

Отчёт по качеству расшифровки
структуры белка глюкокиназы человека
3vev методом RSA
выполнила студентка 4ого курса ФББ
Анастасия Теплова

Москва, 2018

Аннотация

В работе проанализировано качество расшифровки структуры глюкокиназы человека Зеве. Были рассмотрены параметры качества модели в целом. Были детально проанализированы некоторые маргинальные остатки. Также было проведено сравнение исходной модели с моделью из PDB-redo.

Введение

Глюкокиназа человека (ГК) (EC 2.7.1.2) катализирует реакцию превращения глюкозы в глюкозо-6-фосфат (рисунок 1) и является сенсором уровня глюкозы в плазме крови и важнейшим регулятором гомеостаза глюкозы [1, 2]. Мутации, которые инактивируют ГК связаны с диабетом, а мутации, которые активируют ГК ассоциированы с гипергликемией. ГК обладает уникальными кинетическими особенностями, хотя она имеет низкую базовую аффинность к глюкозе, она обладает положительной кооперативностью – связывание ГК с глюкозой быстро повышает ее активность, когда концентрация глюкозы в крови поднимается выше пороговой.

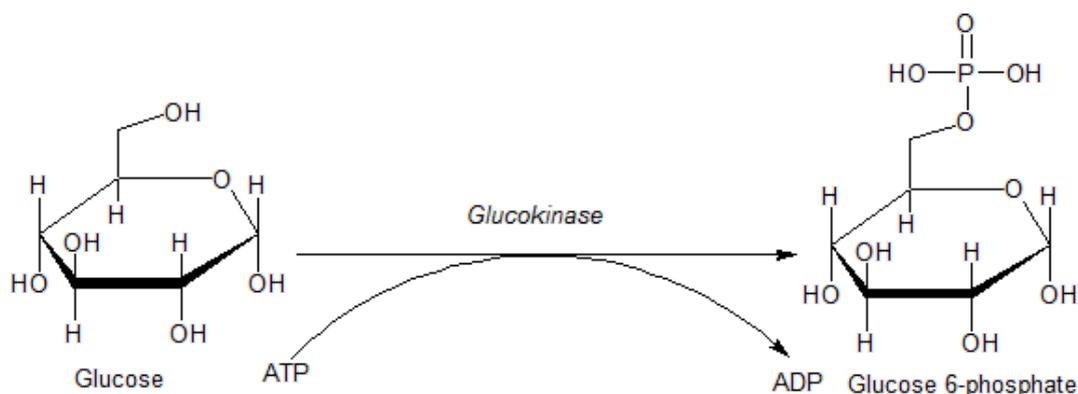


Рисунок 1. Реакция катализируемая глюкокиназой: превращение D-глюкозы в D-глюкозо-6-фосфат, которое сопровождается гидролизом АТФ.

ГК состоит из двух доменов (Рисунок 2). Малый домен включает остатки 66-202 и цепь С (остатки 442-465 на С-конце). Большой домен – это весь остальной белок. В большинстве структур ГК принимает активную закрытую конформацию, в которой глюкоза связана в активном сайте, который формируется обоими доменами. В таком состоянии ГК активно взаимодействует с глюкозой, так что все ОН-группы глюкозы вовлечены в водородные связи. Исследования показывают, что ГК довольно подвижный белок, который может принимать много различных конформаций, обладающих разным сродством к глюкозе.

С практической точки зрения интересны активаторы ГК, которые снижают уровень глюкозы в крови и повышают стимулируемую глюкозой секрецию инсулина. Работа в рамках которой была получена модель Зеве как раз была направлена на изучение механизма активации ГК и изучение конформационных состояний ГК в растворе в процессе протекания реакции.

Общая информация о модели

Модель 3vev представляет собой структуру глюкокиназы человека в активной закрытой конформации в комплексе с активатором и глюкозой и атомом Na^+ . Структура была получена авторами Shenping Liu, Mark J. Ammirati, Xi Song, John D. Knafels, Jeff Zhang, Samantha E. Greasley, Jeffrey A. Pfefferkorn, and Xiayang Qiu и загружена 08 Feb 2012. По результатам расшифровки авторы опубликовали статью [1], в которой исследовали механизм активации глюкокиназы. В рамках этой работы авторы получили другие структуры данного белка также в активной закрытой конформации в комплексе с другими веществами активаторами (PDB IDs: 4DHY, 3VF6,), в комплексе с $\text{ATP}\gamma\text{S}$ – негидролизуемым аналогом АТФ (3VEY), и структуру открытой конформации (4DCH).

Для решения фазовой проблемы использовался метод молекулярного замещения, за основу была взята структура глюкокиназы поджелудочной железы человека 3F9M.

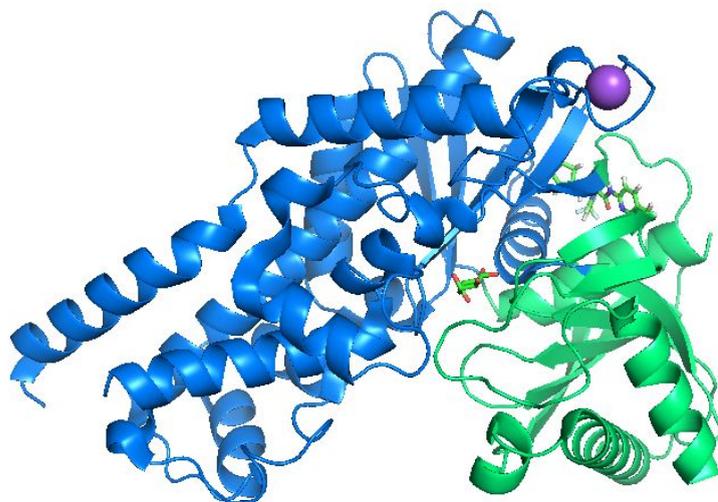


Рисунок 2. Структура 3vev общий вид. Синим покрашен большой домен, зеленым – малый домен.

