

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ имени
М.В.ЛОМОНОСОВА**

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**Отчёт по качеству расшифровки структуры рибонуклеотид
редуктазы (PDB ID: 3R1R)**

**Выполнил студент 4-го курса ФББ МГУ
Кравченко Павел Андреевич**

**Преподаватели:
Алексеевский Андрей Владимирович
Спирин Сергей Александрович
Маслова Валентина**

Москва 2019 г.

Содержание

Аннотация	2
Введение	2
Результаты и обсуждение	5
Общие сведения о модели	5
Индикаторы качества модели	6
Анализ качества модели с помощью сторонних сервисов	8
Molprobity	8
EDS	9
WHAT_CHECK	10
Визуализация выбранных маргинальных остатков	11
ARG (323)	12
GLY (270) A	13
GLY (271) A	13
ALA (74) A	15
LEU (6) A	16
Сравнение модели PDB с моделью PDB_redo	17
Заключение	23
Ссылки на литературу и источники информации	24

Аннотация

В данной работе был произведен анализ качества расшифровки структуры рибонуклеотид редуктазы (PDB ID: 3R1R), полученной методом рентгеноструктурного анализа (РСА). Он включает в себя некоторые показатели качества структуры, такие как R-фактор, R free, карты Рамачандрана, RSR, Z-score. Целью анализа была оценка указанных выше показателей качества модели, а также подробное рассмотрение некоторых маргинальных аминокислотных остатков. В результате работы был сделан вывод о качестве расшифровки структуры белка.

Введение

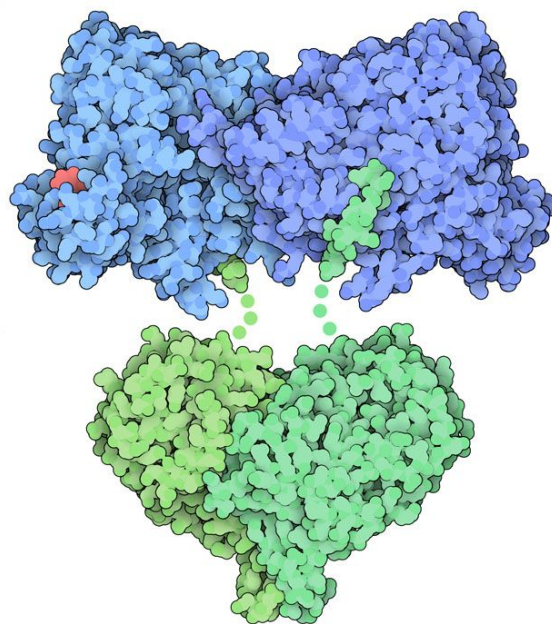


Рисунок 1. Устройство рибонуклеотид редуктазы. [6]

Рибонуклеотидредуктаза необходима для синтеза дезоксирибонуклеотидов, образующих ДНК. Фермент модифицирует рибонуклеотиды, связав АТФ в качестве позитивного аллостерического регулятора (показан красным на рис. 1). Ингибирование фермента происходит в том же аллостерическом сайте при связывании dATP, чтобы предотвратить образование токсичных уровней дезоксирибонуклеотидов.

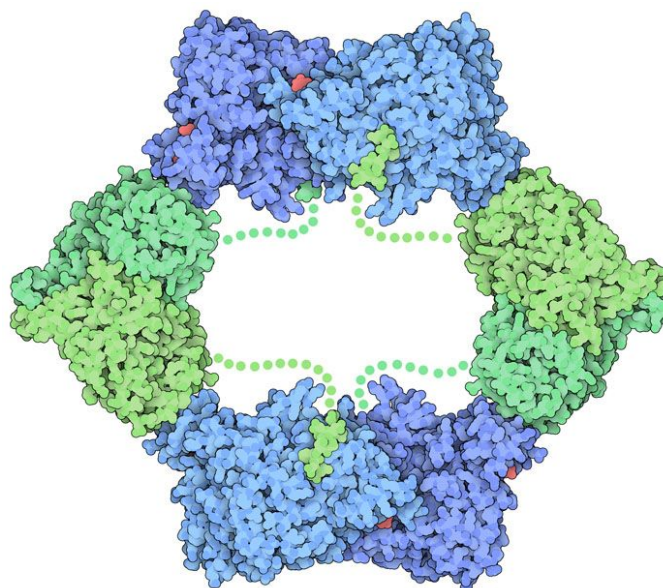


Рисунок 2. Дезактивированный фермент. [6]

Интересен способ, которым достигается дезактивация фермента. Если восстановить соседние белки в составе кристалла, то видно, что соседние молекулы “стыкуются” друг с другом и образуют кольца, блокируя друг друга по-цепочке (рис 2.). Механизм реакции основан на использовании свободных радикалов (рис. 3), а такая молекулярная перестройка нарушает путь для свободного радикала между субъединицами.

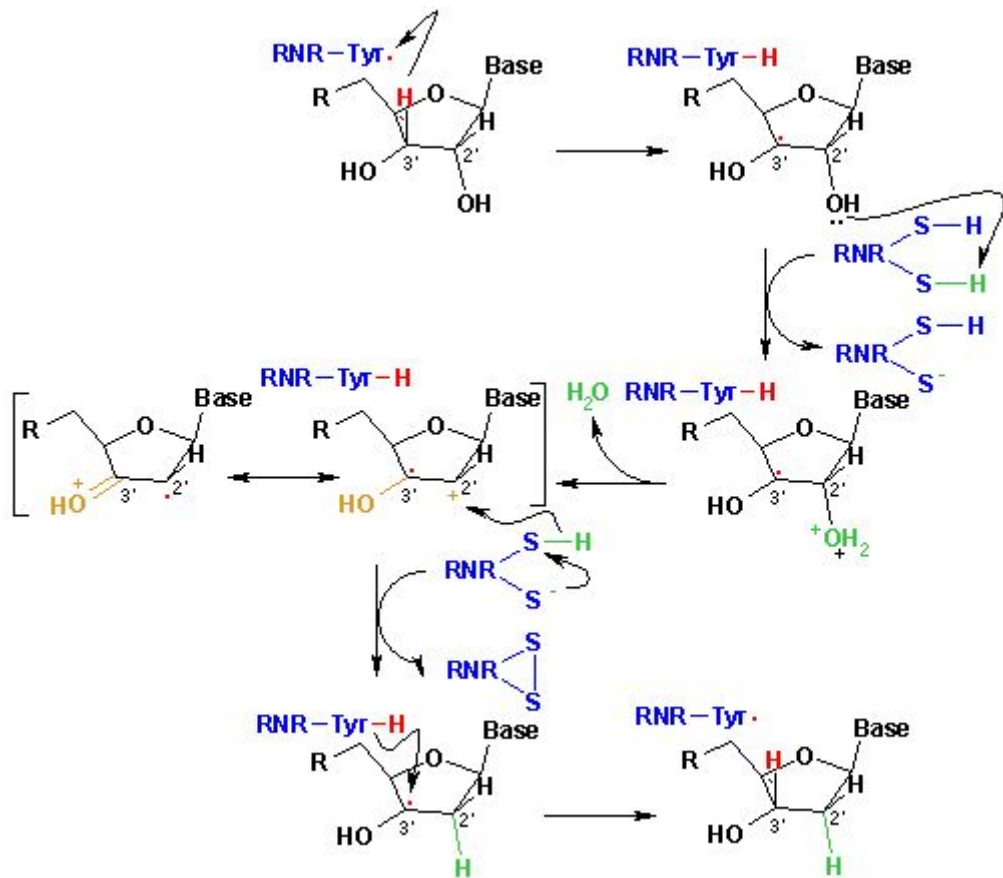


Рисунок 3. Механизм реакции, проводимой рибонуклеотидредуктазой [5]

Результаты и обсуждение

Общие сведения о модели

Из базы данных PDB был скачан файл с описанием структуры белка.

