

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

Отчёт по качеству расшифровки структуры белка галектина-1 *Rattus norvegicus* методом PCA

студентки 4 курса
Анны Образцовой

Москва

2015

Аннотация

В данном отчете приводится краткая информация об исследуемой белке, а также проводится оценка качества расшифровки структуры 3M2M в целом, анализ нескольких маргинальных остатков и сравнение модели из PDB и из PDB_redo.

Введение

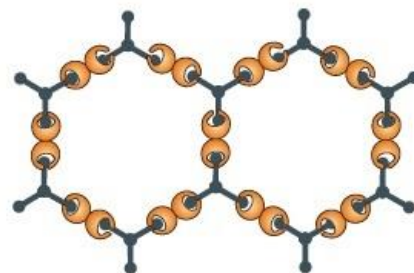
Галектин-1 является представителем семейства галектинов – небольших белков, связывающих углеводы с β -галактозидным остатком. Эти животные белки имеют сходную структуру и как минимум один консервативный углевод распознающий домен (carbohydrate recognition domain, CRD), который состоит примерно из 130 аминокислотных остатков и отвечает за связывание углеводов. [1]

Связывая гликаны клеточной поверхности, галектины могут оказывать влияние на клеточный цикл, адгезию, дифференциацию и апоптоз. Эти белки представляют особый интерес для иммунологии, так как участвуют в регуляции функциональной активности клеток иммунной системы. [2]

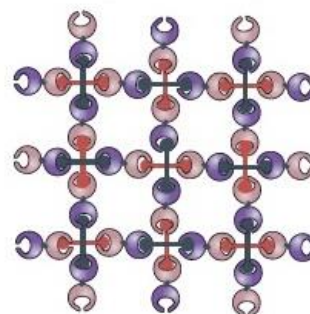
Галектины подразделяют на три типа:

- прототип (содержит один CRD и функционирует в димерном виде),
- тандем-тип (содержит два CDR),
- химерный тип (состоит из одного CDR и дополнительной части, позволяющей образовывать пентамеры).

Так как биологически активная единица галектина содержит более одного углевод связывающего домена, эти белки позволяют образовывать сшивки между разными углеводными молекулами. Благодаря этому на поверхности клеток образуются упорядоченные



трёхвалентный лиганд



четырёхвалентный лиганд

Рис. 1. Схематичное изображение различных ячеек, которые могут быть образованы двухвалентными галектинами и мультивалентными лигандами. [1]

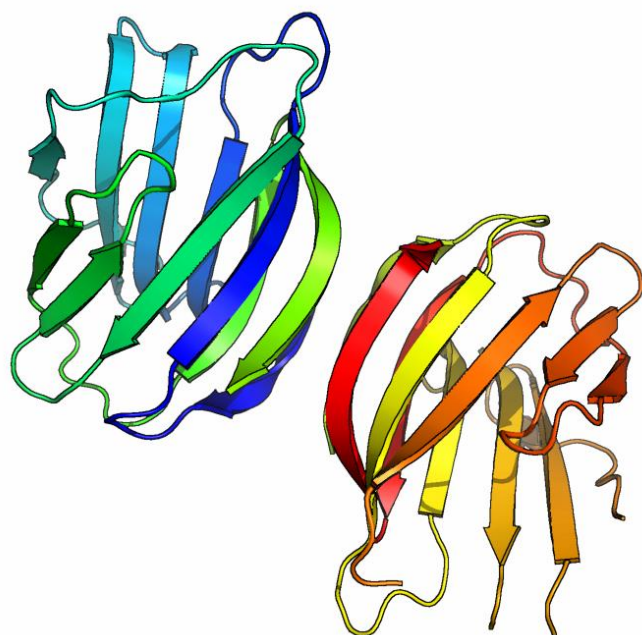


Рис. 2. Структура гомодимера галектина-1 крысы (PDB код 3M2M).

структуры, также называемые ячейками, вид которых может отличаться в зависимости от валентности лиганда (**рис. 1**).

Галектин-1 относится к прототипу. Его структура представляет собой β -сэндвич из антипараллельных слоёв (**рис. 2**). Каждая субъединица галектина-1 крысы имеет массу 14.9 кДа и состоит из 135 аминокислотных остатков [3].

При связывании углеводов галектином-1 наблюдается заметный «гликозидный кластерный эффект» - изменение аффинности связывания мультивалентного лиганда по сравнению с моновалентным [4]. Чтобы понять механизм этого явления Emma Salomonsson с коллегами исследовали структуру галектина-1 в комплексе с β -лактозой, а также различия в связывании моно- и мультивалентных лигандов нативным димерным галектином-1 и мутантным мономерным. [5]

Оказалось, что димеры и мономеры галектина-1 связывают лиганды с одинаковой аффинностью. Следовательно, кластерный эффект не связан с взаимодействием лиганда с дополнительным углеводов связывающим сайтом, имеющимся у димерной формы. Возможно, мультивалентные лиганды имеют дополнительные взаимодействия около связывающего сайта, что и повышает аффинность их связывания.

