

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

**Отчет о качестве расшифровки структуры белка 2пх9
методом рентгеноструктурного анализа**

Работу выполнила студентка 4 курса
Климчук Олеся Игоревна

Москва, 2014

Оглавление

1. Введение.....	3
2. Результаты и обсуждение	4
2.1. Метод расшифровки	4
2.2. Симметрия	4
2.3. Индикаторы качества	4
2.4. Маргинальные остатки	9
2.5. Активный сайт.....	14
2.6. Сравнение с моделью из PDB_REDO	15
3. Заключение.....	16
4. Список литературы.....	17

Аннотация

В данной работе произведена оценка индикаторов качества структуры белка 2пх9, расшифрованной методом рентгеноструктурного анализа, и проанализированы некоторые маргинальные остатки структуры. Произведено сравнение структуры с моделью из PDB_REDO.

1. Введение

Структура 2px9 была расшифрована в 2006 году для карбоксилтрансферазного домена α -субъединицы оксалоацетат декарбоксилазы *Vibrio cholerae* (мономер α -субъединицы оксалоацетат декарбоксилазы состоит из трех функциональных доменов: карбоксилтрансферазного домена и коротких биотин-связывающего и линкерного доменов) [1].

Оксалоацетат декарбоксилаза обнаружена только у анаэробных микроорганизмов и играет ключевую роль при сохранении ими энергии [1]. Данный фермент относится к карбокси-лиазам (ЕС 4.1.1.3) и способен сопрягать биотин-зависимое декарбоксилирование оксалоацетата с транспортом ионов Na^+ из клетки. Собственно биотин-зависимое декарбоксилирование оксалоацетата с освобождением пирувата и углекислоты осуществляет цитоплазматическая α -субъединица. Транспорт ионов Na^+ из клетки за счет выделившейся при декарбоксилировании энергии возможен лишь при наличии в оксалоацетат декарбоксилазном комплексе мембраносвязанной β -субъединицы. Транспорт ионов Na^+ из клетки ведет к увеличению трансмембранного электрохимического градиента Na^+ , используемого для синтеза АТФ Na^+ -зависимой АТФ-синтазой.

До 2006 года оксалоацетат декарбоксилаза изучалась экспериментально на протяжении 25 лет [2]. Посредством электрофоретического анализа было установлено, что α -субъединица функционирует в виде гомодимера. Соответственно, PDB-файл данной структуры содержит информацию об одной биомолекуле, состоящей из двух цепей, имеющих одинаковую аминокислотную последовательность. Посредством масспектрометрического анализа было показано связывание одного катиона Zn^{2+} мономером α -субъединицы. Также экспериментально было показано, что оксалоацетат декарбоксилаза полностью инактивируется специфическим мутагенезом остатков Asp17, His207 и His209, являющихся консервативными лигандами катиона Zn^{2+} (здесь и далее нумерация остатков дана по структуре 2px9). При расшифровке данной структуры авторов интересовала роль в катализе консервативного остатка Lys178. В результате анализа расшифрованной структуры и специфического мутагенеза остатка Lys178 авторы делают вывод о том, что данный остаток очень важен для катализа, однако механизм его участия в катализе по-прежнему не ясен [1].

2. Результаты и обсуждение

2.1. Метод расшифровки

Структура 2nx9 была расшифрована методом молекулярного замещения. В качестве пробной модели была использована структура 1rqh, расшифрованная для 5S-субъединицы транскарбоксилазы *Propionibacterium shermanii* [1, 3]. После нахождения оптимального размещения пробной модели был проведен комбинированный синтез Фурье с использованием рассчитанных по оптимально размещенной модели фаз и с использованием модулей структурных факторов, рассчитанных из измеренных в рентгеновском эксперименте интенсивностей рефлексов.

2.2. Симметрия

Структура 2nx9 имеет примитивную кристаллическую решетку и относится к пространственной группе $2_12_12_1$, т.е. существует движение, переводящее все атомы кристалла в себя при повороте на 180° и вращении на $\frac{1}{2}$ периода. Биологическая единица совпадает с асимметрической единицей. Некристаллографических симметрий асимметрической единицы не наблюдается [4].

2.3. Индикаторы качества

2.3.1. R-value и R-free

В рентгеновском эксперименте было измерено 107243 рефлекса [5], в том числе 100388 рефлексов (working set) были использованы для расшифровки структуры и 5272 контрольных рефлекса (test set) были использованы для расчета R-free [4]. Значения R-free и R-value (working set) равны 28.4% и 23.2% соответственно [4]. Значение R-value (working + test set) равно 23.5% [4]. Полученные значения R-value (working set) и R-free отличаются друг от друга незначительно (на 5.2%), что позволяет исключить вероятность переподгонки структуры, однако данные значения R-value (working set) и R-free свидетельствуют о том, что расшифровка структуры находится на грани среднего и низкого качества.

Среди всех измеренных в рентгеновском эксперименте рефлексов 91692 рефлекса (85.5%) имели интенсивность, в три раза превышающую стандартное отклонение.

2.3.2. Разрешение

Минимальное разрешение структуры 2nx9 равно 71.982 Å [5] или 35.000 Å [4]. Максимальное разрешение структуры 2nx9 равно 1.700 Å [4, 5], т.е. на восстановленной функции электронной плотности не различимы детали меньшие, чем 0.850 Å. Полнота набора рефлексов, использованных для расшифровки структуры, равна 98,5% [4].

2.3.3. Сводная статистика сервиса MolProbity

С помощью сервиса MolProbity к структуре 2nx9 были добавлены атомы водорода, затем были проанализированы контакты атомов и геометрия структуры. Сводная статистика анализа структуры приведена в таблице 1.

