

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**имени М.В.ЛОМОНОСОВА**

---

**ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ**

**Отчет о качестве RSA-расшифровки кристаллической  
структуры N-концевой области субъединицы когезина  
человека STAG1 в комплексе с пептидом RAD21.**

**(PDB ID: 6RRC)**

Подготовила:

Ралдугина Василиса Николаевна

---

## Аннотация

---

В этой работе было проведено исследование качества расшифровки кристаллической структуры N-концевой области субъединицы когезина человека STAG1 в комплексе с пептидом RAD21. Рассматривались как показатели качества приведенной модели в целом, так и отдельные остатки. Было проведено сравнение данной структуры со структурой 6RRK, отредактированной сервисом PDB-REDO.

---

## Введение

---



Рисунок 1. Структура N-концевой области субъединицы когезина человека STAG1 в комплексе с пептидом RAD21. (PDB ID: 6RRC).

Когезин — это комплекс белков, который регулирует процесс разделения сестринских хроматид в ходе деления клетки (как мейоза, так и митоза). Он представляет из себя белковый комплекс, играющий ключевую роль в репарации ДНК путем гомологичной рекомбинации, а также в когезии и сегрегации хромосом во время клеточного деления.

RAD21 и STAG1 являются неотъемлемой частью когезинового комплекса, считается, что они могут влиять на локализацию и экспрессию генов.

Когезиновый комплекс необходим для сцепления сестринских хроматид после репликации ДНК. По-видимому, он образует большое белковое кольцо, внутрь которого могут быть захвачены сестринские хроматиды. В анафазе комплекс расщепляется и диссоциирует от хроматина, позволяя сестринским хроматидам сегрегировать. Когезиновый комплекс может также играть роль в сборке полюсов веретена деления во время митоза.

Когезин играет ключевую роль в трехмерной упаковке хроматина в трехмерном ядерном пространстве. Организация хроматина обеспечивает эффективную упаковку генома и, при этом, правильную экспрессию и репликацию генетического материала. Показано, что пространственная организация генома играет важную роль в транскрипционном контроле генов. Важные хромосомные активности связаны как со структурными свойствами, так и с пространственными конформациями хромосом.

Было выявлено, что геном млекопитающих организован в топологически ассоциированные домены (ТАДы): регионы, которые показывают высокий уровень взаимодействия внутри себя и практически не имеют взаимодействия с соседними регионами. Центральные области ТАДов показывают высокий уровень взаимодействия хроматина между собой и совпадают с наличием тканеспецифических генов и связанных с ними энхансеров, взаимодействие которых с их родственными

промоторами поддерживается и облегчается наличием когезина и CTCF (CCCTC-binding factor). Когезин и CTCF играют ключевую роль в организации ТАДов и показывают высокий уровень связывания в их граничных областях. Нокаут или нокаун когезина, уменьшает плотность взаимодействий внутри ТАДов, а понижение уровня CTCF приводит также и к повышению количества взаимодействий между разными ТАДами.

Также в хроматине часто наблюдаются петли промотор-энхансер, связанные с активацией генов. На формирование этих петель оказывают влияние как CTCF, так и когезиновый комплекс, частью которого является RAD21.

---

## Результаты и обсуждение

---

### Общая информация о модели

Данные с сайта RCSB были проанализированы и преобразованы в таблицу:

Таблица 1. Общая информация о структуре 6RRC.

Состав комплекса	<ul style="list-style-type: none"><li>• Субъединица когезина SA-1 две цепи (A , C)</li><li>• Гомолог белка репарации двухцепочечного разрыва rad21 две цепи (B , D)</li><li>• Лиганд SO4</li></ul>
Год	2019
Авторы	Newman, J.A., Katis, V.L., von Delft, F., Arrowsmith, C.H., Edwards, A., Bountra, C., Gileadi, O.
Статья	Lelij, Petra van der, Joseph A. Newman, Simone Lieb, Julian Jude, Vittorio Katis, Thomas Hoffmann, Matthias Hinterdorfer, et al. 2020. "STAG1 Vulnerabilities for Exploiting Cohesin Synthetic Lethality in STAG2-Deficient Cancers." Life Science Alliance 3 (7). doi: 10.26508/LSA.202000725.
Метод решения фазовой проблемы	Молекулярное замещение

Число измеренных рефлексов	46073
Разрешение	2.37 Å
Полнота набора структурных факторов	99.9%
Параметры кристаллографической ячейки	a = 117.077 Å, b = 151.662 Å, c = 126.656 Å $\alpha = 90.00^\circ$ , $\beta = 90.0^\circ$ , $\gamma = 90.00^\circ$
Кристаллографическая группа	C 2 2 2

### Значения индикаторов качества модели в целом

Параметры для оценки качества полученной модели представлены ниже.