Результаты и обсуждение

Общая информация о модели

Модель была добавлена в PDB в 2010 году, авторами обозначены Lobsanov Y.D., Rini J.M. и Leffler H.

В файле структурных факторов содержится 18325 записи, для оптимизации модели использовалось 17063 структурных факторов. Фазовую проблему решали методом молекулярного замещения с помощью программы AMoRe, используя структуру человеческого галектина-2 (PDB код - 1HLC) [5].

Одна кристаллографическая ячейка содержит восемь молекул белка (цепи A, B, ... H), которые образуют канонические димеры (AB, CD, EF и GH). В комплексе присутствуют пять молекул лактозы, связанные с цепями A, C, D, E и H.

В **таблице 1** приведены прочие данные о модели, полученные на сервере EDS [6].

Таблица 1. Некоторая информация о модели 3M2M.

Характеристика	Значение
Заявленное разрешение	2.95 Å
Диапазон разрешения	36.01 - 2.83 Å
Полнота набора структурных факторов	67.2 %
Кристаллографическая группа	P 1
Длины направляющих векторов ячейки (Å)	a=40.1, b=58.6, c=121.6
Углы между векторами (°)	$\alpha=101.1$, $\beta=91.6$, $\gamma=110.9$

Разрешение структуры можно назвать удовлетворительным, полнота данных - достаточно низкая.

Значения индикаторов качества модели в целом

Значения R-фактора и R_{free} - 0.206 и 0.248 соответственно. Это говорит об умеренном соответствии модели наблюдаемым значениям структурных факторов, оптимальное значение этих характеристик должно быть не больше 0.2. Разница между индикаторами не превышает 0.1, значит, свидетельств переоптимизации модели нет.

С помощью сервиса MolProbity [7] к модели были добавлены атомы водорода, а также приняты некоторые из предложенных инверсий боковых цепей. Для полученной модели была построена карта Рамачандрана (**рис. 3**) и получены некоторые характеристики (**таблица 2**).