Таким образом, Clashscore структуры 2nx9 равен 5.82 на 1000 атомов структуры, т.е. лишь 0.582% атомов структуры имеют недопустимые наложения (т.е. их Ван-дер-Ваальсовы радиусы перекрываются больше, чем на 0.4 Å). Структура 2nx9 имеет 92^{ую} перцентиль, т.е 92% структур подобного разрешения имеют больший Clashscore, из чего следует, что относительно мало атомов структуры 2nx9 имеют недопустимые наложения. Структура 2nx9 содержит 2.93% остатков, конформация боковых цепей которых существенно отличается от ротамеров (в идеале < 1% остатков). Структура 2nx9 содержит 0.44% остатков, находящихся в запрещенной области на карте Рамачандрана (в идеале < 0.05% остатков) и 96.01% остатков, находящихся в предпочитаемой области на карте Рамачандрана (в идеале > 98% остатков).

Таблица 1. Сводная статистика анализа структуры 2nx9 с помощью сервиса MolProbity (добавлены атомы водорода).

All-Atom Contacts	Clashscore, all atoms:	5.82	92 nd percentile* (N=819, 1.70Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	22	2.93% Goal: <1%
	Ramachandran outliers	4	0.44% Goal: <0.05%
	Ramachandran favored	866	96.01% Goal: >98%
	MolProbity score [^]	1.95	60 th percentile* (N=9248, 1.70Å ± 0.25Å)
	Cβ deviations >0.25Å	4	0.47% Goal: 0
	Bad backbone bonds:	5 / 7114	0.07% Goal: 0%
	Bad backbone angles:	13 / 9654	0.13% Goal: <0.1%

2.3.4. Карта Рамачандрана

С помощью сервиса MolProbity была построена карта Рамачандрана структуры 2nx9 (предварительно к структуре были добавлены атомы водорода). Построенная карта Рамачандрана приведена на рисунке 1.

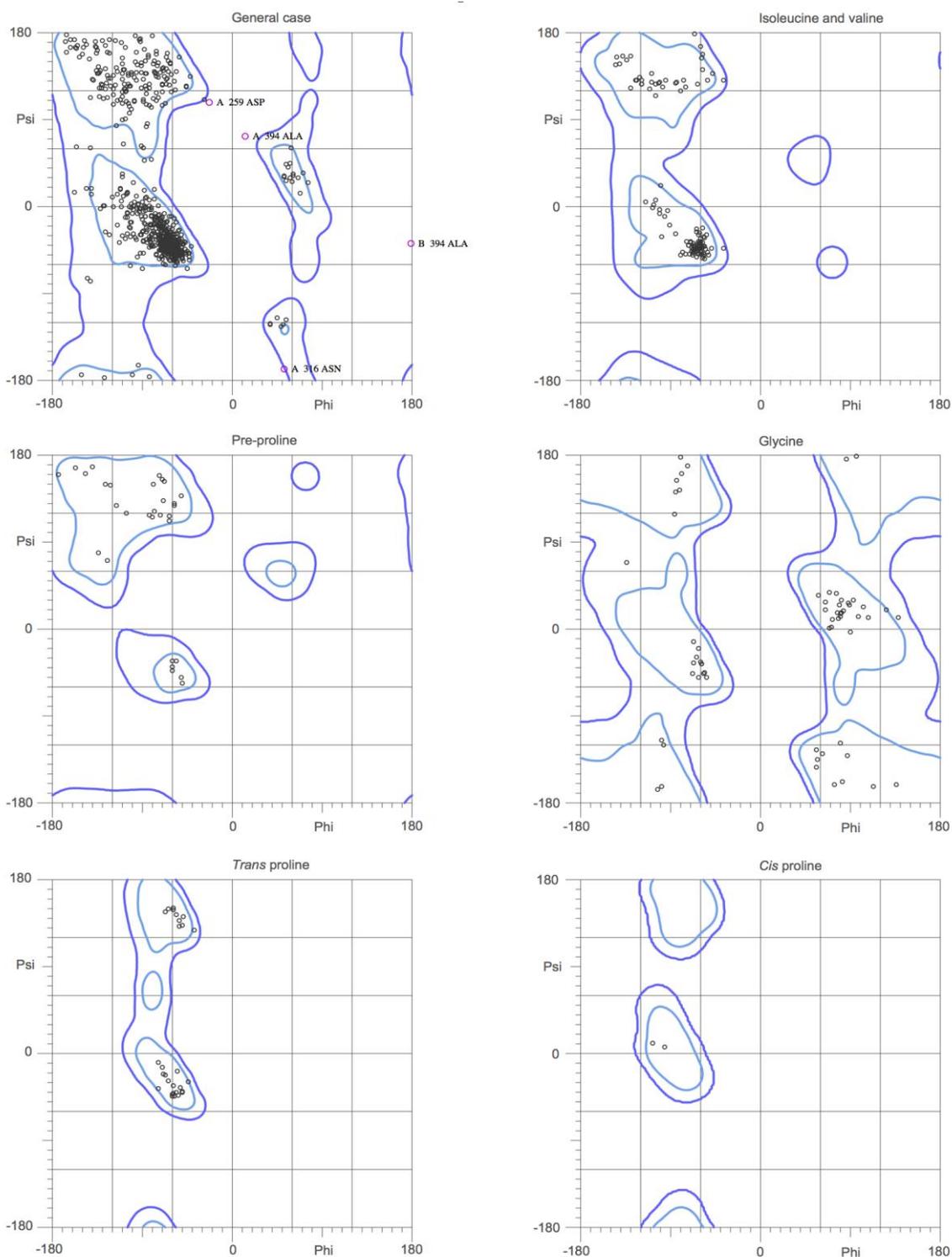


Рисунок 1. Карта Рамачандрана структуры 2nx9 (добавлены атомы водорода).