Асимметрическая единица – мономер, биологическая единица – также мономер. Полипептидная цепь глюкокиназы состоит из 470 аминокислотных остатков, молекулярная масса 53 кДа. Авторам удалось расшифровать пространственную структуру с 12 по 465 остаток. Кроме остатков на N- и C-конце не расшифрованы остатки Gly94-Glu95-Glu96-Gly97-Gln98 в подвижной петле внутри структуры.

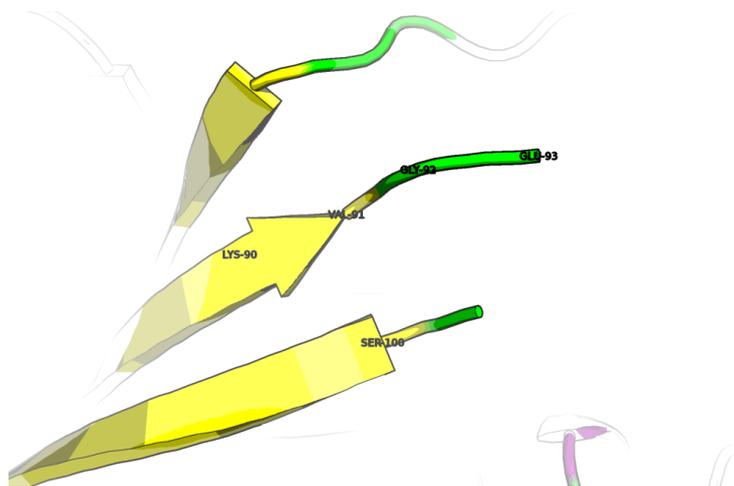


Рисунок 3. Пропущенные остатки в структуре Znev.

Общая информация о модели собрана в таблице 1.

Таблица 1. Характеристики модели Znev.

Параметр	Значение
число измеренных рефлексов	44808
разрешение	1.8Å
полнота набора структурных факторов	100%
диапазон разрешений структурных факторов	1.8Å - 17.89Å
параметры кристаллографической ячейки	a: 67.17 Å b: 82.11 Å c: 86.22 Å α: 90° β: 90° γ: 90°
группа симметрии	<i>P</i> 21 21 21
наличие некристаллографических симметрий в асимметрической ячейке	нет (псевдотрансляции не обнаружены)

Значения индикаторов качества модели

Основными показателями качества модели являются R-фактор и R-free. R-фактор отражает, насколько хорошо модель соответствует экспериментальным данным. R-фактор минимизируется в процессе оптимизации модели. Хорошими считаются показатели меньше 0.25. R-фактор модели Znev = 0.173, то есть модель хорошо соответствует эксперименту.

R-free вычисляется для контроля переоптимизации. Измеренные рефлекссы делятся на 2 части: рабочие – по ним модель оптимизируется и контрольные – по ним считается R-free. Хорошими считаются значения R-free < 25% – у нашей модели хорошее значение R-free = 20%. Также важна разница между значениями Rfree - R была маленькой (в идеале 0.05). По этому параметру наша модель тоже хорошая.

Таблица 2. Значения R-фактора и R-free структуры Znev

Параметр	Значение
R-фактор	0.173
R-free	0.2
R-free - R	0.027
Доля структурных факторов для вычисления R-free	5.05%

Также качество модели характеризуют геометрические параметры структуры: длины и углы связей, карта Рамачандрана, перекрытие Ван-дер-Ваальсовых радиусов атомов, ротамеры (конформации боковых цепей).

Таблица 3. Геометрические характеристики качества модели Znev

Параметр	Количество маргинальных остатков (outliers)
Углы между связями	0 из 4931
Длины связей	0 из 3672
RSR (соответствие модели группы атомов экспериментальной электронной плотности)	24 из 460 (5.3%)

Карта Рамачандрана	1 из 456 (0.2%)
Ротамеры боковых цепей	13 из 395 (3.4%)
Перекрытия (clashes)	15 (clashscore* = 2)

(*) Clashscore – параметр, показывающий число стерических перекрытий на 1000 атомов. Перекрытием считается пересечение Ван-дер-Ваальсовых радиусов более, чем на 0.4 Å.

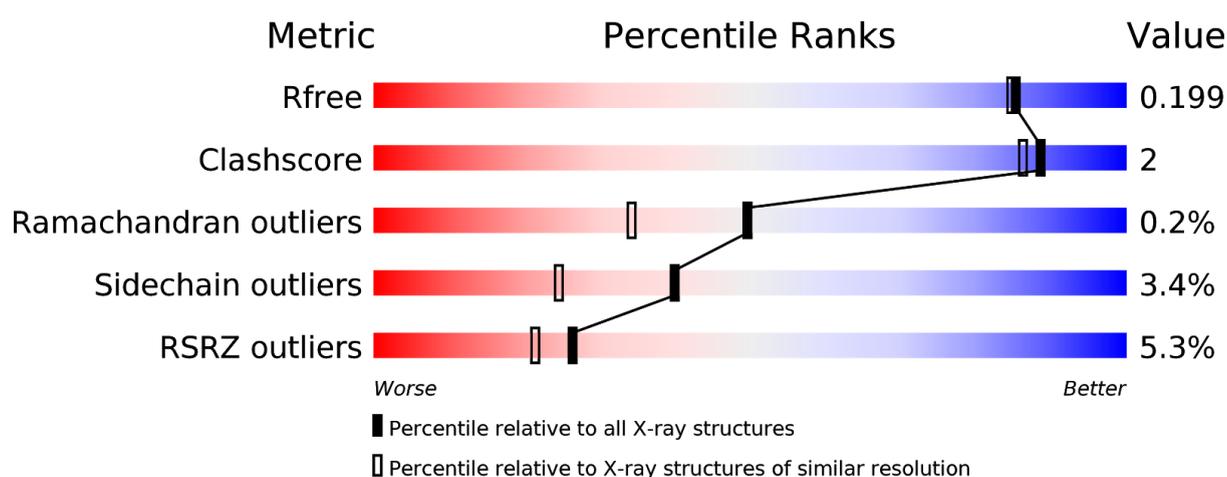


Рисунок 4. Качество модели 3vev по сравнению с моделями с таким же разрешением.