MolProbity Ramachandran analysis

3m2mFH_reg.pdb, model 1

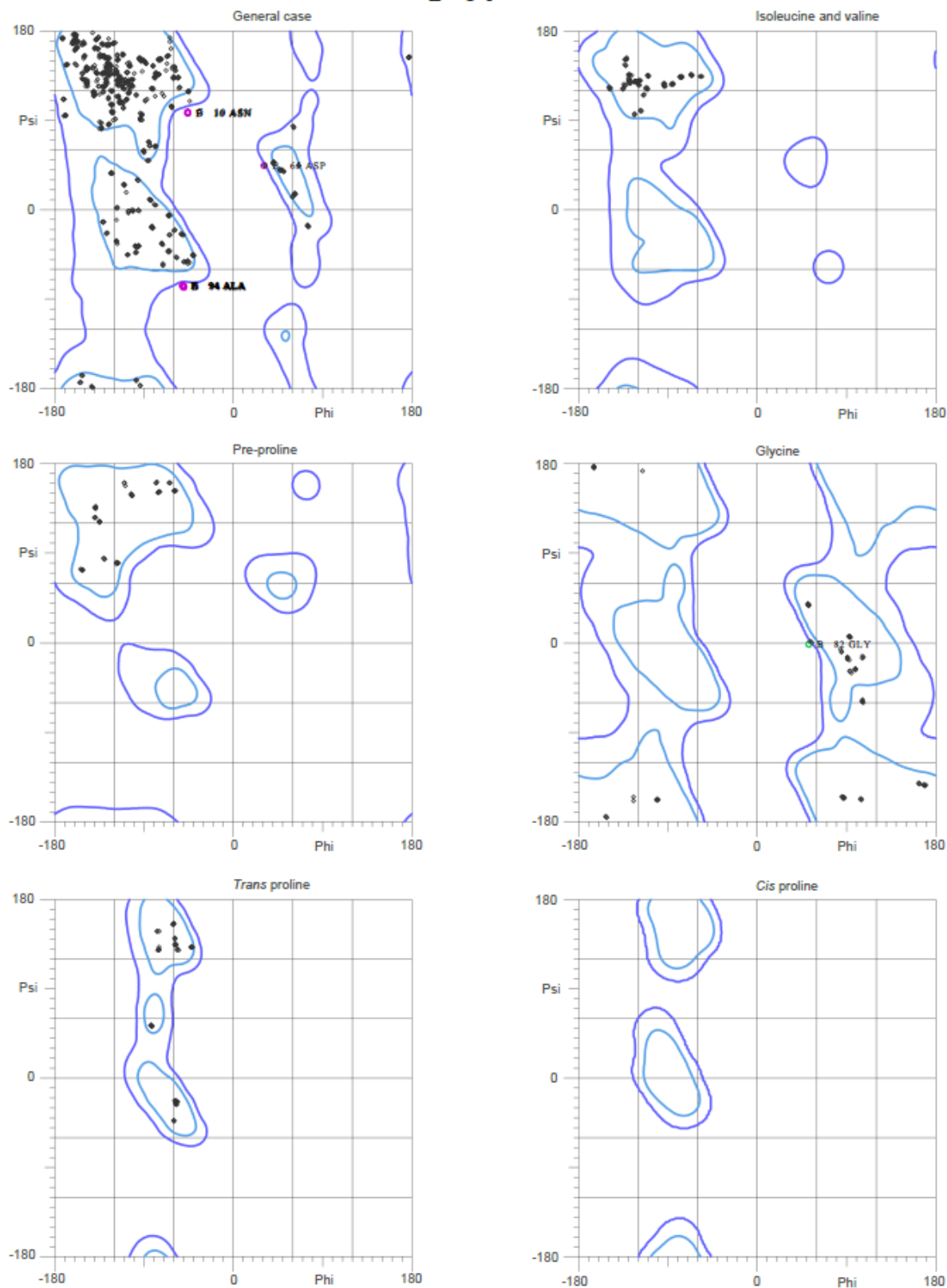


Рис. 3. Карты Рамачандрана для различных подмножеств аминокислотных остатков, построенные сервисом MolProbity. Голубыми линиями выделены предпочтительные области значений углов ϕ и ψ . Положение маргинальных остатков на карте выделено фиолетовым и зелёным цветом. Все 10 маргинальных остатков плохо различимы, так как некоторые накладываются друг на друга на карте.

Таблица 2. Характеристики модели, полученные с помощью сервиса MolProbity. Зелёным выделены характеристики, говорящие в пользу рассматриваемой модели, красным – обратное.

Характеристика	Значение в абсолютных единицах и в процентах (если возможно)		Оптимум
Clashscore	15.83		95th percentile (N=94, 2.95Å ± 0.25Å)
Poor rotamers	40	4.70%	Goal: <0.3%
Favored rotamers	724	85.08%	Goal: >98%
Ramachandran outliers	10	1.00%	Goal: <0.05%
Ramachandran favored	938	93.99%	Goal: >98%
MolProbity score	2.62		90th percentile (N=3863, 2.95Å ± 0.25Å)
C β deviations >0.25Å	0	0.00%	Goal: 0
Bad bonds	18 / 8119	0.22%	Goal: 0%
Bad angles	13 / 10966	0.12%	Goal: <0.1%

Как видим, всего три характеристики находятся в хорошем диапазоне – число близких контактов на тысячу атомов, интегральная оценка MolProbity и число отклонений C β атомов. Число отклоняющихся ротамеров и маргинальных остатков на карте Рамачандрана слишком большое (10 штук или 1%), а число предпочтительных – слишком маленькое. Кроме того, наблюдается увеличенное число ковалентных связей и валентных углов, существенно отклоняющихся от теории.

Стоит отметить, что до добавления атомов водорода и поворотов некоторых боковых цепей Clashscore был равен 18. Все наложения атомов находятся внутри ячейки (значение Symm-Clashes для всех цепей равно 0).

Вернёмся к данным с сервера EDS. Среднее значение пространственного R-фактора (RSR) для модели составляет 0.144 со стандартным отклонением 0.039, что говорит о приемлемом соответствии построенной модели «экспериментальной» электронной плотности. Рассмотрим график z-score RSR для остатков различных цепей (**рис. 4**) с сервера EDS [6].

Z-score – это нормализованное значение RSR для остатков данного типа при данном разрешении, позволяющее оценить RSR остатка со статистической точки зрения. Остаток считается маргинальным, если это значение превышает 2 [8]. Почти все остатки имеют хорошее значение Z-score, меньше нуля (то есть их RSR лучше среднего), однако есть два маргинальных остатка – Asn50 цепей B и F.