- R-Value Free (тестовая выборка): 0.248
- R-Value Work (тренировочная выборка): 0.214
- R-Value Observed (весь набор рефлексов): 0.215

R-фактор характеризует соответствие модели ее экспериментальным значениям и равен 0.214 (хорошие значения  $< 0.25$ ).

R-free характеризует контроль модели и равен 0.248. Значение чуть больше допустимого, но я не считаю это критичным (хороший R-free  $< 0.2$ ).

Также оценка структуры проводилась с помощью сервера MolProbity.

Чтобы улучшить оценку на самом сервере были добавлены водороды и произведена оптимизация связей.

Таблица 2. Оценка качества структуры 6RRC с сервера MolProbity.

All-Atom Contacts	Clashscore, all atoms:	2.88		100 <sup>th</sup> percentile* (N=334, 2.37Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	19	3.11 %	Goal: <0.3%
	Favored rotamers	546	89.51 %	Goal: >98%
	Ramachandran outliers	2	0.29 %	Goal: <0.05%
	Ramachandran favored	662	97.64 %	Goal: >98%
	Rama distribution Z-score	-1.43 ± 0.28		Goal: abs(Z score) < 2
	MolProbity score <sup>^</sup>	1.53		99 <sup>th</sup> percentile* (N=8020, 2.37Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00 %	Goal: 0
	Bad bonds:	0 / 5674	0.00 %	Goal: 0%
	Bad angles:	0 / 7646	0.00 %	Goal: <0.1%

По карте Рамачандрана были найдены 2 аутлаера: C:192:Ser и A:158:Asp. 97,6% всех остатков находились в предпочтительных

(98%) областях. 99,7% всех остатков находились в разрешенных (> 99,8%) областях.

