – аспартат-247 (Рисунок 4). Он попал в маргиналы из-за плохого угла ω и большого В-фактора. На рисунке видно, что остаток достаточно хорошо описывается электронной плотностью на уровне срезки 2σ (что совсем не обыденность для белка в таком разрешении). Zscore у этого остатка – 0.71, что не выходит за рамки порогов. В белке NeuB этот остаток образует водородную связь с восстановленной N-ацетилнейраминовой кислотой.

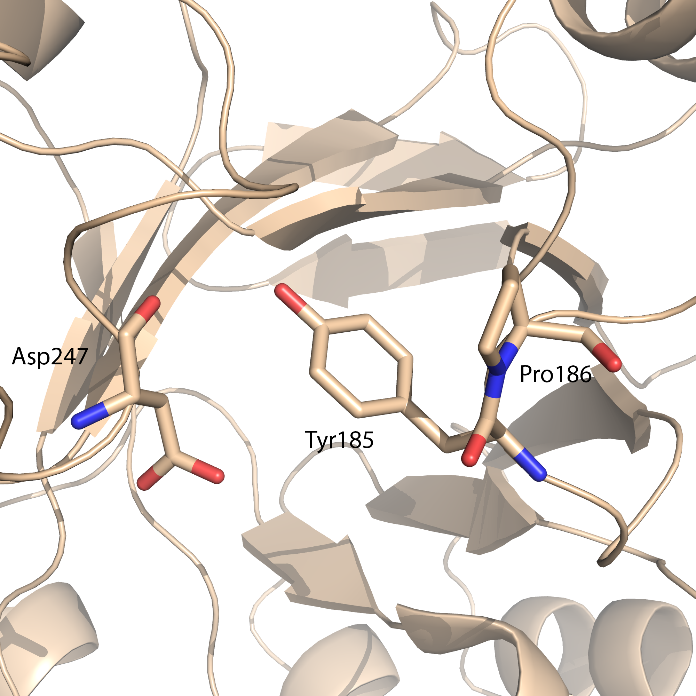
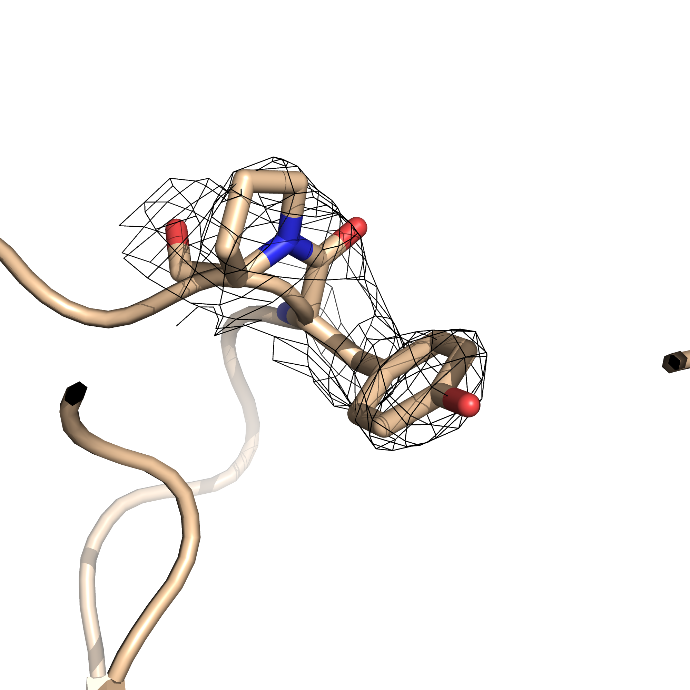
Этот остаток находится совсем рядом с двумя предыдущими рассмотренными маргиналами. И если принять гипотезу о том, что этот остаток необходим для ферментативной активности белка, то тогда и предыдущие маргиналы имеют такие торсионные углы, чтобы обеспечить белку нормальное функционирование. Эти остатки расположены на петле, находящейся над бета-бочонком, присущим всем белкам из суперсемейства альдолаз.