- Запись была сделана 28 января 1998 года. В 1997 году в журнале Structure вышла статья [1] с отчетом о расшифровке структуры, целью которой являлось установление значения редукции остатков цистеина в каталитическом центре фермента. Также расшифровка структуры уточнила представления о возможном каталитическом механизме реакции и возможных вариантах ингибирования.
- Разрешение структуры - 3Å с довольно хорошей поддержкой (полнота набора гармоник 96.1%) и диапазоном разрешений от 2.99Å - 39.7Å. Всего было измерено 65054 рефлекса. Файл содержит 18634 атома, образующих две белковые цепи: А и В, состоящих в сумме из 2363 аминокислотных остатков.
- Кристаллизация проводилась в растворе 1.7 М сульфата лития, 10 ММ сульфате магния в 25 ММ и цитратном буфере (Ph=6). Концентрация белка в растворе для кристаллизации составила 17 МГ/МЛ. Также в раствор были добавлены 10 ММ АМР РNР и 10 ММ СDР. Создание модели проводилось de novo с помощью программ SCALEPACK, TNT, REFMAC, DENZO.

Основная информация о полученной модели представлена в таблице 1.

Таблица 1. Основные параметры модели. [3, 4]

Параметр	Значение
Параметры записи	
PDB ID:	3R1R
Организм	Escherichia coli (strain K12)
Название публикации	BINDING OF ALLOSTERIC EFFECTORS TO RIBONUCLEOTIDE BINDING OF ALLOSTERIC EFFECTORS TO RIBONUCLEOTIDE REDUCTASE PROTEIN R1: REDUCTION OF ACTIVE-SITE CYSTEINES PROMOTES SUBSTRATE BINDING.
Авторы	M.ERIKSSON, U.UHLIN, S.RAMASWAMY, M.EKBERG, K.REGNSTROM, B.M.SJOBERG, H.EKLUND

Год публикации	1997
Параметры белка	
Вес структуры	268238.38
Число атомов в структуре	17943
Число аминокислотных остатков	2363
Число цепей	2
Симметрия	Asymmetric - C1
Стехиометрия белка	Monomer - A
Параметры структуры	
Метод определения структуры	X-RAY DIFFRACTION
Число измеренных рефлексов	65054
Разрешение	3 Å
Полнота набора рефлексов с данным разрешением	96.1%
Пространственная группа	H 3 2
Параметры ячейки	a = 224.610, b = 224.610, c = 336.630; $\alpha = 90.00$, $\beta = 90.00$, $\gamma = 120.00$
Используемое для построения программное обеспечение	SCALEPACK, TNT, REFMAC, DENZO
Метрики	
R-Factor	0.263
R-Free	0.287

По классификации Pfam [4] данный фермент относится к семейству рибонуклеотидредуктаз (Ribonucleotide reductase, small chain).

Индикаторы качества модели

Для оценки качества полученной модели в кристаллографии используют несколько основных индикаторов качества:

- **R фактор** – это среднеквадратичное отклонение расстояний между модулями структурных факторов, рассчитанных по построенной модели и полученных в эксперименте. Данный фактор отражает то, насколько в целом модель хорошо описывает результаты PCA. Из таблицы 1 видно, что R фактор имеет значение 0.263, что несколько больше общепринятой нормы в 0.25.
- **R-free** - это тот же R фактор, но для минимизации R-free использовалась только часть полученных в эксперименте структурных факторов, а 5% использовалась для валидации модели. У авторов модели R-free не сильно отличается от R фактора, что может говорить о том, что переоптимизации удалось избежать, однако оба фактора имеют надпороговое значение. Нужно с осторожностью относиться к описанию данной моделью экспериментальных данных.
- Рассчитывают долю **маргинальных остатков** с помощью:
 - построения карт **Рамачандрана** - данные карты отражают разрешенные области углов ϕ и ψ для данного белка. На рисунке 5 показана карта Рамачандрана для модели рибонуклеотид редуктазы.
 - проверки не перекрывания **Ван-дер-Ваальсовых радиусов** атомов
 - измерения длин и углов связей, ротаметров

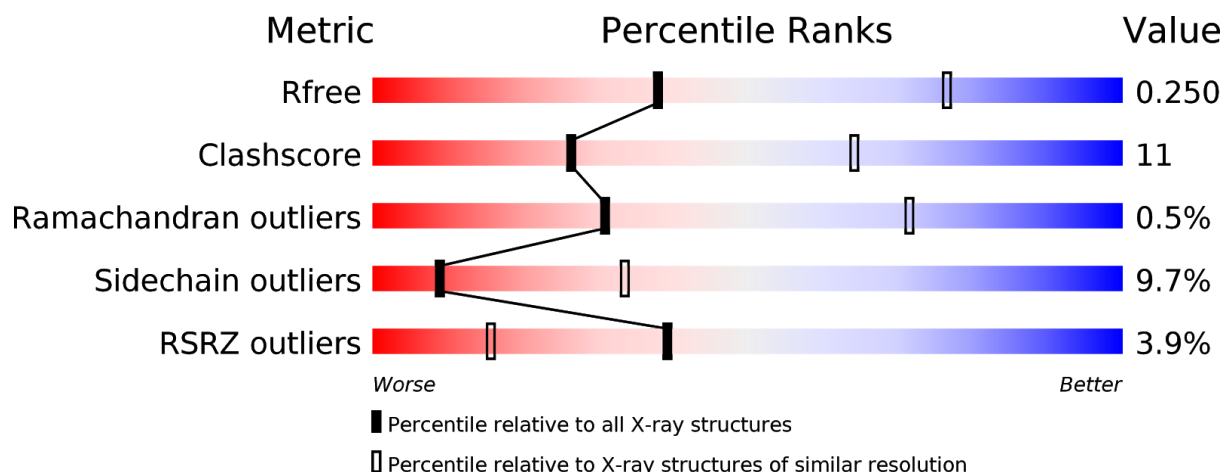


Рисунок 4. Основные показатели качества модели. Черные полоски отражают процентиль структур с таким же значением параметра относительно всех PCA моделей, а белые отражают процентиль структур относительно PCA моделей близкого разрешения. [3]

Анализ качества модели с помощью сторонних сервисов

Molprobity

Был проведен анализ качества модели с помощью сервиса **Molprobity** [7]. К модели были добавлены атомы водорода. При добавлении ротамеров было исправлено 3 “плохих” ротамера. Ниже представлено заключение сервиса о качестве модели.

- **"ClashScore"** — число недопустимых наложений атомов на 1000 (10.49 ~ 1% атомов); перцентиль данной структуры по отношению к структурам примерно такого же разрешения: "97th percentile" значит, что 97% структур ClashScore хуже (больше), чем у данной. Относительно критериев недопустимых наложений можно заключить, что модель имеет хорошее качество.
- **"Poor rotamers"** — число остатков с маргинальными по отклонению от ротамеров боковыми цепями. Рекомендуемое значение <0.3%. Всего нашлось 214 остатков (11.11%). Это может говорить о достаточно плохом качестве модели.
- **"Ramachandran outliers"** — полные маргиналы по карте Рамачандрана, лежат вне допустимой области. Рекомендуемое значение <0.05%. Всего нашлось 12 остатков (0.54%), что также является показателем плохого качества модели.
- **"Ramachandran favored"** — число и процент остатков в предпочитаемой области. Предпочтительным значением параметра является значение > 98%. Сервис определил, что в предпочтительной области находятся 91.47% остатков.