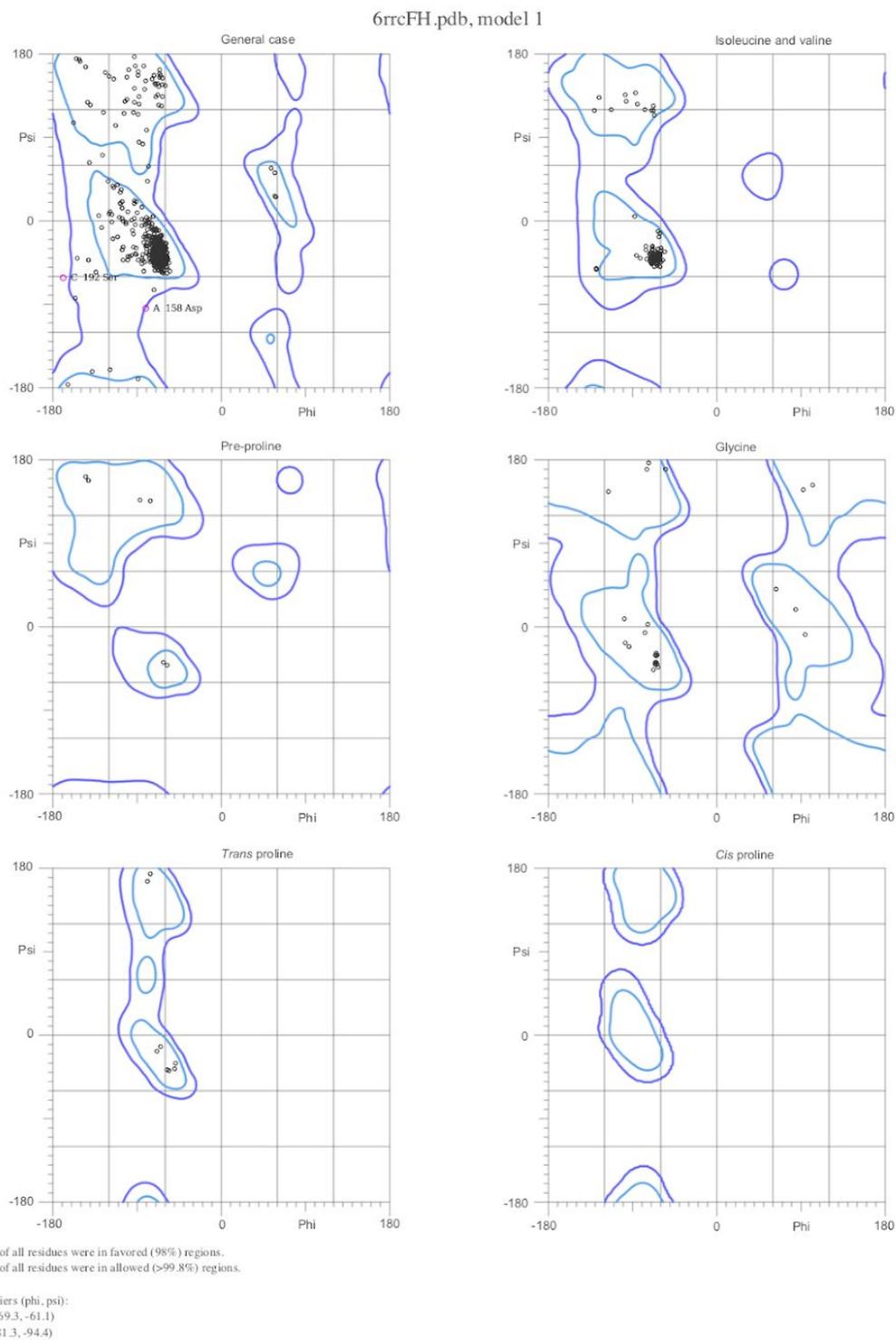


Рисунок 2. Карта Рамачандрана (MolProbity)

### Анализ маргинальных остатков.

Таблица 3. Список маргинальных остатков по различным параметрам.

Остаток	Показатель
C:192:Ser	Аутлаер по карте Рамачандрана Неротамер
A:158:Asp.	Аутлаер по карте Рамачандрана Неротамер RSRZ outliers = 4.2
A:227:MET	Перекрывается с A:314:GLU A:227:MET:HB3 - A:314:GLU:HG2 - 0.640
D:322:ARG	Перекрывается с A:152:THR A:152:THR:O - D:322:ARG:NH2 - 0.592 Перекрывается с A:153:GLU A:153:GLU:HA - D:322:ARG:HH21 - 0.466 Неротамер
D:334:SER	Перекрывается с D:338:ARG D:334:SER:O - D:338:ARG:N - 0.513 Неротамер
C:143:GLN	Неротамер
B:335:LYS	RSRZ = 5.2
A:148:ILE	RSRZ = 4.9
A:150:LYS	RSRZ = 4.6

В случае плохого перекрывания атомов наблюдаемые значения должны быть меньше или равны 0.4 ангстремам.

На следующем этапе работы было принято решение рассмотреть 5 маргиналов более подробно.

### **C:192:Ser**

При рассмотрении этого маргинала было решено обратить внимание на его электронную плотность ( $\text{carve} = 2$ ,  $\text{level} = 0.5$ ). Этот остаток был выявлен как маргинал на карте Рамачандрана (рисунок 2). Вероятно его неротамерность и то, что он был сочтен маргиналом связано с неточной или ошибочной расшифровкой электронной плотности.