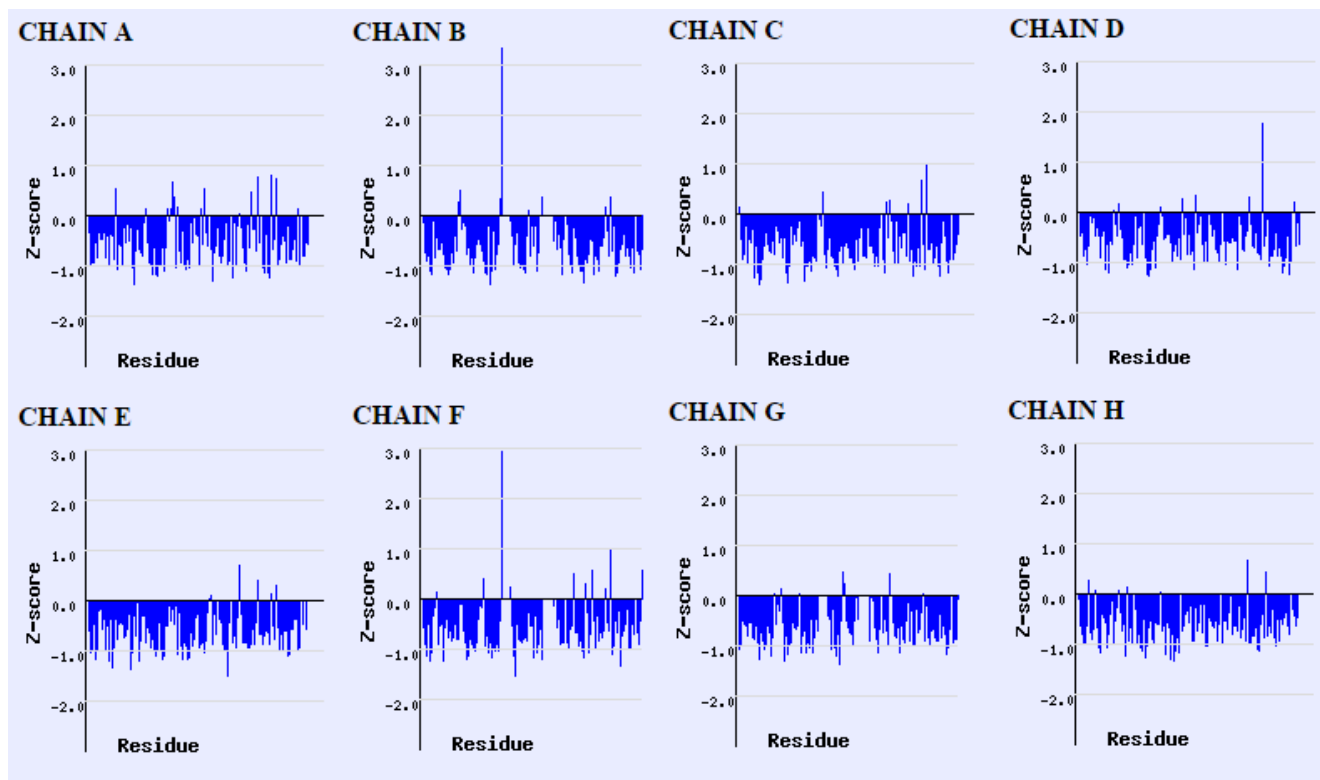


Рис. 4. График Z-score для восьми белковых цепей модели 3M2M.

Список маргинальных остатков

По данным сервисов MolProdit, EDS и wwPDB Validation Report был составлен небольшой список маргинальных остатков, информация о которых приводится в **таблице 3**.

Таблица 3. Список маргинальных остатков.

Цепь	Остаток	Критерий
B	Asn50	Z-score RSR = 3.3, среднее B-фактора = 64.9, слишком близкое положение атомов HB3 и LYS28:HG3 (1.68 Å)
F	Asn50	Z-score RSR = 2.9, среднее B-фактора = 63.1, слишком близкое положение атомов C и ARG48:HB3 (2.36 Å), H и ARG48:C (2.13 Å), CB и LYS28:HG3 (2.48 Å)
B	Phe49	Отклоняющийся ротамер, нехарактерный угол связи N-CA-C (96.09°), слишком близкое положение атомов HA и ASN56:OD1 (2.07 Å)
A, C, D, E, H	Ala94	Выброс на карте Рамачандрана
A	Gln93	Отклоняющийся ротамер, подозрение на необходимость инверсии боковой цепи, слишком близкое положение атомов NE2 и H (расстояние - 1.89 Å), CD и H (2.15 Å), NE2 и N (2.61 Å), HG3 и GLU115:OE2 (2.13 Å), HG3 и GLU115:CD (2.4 Å)
F, G, B	Asn10	Выброс на карте Рамачандрана, подозрение на необходимость инверсии боковой цепи
B	Asn113	Отклоняющийся ротамер, подозрение на необходимость инверсии боковой цепи
B	Asp65	Выброс на карте Рамачандрана
B	Gly82	Выброс на карте Рамачандрана

Детальный анализ маргинальных остатков

Для анализа были выбраны первые 5 маргинальных остатков из таблицы 3.

1) Аспарагин 50 (В)

Характеристики: Z-score RSR = 3.3, среднее В-фактора = 64.9, число наложений атомов: 1.