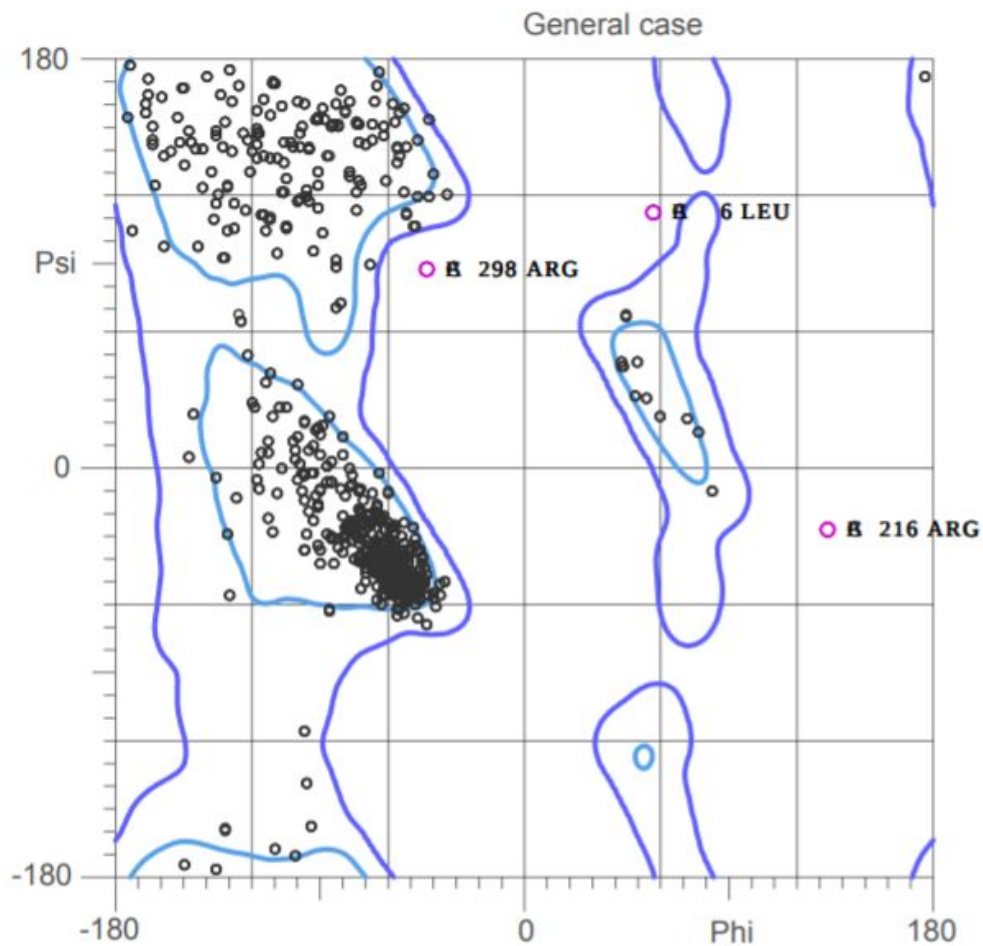


Рисунок 5. Карта Рамачандрана для модели рибонуклеотид редуктазы в общем случае, полученная с помощью сервиса **Molprobrity**. Диапазон углов от -180° до 180° . По оси абсцисс ϕ , а по оси ординат – ψ . Сине-фиолетовым цветом выделены разрешенные области значений углов ϕ и ψ , а голубым – наиболее предпочтительные области. Можно заметить, что на первый взгляд большинство остатков имеют углы, попадающие в наиболее предпочтительные области. [4]

EDS

- **RSR**. Помимо R фактора и R-free можно оценивать качества построения модели с помощью пространственного R фактора. Он позволяет оценить соответствие рассчитанных и полученных в эксперименте электронных плотностей. Для анализа RSR использовался сервис EDS [8]. Значение RSR для всей модели составило 0.227, что является достаточно хорошим показателем для данного разрешения. Данный фактор можно использовать для локальной оценки качества.
- **RSRZ**. Также был определен Z-score. Данный параметр является нормированной и центрированной на RSR по PDB величиной, то есть,

показателем того, насколько RSR для остатка в данной модели отличается от среднего значения RSR по структурам из PDB с близким разрешением. Пороговым значением для данной величины является значение 2. Для целой цепи A значение RSRZ составило 0.01. Некоторые маргинальные по данному критерию остатки указаны в таблице 2.

WHAT_CHECK

- **Coarse Packing Quality Control.** Была посчитана общая оценка правильности окружения остатков. Окружение считается плохим, если скор < -5.0. Значение параметра составило -0.764, что отражает достаточно хорошее окружение остатков в модели.
- **Anomalous bond lengths.** Все связи в модели оказались стандартной длины.
- **Peptide flip validation.** Были обнаружены 6 остатков, два из которых указаны в таблице 2. Данные остатки поворачиваются на 180 градусов при релаксации связи C-C.
- **Bond angles that deviate more than 4 sigma.** Не было найдено аномальных углов.
- **Ramachandran outlier.** Было найдено 12 остатков, находящихся вне оптимальных значений углов.

Можно заметить, что, в основном, при построении модели ошибки возникают в остатках с длинными подвижными боковыми цепями.

Таблица 2. Примеры аминокислотных остатков, маргинальных по одному или более критериям.

№	Остаток	Критерий
1	ARG (323 A)	Coarse Packing Quality Control = -7.734. RSRZ = 4.6
2	ARG (269)A	Coarse Packing Quality Control = -7.912
3	GLY (270) A	RSRZ = 11.7
4	GLY (271) A	RSRZ = 7.3
5	ASP (11) A	RSRZ = 6.6
6	ASN (35) A	Peptide flip validation
7	ALA (74) A	Peptide flip validation. RSRZ = 2.7
8	ARG (298) A	Ramachandran outlier. RSRZ = 3.3
9	LEU (6) A	Ramachandran outlier. RSRZ = 4.2

В отчете по валидации структуры белка [10] содержится информация о маргинальности всех остатков как в визуальном, так и в табличном формате. Рисунок 6 показывает карту маргинальности остатков цепи А. Качество остатков отражено следующим образом:

- Зеленый цвет остатка означает отсутствие маргинальности по всем показателям
- Желтый цвет остатка означает маргинальность по одному показателю
- Оранжевый цвет остатка означает маргинальность по двум показателям
- Красный цвет остатка означает маргинальность по более, чем двум показателям
- Серыми выделены остатки, не представленные в структуре
- Красными точками указаны остатки, имеющие RSRZ > 2