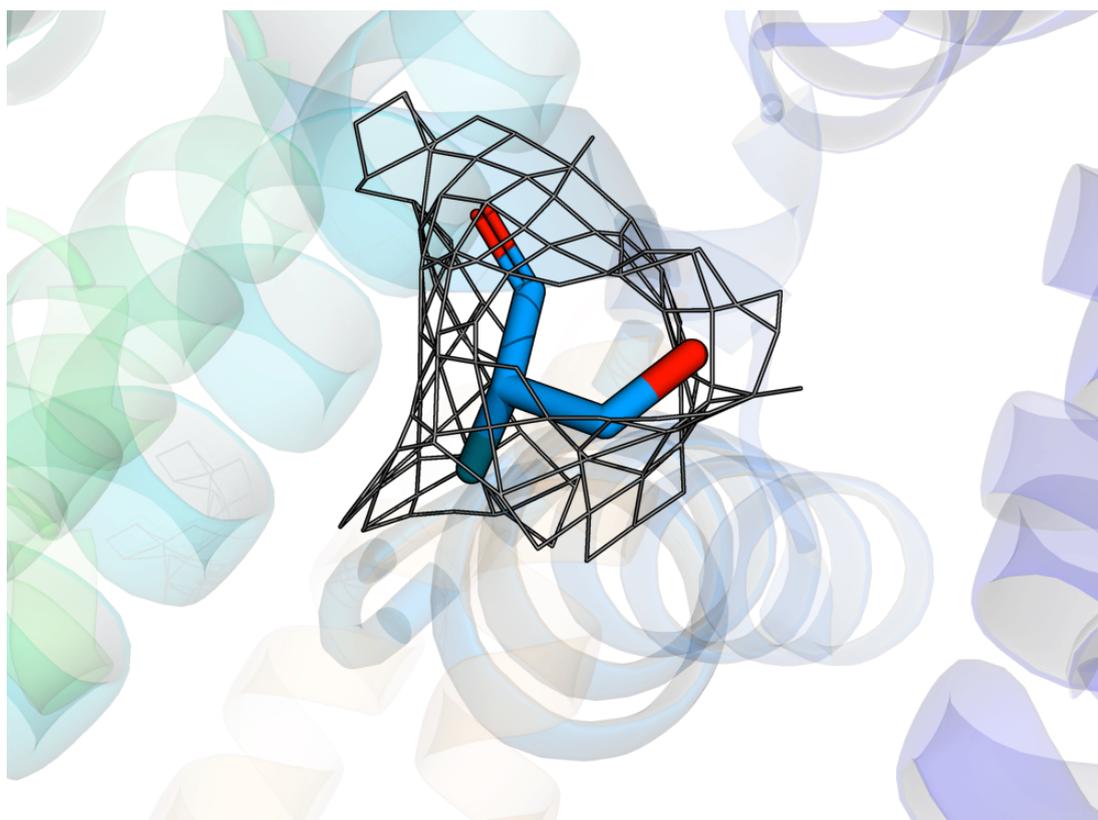


Рисунок 3. Электронная плотность Ser 192

Было решено, проверить, связана ли подвижность остатка с тем, что он плохо вписан в электронную плотность. Была произведена окраска по B-фактору, которая выявила, что данный маргинал не относится к числу подвижных. Следовательно, можно сделать

Вывод о том, что электронная плотность довольно низкого качества и не зависит от подвижности.

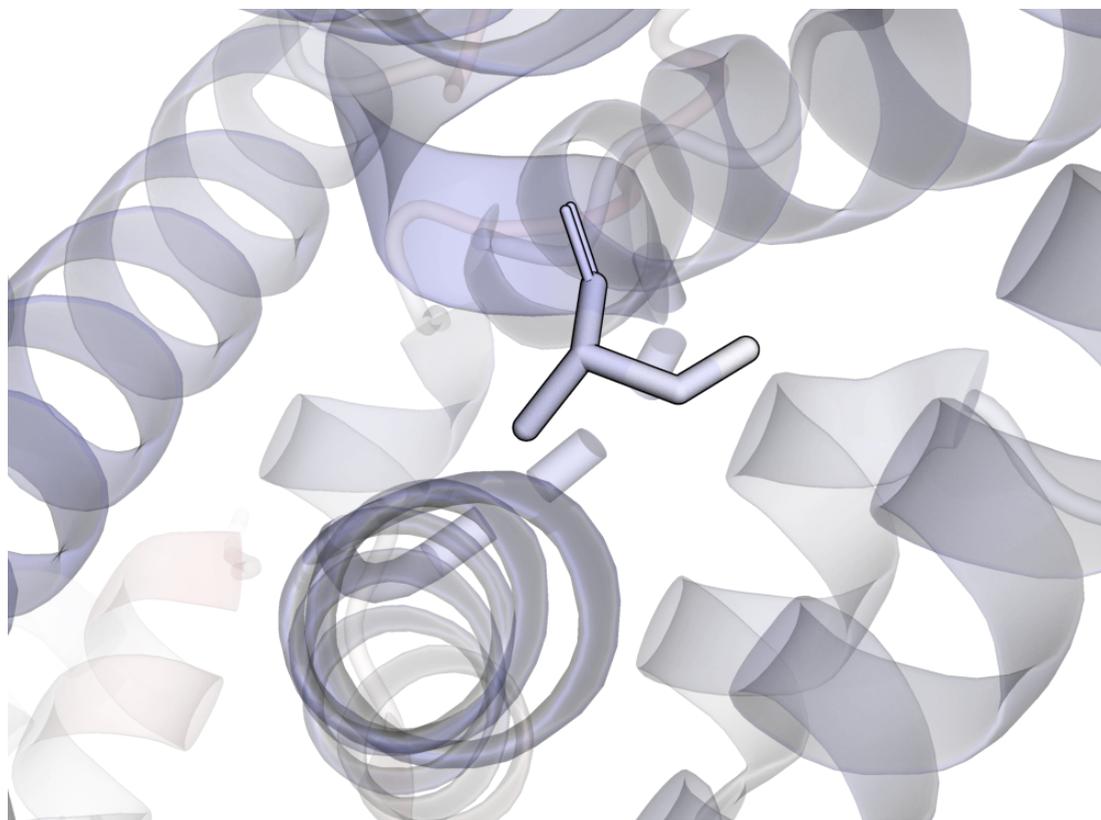


Рисунок 4. Ser 192, окрашенный по B-фактору

#### **A:158:Asp.**

Этот остаток был выбран из-за его слишком большого значения RSRZ, равняющегося 4.2. Такое значение означает, что остаток плохо вписан в электронную плотность.