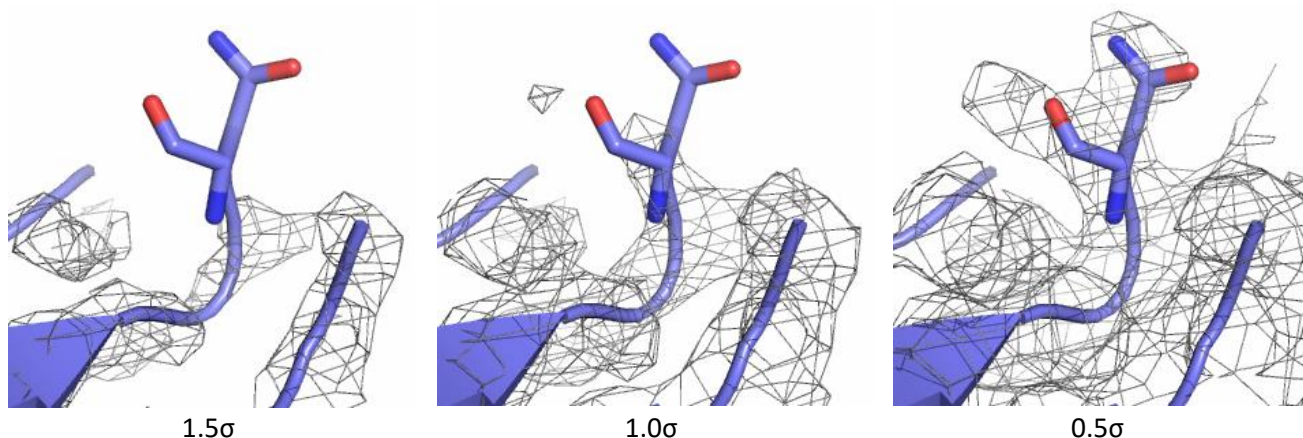


Рис. 5. Отображение электронной плотности вокруг остатка Asn50 цепи В и нескольких ближайших остатков на разных уровнях подрезки.

Для всех рассматриваемых остатков электронная плотность визуализировалась с помощью PyMOL. На **рис. 5** видно, что электронная плотность Asn50 отображается только при уровне подрезки 0.5, и координаты атомов вписываются в сгущение ЭП не очень хорошо. В цепи В отсутствует 12 остатков, четыре из которых следуют сразу за Asn50 (с 51 по 54). По положению остатков Asn50 и Ala55 логично предположить, что остатки 51-55 образуют петлю на поверхности белка (**рис. 6**).

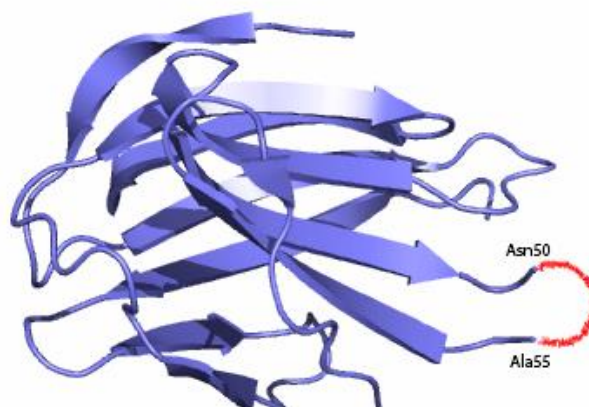


Рис. 6. Цепь В структуры 3M2M. Красным цветом показано возможное положение петли из остатков 51-54.

Возможно, эта структура слишком подвижна даже в кристалле, поэтому остатки 51-54 вообще отсутствуют в PDB файле, а остаток Asn50 имеет «слабую» электронную плотность и, как следствие, отклоняющиеся параметры.

2) Аспарагин 50 (F)¹

Характеристики: Z-score RSR = 2.9, среднее В-фактора = 63.1, число наложений атомов: 3.

Изображение электронной плотности для остатка Asn50 цепи F приведено на **рис. 7**. В целом для этого остатка общая картина аналогична предыдущему случаю. Изображение электронной

¹ Очевидно, что остаток Asn50 из цепи F очень похож по своим характеристикам на остаток Asn50 из В, рассмотренный выше. Этот остаток был взят в надежде, что наложения атомов, отсутствующие в предыдущем случае, могут значительно изменить картину, что, однако, не подтвердилось.

плотности при уровне подрезки 1.0 покрывает только часть атомов, при уровне подрезки 1.5 – больше, но остаток по-прежнему очень плохо вписан. В цепи F отсутствуют те же остатки, что и в цепи B, расположение одинаковое, поэтому предположения о причине маргинальности остатка те же.