Структура 2пх9 содержит 4 остатка, находящихся в запрещенной области на карте Рамачандрана (Asp259A, Asn316A, Ala394A, Ala394B). Вероятно, эти остатки являются маргинальными, поскольку имеют недопустимые значения торсионных углов ϕ и ψ . Остальные 99.56% остатков находятся в допустимой области, в том числе 96.01% остатков находятся в предпочитаемой области.

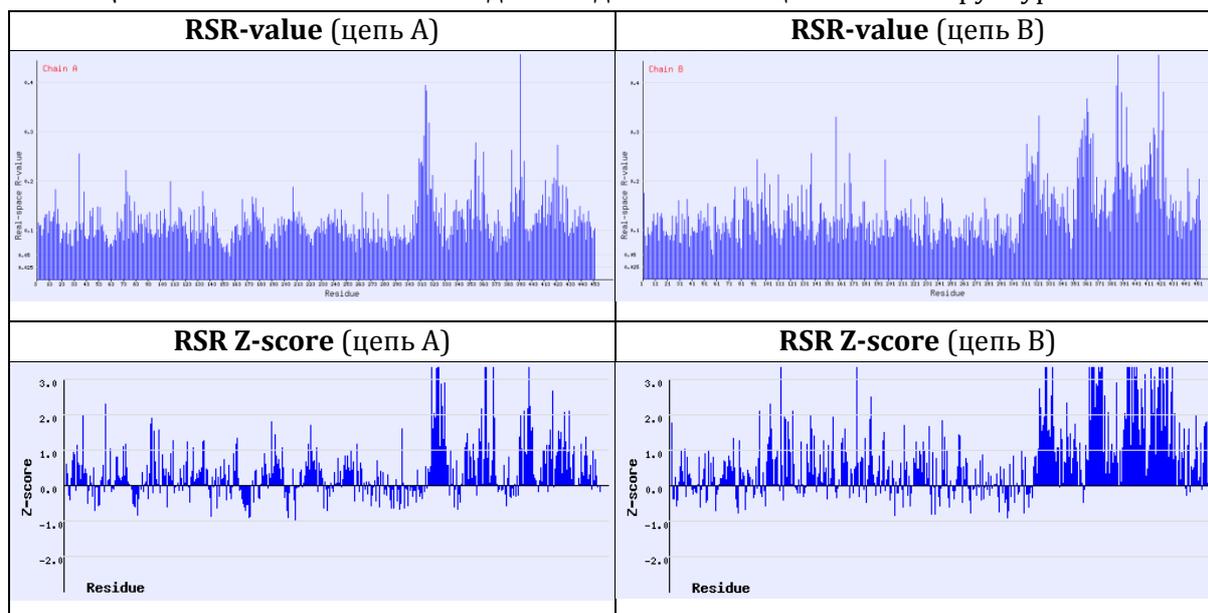
2.3.5. RSR-value и RSR Z-score

С помощью сервера EDS в структуре 2пх9 было проанализировано качество восстановления электронной плотности. В таблице 2 приведены RSR-value и RSR Z-score для каждого остатка цепей А и В структуры 2пх9.

RSR-value позволяет оценить соответствие восстановленной по модели электронной плотности и «экспериментальной» электронной плотности. Чем больше RSR-value, тем хуже расшифрованы координаты атомов в модели. RSR Z-score позволяет сравнить RSR-value остатка со средним RSR-value остатков того же типа в структурах подобного разрешения. Чем больше RSR Z-score, тем хуже расшифрованы координаты атомов в модели по сравнению с координатами атомов остатков того же типа в структурах подобного разрешения.

По-видимому, цепи А и В структуры 2пх9 на N-конце имеют качество восстановления лучше, чем на С-конце. Возможно, это связано с тем, что N-конец каждой из двух цепей имеет архитектуру ТИМ-барреля, в то время как С-конец каждой из двух цепей представляет собой боковую цепь (Рисунок 2).

Таблица 2. RSR-value и RSR Z-score для каждого остатка цепей А и В структуры 2пх9.



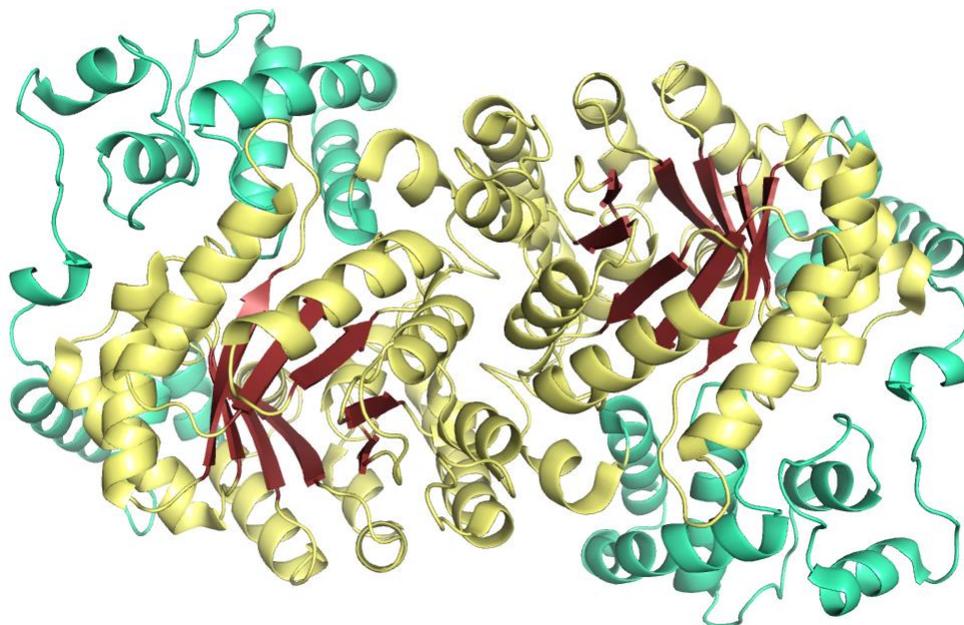


Рисунок 2. Структура 2nx9. N-конец каждой из двух цепей имеет архитектуру ТИМ-барреля (показан желтым и коричневым цветами, ~1-300 остатки). С-конец каждой из двух цепей представляет собой боковую цепь (показан циановым цветом, ~301-454 остатки).