На рисунке 5 мы видим подтверждение плохой вписанности остатка в электронную плотность (уровень подрезки 0.5 и  $\sigma_{cut} = 2$ ). Это происходит по причине, что остаток очень подвижный (что выявляется при покраске по B-фактору), т.к. находится в петле на периферии белковой глобулы. Данный факт и объясняет такое значение RSRZ.

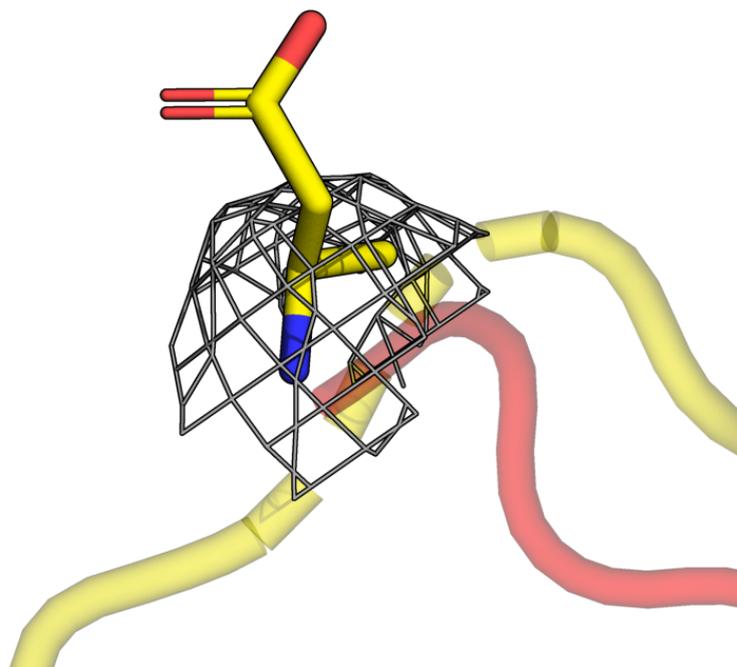


Рисунок 5. Электронная плотность Asp 158

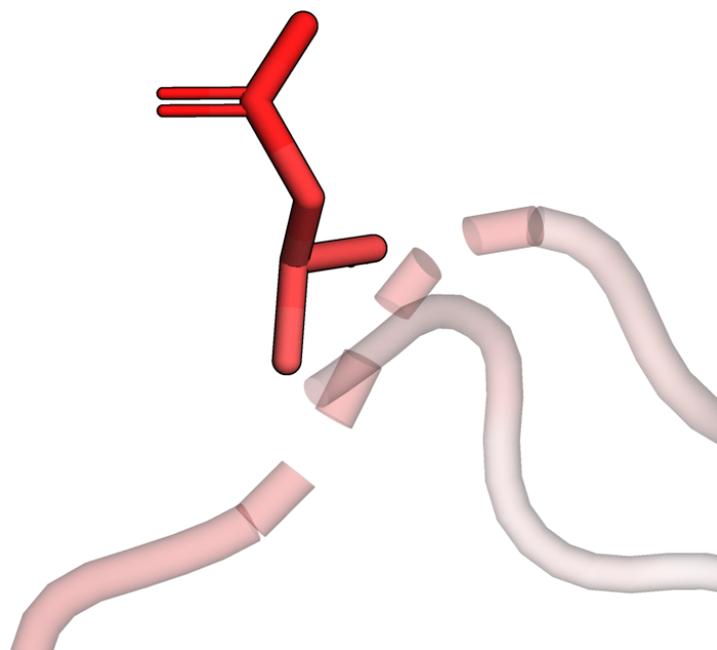


Рисунок 6. Asp 158, окрашенный по B-фактору

### **A:227:MET**

Даже невооруженным взглядом можно заметить, что атомы НВ3 и НG2 находятся очень близко друг к другу, из-за чего и происходит их перекрывание. Однако такое расположение Met 227 обусловлено тем, что он находится в альфа-спирали – стабильной и жесткой структуре.