На рисунке 4 приведено сравнение показателей качества модели со средними значениями этих параметров в моделях с таким же разрешением. Видно, что качество модели 3vev высокое относительно других структур по R-free и Clashscore, но по маргинальным остаткам и по RSR ниже среднего.

Для дальнейшего анализа качества модели использовали сервис MolProbity [3]. В структуру были добавлены атомы водорода, было рекомендовано изменение положения остатков (flip) His317 и Gln337. Результаты анализа контактов между атомами и геометрии представлены в таблице 4.

Таблица 4. Значения геометрических параметров качества модели 3vev, определенные сервисом MolProbity.

All-Atom Contacts	Clashscore, all atoms:	2.22		99 th percentile* (N=837, 1.80Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	13	3.29%	Goal: <0.3%
	Favored rotamers	365	92.41 %	Goal: >98%
	Ramachandran outliers	1	0.22%	Goal: <0.05%
	Ramachandran favored	440	98.65 %	Goal: >98%
	MolProbity score [^]	1.39		97 th percentile* (N=11444, 1.80Å ± 0.25Å)
	Cβ deviations >0.25Å	7	1.66%	Goal: 0
	Bad bonds:	5 / 3673	0.14%	Goal: 0%
	Bad angles:	13 / 4951	0.26%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 10	0.00%	Expected: ≤1 per chain, or ≤5%

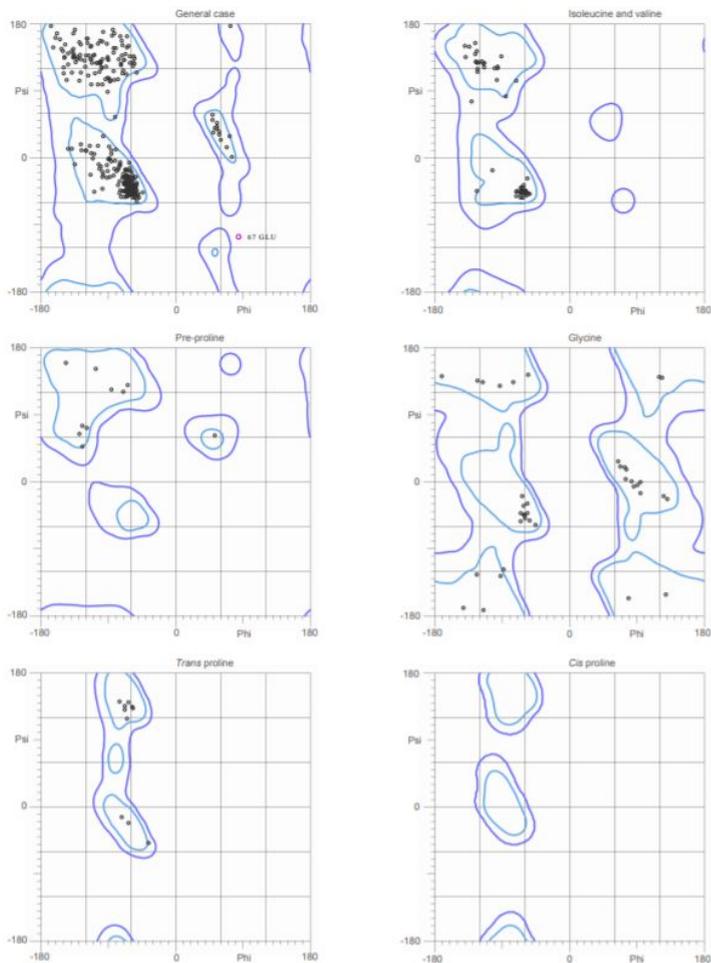


Рисунок 5. Карты Рамачандрана построенные сервисом MolProbity. Голубым показаны благоприятные (favorable) комбинации торсионных углов, а синим — разрешенные (allowed) комбинации.

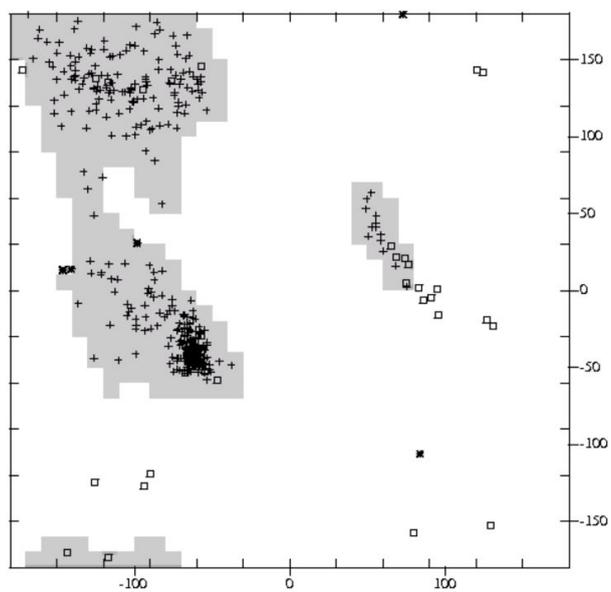


Рисунок 6. Карта Рамачандрана для всего белка построенная сервисом EDS.

Карта Рамачандрана, построенная сервисом MolProbity показана на рисунке 5. 98.5% (455/462) остатков попадают в благоприятные (favoured) области (98%). 99.8% (461/462) остатков попадают в разрешенные области (>99.8%). Был обнаружен 1 маргинальный остаток GLU 67 ($\phi = 84.1$, $\psi = -106.3$). Для сравнения карта Рамачандрана была также построена сервисом EDS [4] (рисунок 6). В данном случае список маргинальных остатков оказался больше (5 остатков).