**Рисунок 4 Изображение аспартата-247. Также показан срез электронной плотности с уровнем 2σ**

### Тирозин-185, маргинал по плохому углу ω и странному окружению

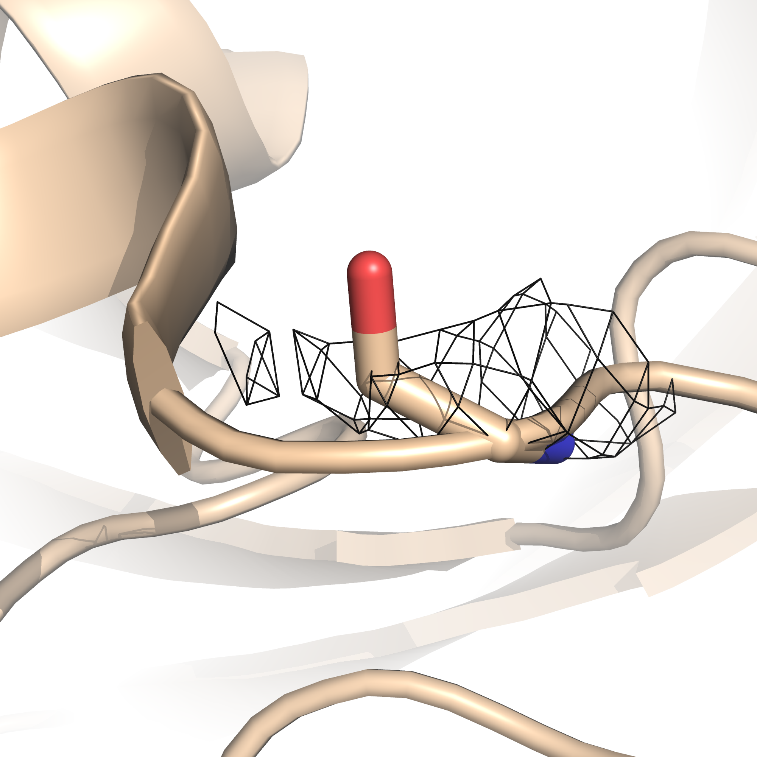
Остаток тирозина-185 также указан в NeuB как каталитически активный. На Рисунок 5 показаны сами остатки вместе с электронной плотностью уровня подрезки 2σ. Видим, что электронная плотность довольно неплохо описывает модель (особенно в контексте других аминокислотных остатков). Также на рисунке изображено взаимное расположение потенциальных каталитических остатков – аспарагина-247 и тирозина-185. В белке NeuB эти остатки координируют восстановленую N-ацетилнейраминовую кислоту с образованием водородных связей [[2](#_ENREF_2)]. Видимо, все эти странности в углах и окружении нужны для успешной координации соединения.



**Рисунок 5 Слева - изображение остатков тирозина-185 и пролина-186. Также черным показан срез электронной плотности уровнем 2σ. Справа – взаимное расположение потенциальных каталитических остатков – Asp247 и Tyr185.**

### Глицин 296, маргинал по Zscore

На Рисунок 6 изображен сам остаток и описывающая его электронная плотность. Видно, что даже на маленьком уровне подрезки ЭП плохо описывает остаток, что говорит о несоответствии модели ЭП. Очень высокий Zscore = 5.995 (что почти в три раза больше условного порога 2) говорит о том, что этот остаток отклоняется от среднего остатка глицина среди белков с таким же разрешением больше, чем на 3 сигмы (сигма равна 1.18).



**Рисунок 6 Изображение глицина-296 и электронной плотности с уровнем подрезки 1σ**

## Сравнение модели из PDB с моделью из PDB\_redo

На сервере PDB REDO [[5](#_ENREF_5)] была построена структура 1vli\_final по имеющимся экспериментальным данным о 1vli. В Таблица 3 перечислены некоторые характеристики для 1vli и 1vli\_final.

Значения R и Rfree ухудшились, и разница их также увеличилась. Значит, новая модель хуже подогнана под экспериментальные данные. При этом говорится, что RSR 17 остатков значительно улучшился, а 3 остатков – значительно ухудшился. RSRZ улучшился у 20 остатков, а ухудшился – у 15 остатков.

Таблица Некоторые характеристики PDB и PDB\_final

|  |  |  |
| --- | --- | --- |
| Характеристика | 1vli | 1vli\_final (PDB REDO) |
| R | 0.1900 | 0.2003 |
| Rfree | 0.2380 | 0.2366 |
| Число маргиналов по длинам связей | 4 | 0 |
| Число маргиналов по значениям валентных углов | 14 | 14 |

На основе имеющихся данных можно сделать вывод, что PDB REDO практически не улучшил структуру 1vli.

# Заключение

Проведенный анализ структуры 1vli выявил несколько маргинальных остатков. Маргинальность некоторых из них можно объяснить потенциальностью образования активного сайта фермента, маргинальность других является ошибкой построения модели. Из-за отсутствия каких-либо статей на счет этой структуры, выводы мои являются субъективными и не подкреплены ничьими другими наблюдениями.

В целом разрешение структуры – среднее, так как много остатков плохо описываются электронной плотностью, слишком много маргинальных остатков по Zscore и другим показателям качества.

# Список литературы

1. Szklarczyk, D., et al., *The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.* Nucleic Acids Res, 2011. **39**(Database issue): p. D561-8.

2. Gunawan, J., et al., *Structural and mechanistic analysis of sialic acid synthase NeuB from Neisseria meningitidis in complex with Mn2+, phosphoenolpyruvate, and N-acetylmannosaminitol.* J Biol Chem, 2005. **280**(5): p. 3555-63.

3. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

4. Kleywegt, G.J., et al., *The Uppsala Electron-Density Server.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2240-9.

5. Joosten, R.P., et al., *PDB\_REDO: constructive validation, more than just looking for errors.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 4): p. 484-96.