Сводная информация об RSR Z-score для цепей А и В структуры 2nx9 представлена в таблице 3. Таким образом, цепь А содержит 4.66% остатков, которые имеют RSR Z-score > 2, в то время как цепь В содержит 11.92% остатков, которые имеют RSR Z-score > 2. Несмотря на то, что цепи А и В идентичны по аминокислотной последовательности (каждая цепь длиной 454 остатка), они входят в состав одной биомолекулы, и авторы проделали двойную работу, оптимизируя данную структуру.

Таблица 3. Сводная информация об RSR Z-score для цепей А и В структуры 2nx9.

Chain	Nres	<Z>	Sigma	%<-1	%<0	%>1	%>2
A	451	0.41	1.11	0.22	35.25	16.41	4.66
B	453	0.79	1.48	0.00	27.81	29.58	11.92
All	904	0.60	1.31	0.11	31.53	23.01	8.30

2.4. Маргинальные остатки

При поиске маргинальных остатков в структуре 2nx9 была проанализирована выдача сервиса MolProbity, выдача сервера EDS, а также протокол WHAT_CHECK. Некоторые найденные маргинальные остатки приведены в таблице 4, с указанием критерия маргинальности.

В первую очередь в таблицу 4 были внесены остатки, находящиеся в запрещенной области на карте Рамачандрана (всего 4 остатка по данным сервиса MolProbity). Затем в таблицу 4 были внесены остатки, имеющие cis-конформацию (в результате визуальной проверки всех остатков, имеющих маленькое значение торсионного угла ω по протоколу WHAT_CHECK, было найдено 5 таких остатков) и остатки, имеющие наибольшее RSR-value (всего 3 остатка по данным сервера EDS). Также в таблицу 4 были внесены остатки, нуждающиеся в инверсии боковой цепи и остатки с неблагоприятной конформацией боковой цепи (всего 3 остатка, хотя по данным сервиса MolProbity структура 2nx9 содержит 10 остатков, для которых имеются свидетельства в пользу инверсии боковой цепи: His19A, Gln130A, Gln134A, Asn307A, Gln311A, Gln425A, His19B, Gln166B, Gln314B, Gln346B). В последнюю очередь в таблицу 4 были внесены остатки, имеющие наименьший RMS Z-score, т.е. имеющие наиболее неблагоприятное окружение (всего 2 остатка, хотя по данным протокола WHAT_CHECK структура 2nx9 содержит 21 остаток, имеющих RMS Z-score < -5). Некоторые остатки являются маргинальными по нескольким критериям одновременно.

Большая часть маргинальных остатков принадлежит С-концу каждой из двух цепей структуры 2nx9.

Таблица 4. Некоторые маргинальные остатки в структуре 2nx9.

Маргинал	Критерий
Asp259A	Находится в запрещенной области на карте Рамачандрана
Asn316A	Находится в запрещенной области на карте Рамачандрана, RSR-value = 29.1%, RSR Z-score = 3.302
Ala394A (Рисунок 3)	Находится в запрещенной области на карте Рамачандрана, RSR-value = 67.1%, RSR Z-score = 12.299, Cis-конформация
Ala394B	Находится в запрещенной области на карте Рамачандрана, RSR-value = 35.0%, RSR Z-score = 5.541, Cis-конформация

Gln315B (Рисунок 4)	Cis-конформация
Asn316B (Рисунок 4)	Cis-конформация
Thr386B (Рисунок 5)	Cis-конформация, RSR-value = 39.4%, RSR Z-score = 6.305
Glu387B (Рисунок 5)	RSR-value = 46.2%, RSR Z-score = 4.808
Gln420B (Рисунок 6)	RSR-value = 55.5%, RSR Z-score = 7.162
His19A	Явно нуждается в инверсии боковой цепи
Gln311A	Явно нуждается в инверсии боковой цепи
Gln425A	Имеет неблагоприятную конформацию боковой цепи, возможно, нуждается в инверсии боковой цепи
Tyr94B (Рисунок 7)	Имеет неблагоприятное окружение (RMS Z-score = -7.69), RSR-value = 24.4%, RSR Z-score = 4.003
Tyr94A	Имеет неблагоприятное окружение (RMS Z-score = -7.52)