Согласно MolProbity, в структуре 13 остатков с неблагоприятным положением атомов боковой цепи (3.3%) Хорошее значение для этого параметра <0.3%. Также сервисом MolProbity было проверено перекрывание остатков. У структуры получился хороший Clashscore (2.2 99ый перцентиль).

Также в EDS были проверены значения RSR и RSRZ для аминокислотных остатков (рисунок 7 и 8). RSR показывает соответствие модели экспериментальным данным.

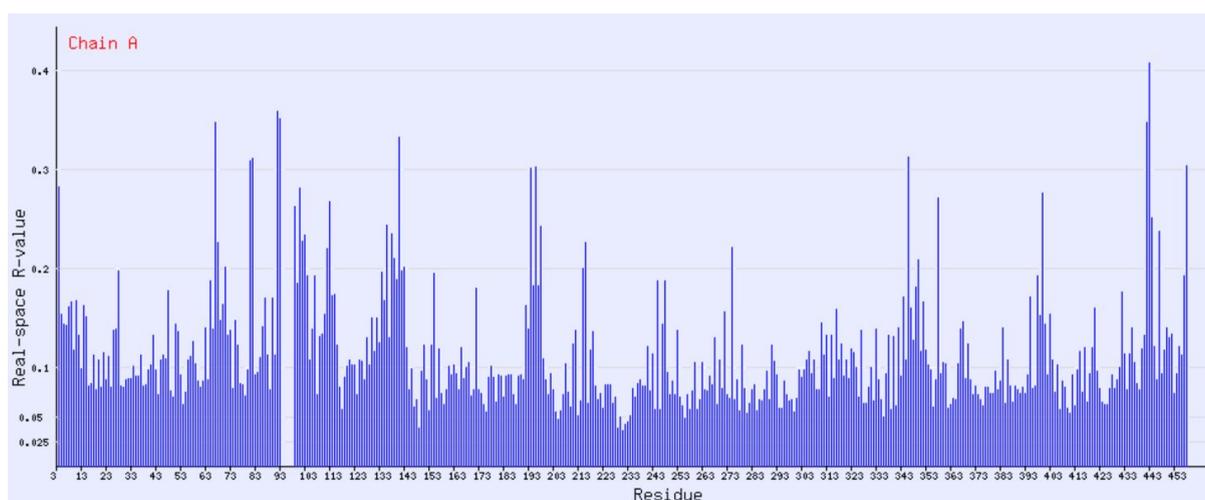


Рисунок 7. Значения RSR для остатков структуры Znev. Источник EDS.

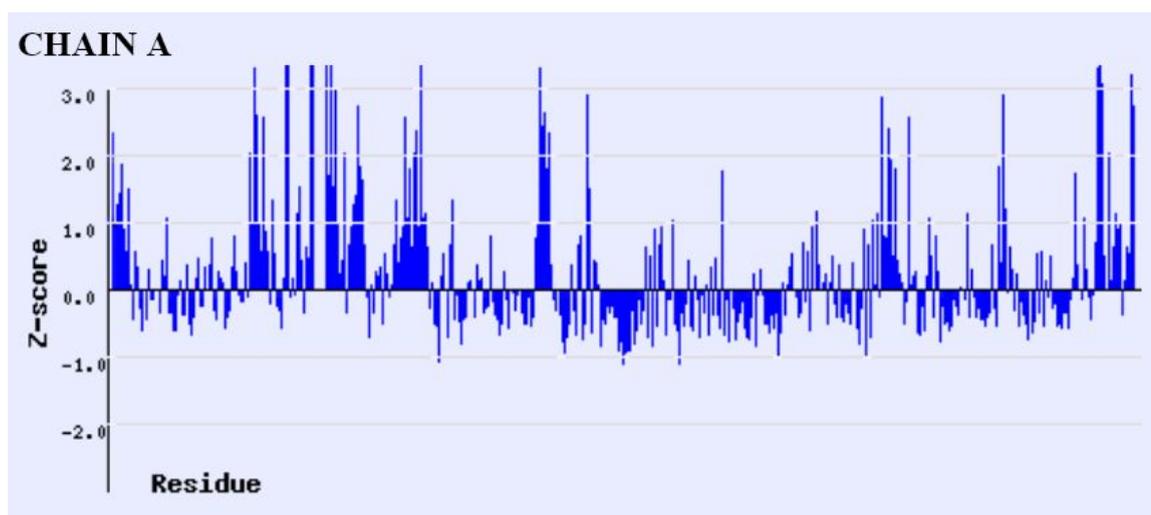


Рисунок 8. Значения RSRZ для остатков структуры Znev. Приемлемые значения <2. Источник EDS.

Маргинальные остатки были внесены в таблицу 5. Так как их оказалось много больше 15, в таблице приведены не все маргинальные остатки, а только те, которые проанализированы ниже в соответствующем разделе.

Таблица 5. Некоторые маргинальные остатки в структуре 3vev.

Остаток	Обоснование маргинальности
Glu67	Попадает в запрещенную область карты Рамачандрана.
Gln106	Ван-дер-Ваальсовы радиусы атома HE22 данного остатка и атома HD2 из Lys136 перекрываются на 0.91 Å
Lys136	Ван-дер-Ваальсовы радиусы атома HD2 данного остатка и атома HE22 из Gln106 перекрываются на 0.91 Å
His50	Атом азота NE2 His50 перекрывается с атомом азота NH1 A_Arg155, радиус перекрытия 0.497 Å.
A_Arg155	Атом азота NH1 A_Arg155 перекрывается с атомом азота NE2 His50, радиус перекрытия 0.497 Å.
Lys104	Неблагоприятный ротамер. Также водород HG2 Lys104 перекрывается с атомом углерода CD2 остатка His137. Радиус перекрытия 0.52 Å.
His317	Подозрение на инверсию боковой цепи (по данным MolProbe)
Met139	Плохое соответствие экспериментальной электронной плотности (RSRZ = 4.3 > 2)
His141	Плохое соответствие экспериментальной электронной плотности (RSRZ = 2.4 > 2)

Анализ некоторых маргинальных остатков

Остатки с подозрением на необходимость инверсии боковой цепи

MolProbity предложил перевернуть остаток His317, но если визуализировать карту электронной плотности (рисунок 9), то можно сделать вывод, что исходная конформация была правильной, так как при визуализации ЭП на уровне подрезки 1.5 σ по наблюдаемым сгусткам ЭП около предполагаемых атомов O. Кроме того, His317 в исходной конформации образует водородную связь с молекулой воды (рисунок 10), в то время как объяснить целесообразность этого поворота взаимодействием с какими-то другими остатками также не получается (рис).

