

Отчет о качестве расшифровки структуры
рибонуклеазы Z (1Y44) *Bacillus Subtilis*
методом рентгеноструктурного анализа

Студентки 4 курса

Факультета биоинженерии и биоинформатики

МГУ им. М.В.Ломоносова

Федоровой А.Д.

Москва, 2015

СОДЕРЖАНИЕ

Аннотация.....	3
Введение	3
Результаты и обсуждения	5
Общая информация о модели. Значения индикаторов качества модели в целом.	5
Анализ маргинальных остатков	7
Сравнение модели из PDB с моделью из PDB_redo.....	13
Заключение	14
Список литературы.....	15

АННОТАЦИЯ

Отчет содержит информацию о модели структуры 1Y44 и ее качественные характеристики. Была проведена работа с серверами, с помощью которых можно оценить качество модели.

ВВЕДЕНИЕ

Известно, что транспортные РНК синтезируются в виде первичного транскрипта, который затем должен процессироваться. Эндонуклеолитический путь процессинга с 3'-конца выполняет фермент рибонуклеаза Z (РНКаза Z). Она разрезает прекурсор тРНК сразу после добавления неспаренного нуклеотида к последней паре оснований акцепторного стебля на 3'-конец (он используется как детерминанта для многих аминоксил-тРНК синтетаз). Это способствует добавлению мотива ССА к тРНК с помощью нуклеотидил-трансферазы.

РНКаза Z (ЕС 3.1.26.11) широко представлена в эукариотах, археях и почти в половине бактерий с известными секвенированными геномами^[1]. Она является необходимой для роста, мутации в одном из двух генов, кодирующих РНКазу (*ELAC2*) приводят к раку простаты^[2].

РНКаза Z принадлежит к β -лактамазному семейству Zn-зависимых металло-гидролаз. Кристаллическая структура была разрешена для селенометионин-помеченной РНКазы Z путем MAD (multiple anomalous dispersion - множественная аномальная дисперсия). Был обнаружен один гомодимер ассиметрической единицы рибонуклеазы. Основной домен каждой субъединицы состоит из двух $\beta\beta\beta\alpha\beta\alpha\beta$ -мотивов (рис.1). Основные водородные взаимодействия - Ala 41, Gln 43, His 44, Ser 103 and Thr 107 из одного мономера и Asn 18, Glu 40 and Tyr 70 из другого.

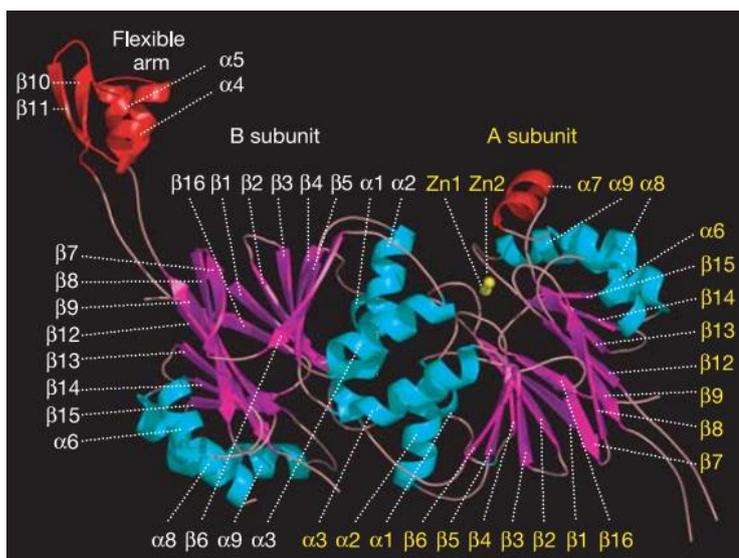


Рис.1. Структура РНКазы Z с разрешением 2.1Å (ribbon), α -спирали показаны бирюзовым, β -листы – фиолетовым.

Был также установлен возможный механизм. Два атома цинка координируются с помощью пяти молекул гистидина, двух остатков аспартата, воды и фосфат-иона.

Asp67 действует как основной участник в удалении протона с молекулы воды, позволяя гидроксид-иону совершить нуклеофильную атаку на фосфат-группу, поляризованную двумя Zn^{2+} -ионами. В результате этого, происходит разрезание фосфодиэфирной связи между 73 и 74 основаниями прекурсора тРНК. 3'-концевой кислород 73 основания стабилизируется за счет протона из рядом находящейся воды, превращаясь в 3'-ОН группу (рис.2) [1].