• Molecule 1: RIBONUCLEOTIDE REDUCTASE R1 PROTEIN

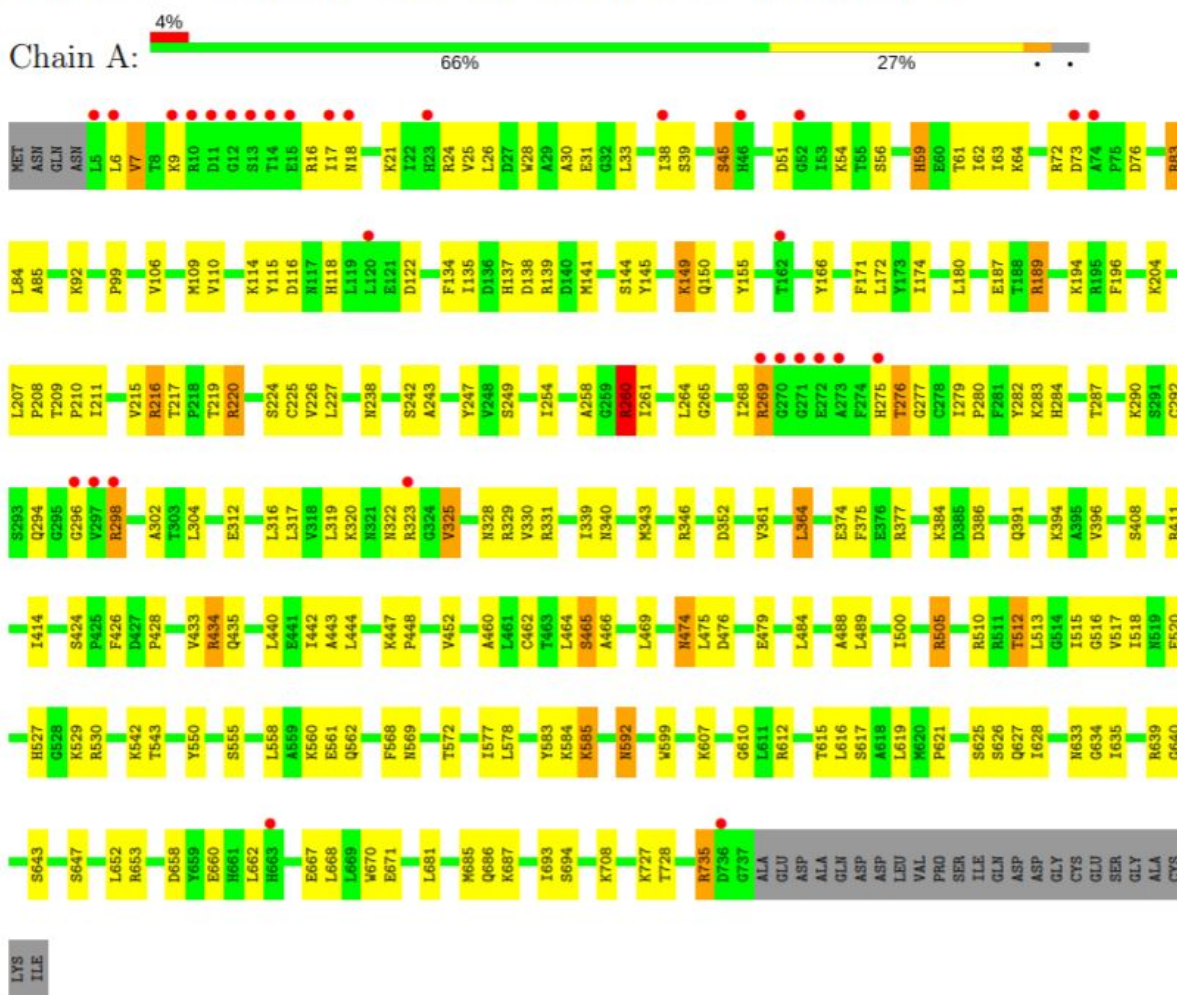


Рисунок 6. Карта маргинальности остатков цепи A структуры 3R1R. Описание в тексте. Видно, что модель содержит достаточно большое количество остатков с $RSRZ > 2$ и маргинальностью по одному или более критериям. [10]

Далее была проведена визуализация выбранных маргинальных остатков.

Визуализация выбранных маргинальных остатков

Для визуализации были выбраны следующие маргинальные остатки.

Таблица 3. Аминокислотные остатки, выбранные для более подробного изучения.

№	Остаток	Критерий
1	ARG (323) A	Coarse Packing Quality Control = -7.734. RSRZ = 4.6
3	GLY (270) A	RSRZ = 11.7
4	GLY (271) A	RSRZ = 7.3
7	ALA (74) A	Peptide flip validation. RSRZ = 2.7
9	LEU (6) A	Ramachandran outlier. RSRZ = 4.2

ARG (323)

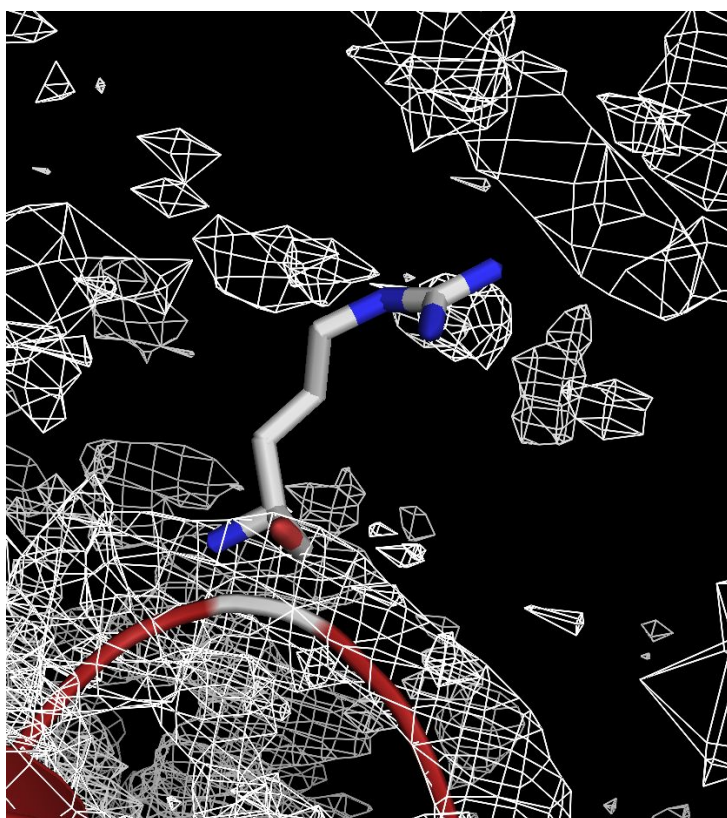


Рисунок 7. Визуализация электронной плотности остатка ARG (323) A на уровне подрезки (сигма) = 1.

Из рисунка 7 видно, что остаток аргинина совершенно не покрыт электронной плотностью и координаты его атомов, видимо, были взяты из общих соображений о длинах связей и углах. Данный остаток был выбран из-за крайне низкого значения параметра правильности окружения (-7.734) и достаточно высокого значения RSRZ (4.6). Теперь мы видим, что электронной плотности на боковой цепи нет совсем.

GLY (270) A

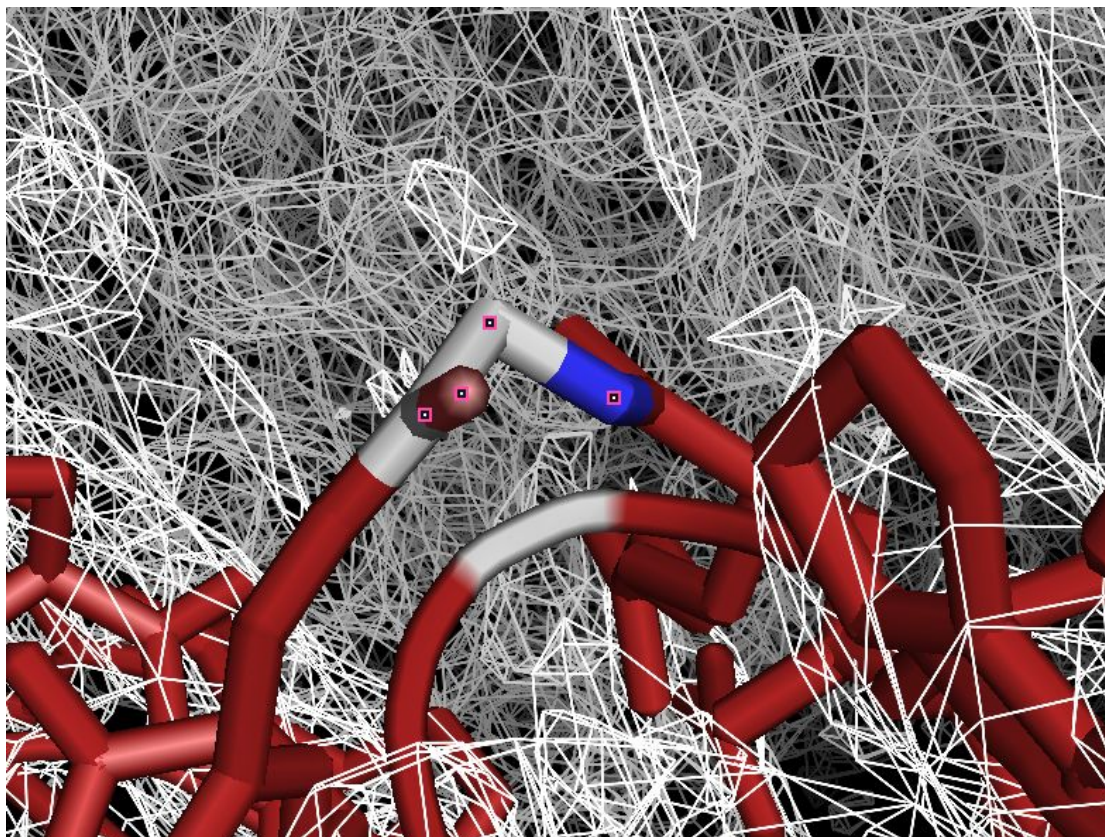


Рисунок 8. Визуализация электронной плотности остатка GLY (270) A на уровне подрезки (сигма) = 1.