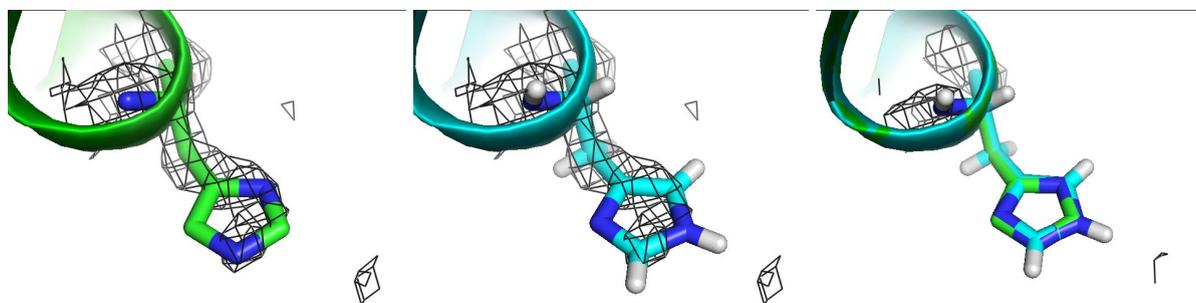


Рисунок 9. Визуализация электронной плотности около остатка His317. Слева направо: Исходное расположения атомов His317 (цепь окрашена в зеленый) с визуализацией ЭП на уровне подрезки 1.5 σ ; Положение атомов, которое предлагает MolProbity с визуализацией ЭП на уровне подрезки 1.5 σ ; совмещение структур и визуализация ЭП на уровне подрезки 2.5 σ ЭП не покрывает атомы.

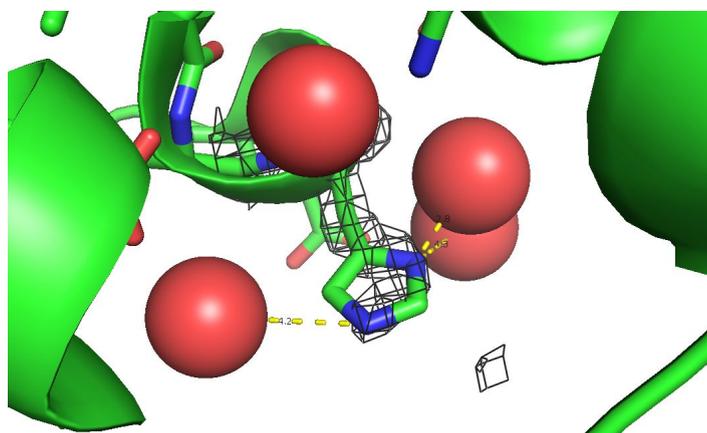


Рисунок 10. Водородная связь между молекулой воды и His317.

Остатки из запрещенной области карты Рамачандрана

Остаток Glu 67 попадает в запрещенную область карты Рамачандрана ($\phi = 84.1$, $\psi = -106.3$). Это единственный остаток, который *MolProb* отнес к маргинальным по данному параметру. EDS этот остаток тоже выделяет. Glu67 находится в экспонированной подвижной петле. На рисунке 11 показана карта электронной плотности на участке Thr60-Glu70, а на рисунках 12 А и В остаток Glu 67 показан крупно. При уровне

подрезки 1.5σ вокруг этого остатка нет сгущения электронной плотности. Важно заметить, что для Glu67 не расшифрованы атомы боковой цепи (CB, CG, CD, OE1, OE2). Исходя из этого можно сказать, что положение этого остатка не удалось расшифровать.