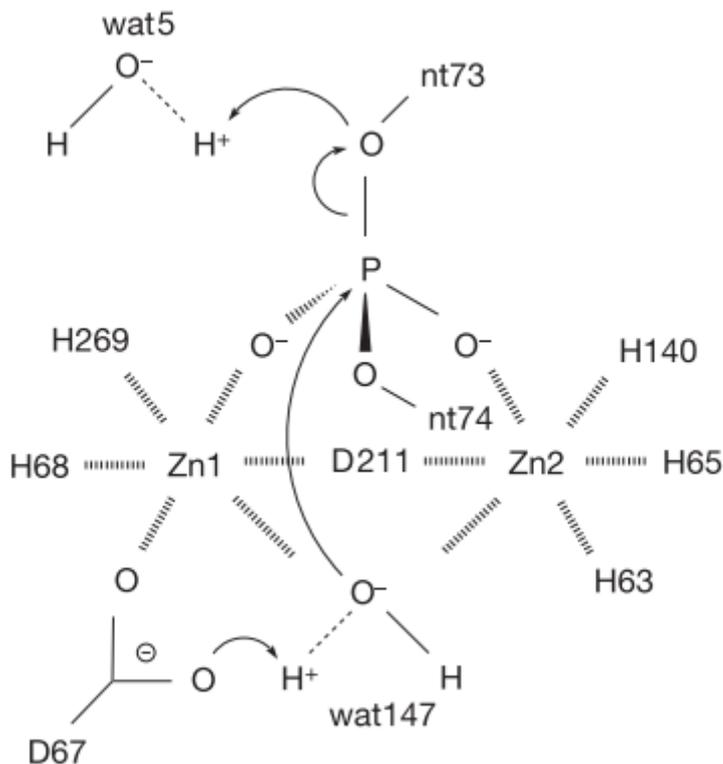


Рис.2. Предполагаемый механизм работы РНКазы Z (описание в тексте).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

1. Общая информация о модели. Значения индикаторов качества модели в целом.

Белок представляет собой гомодимер из двух цепей А и В длиной в 307 аминокислотных остатков и массой 34023 Да (этот же состав комплекса верен и для биологической единицы). В составе также есть такие лиганды, как глицерин, 2-(N-морфолино)-этансульфоновая кислота (MES), фосфат-ион и 2 иона цинка.

Модель структуры 1Y44 получена de la Sierra-Gallay, I.L., Pellegrini, O., Condon, C. в 2005 году. Методом разрешения фазовой проблемы является мультиволновая аномальная дисперсия (MAD) с использованием селенометионина. Некоторая информация о параметрах модели представлена в таблице 1.

Параметр	Значение
Число измеренных рефлексов	85742
Разрешение	2.06 Å
Полнота набора структурных факторов	100%
Диапазон разрешений структурных факторов	34.78 - 2.06 Å
Параметры кристаллографической ячейки	a=42.98 Å, b=188.33 Å, c=177.91 Å alpha=90.00, beta=90.00, gamma=90.00
Тип кристаллографической ячейки	C 2 2 2 1

Табл.1. Некоторые параметры модели.

Значения индикаторов качества модели в целом приведено в таблице 2. R-фактор и R_free вычисляются по оптимизированной модели и измеренным структурным факторам. Они характеризуют соответствие модели экспериментальным данным, т.е. структурным факторам. Соответственно, чем R-фактор меньше, тем модель лучше. В нашем случае, R-фактор = 21,5% (что меньше 25%), это хорошее значение.

R-фактор	0.215
R_free	0.248
R_free - R	0,025

Табл.2. Значения индикаторов качества модели.

Все рефлексы подразделяются на рабочие (служащие для оптимизации модели) и контрольные. Значение параметра R_free вычисляется только по контрольным рефлексам и только по окончательной модели. Соответственно, если модель

правильная, то R_{free} должен оказаться примерно равным R . В нашем случае, значение $R_{free}-R = 3,3\%$, что говорит нам о хорошем качестве кристаллографической модели 1Y44.

Рассмотрим карту Рамачандрана (рис.3). Она показывает положение пар торсионных углов ϕ/ψ для полипептидной цепи. По ней можно определить, лежат ли остатки в допустимой области, а также найти маргинальные остатки по конфигурации остова.

