

**Московский государственный университет им. М.В.Ломоносова Факультет  
биоинженерии и биоинформатики**

**Отчёт по качеству расшифровки структуры белка 1RDQ,  
полученной с помощью рентгеноструктурного анализа**

Выполнил:  
студент 4-го курса  
Горбатенко Владислав

**Москва, 2018**

## Аннотация

В данном отчете была рассмотрена структура 1RDQ каталитической субъединицы Y204A мутанта мышинной цАМФ-зависимой протеинкиназы, полученная с помощью рентгено-структурного анализа, и проведен анализ её качества. Структура была опубликована в 2004 году группой авторов - Jie Yang, Lynn F. Ten Eyck, Nguyen-Huu Xuong и Susan S. Taylor [1].

## Введение

ЦАМФ-зависимая протеинкиназа (или протеинкиназа А, пКА) — фермент класса трансфераз, осуществляющий перенос фосфата с АТФ на различные субстраты, тем самым регулируя их активность. Молекула пКА представляет собой тетрамер, состоящий из двух каталитических и двух регуляторных субъединиц. В таком состоянии фермент неактивен. Активация происходит за счет связывания цАМФ регуляторными субъединицами и последующей диссоциации всех четырех.

Целью авторов статьи [1] было изучение вклада структуры под названием «P+1 loop» (198-205 остатки), которая опосредует коммуникации между субстратом и каталитическими остатками. Для этого был получен мутант каталитической субъединицы фермента, в котором (предположительно, играющий важную роль) Тут в позиции 204 заменен на Ala, и определена его структура с помощью РСА (рис 1).

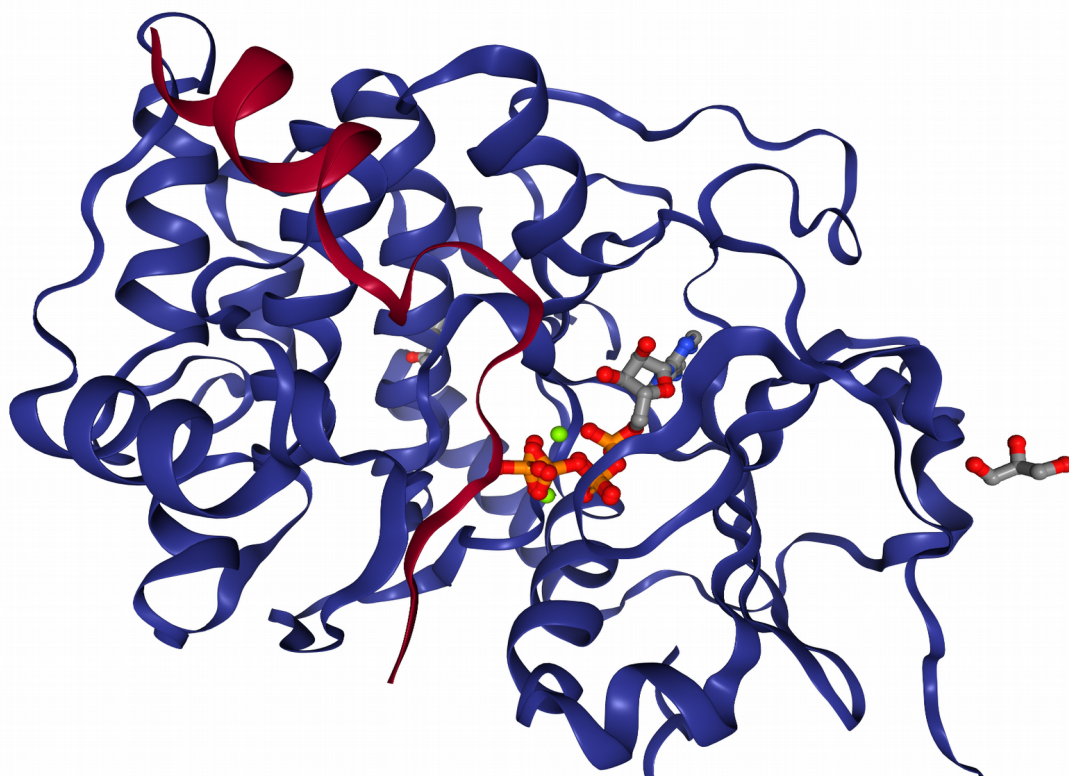


Рис 1. Структура 1RDQ. Синим цветом окрашена сама каталитическая субъединица, бардовым — ингибиторный пептид. В центре можно увидеть серо-красно-оранжевую молекулу АТФ и два зеленых иона  $Mg^{2+}$ . Также в структуре присутствуют молекулы MPD (на заднем плане) и глицерола (справа), входящие в состав раствора кристаллизации.