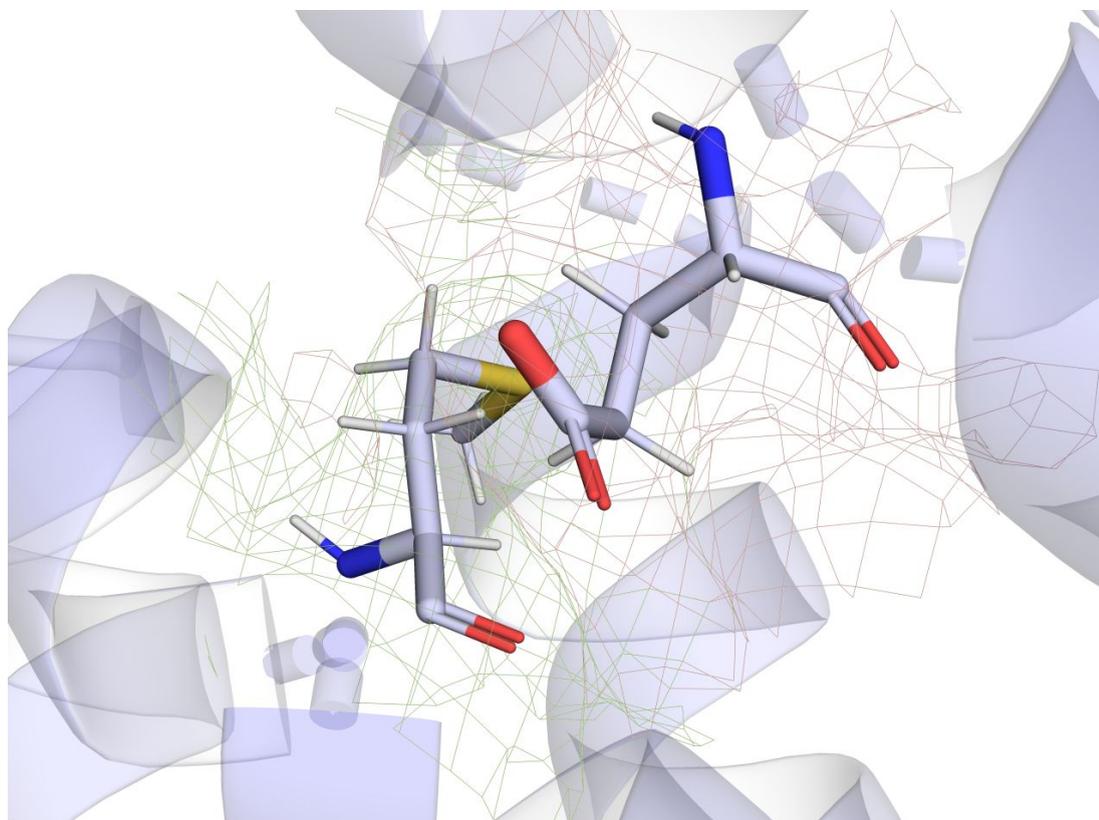


Рисунок 7. Наложение электронной плотности Met 227 и Glu 314

### **D:322:ARG**

Как и в предыдущем случае, здесь происходит перекрывание электронных плотностей. Этот аминокислотный остаток является неротамером, чем можно объяснить подобный эффект наложения электронных плотностей.