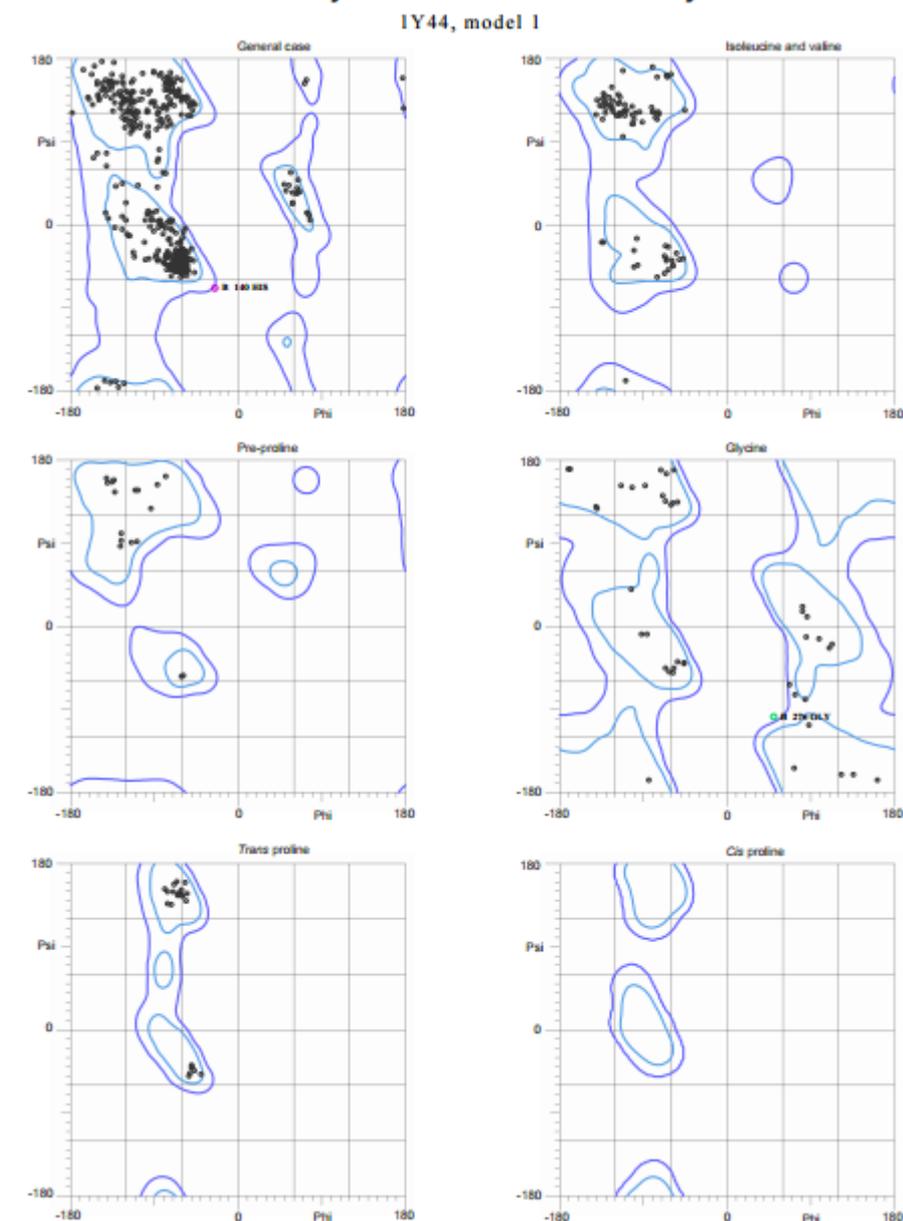


Рис.3. Карта Рамачандрана, полученная с помощью сервиса MolProbity. Слева-направо, сверху-вниз: для всех остатков, кроме перечисленных далее; для изолейцина и валина; для остатков перед пролином; глицина; транс-пролина; цис-пролина.

Согласно рис.3., большинство аминокислотных остатков лежат в допустимой области. Всего 95% всех остатков лежат в предпочитаемой области; 99,6% остатков в разрешенной области. Таким образом, модель можно считать хорошей по этому параметру.

2. Анализ маргинальных остатков

С помощью сервера EDS была построена карта Рамачандрана (рис.4).