## Результаты и обсуждение

### Общая информация о модели

Биологическая единица совпадает с ассиметрической и представляет гетеродимер, состоящий из каталитической субъединицы фермента (цепь — E, 350 аминокислот) и ингибиторного пептида (цепь — I, 20 аминокислот). Кроме того, в комплекс входят 2 иона  $Mg^{2+}$ , АТФ в 26% молекул или продукты гидролиза АТФ (АДФ + фосфат) в 74% молекул и входящие в состав раствора кристаллизации MPD (2-метилпентан-2,4-диол) и глицерол.

Разрешение модели — 1,26 Å. Диапазон разрешений — 1.26-50 Å. Число измеренных рефлексов — 122674; полнота набора структурных факторов — 91.4%. Метод решения фазовой проблемы — молекулярное замещение (гомологичная структура — 1АРМ).

Параметры кристаллографической ячейки:

- Длины направляющих векторов ячейки:  $a = 57.760\text{Å}$  ,  $b = 79.513\text{Å}$  ,  $c = 97.739\text{Å}$ ;
- Углы между направляющими векторами:  $\alpha = \beta = \gamma = 90^\circ$ ;
- Кристаллографическая группа:  $P 2_12_12_1$  .
- Некристаллографическая симметрия присутствует.

### Индикаторы качества модели

В первую очередь, для характеристики соответствия модели экспериментальным данным используют R-фактор и R-free:  $R = 13.2\%$ ;  $R\text{-free} = 16.2\%$ . R-free построен на основе ~5% (5630 из 122674) рефлексов. То, что оба фактора меньше 20% говорит о хорошем соответствии модели эксперименту, а то, что их разность меньше 10% указывает на отсутствие переоптимизации. Эти данные можно найти в pdb-файле или на странице, посвященной экспериментальным данным структуры на сайте RCSB [2].

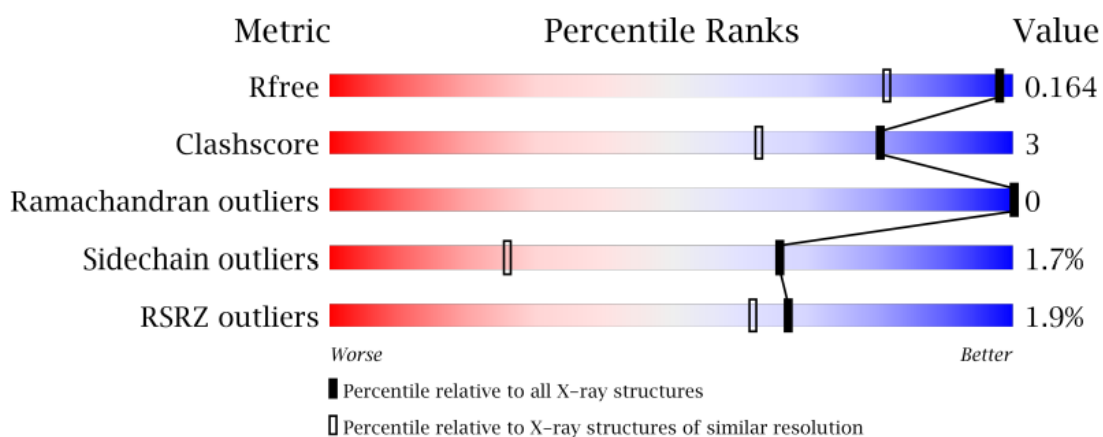


Рис 2. Percentile plot для структуры 1RDQ. Градиент цвета указывает на то, лучше (синий) или хуже (красный) параметры модели относительно средних значений. Закрашенные прямоугольники показывают качество относительно всех PCA структур, прозрачные — относительно структур с аналогичным разрешением.

В полном отчете по валидации [3] есть Percentile plot (рис 2), который отражает качество модели относительно других моделей PDB. В целом, показатели модели среднего (даже выше среднего) качества. Количество маргинальных остатков по параметру RSRZ – 1.9%; по конформации боковой цепи — 1.7%. Эти значения много меньше 10% и, следовательно, довольно хорошие, даже несмотря на то, что у структур с аналогичным разрешением маргиналов по конформации боковой цепи меньше.

Значения других параметров были получены с помощью сервиса MolProbity [4]. При этом были добавлены водороды и инвертированы боковые цепи остатков 36ASN, 62HIS, 113ASN, 181GLN B, 245GLN, 283ASN.

Таблица 1. Результат анализа контактов и геометрии всех атомов 1RDQ сервисом MolProbity.

Clashscore, all atoms:	2.88		97 <sup>th</sup> percentile* (N=365, 1.26Å ± 0.25Å)
Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Poor rotamers	9	2.97%	Goal: <0.3%
Favored rotamers	286	94.39%	Goal: >98%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	342	97.71%	Goal: >98%
MolProbity score <sup>^</sup>	1.49		78 <sup>th</sup> percentile* (N=2027, 1.26Å ± 0.25Å)
Cβ deviations >0.25Å	0	0.00%	Goal: 0
Bad bonds:	10 / 3074	0.33%	Goal: 0%
Bad angles:	14 / 4175	0.34%	Goal: <0.1%
Cis Prolines:	0 / 13	0.00%	Expected: ≤1 per chain, or ≤5%

Clashscore показывает, что в данной структуре всего 3 атома на 1000 имеют недопустимые наложения (Ван-дер-Ваальсовы радиусы пересекаются больше, чем на 0.4Å), что лучше, чем у 97% структур с аналогичным разрешением (1.26Å ± 0.25Å).