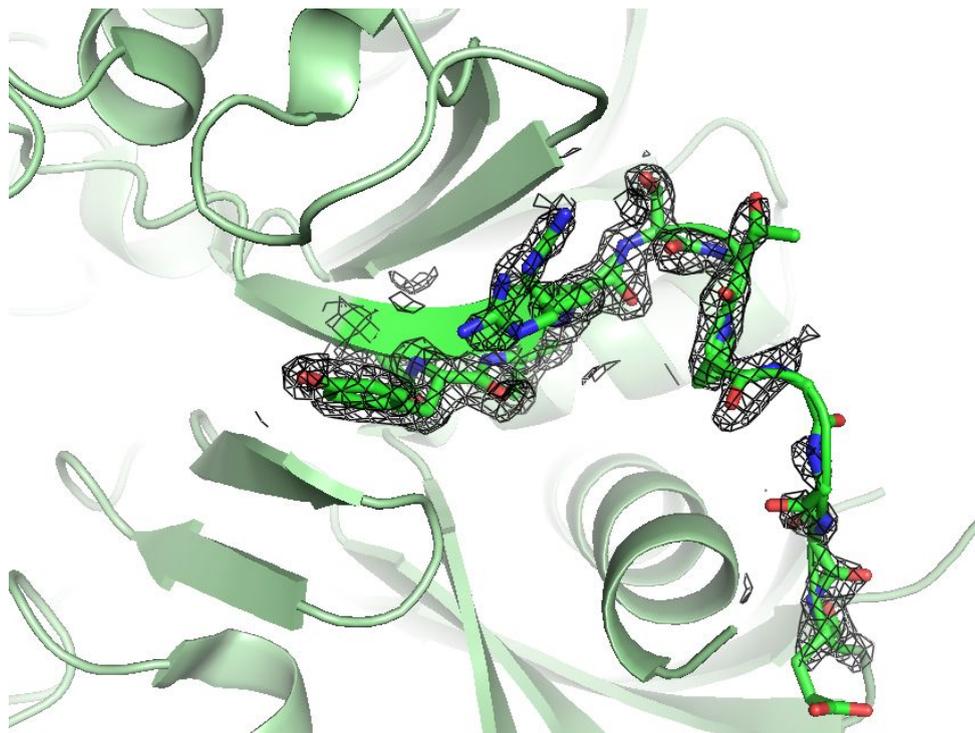


Рисунок 11. Визуализации карты ЭП на участке Thr60-Glu70. Уровень подрезки 1.5σ

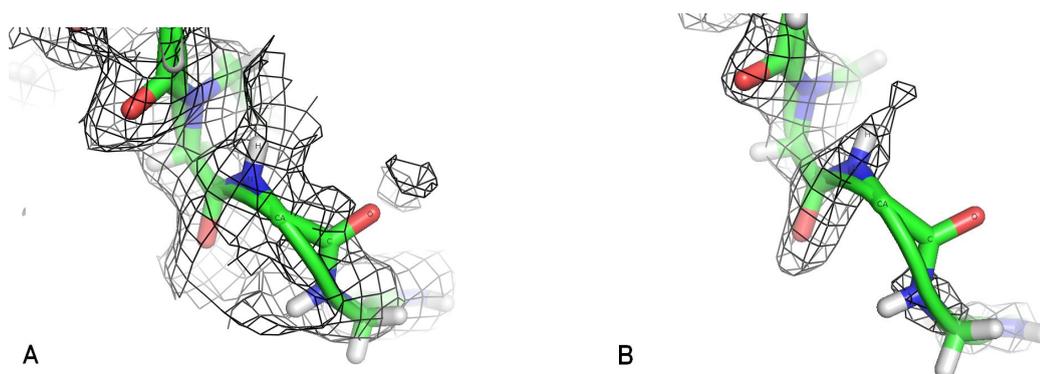


Рисунок 12. Визуализация карты ЭП для остатка Glu67 на уровне подрезки 0.5σ (рисунок А) и 1.5σ (рисунок В). На уровне подрезки 1.5σ вокруг остатка Glu67 нет сгущения ЭП. Также для него не расшифрованы атомы боковой цепи (на рисунке показаны все атомы этого остатка, которые есть в исходном файле PDB).

Перекрывающиеся остатки

Lys136 и Gln106

Сервис *MolProbe* сообщает о перекрывающихся атомах водорода остатков Lys136 (HD2) и Gln 106 (HE22), радиус перекрытия 0.91Å. Между этими остатками образуется водородная связь (показана на рисунке 13 А желтой линией, расстояние 2.9 Å). Остатки довольно хорошо вписываются в карту электронной плотности (также показана на рисунке 13А), хотя возможно остаток Gln106 можно вписать немного по-другому, чтобы избежать перекрытия атомов водорода.

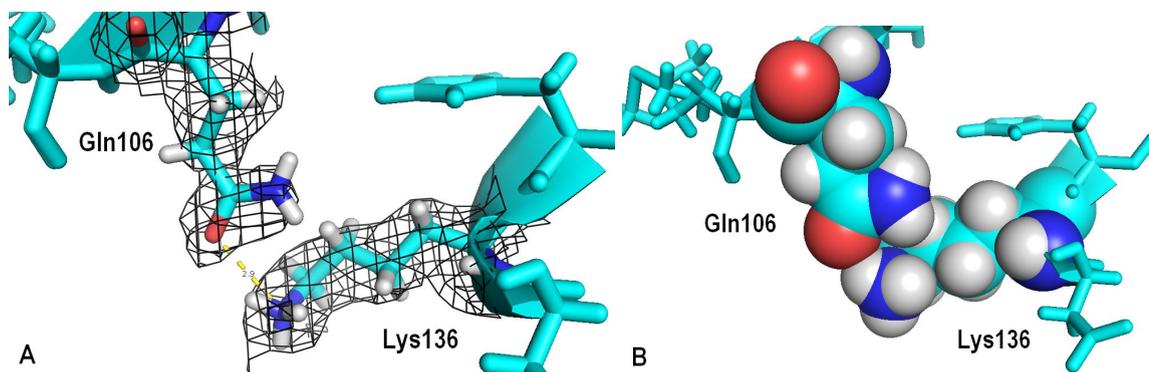


Рисунок 13. Визуализация перекрывающихся остатков Lys136 (HD2) Gln106 (HE22). Радиус перекрытия 0.91 Å. (А) Карта ЭП с уровнем подрезки 1.5σ. (В) Ван-дер-Ваальсовы радиусы.

His50 и A_Arg155

Для Arg155 occupancy не равно 1 (равно 0.5 для альтернативных положений) и возможно 2 альтернативных положения. Но в одном из этих положений атом азота NH1 Arg155 перекрывается с атомом NE2 His50. На рисунке 14А показана карта электронной плотности с уровнем подрезки 1.5σ, а на рисунке ХВ ван-дер-ваальсовы радиусы атомов. В_Arg155 не перекрывается с остатком His50 и, скорее всего, правильным будет именно такое положение.