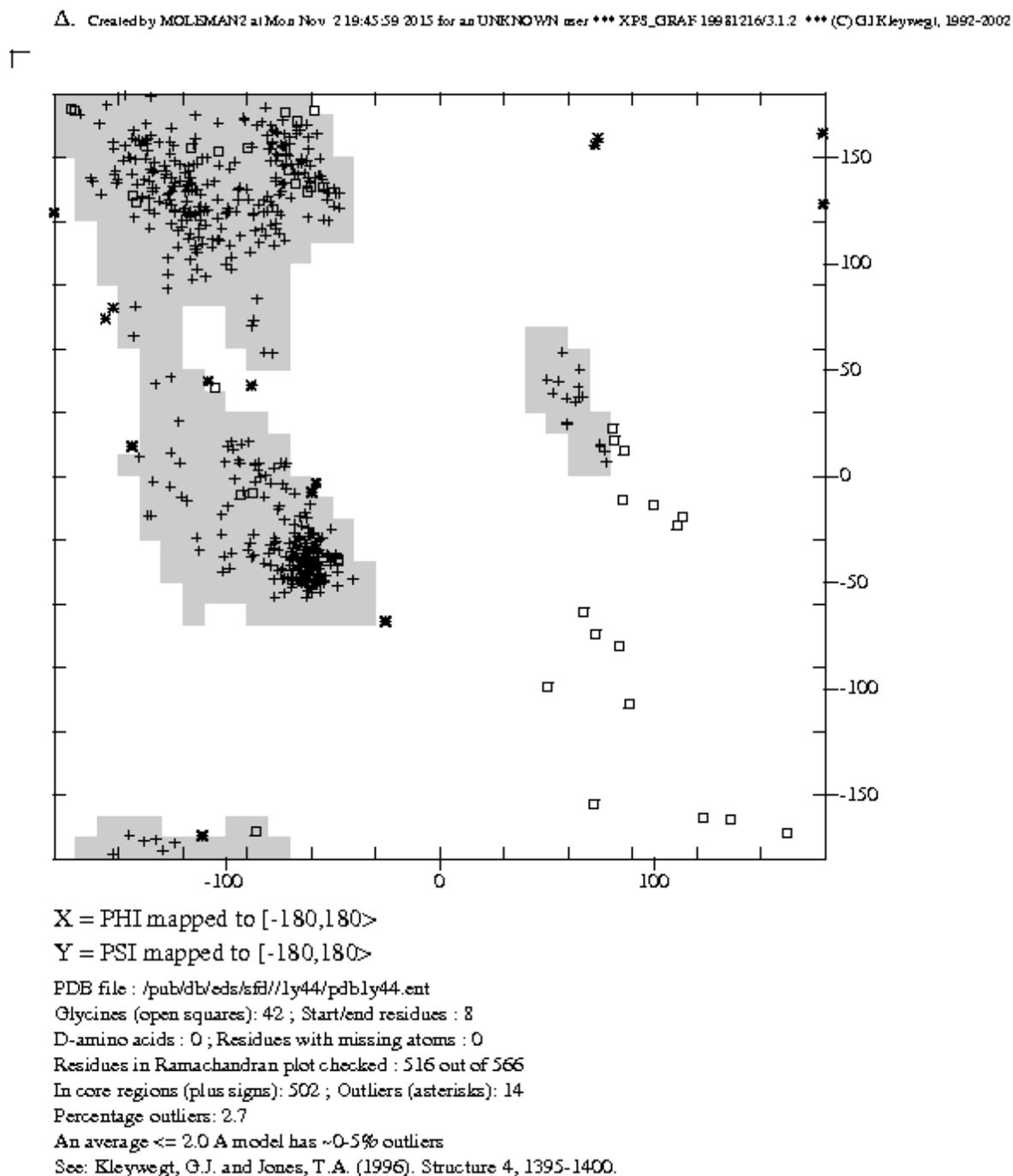


Рис.4. Карта Рамачандрана, построенная с помощью сервера EDS.

Хотя на карте Рамачандрана, полученной с помощью сервера MolProbity, маргинальных остатков выявлено не было, на карте Рамачандрана, построенной с помощью сервера EDS были выявлены следующие маргинальные остатки (табл.3). Здесь можно предположить два варианта: либо координаты остовных атомов существенно неточны, либо это крайне редкие особенности.

Номер и цепь остатка	Тип остатка	Угол φ	Угол ψ
A21	SER	-179.9	124.0
A37	ASP	72.6	155.9
A38	CYS	-152.3	79.2
A138	VAL	-110.8	-169.3
A140	HIS	-143.6	14.1
A212	THR	178.9	161.2
A238	ASP	-108.1	44.6
A272	ALA	-57.5	-3.2
B21	SER	179.5	128.1
B37	ASP	74.0	159.0
B38	CYS	-156.1	74.1
B70	TYR	-59.5	-7.6
B140	HIS	-25.2	-68.4
B143	GLU	-87.7	42.7

Таблица 3. Перечень маргинальных остатков, полученные с помощью сервера EDS.

Далее рассмотрим некоторые параметры оценки качества модели, приведенные в EDS. RSR - пространственный R-фактор для всех остатков, он позволяет найти маргиналов по соответствию боковой цепи «экспериментальной» электронной плотности. Он приведен на рис.5. Особые участки с повышенным RSR-фактором (подозрительные на наличие неточностей в координатах атомов) представлены на рис.6. Хорошими значениями RSR являются те, что меньше 10%. Известно, что атомы в кристалле находятся в состоянии теплового движения, что снижает их рассеивающую способность. Для того чтоб это учитывать, существует некая поправка – температурный фактор. Он представлен на рис.7. Z-score является относительной оценкой RSR, он представлен на рис.8. Высокие положительные значения $Z > 2$ говорят о том, что остаток плохо вписан в электронную плотность по сравнению с другими структурами с тем же разрешением, т.е. его можно причислить к маргинальным.