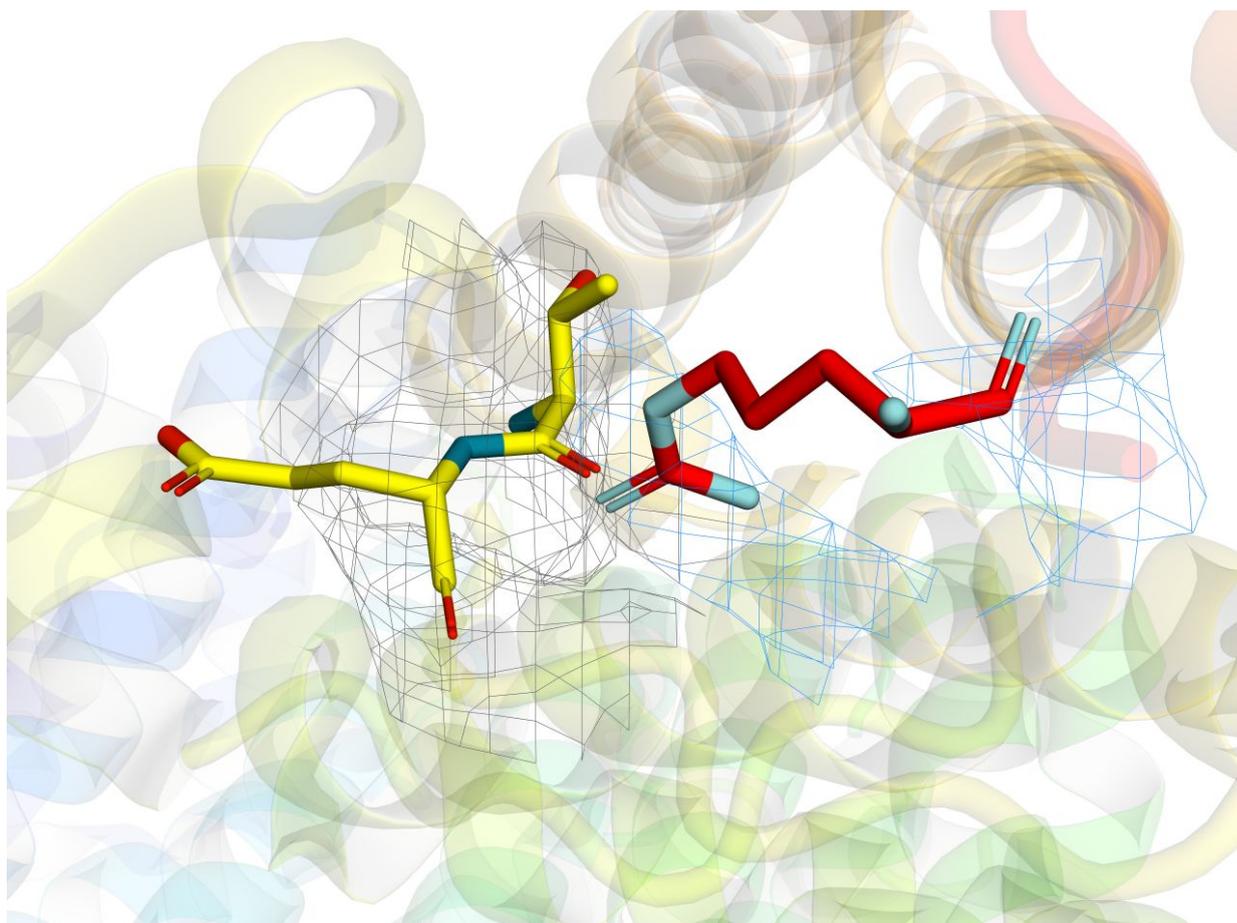


Рисунок 7. Наложение электронной плотности Arg 322 и Thr 152 и Glu 153

### **B:335:LYS**

Этот остаток был выбран для рассмотрения в связи с его большим значением  $RSRZ = 5.2$ . Как уже было упомянуто ранее, большое значение данного параметра означает, что остаток будет плохо вписываться в электронную плотность. Такое может происходить потому что данный остаток является одним из концевых остатков цепи, удаленных от центра молекулы белка.