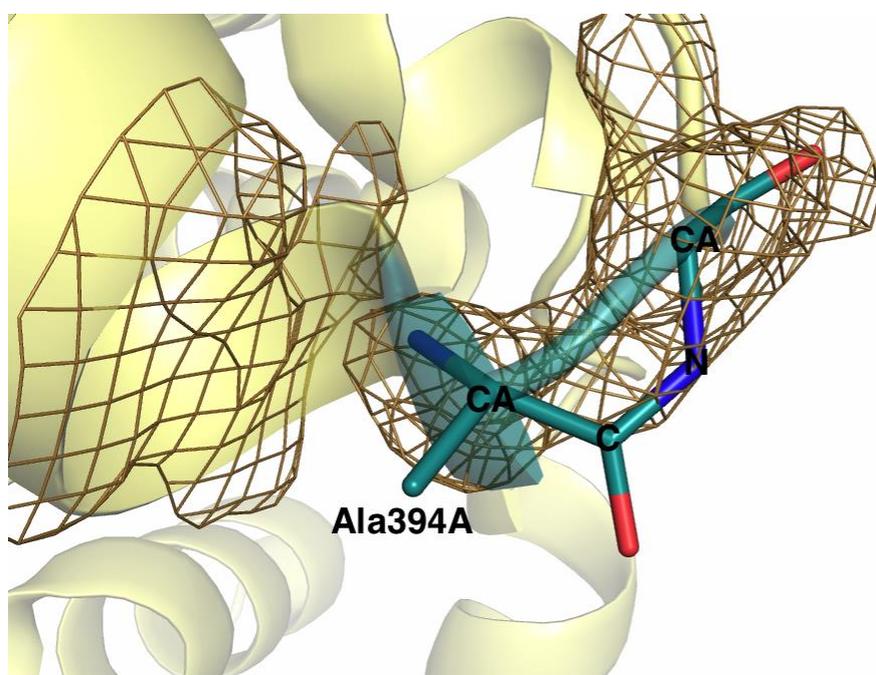


Рисунок 3. Структура 2px9, цепь А, остаток Ala394. Данный остаток является маргинальным по критериям, описанным в таблице 4. Подрезка «экспериментальной» электронной плотности данного остатка на уровне 1 показана коричневой сеткой.

Электронная плотность остатка Ala394A (Рисунок 3), рассчитанная по модели, плохо согласуется с «экспериментальной» электронной плотностью, из-за чего при расшифровке координат атомов в структуре 2nx9 данный остаток оказался в *cis*-конформации. Кроме того, данный остаток имеет недопустимые значения торсионных углов φ и ψ и находится в запрещенной области на карте Рамачандрана. Таким образом, маргинальность данного остатка связана с ошибочной расшифровкой координат его атомов.

Остатки Gln315B и Asn316B (Рисунок 4) имеют *cis*-конформацию, однако других признаков маргинальности для них не наблюдается. Возможно, такая странная конформация данных остатков связана с изгибом полипептидной цепи.

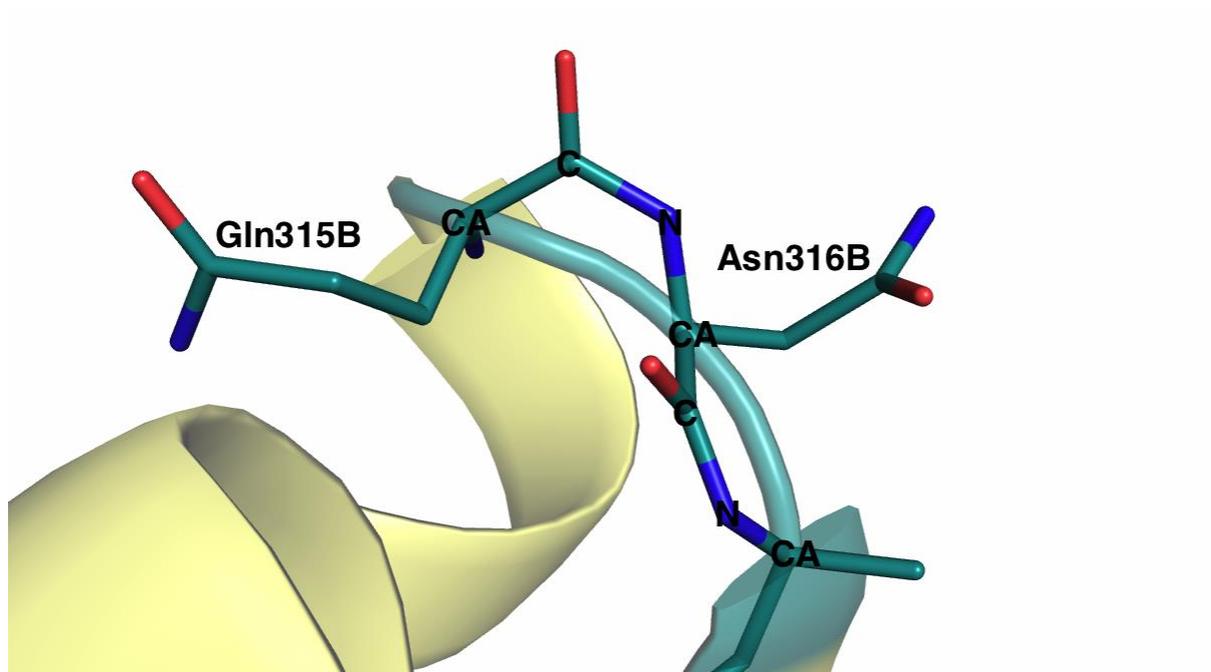


Рисунок 4. Структура 2nx9, цепь B, остатки Gln315 и Asn316. Данные остатки являются маргинальными, т.к. имеют *cis*-конформацию.

Электронная плотность остатков Thr386B и Glu387B (Рисунок 5), рассчитанная по модели, плохо согласуется с «экспериментальной» электронной плотностью. Даже при подрезке «экспериментальной» электронной плотности на уровне 1, боковые цепи данных остатков лишены электронной плотности. Таким образом, маргинальность данных остатков связана с ошибочной расшифровкой координат их атомов. Возможно, *cis*-конформация остатка Thr386B связана с изгибом полипептидной цепи.