В структуре содержатся 9 плохих ротамеров (~3% от общего числа), что больше порога (<0.3%), свойственного хорошим моделям. В предпочитаемых конформациях находятся 94.39% ротамеров. Это тоже хуже, чем у хороших структур (>98%).

Маргиналы по карте Рамачандрана (рис 3) отсутствуют (на Percentile plot это тоже отображено). В предпочитаемой области находятся 97.71% остатков (это почти хорошо; порог >98%). В их числе изолейцины, валина, транс-пролины, препролиновые остатки и почти все глицины (один располагается в допустимой области). Цис-пролины отсутствуют.

MolProbity score, объединяющий три описанные выше оценки в одну, говорит о том, что эта модель является собой 78-й перцентиль, то есть, что она лучше 78% моделей с аналогичным разрешением.

Кроме того, в модели отсутствуют Cβ-атомы с неприемлимым отклонением от ожидаемого положения, но присутствуют 10 ковалентных связей и 14 валентных углов, существенно отклоняющихся от теории.

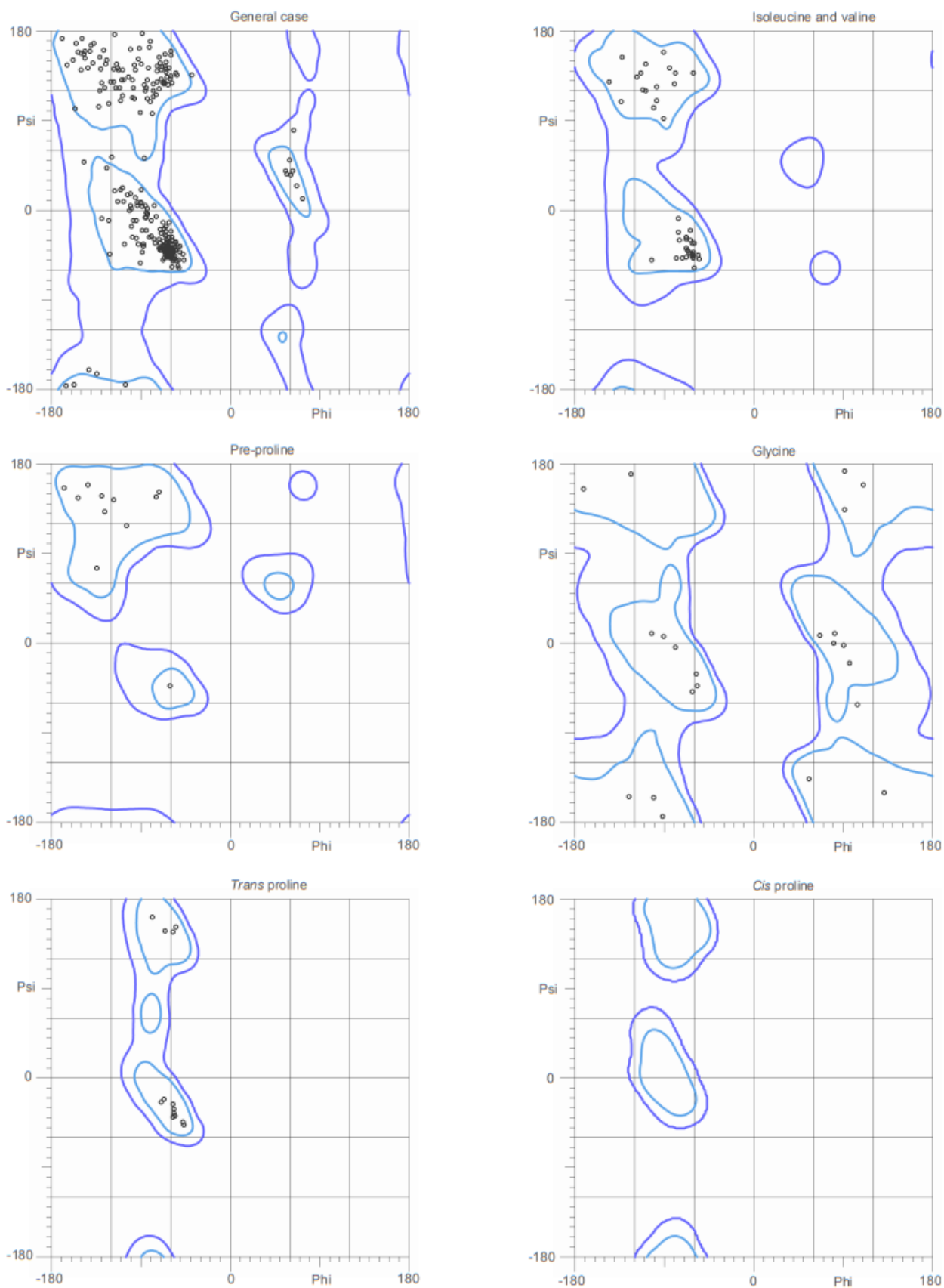


Рис 3. Карта Рамачандрана структуры 1RDQ, построенная сервисом MolProbity. Границы допустимых торсионных углов обозначены синим, предпочитаемых — голубым.

## Маргинальные остатки

Таблица 2. Список нескольких маргинальных остатков каталитической субъединицы (цепь E) 1RDQ.

№	Аминокислотный остаток	Показатели маргинальности
1	GLN 245	Плохой ротамер: $\chi = 298.1, 130.6, 224.6$ ; Clash $0.41\text{\AA}$ OE1-атом с HB2-атомом GLN 242;
2	HIS 62	Плохой угол: CB-CG-CD2
3	ILE 135	Clash $0.66\text{\AA}$ HG23-атом с O-атомом HIS 131; Clash $0.45\text{\AA}$ CD1-атом с HA-атомом LEU 132; Плохой ротамер: $\chi = 95.1, 295$
4	SEP 338	Clash $0.68\text{\AA}$ HB3-атом с NH2 ARG 336; Плохой угол: CB-OG-P; Длина связи: O3P - P
5	LYS 105	Плохой ротамер: $\chi = 189.1, 126.8, 71.3, 125.4$