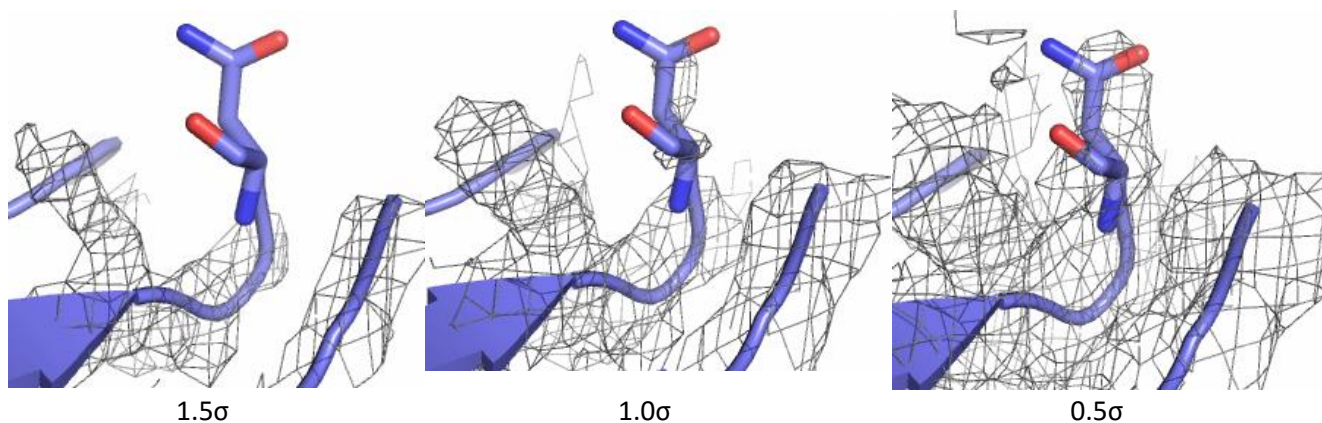


Рис. 7. Отображение электронной плотности вокруг остатка Asn50 цепи F и нескольких ближайших остатков на разных уровнях подрезки.

3) Фенилаланин 49 (B)

Характеристики: Z-score RSR = -0.47, среднее B-фактора = 53.2, число наложений атомов: 1, угол связи N-CA-C = 96.09° (показан на рис. 8), отклоняющийся ротамер.

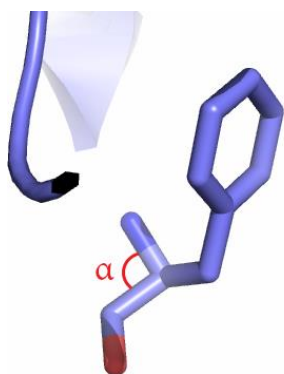


Рис. 8. Остаток Phe49 цепи B. Угол α - отклоняющийся от нормального значения.

На рис. 9 видно, что изображение электронной плотности данного остатка полностью покрывает атомы остатка только при уровне подрезки 0.5. Тем не менее атомы достаточно хорошо вписаны в сгущение ЭП, что подтверждается значением Z-score RSR. Так как этот остаток расположен возле петли, можно предположить, что он тоже претерпевает колебания, что согласуется с большим температурным фактором. Не исключено и просто плохое качество экспериментальных данных.

Можно предположить, что координаты атомов этого остатка определены не точно, что и обуславливает отклоняющееся значение угла связи N-CA-C и положение боковой цепи.

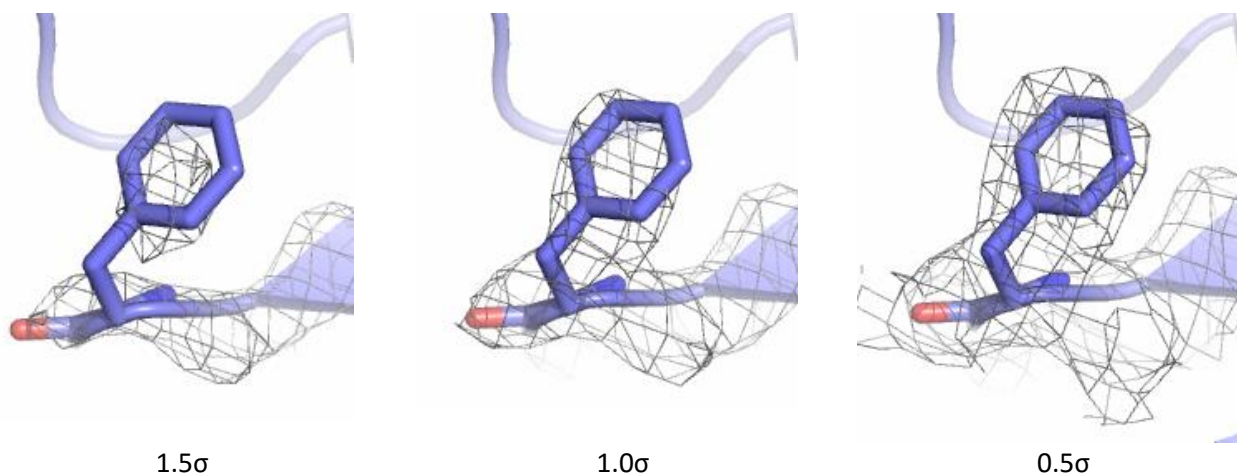


Рис. 9. Отображение электронной плотности вокруг остатка Phe49 цепи В и нескольких ближайших остатков на разных уровнях подрезки.