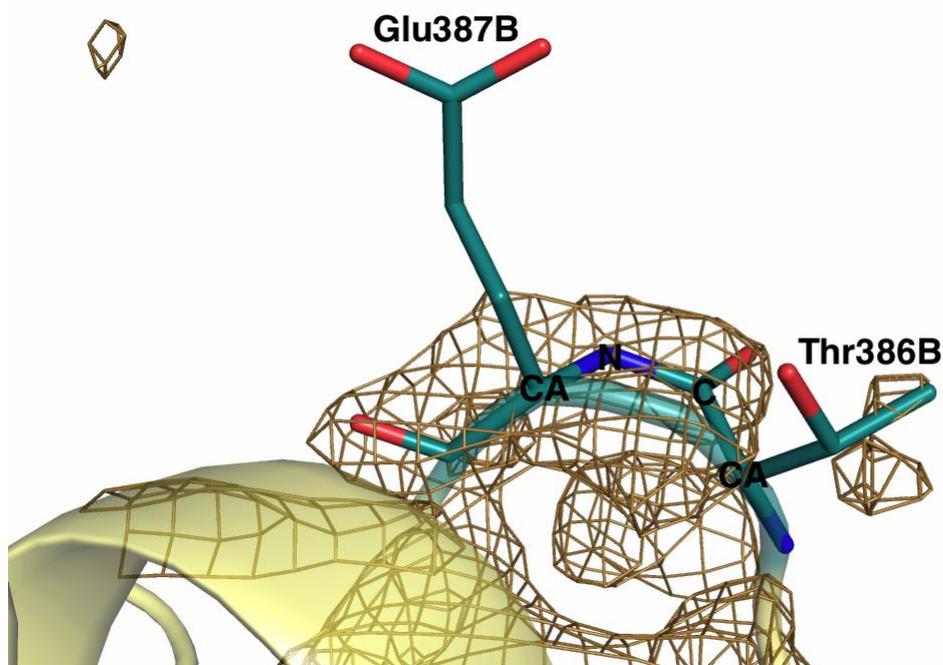


Рисунок 5. Структура 2nx9, цепь В, остатки Thr386 и Glu387. Данные остатки являются маргинальными по критериям, описанным в таблице 4. Подрезка «экспериментальной» электронной плотности данных остатков на уровне 1 показана коричневой сеткой.

Электронная плотность остатка Gln420В (Рисунок 6), рассчитанная по модели, плохо согласуется с «экспериментальной» электронной плотностью. Даже при подрезке «экспериментальной» электронной плотности на уровне 1, данный остаток полностью лишен электронной плотности. Таким образом, маргинальность данного остатка связана с ошибочной расшифровкой координат его атомов.

Маргинальный остаток Tyr94В имеет недостаточно хорошее качество расшифровки атомов, поскольку недостаточно хорошо вписан в «экспериментальную» электронную плотность. Возможно, поэтому данный остаток имеет неблагоприятное окружение – ему некому донировать протон гидроксильной группы. На рисунке 7 видно, что протон гидроксильной группы данного остатка повернут вправо, в то время как слева имеется «свободная» электронная плотность. Возможно, если бы протон был повернут влево, он «вписался» бы в электронную плотность и смог бы образовать водородную связь с кислородом остатка Ala441В.

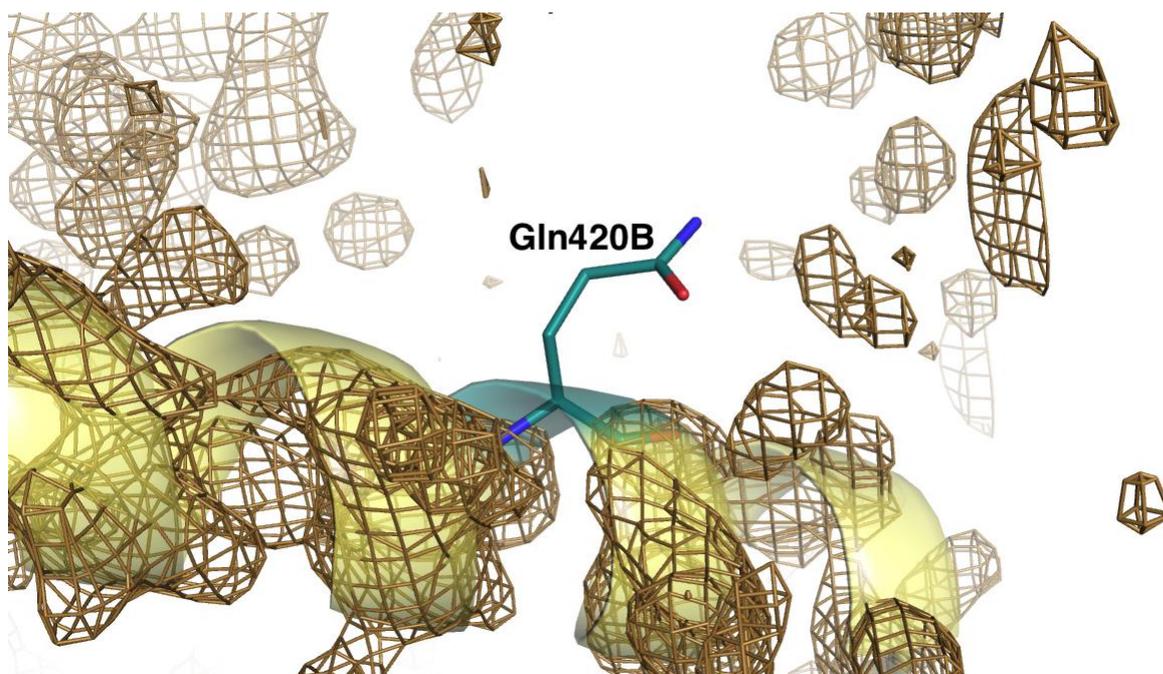


Рисунок 6. Структура 2px9, цепь В, остаток Gln420. Данный остаток является маргинальным, т.к. полностью лишен электронной плотности. Подрезка «экспериментальной» электронной плотности цепи В на уровне 1 показана коричневой сеткой.