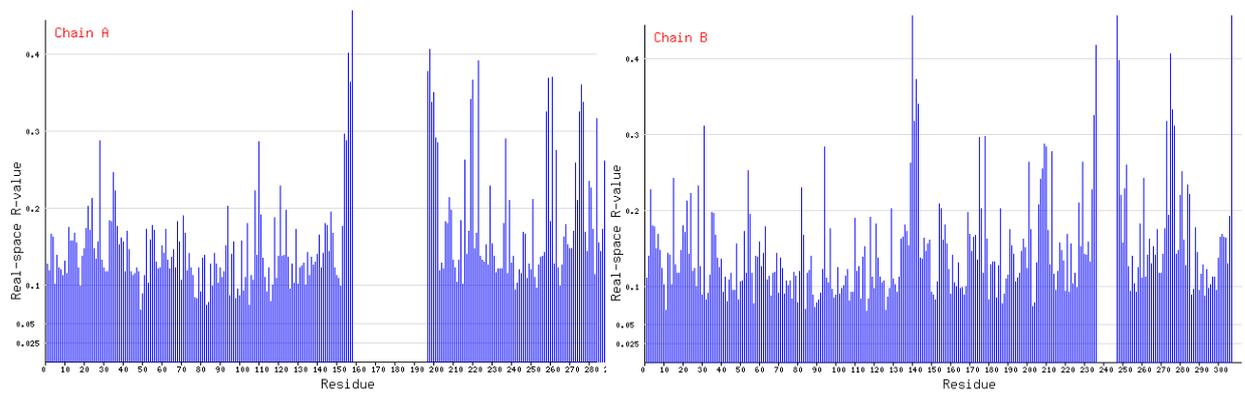


Рис. 5. RSR-значения для цепи остатков А (слева) и цепи В (справа).

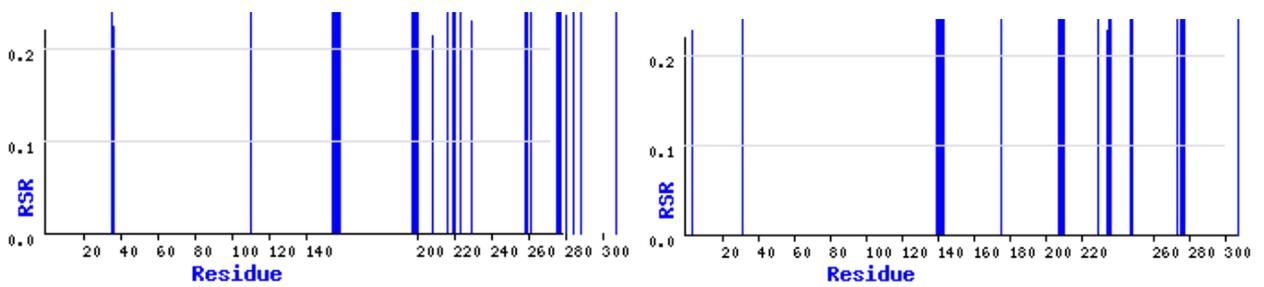


Рис. 6. Особые участки с повышенным RSR-фактором для остатков цепи А (слева) и В (справа).

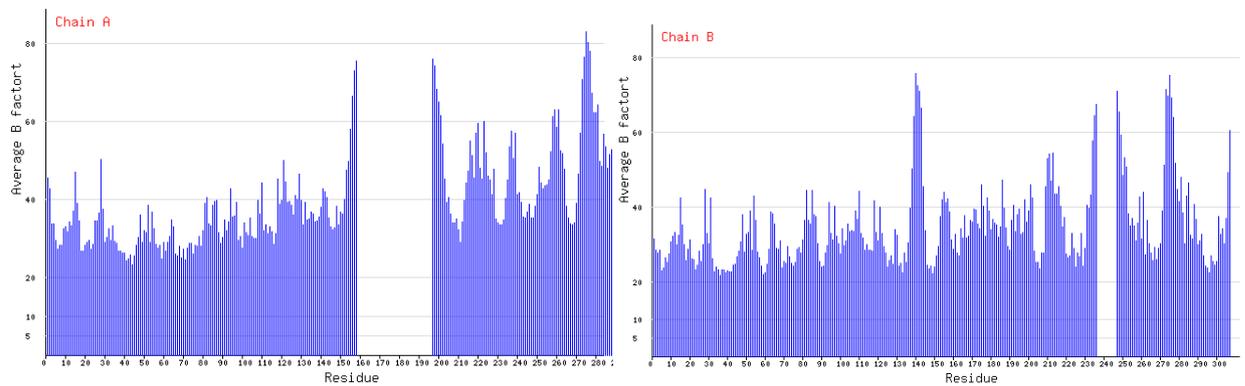


Рис. 7. Температурный фактор для остатков цепи А (слева) и для цепи В (справа).