## Анализ маргинальных остатков

### 1. Gln-245: плохой ротамер, недопустимое наложение (на Gln-242) атомов.

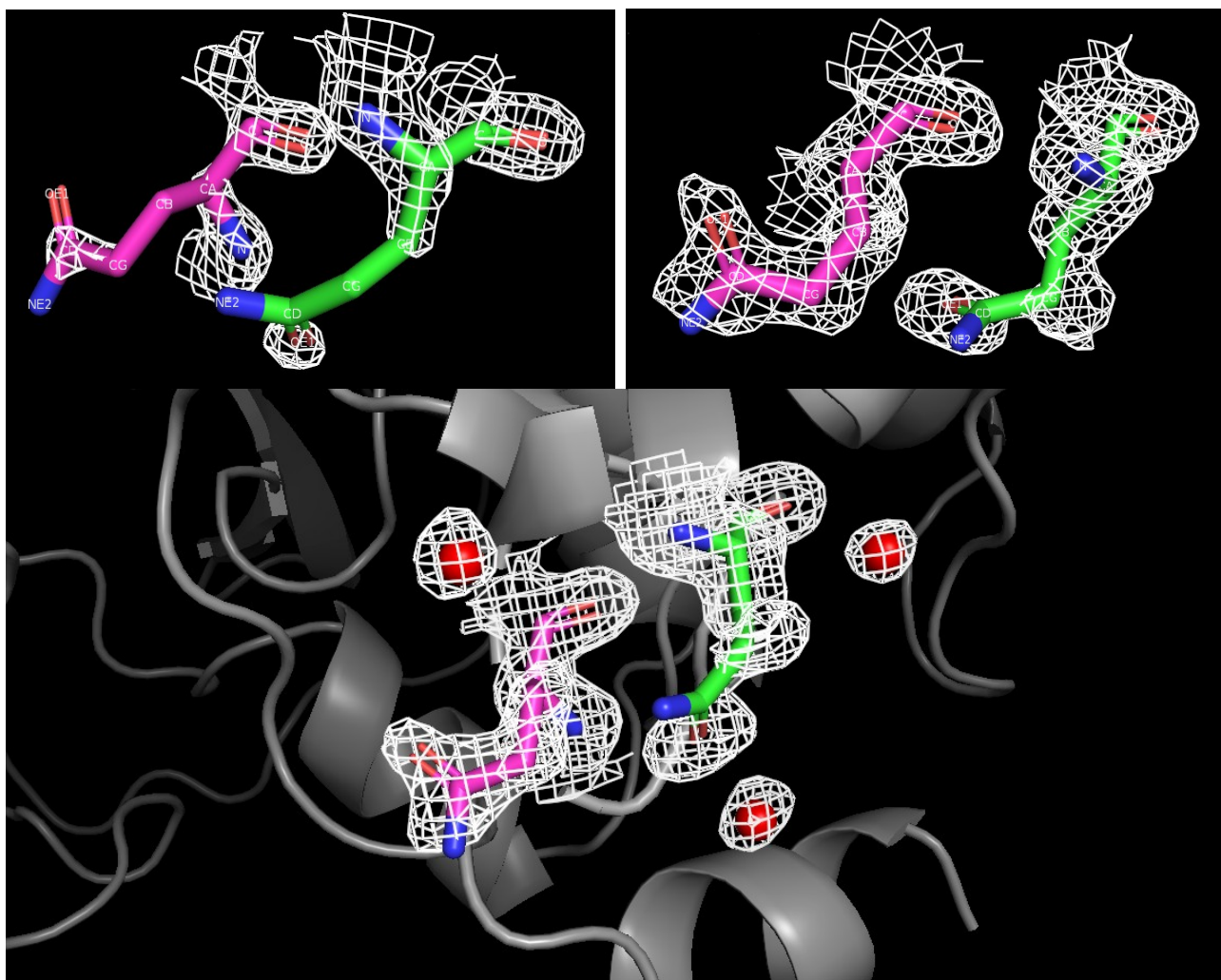


Рис 4. Gln-245 (зеленый) и Gln-242 (пурпурный). Сверху слева отображена электронная плотность с уровнем подрезки 2, справа и снизу — с уровнем подрезки 1. Снизу глутамины отображены в

структуре и с окружающими их кислородами молекул воды.

Этот глутамин является маргинальным по значениям углов  $\chi$  боковой цепи и по недопустимому атомному перекрытию с Gln-242. По рисунку 4 видно, что эта пара аминокислот находится близко к поверхности (Gln-242 в петле, Gln-245 в бета тяже). Поскольку Gln-242 находится в предпочитаемой конформации по данным MolProbity, удобно будет сравнивать Gln-245 с ним.