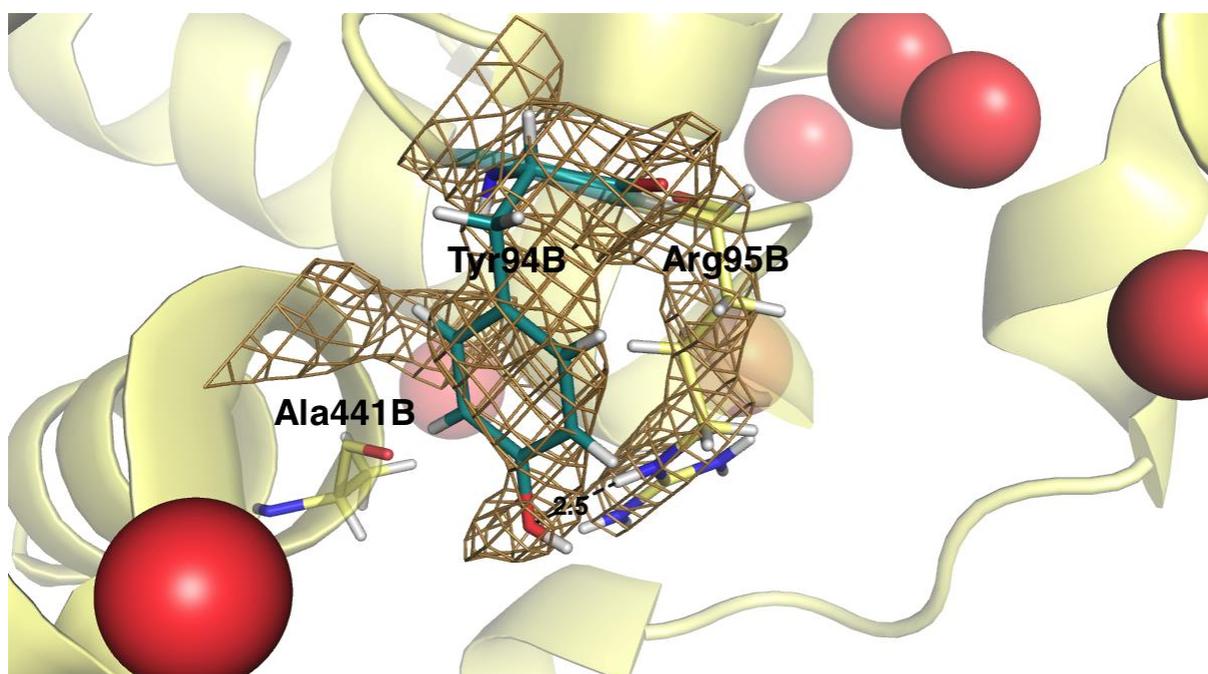


Рисунок 7. Структура 2px9, цепь В, остатки Tyr94, Arg95 и Ala441. Остаток Tyr94 является маргинальным по критериям, описанным в таблице 4. Подрезка «экспериментальной» электронной плотности данного остатка на уровне 1 показана коричневой сеткой. Атомы кислорода молекул воды показаны в виде красных шариков.

2.5. Активный сайт

В статье о расшифровке структуры 2px9 описана роль в катализе аминокислотных остатков Asp17, His207 и His209, координирующих катион Zn^{2+} в активном сайте; обсуждается роль в катализе остатка Lys178 [1].

Для изучения роли в катализе консервативного остатка Lys178 авторы провели наложение структуры 2px9, расшифрованной для карбоксилтрансферазного домена α -субъединицы оксалоацетат декарбоксилазы *Vibrio cholerae* и структуры 1rqh, расшифрованной для 5S-субъединицы транскарбоксилазы *Propionibacterium shermanii*. В результате наложения структур произошло совмещение остатка Lys178 структуры 2px9 и карбоксилированного остатка Lys180 структуры 1rqh. Поскольку остаток Lys180 структуры 1rqh карбоксилирован в отсутствие субстрата и декарбоксилирован в присутствии субстрата, авторы предполагают, что данный остаток облегчает перенос карбоксила с субстрата на биотин. Остаток Lys178 структуры 2px9, декарбоксилирован и образует водородную связь с остатком Cys148, по крайней мере в отсутствие субстрата (Рисунок 8). Специфический мутагенез остатка Lys178 полностью инактивирует оксалоацетат декарбоксилазу, в то время как специфический мутагенез остатка Cys148 лишь ослабляет активность оксалоацетат декарбоксилазы, что говорит о ключевой роли остатка Lys178 в катализе и о стабилизирующей роли остатка Cys148 в катализе. Однако механизм участия остатка Lys178 в катализе по-прежнему не ясен [1].

Некоторые каталитические остатки активного сайта имеют небольшие отклонения длин ковалентных связей и валентных углов от допустимых значений, однако маловероятно, что данные остатки являются маргинальными. Следует отметить, что остаток Lys178 имеет меньшую электронную плотность по сравнению с остальными каталитическими остатками активного сайта, несмотря на то что остаток Lys178 достаточно хорошо «вписан» в электронную плотность (Рисунок 8).