Из рисунка 8 видно, что в окрестности остатка глицина 270 модель цепи A выходит за электронную плотность. Видимо, так было получено из-за наличия небольшого облака электронной плотности недалеко от С-альфа атома остатка глицина. Не удивительно, что RSRZ самый большой для модели - 11.7.

GLY (271) A

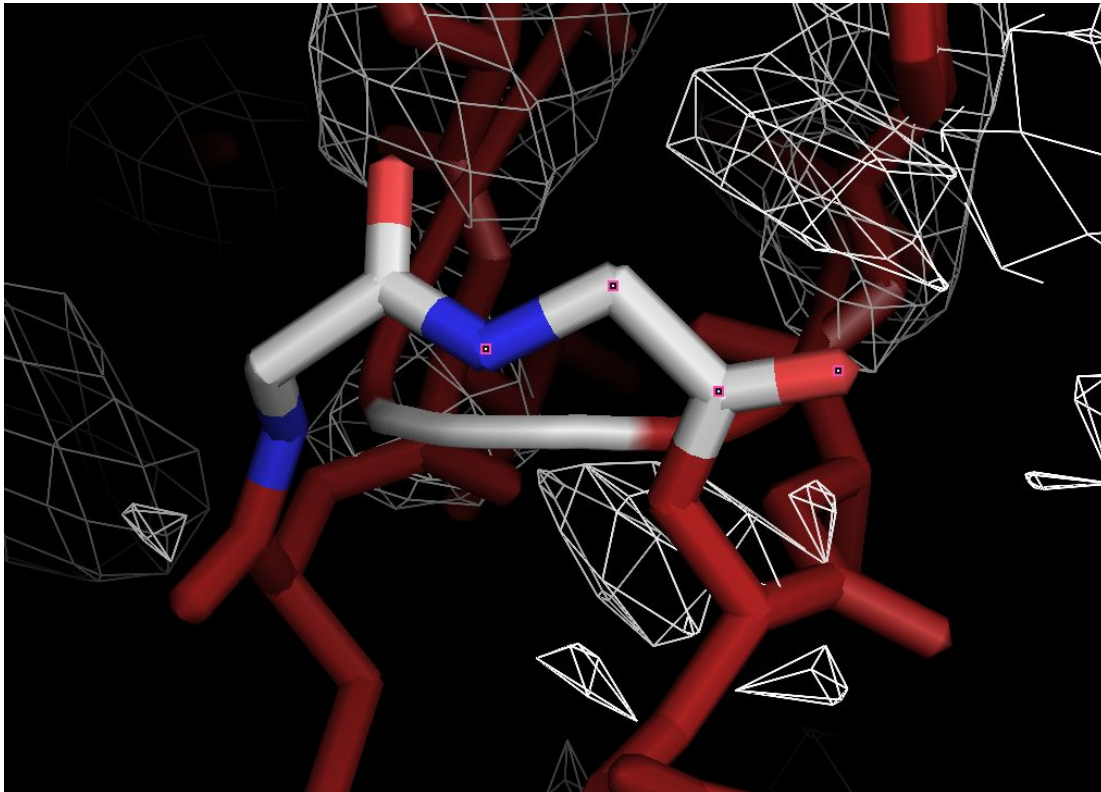


Рисунок 9. Визуализация электронной плотности остатка GLY (271) A на уровне подрезки (сигма) = 1.

Для остатка глицина 271 также верны все замечания, описанные для остатка глицина 270, так как это соседние аминокислоты, образующие поворот в модели. Видно, что плотности так же нет на остатках 269 и 272. RSRZ для остатка 271 оказался равен 7.3, что выше порога в 3.7 раза.

ALA (74) A

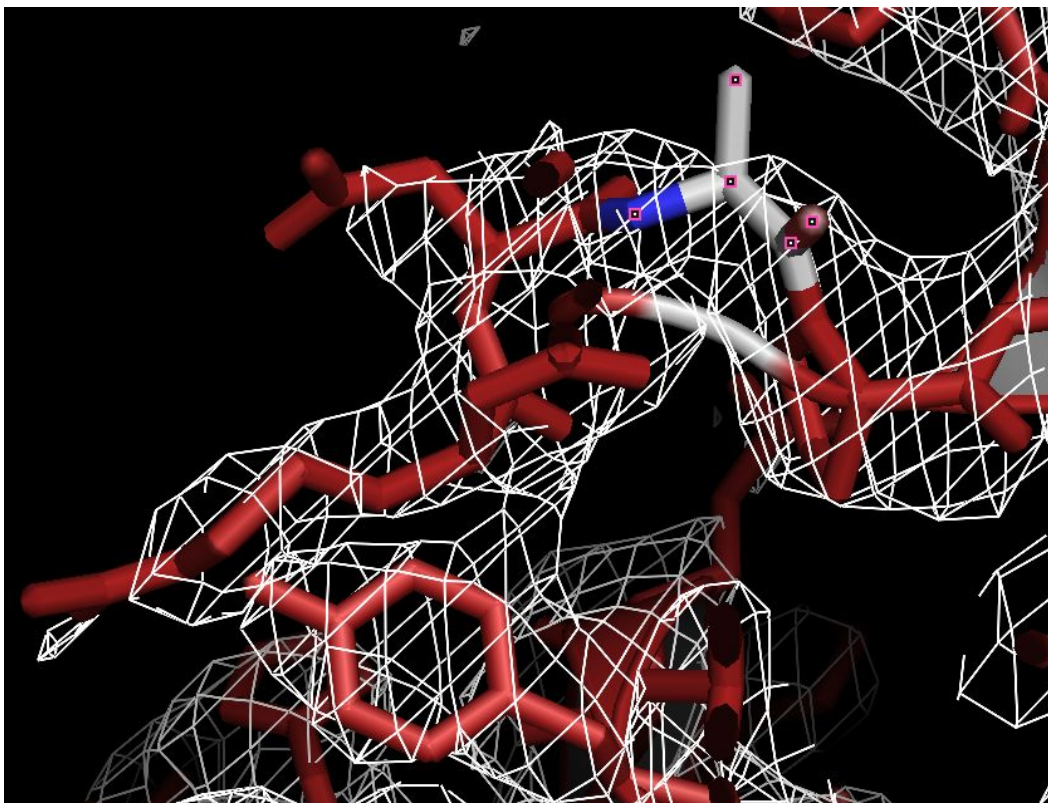


Рисунок 10. Визуализация электронной плотности остатка ALA (74) A на уровне подрезки (сигма) = 1.

Рисунок 10 демонстрирует остаток аланина в составе модели структуры рибонуклеотид редуктазы. Согласно результатам оценки качества модели, данный остаток имеет угловую напряженность и при умозрительном разрезании связи сраелаксирует на 180 градусов. RSRZ скор остатка составил 2.7. Видно, что остов покрыт электронной плотностью, тогда как боковой радикал остался обнажен. В целом, метильная группа аланина редко играет важную роль в составе белка и на неточность можно было бы “закрывать глаза”, если бы отсутствовал угловой фактор.