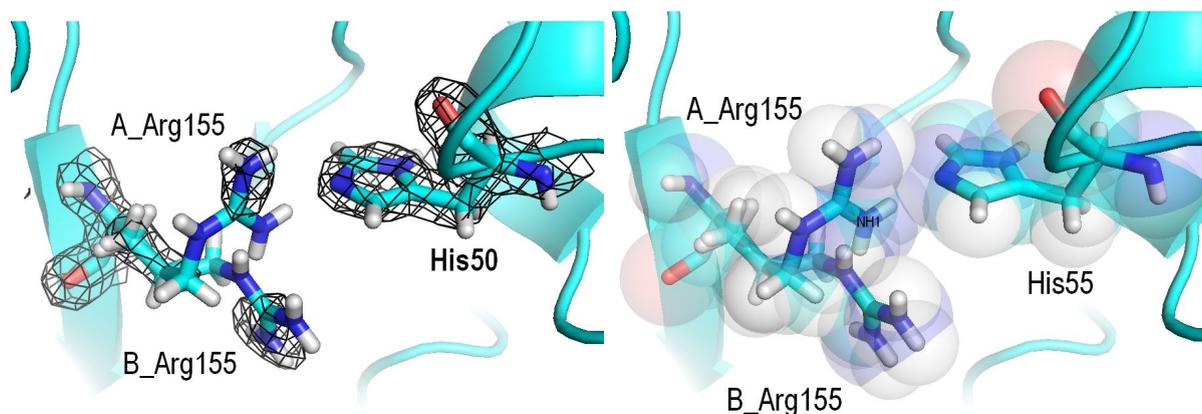


Рисунок 14. Перекрывающиеся остатки His50 и A_Arg155. Атом азота NE2 His50 перекрывается с атомом азота NH1 A_Arg155 радиус перекрытия 0.497 Å. (А) Карта ЭП с уровнем подрезки 1.5σ. (В) Ван-дер-Ваальсовы радиусы.

Ротамеры

Всего согласно сервису MolProbe 13 неблагоприятных ротамеров. Один из них остаток Lys104 (Углы χ_i : 279.5, 99.5, 187.9, 285.7) показан на рисунке 15. Я выбрала для анализа этот остаток, так как он в дополнение перекрывается с остатком His137. Боковая цепь Lys104 не покрывается картой ЭП на уровне подрезки 1.5σ. Возможно, боковая цепь этого остатка расшифрована не совсем правильно

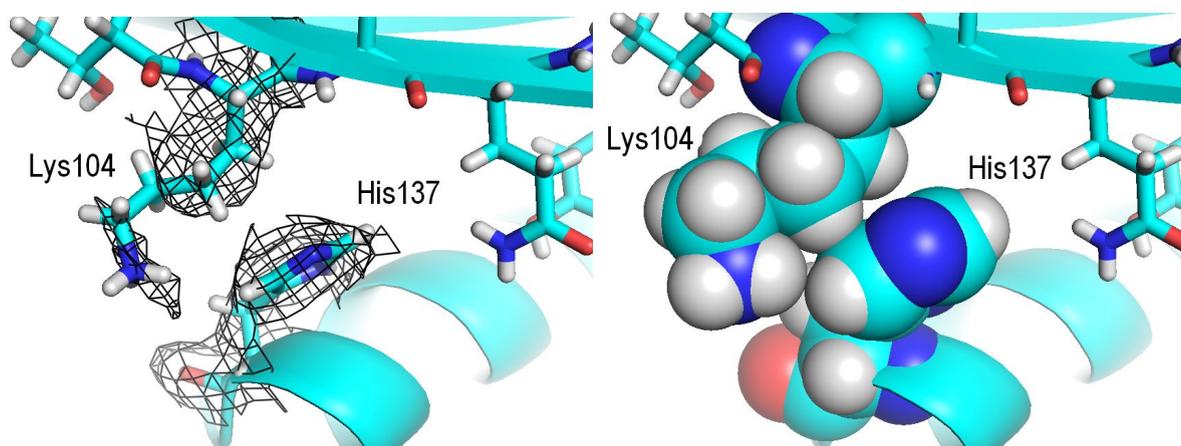


Рисунок 15. Водород HG2 Lys104 перекрывается с атомом углерода CD2 остатка His137. Радиус перекрытия 0.52 Å. (А) Карта ЭП с уровнем подрезки 1.5σ. (В) Ван-дер-Ваальсовы радиусы. на рисунке атомы азота и кислорода боковых цепей получились расположенными близко, но на самом деле в модели расстояние между ними 5.3 Å.

Атомы с высоким RSRZ

По данным со страницы в PDBe [5] в структуре 24 маргинальных остатка по этому параметру. Ранее был приведен график из EDS со значениями RSR b RSRZ для всех аминокислотных остатков полипептидной цепи (см. рисунок 8). Для примера рассмотрены His141 ($Z = 4.337$) и Met139 ($Z = 2.37$). Они попадают в подвижную петлю. His141 совсем не вписывается в карту электронной плотности (Рисунок 16). С Met139 ситуация лучше (рисунок 17)/

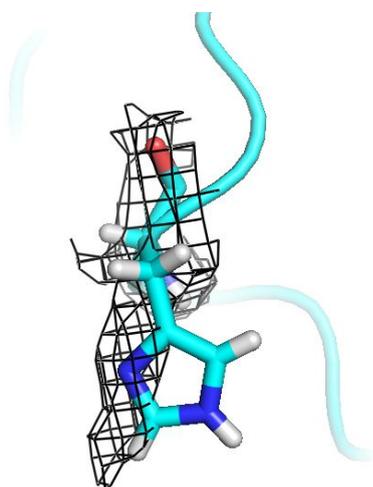


Рисунок 16. Остаток His141 с высоким RSRZ ($Z = 4.337$). Визуализирована электронная плотность на уровне подрезки 1.5σ . Боковая цепь не соответствует карте ЭП.