При уровне подрезки ЭП 2 видно, что в обеих молекулах ЭП сосредоточена у сильно электроотрицательных атомов, при этом ЭП амидной группировки у Gln-242 расположена между кислородом и азотом, а у Gln-245 смещена к кислороду. ЭП с уровнем подрезки 1 выглядит еще интересней: отчетливо видно, что расположение структуры Gln-245 не совпадает с расположением ЭП (при этом Gln-242 полностью покрыт ею). Кроме того, амидный кислород Gln-245 образует водородную связь с находящейся рядом молекулой воды (длина 3.4Å). При этом, электронная плотность соответствующая амидной группе располагается ближе к этой молекуле воды, чем сама амидная группа.

Все это дает основания полагать, что эти маргинальные значения углов  $\chi$ , а также вызванное ими недопустимое наложение атомов, вызваны ошибками расшифровки, нежели каким-то структурными особенностями.

## 2. His-62: плохой валентный угол.

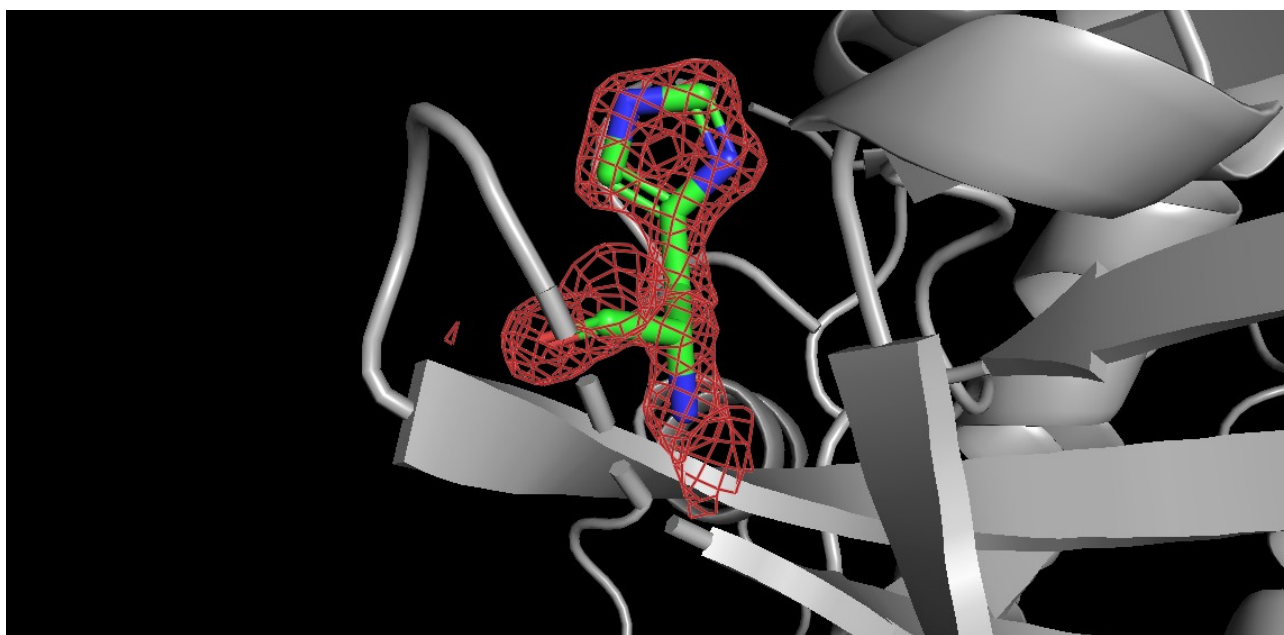


Рис 5. His-62 покрытый электронной плотностью с уровнем подрезки 2.