LEU (6) A

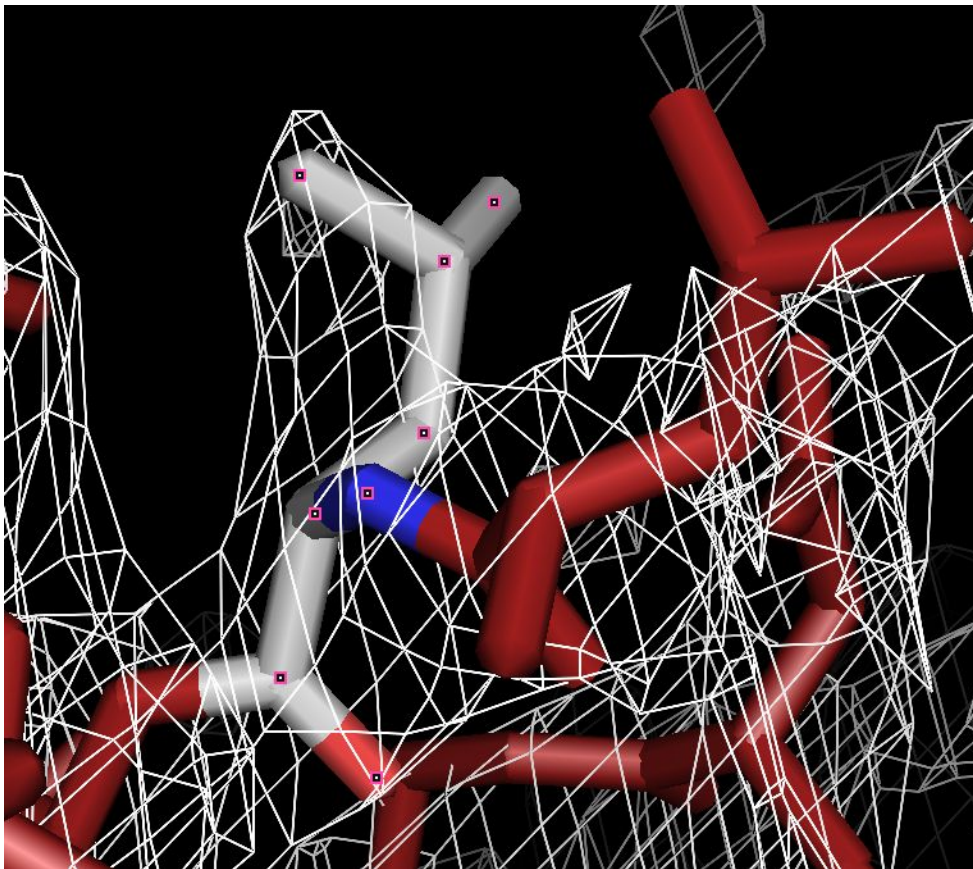


Рисунок 11. Визуализация электронной плотности остатка LEU (6) A на уровне подрезки (сигма) = 1.

Остаток лейцина 6 на цепи A был указан как нетипичный по углу ϕ . Видно, что связь C' - N - C-альфа расположена под острым углом. Также его RSRZ равняется 4.2, видимо, из-за того, что боковой радикал выходит за границу электронной плотности.

Сравнение модели PDB с моделью PDB_redo

База данных PDB_redo [11] содержит оптимизированные по описанным выше параметрам версии моделей PDB.

В модели PDB_redo было изменено 34 ротамера, убрано 33 молекулы воды, 98 остатков стали лучше вписываться в свои электронные плотности, однако для 2х остатков данный параметр стал хуже. Изменение других параметров показано в таблице 4.

Таблица 4. Сравнение качества структур PDB и PDB_redo

Параметр качества	PDB	PDB_redo
R	0,2405	0,2226
R-free	0,2604	0,2595
Z-score для длин связей	0,322	0,497
Z-score для углов	1,081	0,730
Качество карты Рамачандрана	5	26
Число хорошо упакованных остатков	32	56
Нормальность ротамеров (перцентили)	1	34
Достоверность водородных связей (перцентили)	25	25

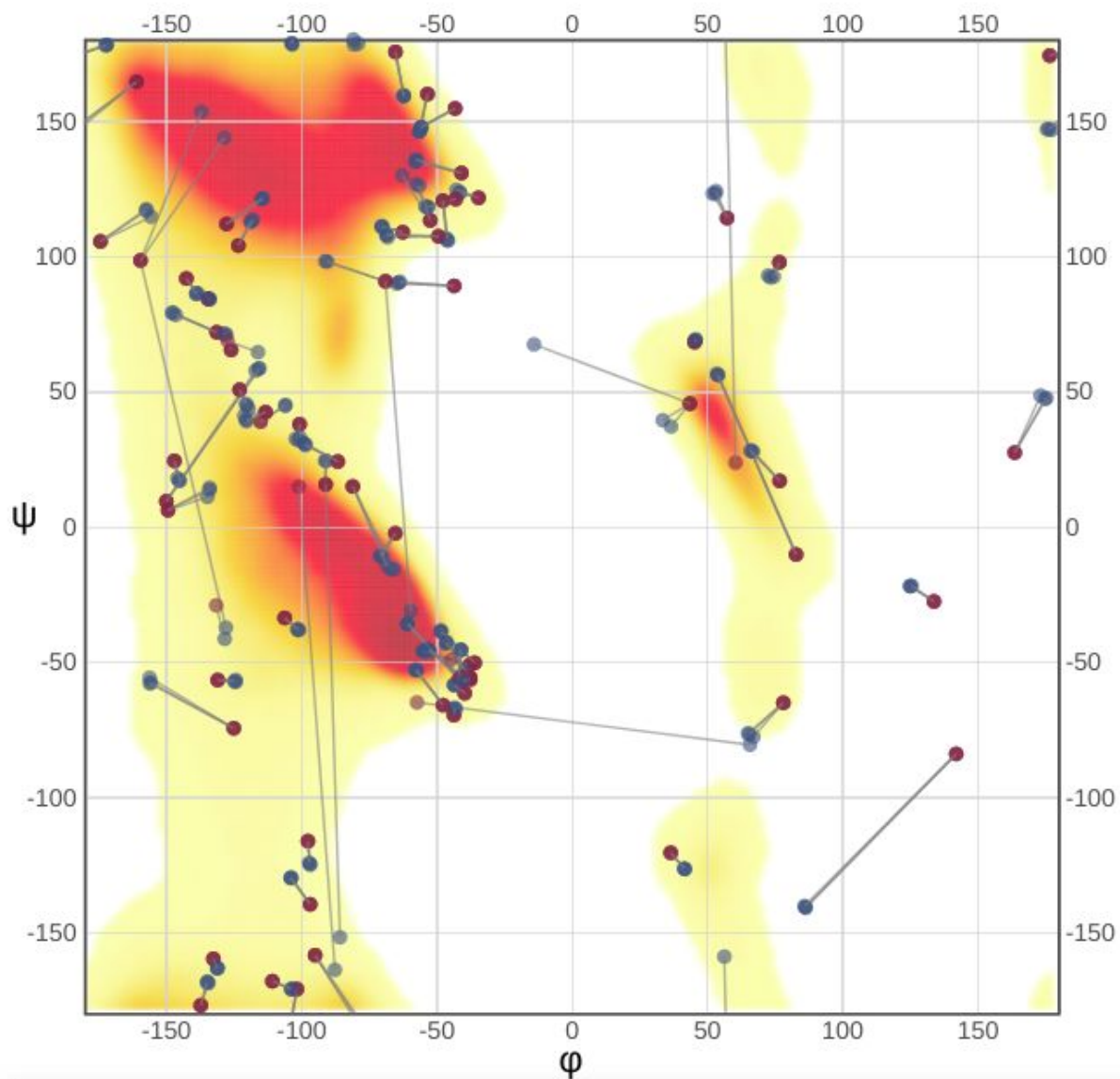


Рисунок 12. Карта Рамачандрана с указанием изменения углов маргинальных остатков в структуре PDB_redo. [11]