4) Аланин 94

Характеристики (для цепи А): Z-score RSR = -0.27, среднее В-фактора = 35.9, число наложений атомов: 0, угол ϕ = -51.4° , угол ψ = -76.4° .

Данный остаток достаточно хорошо вписан в электронную плотность, о чем говорит Z-score RSR чуть меньше нуля и изображение электронной плотности (**рис. 10**). Однако электронная плотность достаточно сильно «размазана», что согласуется с относительно высоким температурным фактором.

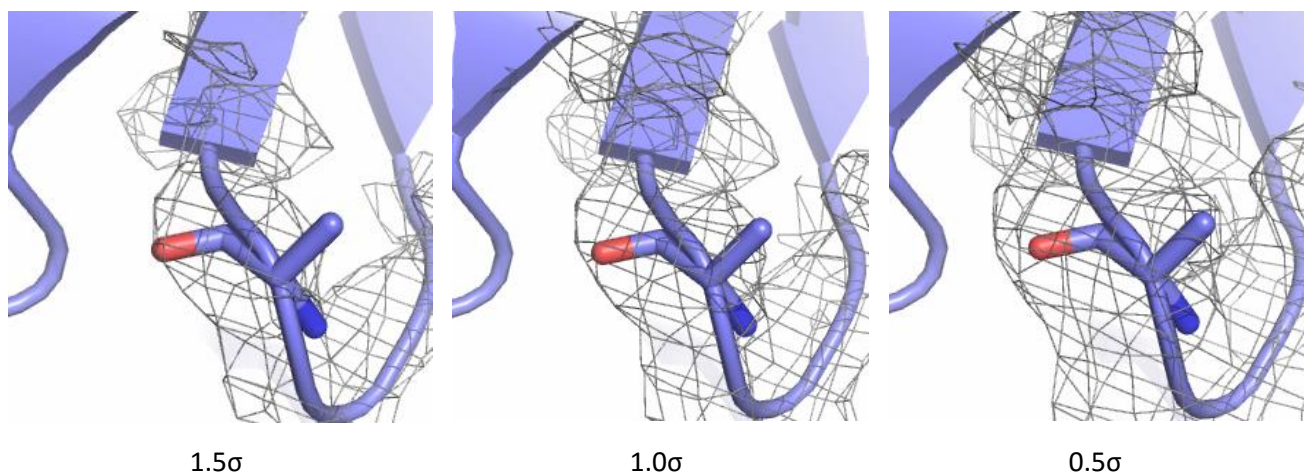


Рис. 10. Отображение электронной плотности вокруг остатка Ala94 цепи А и нескольких ближайших остатков на разных уровнях подрезки.

Этот гидрофобный остаток расположен в петле на поверхности белка. Его боковой радикал повёрнут в сторону белка, возможно, для уменьшения числа контактов с водой. Вероятно, нестандартные торсионные углы могут быть связаны именно с положением этого остатка, то есть это не является ошибкой модели, а действительно происходит в природе. Такое предположение подтверждается тем фактом, что торсионные углы Ala94 одинаково отклонены в пяти цепях белка.

5) Глутамин 93 (A)

Характеристики: Z-score RSR = 0.04, среднее B-фактора = 45.7, число наложений атомов: 5.

RSR остатка Gln93 цепи A близок к среднему для остатков этого типа при схожем разрешении, температурный фактор достаточно высокий. Изображение электронной плотности (**рис. 11**) охватывает все атомы остатка только при уровне подрезки 0.5, при уровне 1.5 вокруг боковой цепи ЭП не отображается вообще.

Всё это снова свидетельствует о том, что остаток мог сильно флуктуировать из-за расположения в петле. Не исключено и низкое качество экспериментальных данных.