Гистидин на рисунке 5 является маргиналом по углу CB-CG-CD2. На рисунке 5 видно, что ЭП идеально покрывает молекулу, так что, видимо, данная маргинальность является особенностью этого гистидина.

### 3. Ile-135: плохой ротамер, недопустимое наложение атомов (His 131; Leu 132)

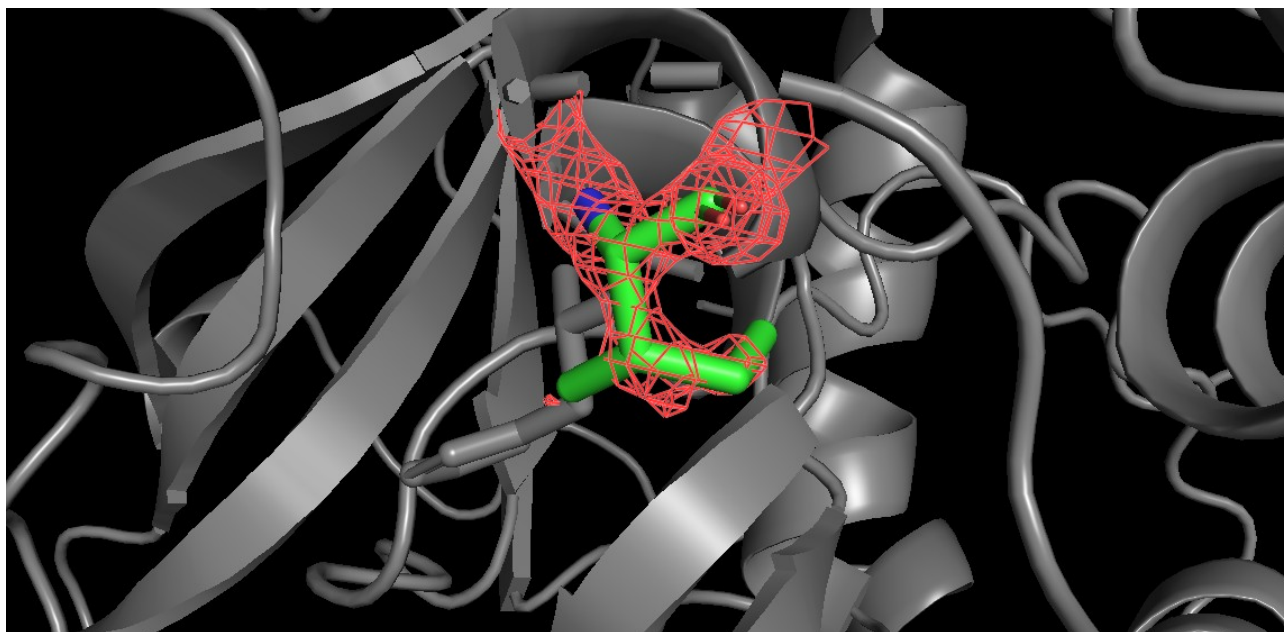


Рис 6. Ile-135 с электронной плотностью на уровне подрезки 2.

На рисунке 6 показан маргинальный остаток изолейцина. Хорошо видно, электронная плотность почти не покрывает CG1 атом и совсем не покрывает CG2 и CD атомы. Скорей всего, это означает, что были допущены ошибки в расшифровке, которые повлияли на значения торсионных углов и, как следствие, на образование недопустимого перекрывания с атомами близлежащих аминокислот: His-131, Leu-132.

Кроме того, следующий за Ile-135 остаток — Gly-136 на карте Рамачандрана располагается не в предпочтительной области, а в допустимой (единственный из всех глицинов). Вероятно, это тоже вызвано влиянием изолейцина.

### 4. Ser-338: выбросы по валентному углу, длине связи и недопустимое перекрывание.

На рисунке 7 изображен фосфосерин. Это и объясняет его маргинальность. Как видно из рисунка, электронная плотность отлично покрывает молекулу, поэтому вряд ли там есть ошибки расшифровки. Если приглядеться к параметрам маргинальности (табл. 2), то можно заметить, что все они связаны с образованием эфирной связи между остатком серина и фосфатом. Выбросы по CB-OG-P углу и длине O3P-P связи связаны напрямую, а уже перекрывание атомов с близлежащим гистидином вследствие изменения структуры аминокислоты.