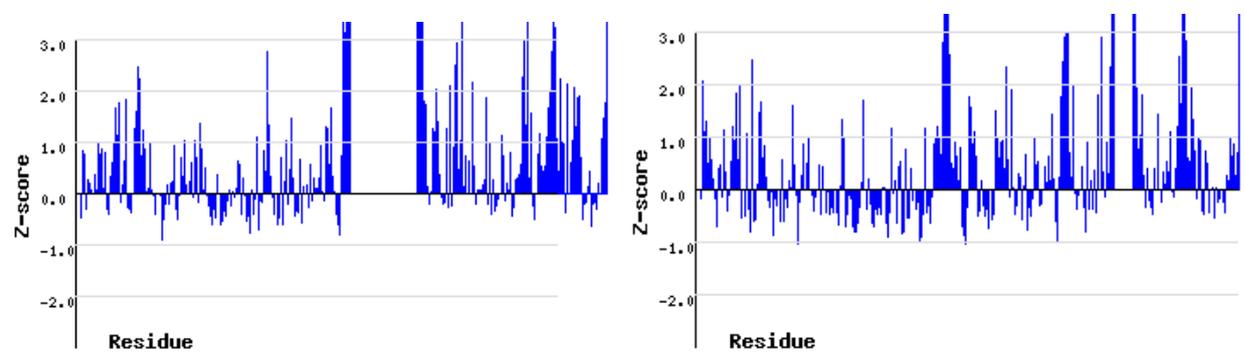


Рис.8. Z-score для остатков цепи А (слева) и цепи В (справа).

Попробуем проанализировать остатки из таблицы 3. Рассмотрим значения факторов для каждого остатка (RSR, Z-score, температурный фактор). Данные приведены в таблице 4.

Остаток	RSR	Z-score	B(температурный фактор)	Можно ли считать маргиналом?
Ser21_A	0,174	0,978	29,08	нет
Asp37_A	0,176	0,729	26,66	нет
Cys38_A	0,153	1,226	26,74	нет
Val138_A	0,126	0,101	34,53	нет
His140_A	0,140	0,31	38,06	нет
Thr212_A	0,104	-0,24	34,17	нет
Asp238_A	0,115	-0,23	50,4	нет
Ala272_A	0,170	1,101	57,05	нет
Ser21_B	0,171	0,926	26,21	нет
Asp37_B	0,168	0,604	23,16	нет
Cys38_B	0,137	0,845	23,3	нет
Tyr70_B	0,091	-0,5	25,43	нет
His140_B	0,567	7,51	75,78	да
Glu143_B	0,340	2,572	66,5	да

Табл. 4. Анализ возможных маргинальных остатков в структуре 1Y44.

Согласно данным таблицы, для остатков His140_B и Glu143_B значение RSR-фактора > 20%. Это может означать, что имеет смысл посмотреть, как эти остатки вписаны в электронную плотность. Аналогичный вывод можно сделать, оценив значения Z-score. Для тех же остатков он превышает 2. Кроме того, для них характерно высокое значение температурного фактора.

На рис.9 и рис.10 изображен остаток гистидина с разным уровнем подрезки электронной плотности. Как можно заметить, при уровне 1.0 σ электронная плотность не расположена вокруг остатка. Для уровня подрезки 1.5 σ также расположение атомов плохо соответствует электронной плотности. Таким образом, этот остаток можно считать маргинальным.

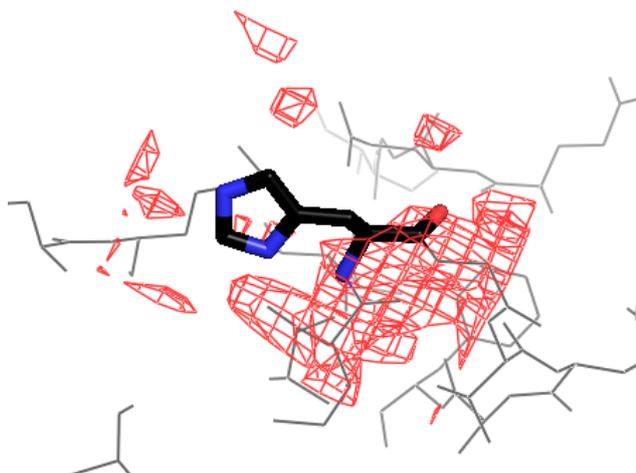


Рис.9. Остаток His140 цепи В и изображение электронной плотности на уровне подрезки 1.0 σ .

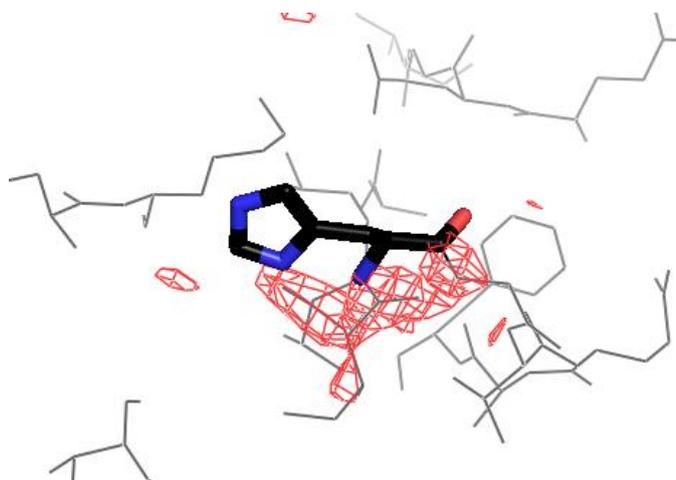


Рис.10. Остаток His140 цепи В и изображение электронной плотности на уровне подрезки 1.5 σ .