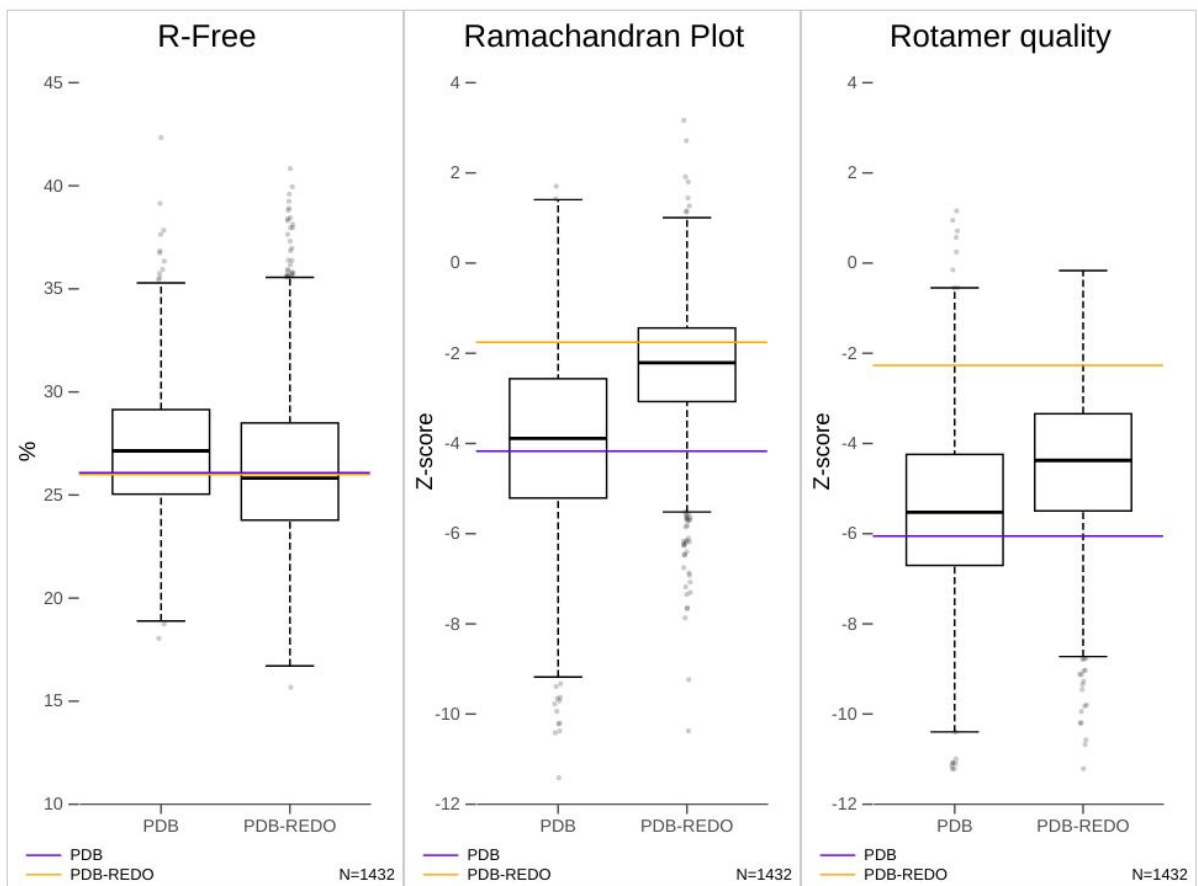


Рисунок 13. Сравнение основных показателей качества расшифровки структуры для моделей из PDB и PDB_redo. Можно заметить, что R-free параметр уменьшился, а Z-score карты Рамачандрана и качество ротамеров увеличилось. [11]

Ниже представлены отличия в визуализации моделей PDB и PDB_redo. Зеленым показана оптимизированная модель PDB_redo, а красным - исходная модель PDB.

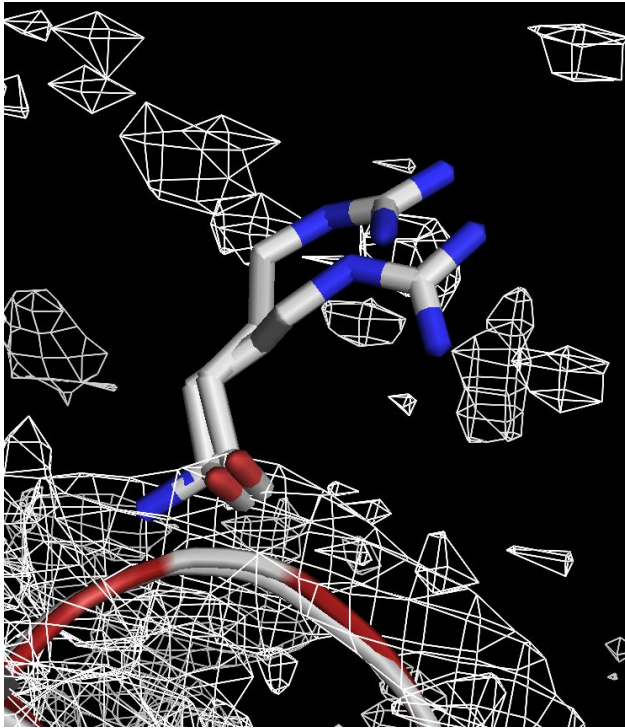


Рисунок 14. Сравнение маргинального остатка ARG (323) A в модели PDB и PDB_redo. Видно, что для данного остатка положение боковой цепи практически не изменилось.

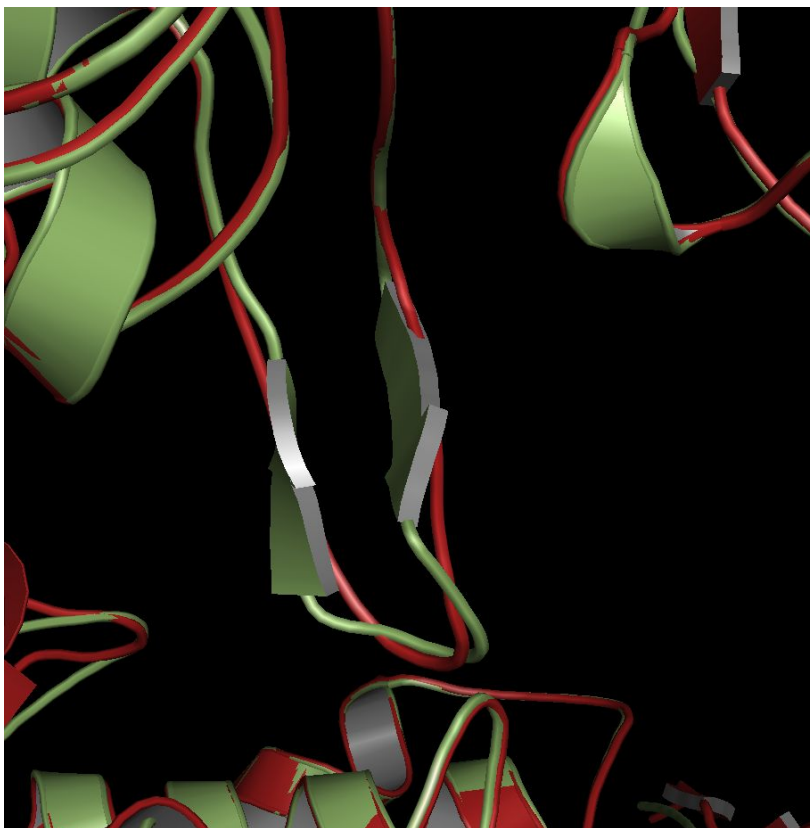


Рисунок 15. Оптимизированная модель PDB_redo приобрела небольшой бета-лист.