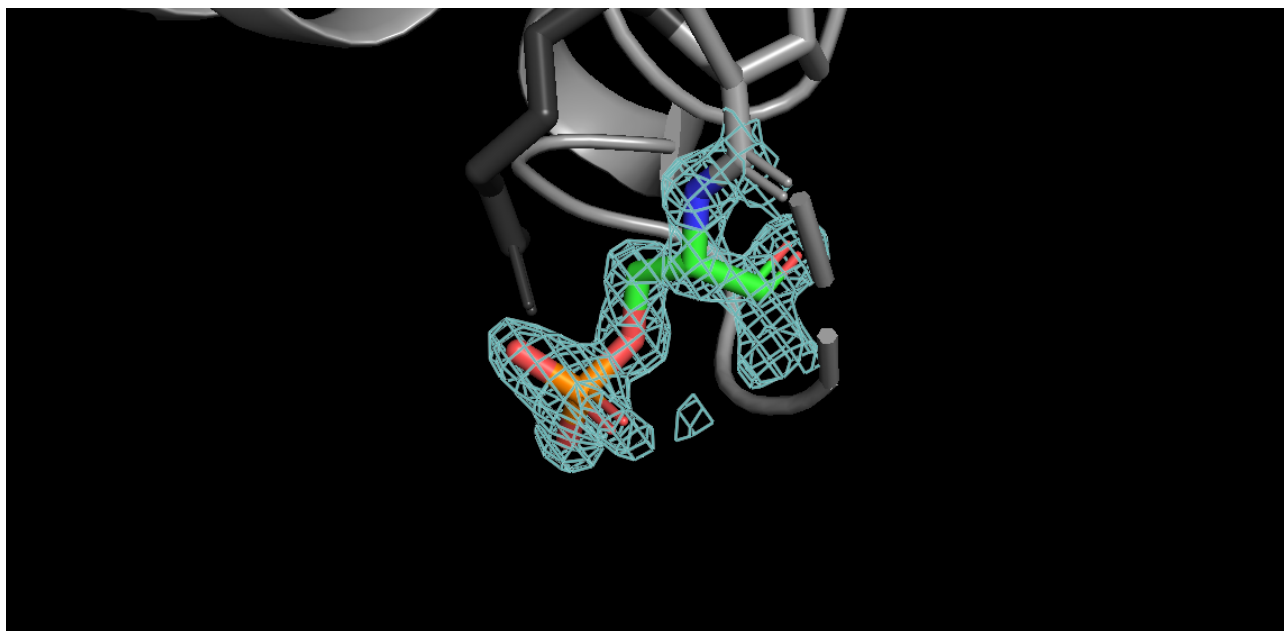


Рис 7. Ser-338 и электронная плотность с уровнем подрезки 2.

### 5. Lys-105: плохой ротамер.

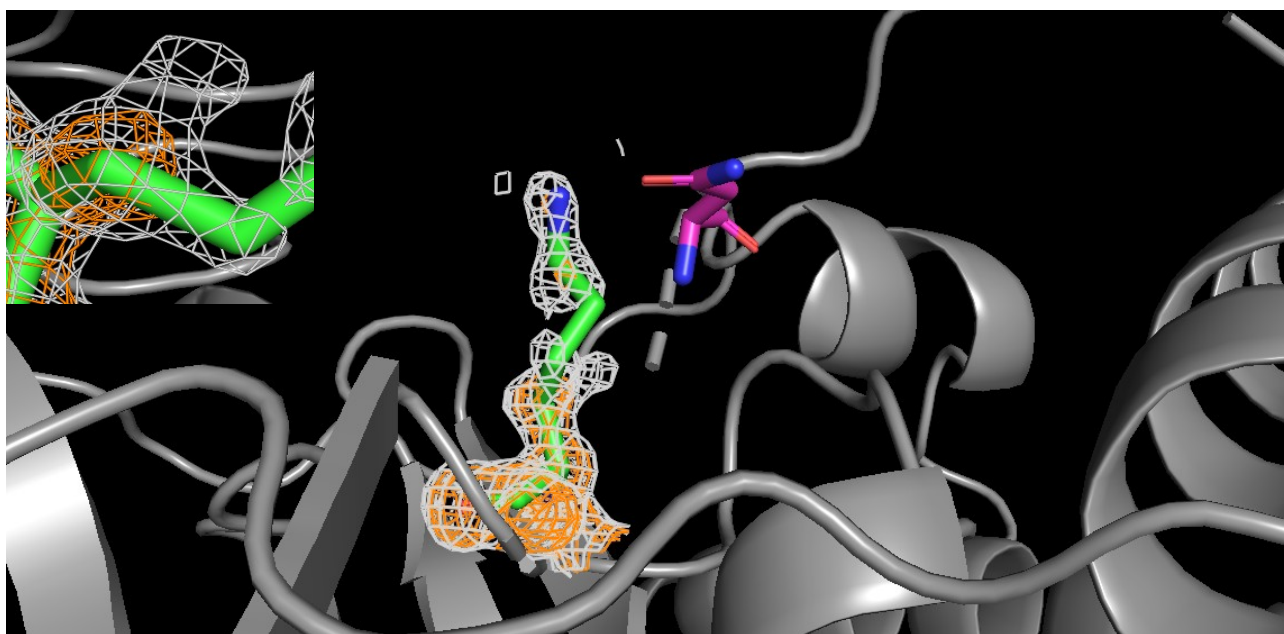


Рис 8. Lys-105 и ЭП с уровнем подрезки 2 (оранжевая) и с уровнем подрезки 1 (белая). Фиолетовым изображен остаток Asn-2. Слева сверху - ЭП вокруг СВ и CG атомов.