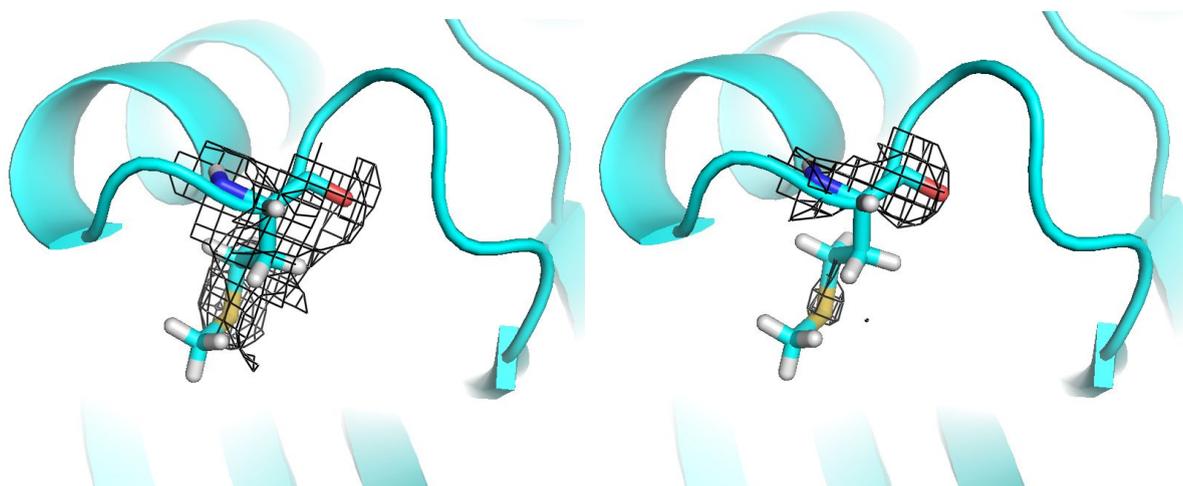


Рисунок 17. Остаток Met139 ($Z = 2.37$). Слева карта ЭП с уровнем подрезки 0.5 , справа с уровнем подрезки 1.5σ . Боковая цепь плохо вписывается в карту ЭП.

Сравнение модели из PDB с моделью из PDB_redo.

В PDB_redo есть улучшенная версия модели Zvev [6]. В таблице 6 приведено сравнение параметров исходной и улучшенной модели. Значения геометрических параметров качества модели полученные с помощью сервиса MolProbity приведены в таблице 7, а на рисунке 18 таблицы 4 и 8 представлены в виде вырезок экрана для удобства сравнения.

Таблица 6. Сравнение показателей качества модели Zvev из PDB и из PDB-redo. Данные со страницы в PDB-redo [6]

Параметр	PDB	PDB-redo
R	0,1715	0,1742
R-free	0,1916	0,2040

Таблица 7. Значения геометрических параметров качества улучшенной модели Zvev (из PDB-redo), определенные сервисом MolProbity.

All-Atom Contacts	Clashscore, all atoms:	2.48		99 th percentile* (N=837, 1.80Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	12	3.02%	Goal: <0.3%
	Favored rotamers	370	92.96%	Goal: >98%
	Ramachandran outliers	1	0.22%	Goal: <0.05%
	Ramachandran favored	439	98.43%	Goal: >98%
	MolProbity score [^]	1.40		97 th percentile* (N=11444, 1.80Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	2 / 3684	0.05%	Goal: 0%

	Bad angles:	7 / 4964	0.14%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 10	0.00%	Expected: ≤1 per chain, or ≤5%

All-Atom Contacts	Clashscore, all atoms:	2.22		99 th percentile* (N=837, 1.80Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	13	3.29%	Goal: <0.3%
	Favored rotamers	365	92.41%	Goal: >98%
	Ramachandran outliers	1	0.22%	Goal: <0.05%
	Ramachandran favored	440	98.65%	Goal: >98%
	MolProbity score [^]	1.39		97 th percentile* (N=11444, 1.80Å ± 0.25Å)
	Cβ deviations >0.25Å	7	1.66%	Goal: 0
	Bad bonds:	5 / 3673	0.14%	Goal: 0%
	Bad angles:	13 / 4951	0.26%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 10	0.00%	Expected: ≤1 per chain, or ≤5%

All-Atom Contacts	Clashscore, all atoms:	2.48		99 th percentile* (N=837, 1.80Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	12	3.02%	Goal: <0.3%
	Favored rotamers	370	92.96%	Goal: >98%
	Ramachandran outliers	1	0.22%	Goal: <0.05%
	Ramachandran favored	439	98.43%	Goal: >98%
	MolProbity score [^]	1.40		97 th percentile* (N=11444, 1.80Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	2 / 3684	0.05%	Goal: 0%
	Bad angles:	7 / 4964	0.14%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 10	0.00%	Expected: ≤1 per chain, or ≤5%

Рисунок 18. Таблицы, полученные в MolProbity в виде вырезок экрана для более удобного сравнения. Вверху для исходной структуры Zvev, внизу для структуры из PDB-redo.

Геометрия модели улучшена не сильно. Изменений сделано не много: удалены 4 молекулы воды, изменены 3 ротамера, боковая цепь 3х остатков инвертирована, 6 остатков стали лучше вписываться в электронную плотность (но в то же время 24 стали вписываться хуже. R-free модели после улучшения выше чем до (то есть этот показатель ухудшился). Таким образом, модель из PDB-redo хуже соответствует экспериментальным данным, чем исходная модель, а ее геометрия сильно не улучшилась.



Заключение

Исходная расшифрованная структура имеет достаточно хорошее качество: в ней всего 1 из запрещенной области карты Рамачандрана, 13 ротамеров, есть перекрывающиеся остатки, но Clashscore хороший. Модель хуже других моделей из PDB по параметру RSRZ, который показывает соответствие модели экспериментальным данным. Попытка улучшить модель не привела к существенному улучшению геометрии и снизила RSRZ, То есть улучшить модель не получилось, хотя как мне кажется, это и не очень нужно, так как модель исходно хорошая.

Источники

[1] Liu S. et. al. Insights into mechanism of glucokinase activation: observation of multiple distinct protein conformations. The Journal of Biological Chemistry. Feb 2012, 287(17):13598-13610. (DOI: [10.1074/jbc.M111.274126](https://doi.org/10.1074/jbc.M111.274126))

[2] [Wikipedia.org/Glucokinase](https://en.wikipedia.org/Glucokinase)

[3] <http://molprobity.biochem.duke.edu>

[4] <https://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=3vev>

[5] <https://www.ebi.ac.uk/pdbe/entry/pdb/3vev/experiment>

[6] <https://pdb-redo.eu/db/3vev>