Рассмотрим другой возможный маргинальный остаток – Glu143 из цепи В. Он изображен на рис.11 и рис.12. Здесь мы наблюдаем аналогичную картину: на двух уровнях подрезки атомы не соответствуют электронной плотности, что говорит нам о маргинальности остатка.

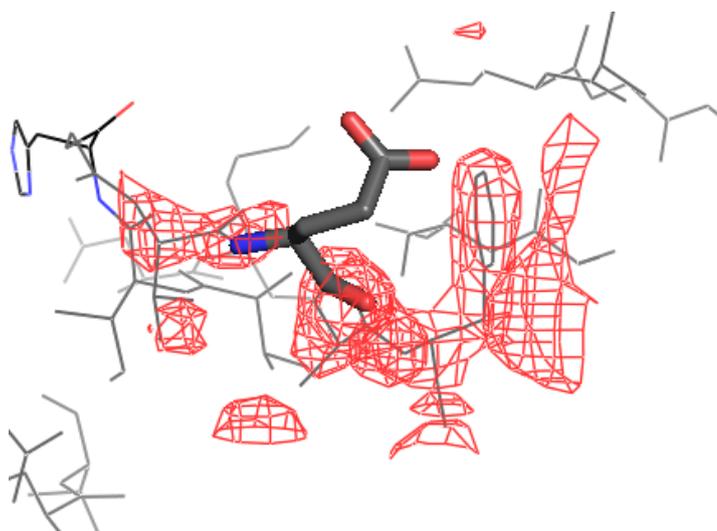


Рис.11. Остаток Glu143 цепи В и изображение электронной плотности на уровне подрезки 1.5 σ .

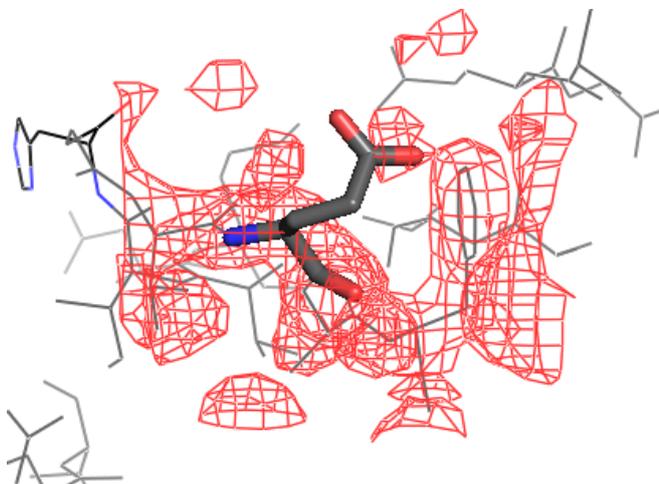


Рис.12. Остаток Glu143 цепи В и изображение электронной плотности на уровне подрезки 1.0 σ .

Воспользуемся сервисом MolProbity. Некоторые параметры приведены в таблице 5. Согласно MolProbity score (интегральной оценки структуры по данным этого сервиса), модель входит в число лучших для структур с подобным разрешением ($2.10\text{\AA} \pm 0.25\text{\AA}$). Однако присутствует некоторое количество отклонений от теории, как и у углов, так и у связей.

Параметр	Значение
Число недопустимых наложений атомов на 1000	8.89
Число остатков с маргинальными по отклонению от ротамеров боковыми цепями	15
Полные маргиналы по карте Рамачандрана	2
MolProbity score	2,2
Ковалентные связи, существенно отклоняющиеся от теории	19/4515
Ковалентные углы, существенно отклоняющиеся от теории	16/6100

Табл.5. Некоторые параметры выдачи сервера MolProbity.

Приведем остатки с маргинальными по отклонению от ротамеров боковыми цепями (их 15): Ile12_A, Leu35_A, Ile96_A, Leu111_A, Lue218_A, Leu221_A, Glu231_A, Tyr274_A, Glu281_A, Glu28_B, Leu35_B, Leu111_B, Leu218_B, Leu221_B, Arg223_B. Проверим, есть ли у них существенно завышенные параметры. По параметру RSR, для остатка Leu35_A RSR=24,6%, что больше 20%. Для него же Z=2,46 (что больше 2). Для остатка Glu28_B RSR=23,2%, что также выходит за 20%. Посмотрим на то, как эти остатки вписаны в электронную плотность (рис.13-15).