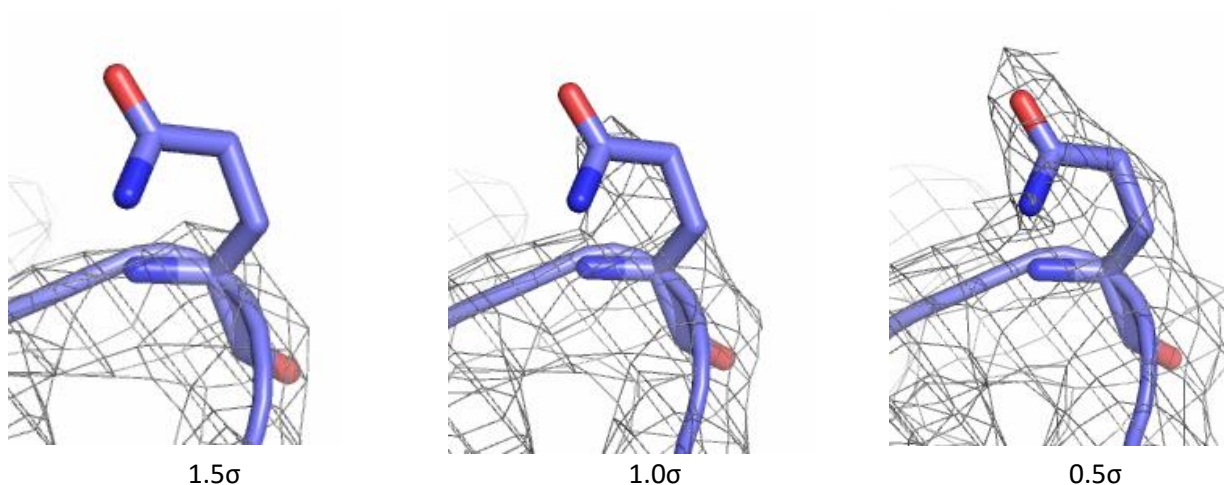


Рис. 11. Отображение электронной плотности вокруг остатка Gln93 цепи A и нескольких ближайших остатков на разных уровнях подрезки.

Сравнение модели из PDB с моделью из PDB_redo

На сервере PDB_redo [9] была получена информация об оптимизированной модели 3M2M. В таблице 4 приводятся различные параметры для изначальной модели и улучшенной. К сожалению, записи с улучшенной структурой нет, так как разрешение исходной структуры слишком низкое, и запись не создавали («The structure was not rebuilt because the resolution was too low»).

Таблица 4. Сравнение характеристик исходной и улучшенной моделей.

Характеристика	PDB	PDB_redo
R	0.2060 (вычисленное по данным - 0.2248)	0.2344
R _{free}	0.2480 (вычисленное по данным - 0.2704)	0.2567
R _{free} - R	0.042	0.0223
Качество карты Рамачандрана	-2.784	-1.450
Нормальность ротамеров χ -1/ χ -2	-3.155	-2.524
Ненасыщенные доноры/акцепторы водородных связей	98	87

Как видим, некоторые параметры улучшились, но не все. Значение R стало больше, значение R_{free} – тоже (в случае вычисленного значения - уменьшилось). Однако разница между этими значениями сократилась. Увеличились показатели качества карты Рамачандрана и нормальности ротамеров по углам χ (оба показателя тем лучше, чем они выше).

Выводы

1. Структура 3M2M имеет не очень высокое разрешение и средние показатели общего качества.
2. В модели присутствует достаточно большое число остатков маргинальных по разным параметрам. Большая часть из них расположена в боковых петлях белковой глобулы, что, хотя бы косвенно, могло повлиять на эти отклонения.
3. С помощью сервиса PDB_redo возможно улучшение некоторых показателей качества модели, однако низкое качество данных не позволяет значительно улучшить модель.

Список литературы

- [1] R.-Y. Yang, G. a Rabinovich, and F.-T. Liu, "Galectins: structure, function and therapeutic potential.," *Expert Rev. Mol. Med.*, vol. 10, no. July 2015, p. e17, 2008.
- [2] В. Д. Якушина, О. А. Васильева, Н. В. Рязанцева, В. В. Новицкий, О. Е. Савельева, Т. С. Прохоренко, Е. Г. Старикова, and А. П. Зима, "Галектин-1 : роль в формировании особенностей врожденного и приобретенного иммунитета," pp. 21–32, 2012.
- [3] "UniProtKB - P11762 (LEG1_RAT)." [Online]. Available: <http://www.uniprot.org/uniprot/P11762>.
- [4] J. J. Lundquist and E. J. Toone, "The cluster glycoside effect," *Chem. Rev.*, vol. 102, no. 2, pp. 555–578, 2002.
- [5] E. Salomonsson, A. Larumbe, J. Tejler, E. Tullberg, H. Rydberg, A. Sundin, A. Khabut, T. Frejd, Y. D. Lobsanov, J. M. Rini, U. J. Nilsson, and H. Leffler, "Monovalent interactions of galectin-1," *Biochemistry*, vol. 49, no. 44, pp. 9518–9532, 2010.
- [6] "EDS: 3m2m." [Online]. Available: <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=3m2m>.
- [7] "MolProbity." [Online]. Available: <http://molprobity.biochem.duke.edu/>.
- [8] "User guide to the wwPDB X-ray validation reports." [Online]. Available: <http://wwpdb.org/validation/ValidationPDFNotesXray>.
- [9] R. P. Joosten, K. Joosten, G. N. Murshudov, and A. Perrakis, "PDB-REDO: Constructive validation, more than just looking for errors," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 68, no. 4, pp. 484–496, 2012.