На рисунке 8 представлен 105-й лизин, боковая цепь которого находится в непосредственной близости к Asn-2 N-конца. Маргинальный этот лизин по значению торсионных углов боковой цепи, а конкретно, от табличных данных [5] отличаются углы  $\chi_2$  и  $\chi_4$ . Если ЭП с уровнем подрезки 2 покрывает только остов и СВ атом, то ЭП с уровнем подрезки 1 покрывает почти все. Интересную структуру, похожую на рогатину, образует ЭП вокруг СВ и CG атомов. Торсионный угол вокруг связи между этими атомами — как раз таки  $\chi_2$ .

Я полагаю, что в данном случае маргинальность отражает особенности проведения эксперимента. Раствор кристаллизации имеет рН=8. Это больше изоэлектрической точки Аспарагина и почти равно рКа протонирования азота амидной группы. В

результате, положительно заряженный лизин меняет конформацию, чтобы подальше отстраниться от положительно заряженного аспарагина (так и остается в кристалле). Но, поскольку рН раствора кристаллизации все таки меньше рКа аспарагина, остаются и нейтральные остатки с  $\Delta^-$  на амидном кислороде. А в таком случае боковая цепь лизина будет ближе к аспарагину, чем первом случае. В итоге в кристалл попадут как молекулы с первым вариантом конформации (в большей степени), так и молекулы со вторым вариантом. Это объясняет «рогатину» из ЭП вокруг СВ-СГ связи.

## ***PDB\_redo***

Из банка PDB\_redo [6] была взята оптимизированная структура 1RDQ после чего по ней был проведен анализ в MolProbity.

Таблица 3. Показатели качества исправленной модели 1RDQ

Clashscore, all atoms:	4.42		90 <sup>th</sup> percentile * (N=365, 1.26Å ± 0.25Å)
Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Poor rotamers	5	1.56%	Goal: <0.3%
Favored rotamers	306	95.33%	Goal: >98%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	343	98.00%	Goal: >98%
MolProbity score <sup>^</sup>	1.37		88 <sup>th</sup> percentile * (N=2027, 1.26Å ± 0.25Å)
Cβ deviations >0.25Å	0	0.00%	Goal: 0
Bad bonds:	12 / 3140	0.38%	Goal: 0%
Bad angles:	7 / 4250	0.16%	Goal: <0.1%
Cis Prolines:	0 / 13	0.00%	Expected: ≤1 per chain, or ≤5%

В таблице 3 представлены показатели качества исправленной структуры. Не могу сказать, что они стали сильно лучше. В основном, изменения довольно незначительные.

Из улучшений:

- Уменьшилось количество плохих ротамеров;
- Увеличилось количество ротамеров в предпочтительной конформации;
- Увеличилось число остатков в предпочтительной области карты Рамачандрана;
- Уменьшилось количество плохих углов

И следовательно улучшилась интегральная оценка MolProbity.

Ухудшения:

- Увеличилось количество недопустимых наложений атомов;
- Увеличилось количество плохих связей.

Количество маргиналов при этом увеличилось.

## **Выводы:**

Структура 1RDQ имеет высокое разрешение и довольно хорошо соответствует результатам РСА. Параметры качества достаточно высоки. В модели присутствует некоторое количество маргинальных остатков. Часть отражает особенности структуры или эксперимента.

Поскольку модель изначально высокого качества, пользоваться оптимизированной структурой не имеет смысла.

## **Список литературы:**

[1] *Jie Yang, Lynn F. Ten Eyck, Nguyen-Huu Xuong and Susan S. Taylor. Crystal Structure of a cAMP-dependent Protein Kinase Mutant at 1.26 Å: New Insights into the Catalytic Mechanism. J. Mol. Biol. (2004) 336, 473–487*

[2] <https://www.rcsb.org/pdb/explore/materialsAndMethods.do?structureId=1RDQ>

[3] [https://files.rcsb.org/pub/pdb/validation\\_reports/rd/1rdq/1rdq\\_full\\_validation.pdf](https://files.rcsb.org/pub/pdb/validation_reports/rd/1rdq/1rdq_full_validation.pdf)

[4] <http://molprobitry.biochem.duke.edu/>

[5] <http://xray.bmc.uu.se/gerard/supmat/chi.html>

[6] <https://pdb-redo.eu/>