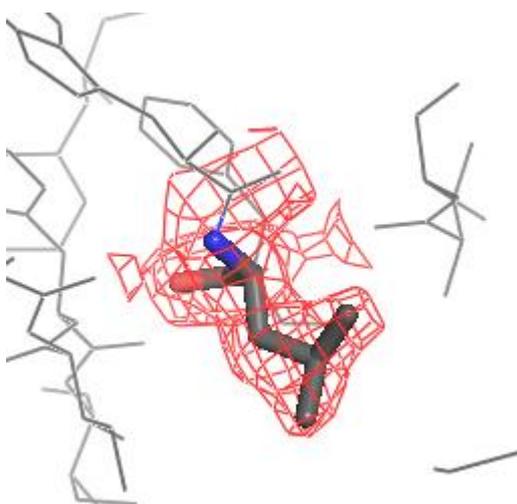


Рис.13. Остаток Leu35 цепи А и изображение электронной плотности на уровне подрезки 1.0 σ .

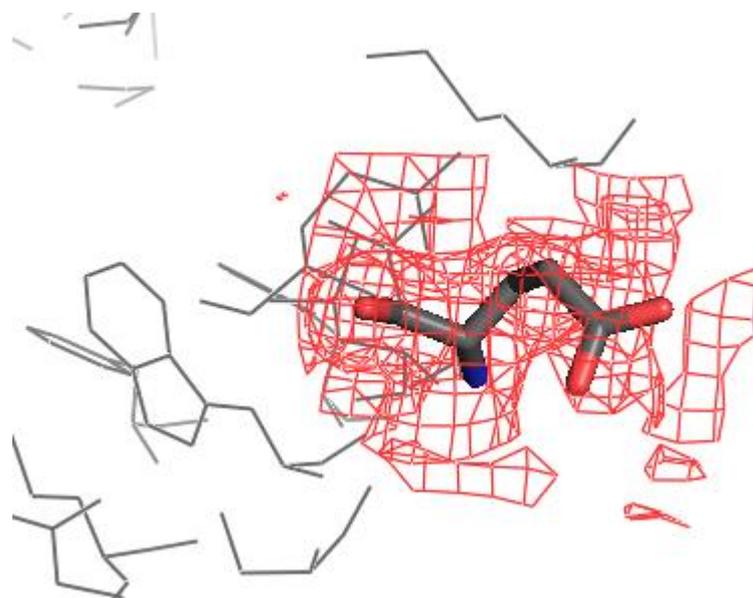


Рис.14. Остаток Glu28 цепи В и изображение электронной плотности на уровне подрезки 0.5 σ .



Рис.14. Остаток Glu28 цепи В и изображение электронной плотности на уровне подрезки 1.0 σ .

Согласно рис. 13, остаток Leu35 цепи А довольно хорошо вписан в электронную плотность и причин считать его маргинальным нет. Однако, для остатка Glu28 цепи В можно отметить, что атомы OE1 и OE2 вписаны в электронную плотность плохо. Таким образом, остаток Glu28 цепи В можно считать маргинальным.

3. Сравнение модели из PDB с моделью из PDB_redo.

Некоторые параметры сравнения приведены в таблице. Как можно заметить, после оптимизации значения параметров улучшились.

Например, это углы χ_1 и χ_2 (для них значение выросло), упало количество недостающих водородных связей, лучше стала характеристика карты Рамачандрана, однако такие параметры, как RMS Z-score для углов и длин связей стало выше, хотя по-прежнему меньше 1.

Параметр	Первоначальная модель	Полностью оптимизированная модель
Chi-1/Chi-2 нормальность ¹	-1.093	-0.920
RMS Z-score для длин связей ²	0.297	0.428
RMS Z-score для углов ²	0.657	0.691
Количество недостающих водородных связей (донор/акцепор) ³	29	28
Характеристика карты Рамачандрана ¹	-1.075	-0.996

Табл. 6. Параметры для первоначальной и улучшенной модели 1Y44. ¹-чем выше значение, тем лучше, ²-значение должно быть ниже 1.000, ³-чем ниже значение, тем лучше.

ЗАКЛЮЧЕНИЕ

Структура 1Y44 имеет среднее разрешение (2.05 Å), хорошие значения R-фактора и R-free в целом. Всего 95% остатков лежат в предпочитаемой области на карте Рамачандрана, 99.6% остатков лежат в разрешенной области, т.е. модель можно считать хорошей по этому параметру (согласно серверу MolProbity). Согласно же серверу EDS, маргинальных остатков по карте Рамачандрана находится довольно много, из которых, согласно сильно отклоняющимся значениям RSR –фактора и Z-score находятся 2 остатка, которые очень плохо вписываются в электронную плотность (и они не единственные!) Кроме того (согласно MolProbity), модель содержит остатки с маргинальными боковыми цепями (15), а также некоторое количество остатков с отклоняющимися от теории ковалентными углами и длинами связей. Таким образом, можно говорить о среднем качестве модели.

СПИСОК ЛИТЕРАТУРЫ

1. *“Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z.”* Li de la Sierra-Gallay I1, Pellegrini O, Condon C. Nature. 2005 Feb 10;433(7026):657-61
2. *“Insights into DNA recombination from the structure of a RAD51–BRCA2 complex”* Pellegrini, L.et al. Nature 420, 287–293 (2002).