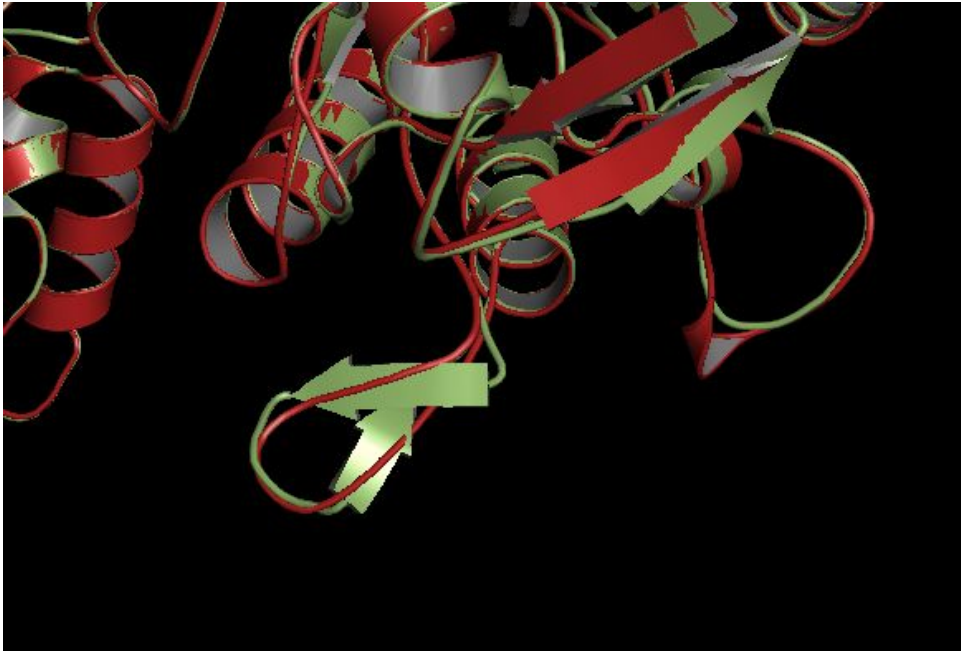


Рисунок 16. Оптимизированная модель PDB_redo приобрела второй небольшой бета-лист.

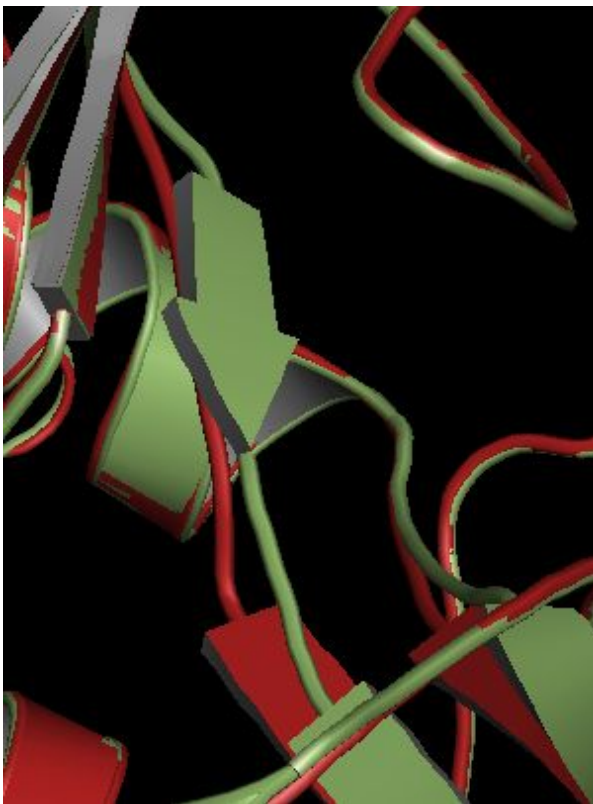


Рисунок 17. Оптимизированная модель PDB_redo приобрела третий небольшой бета-лист.

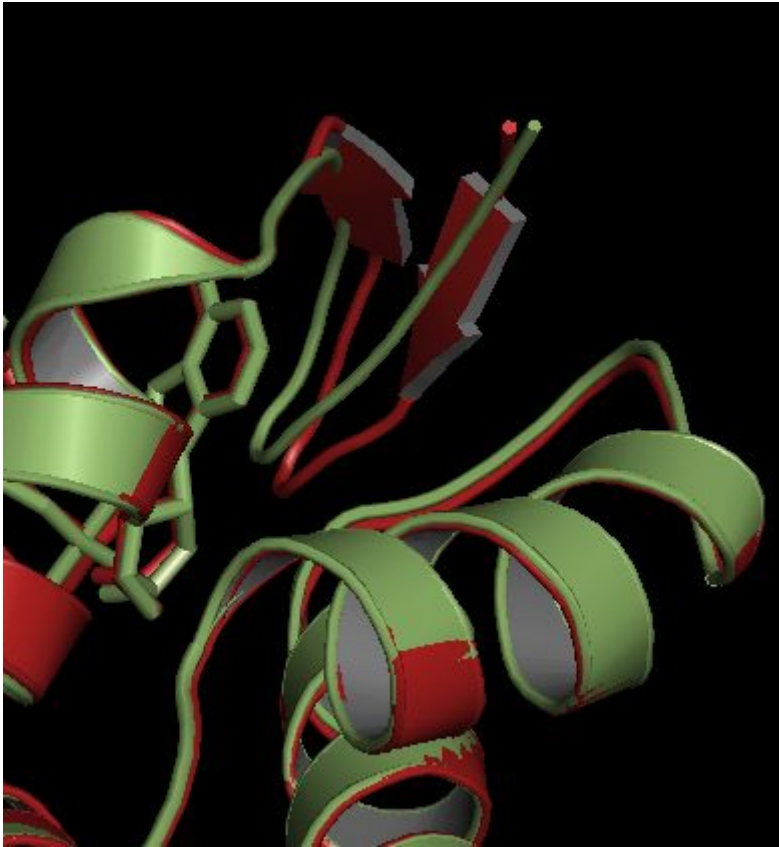


Рисунок 18. Оптимизированная модель PDB_redo утратила бета-складку.

Заключение

Исходя из приведенного выше анализа, можно сделать вывод, что качество расшифровки структуры рибонуклеотид редуктазы (PDB ID: r3r1) можно считать удовлетворительным, а в каких-то аспектах даже плохим. Для многих остатков было показано, что модель не полностью вписывается в электронную плотность, а местами присутствует нарушение допустимых углов и напряжений. Оптимизация модели (PDB_redo), в целом, улучшила ряд параметров, но некоторые проблемные точки все же сохранились, видимо, в силу низкого качества данных. Будет интересно сравнить результаты рентгеноструктурного анализа 1997 года и структуру, полученную с помощью РСА или криоэлектронной микроскопии в наши дни.

Ссылки на литературу и источники информации

1. 3r1r: Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnstrom, K., Sjoberg, B.M., Eklund, H. (1997) Binding of allosteric effectors to ribonucleotide reductase protein R1: reduction of active-site cysteines promotes substrate binding. Structure 5: 1077-1092
2. 1mrr: Atta, M., Nordlund, P., Aberg, A., Eklund, H., Fontecave, M. (1992) Substitution of manganese for iron in ribonucleotide reductase from Escherichia coli. Spectroscopic and crystallographic characterization. J.Biol.Chem. 267: 20682-20688
3. Запись PDBе 3R1R <https://www.ebi.ac.uk/pdbe/entry/pdb/3r1r/>
4. Структура RCSB PDB 3R1R <https://www.rcsb.org/structure/3R1R>
5. Статья в википедии о рибонуклеотид резуктазе https://en.wikipedia.org/wiki/Ribonucleotide_reductase
6. Статья о молекуле месяца - рибонуклеотид резуктазе <http://pdb101.rcsb.org/motm/238>
7. Сервис molprobity <http://molprobity.biochem.duke.edu/>
8. Сервис EDS <http://eds.bmc.uu.se/eds/>
9. Сервис WHAT IF <https://swift.cmbi.umcn.nl/servers/html/index.html>
10. Отчет о стректуре PDB 3R1R https://files.rcsb.org/pub/pdb/validation_reports/r1/3r1r/3r1r_full_validation.pdf
11. PDB-REDO 3R1R <https://pdb-redo.eu/db/3r1r>