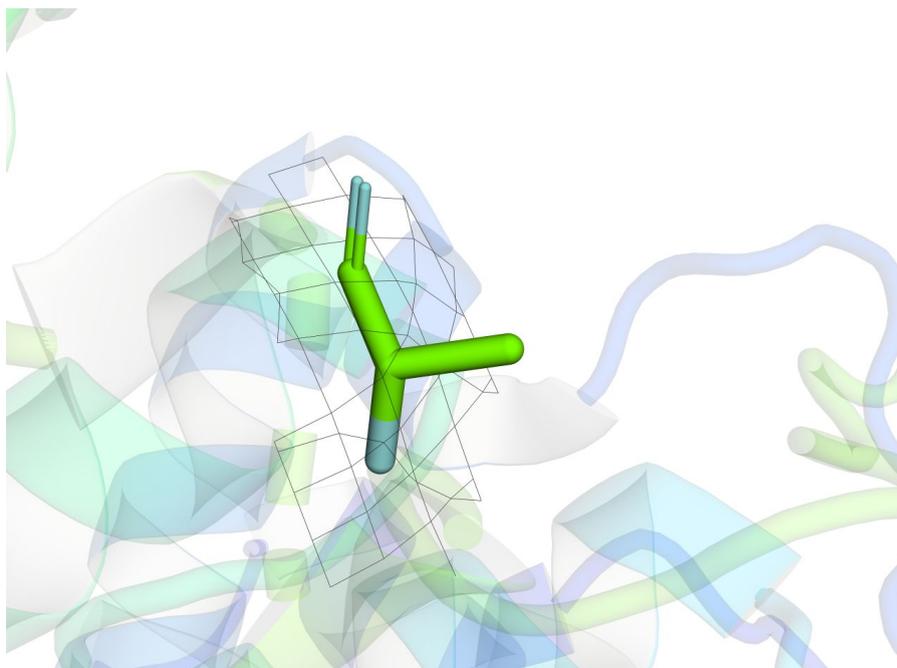


Рисунок 5. Электронная плотность Lys 335

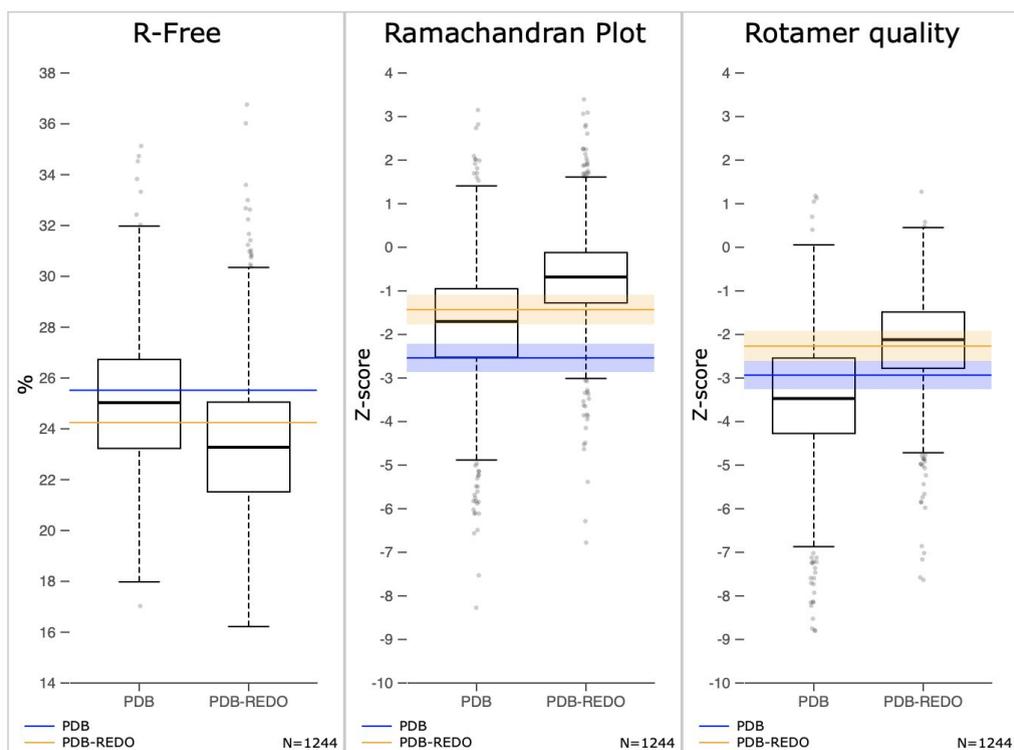
### **Сравнение модели PDB и PDB\_redo**

PDB\_redo – это база данных, где хранятся оптимизированные модели структур из PDB.

После оптимизации нашей структуры значение R-фактора почти не изменилось (было 0.214, а стало 0.22). R-free уменьшился (было 0.248, а стало 0.242). Также были изменены 12 ротамеров, перевернуты 16 цепей и 3 пептида, удалено 26 молекул воды.

С помощью PDB\_redo было произведено некое улучшение структуры, однако нельзя назвать его существенным.

В рассмотренных маргиналах изменений не произошло.




---

## Заключение

---

Несмотря на большее количество маргинальных остатков, структура тем не менее обладает хорошим разрешением. Она дает нам представление о том, какую структуру имеет N-концевая область субъединицы когезина человека STAG1 в комплексе с пептидом RAD21.

---

## СПИСОК ЛИТЕРАТУРЫ

---

1. Lelij, Petra van der, Joseph A. Newman, Simone Lieb, Julian Jude, Vittorio Katis, Thomas Hoffmann, Matthias Hinterndorfer, et al. 2020. "STAG1 Vulnerabilities for Exploiting Cohesin Synthetic Lethality in STAG2-Deficient Cancers." *Life Science Alliance* 3 (7). doi: 10.26508/LSA.202000725.
2. Сервис PDB\_redo: <https://pdb-redo.eu>
3. Сервис MolProbity: <http://molprobity.biochem.duke.edu/>
4. Запись PDB ID: <https://www.rcsb.org/structure/6RRC>