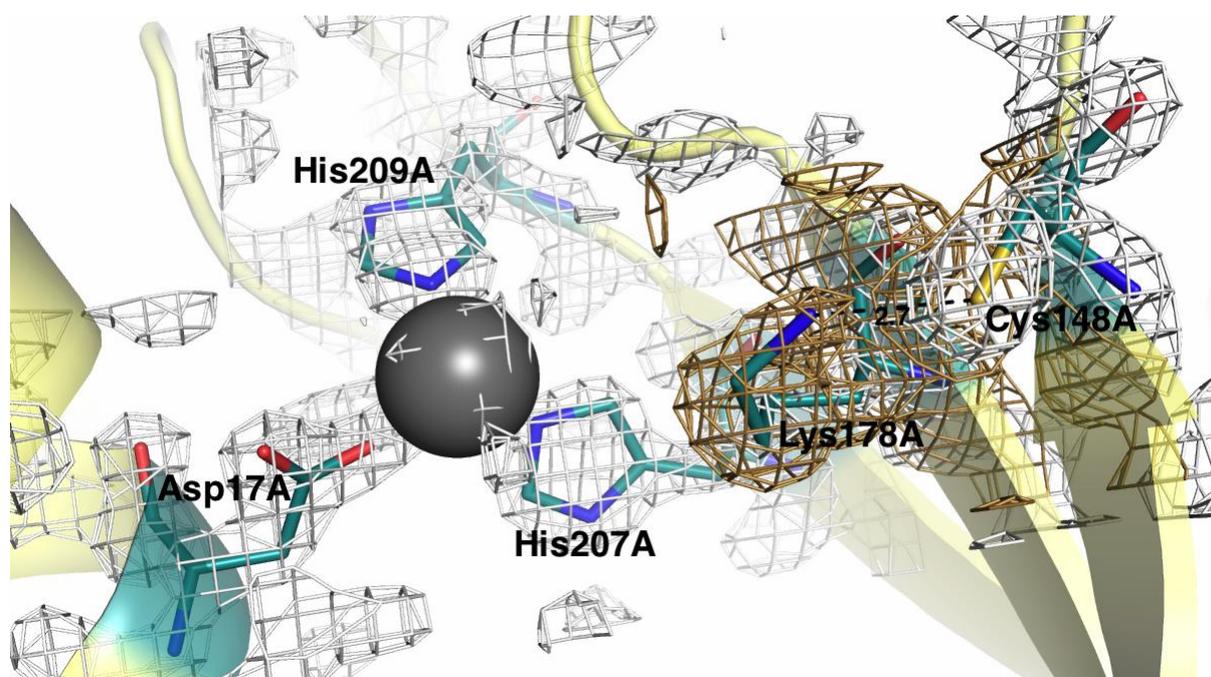


Рисунок 8. Структура 2nx9, цепь А, остатки Asp17, His207, His209, Lys178 и Cys148. Подрезка «экспериментальной» электронной плотности данных остатков на уровне 3 показана белой сеткой. Подрезка «экспериментальной» электронной плотности остатка Lys178 на уровне 1 показана коричневой сеткой. Катион Zn^{2+} показан в виде серого шарика.

2.6. Сравнение с моделью из PDB_REDO

В базе данных PDB_REDO была найдена стандартно оптимизированная по экспериментальным данным структура 2nx9. В результате «оптимизации» R-value (working set) и R-free структуры увеличились на 1.24% и 0.2% соответственно, что говорит если не об ухудшении модели, то, по крайней мере, об отсутствии значительной оптимизации. RSR-value для 103 остатков улучшилось, зато для 27 остатков RSR-value ухудшилось. Судя по протоколу WHAT_CHECK, были несколько улучшены некоторые параметры структуры, однако число недопустимых наложений атомов увеличилось почти в 2 раза и число доноров/акцепторов водородной связи, не образующих ее, увеличилось почти на 70%. Таким образом, проведенная оптимизация структуры 2nx9 улучшила лишь часть параметров модели, другие параметры практически не изменились или даже ухудшились. Вероятно, такой результат оптимизации связан с качеством экспериментальных данных.

3. Заключение

Структура 2nx9 имеет высокое разрешение, однако R-value свидетельствует о том, что расшифровка данной структуры находится на грани среднего и низкого качества. Большая часть маргинальных остатков данной структуры принадлежит С-концу каждой из двух цепей.

Расшифровка структуры карбоксилтрансферазного домена α -субъединицы оксалоацетат декарбоксилазы *Vibrio cholerae* позволила уточнить роль в катализе консервативного остатка Lys178 и позволила предположить различный каталитический механизм оксалоацетат декарбоксилазы и транскарбоксилазы, несмотря на то, что данные ферменты имеют одинаковую субстратную специфичность, имеют идентичные каталитические остатки активного сайта и являются биотин-зависимыми ферментами.

4. Список литературы

- [1] R. Studer, P. Dahinden, W.W. Wang, Y. Auchli, X.D. Li, P. Dimroth, Crystal Structure of the Carboxyltransferase Domain of the Oxaloacetate Decarboxylase Na⁺ Pump from *Vibrio cholerae*, *Journal of molecular biology* 367 (2007) 547-557.
- [2] P. Dimroth, Characterization of a Membrane - Bound Biotin - Containing Enzyme: Oxaloacetate Decarboxylase from *Klebsiella aerogenes*, *European Journal of Biochemistry* 115 (1981) 353-358.
- [3] P.R. Hall et al., Transcarboxylase 5S structures: assembly and catalytic mechanism of a multienzyme complex subunit, *The EMBO journal* 23 (2004) 3621-3631.
- [4] 2NX9.pdb – PDB-файл структуры 2nx9.
- [5] 2nx9-sf.cif – файл структурных факторов структуры 2nx9.