

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

Отчёт по качеству расшифровки структуры мутантного фермента глюкансукразы из
Streptococcus mutans
(PDB ID 3A1C) методом
рентгеноструктурного анализа

Нуждина Екатерина,
студентка 4 курс

Москва,
2015г

Аннотация

В отчёте рассмотрена общая информация о ферменте глюкансукразе, некоторые индикаторы качества модели структуры 3А1С фермента глюкансукразы, а также проведён анализ маргинальных аминокислотных остатков.

Введение

Актуальность

Мы чистим наши зубы дважды в день фтористой зубной пастой, используем средство для полоскания рта, ограничиваем количество употребляемого сахара...но все равно у нас появляется кариес. Кариес появляется из-за бактерий, которые поедают некоторые сахара из нашей пищи, сбраживают их и выделяют кислоты. Эти кислоты разъедают полезные минералы наших зубов. Казалось бы, что проще просто вычистить эту бактерию и избавиться от них навсегда. Но все не так-то просто, у этих бактерий есть способы избежать этого.

Бактерии используют фермент глюкансукразу, что бы строить длинные, липкие цепи сахаров, которые называются глюканы. Эти глюканы приклеивают бактерии на поверхность наших зубов, формируя биопленку, от которой трудно избавиться. Таким образом, изучение фермента глюкансукразы может помочь в создании ингибиторов этой структуры, чтобы предотвратить возможность формирования биопленки бактериями и защитить зубы от кариеса.^[1]

Общие сведения

Глюкансукраза принадлежит к семейству гликозид гидролаз 70. Они катализируют формирование глюканов с разнообразными гликозидными связями из сукрозы посредством реакции трансгликозилирования. Глюкансукразы делятся на 4 разных типа на основе связей, которые они катализируют: mutansucrase, dextransucrase, alternansucrase и reuteransucrase [названия не переведены на русский, так как корректного перевода в каких-либо источниках я не нашла], которые катализируют $\alpha(1-3)$, $\alpha(1-6)$, $\alpha(1-3$ и $1-6)$ и $\alpha(1-4$ и $1-6)$ гликозидные связи, соответственно.^[3]

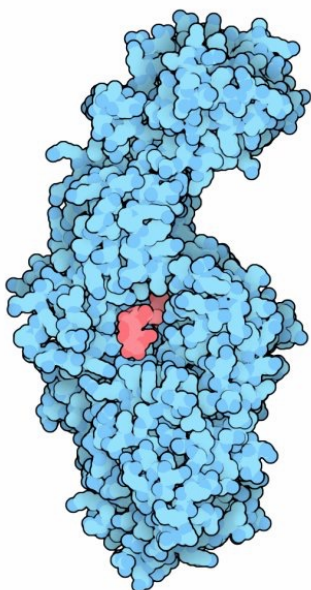


Рис.1. Структура бактериальной глюкансукразы с короткой цепью глюкана. ^[1]

Фермент обозначен голубым цветом, глюкан - красным.

При сравнение аминокислотной последовательности глюкансукраз с другими гликозид-гидролазами было найдено, что они очень похожи на альфа-амилазы. Замечены “перестановки последовательности”, то есть часть в начале последовательности глюкансукразы соответствует части в середине последовательности альфа-амилазы.

Позже по кристаллической структуре подтвердили циклические перестановки в глюкансукразе по сравнению с альфа-амилазами, домены A, B и IV состоят из двух отдельных полипептидных цепей. Сравнение глюкансукразы (3aie) и альфа-амилазы (1smd) представлены на рисунке.

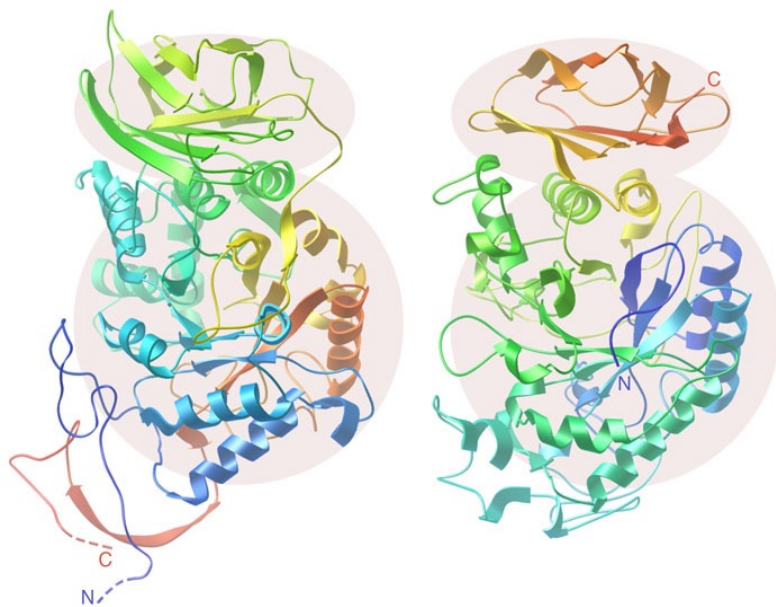


Рис.2. Укладка цепей и доменная структура глюкансукразы (слева) и альфа-амилазы (справа).^[1]
Цепи покрашены от красного с одного конца до синего до другого конца.

Кристалл

Структуру рекомбинантной и селенометионин-замещенной гликозилтрансферазы из *Streptococcus mutans* (GTF-SI) экспрессировали, очистили и закристаллизовали.^[4]

GTF-SI это mutansucrases, которая в основном синтезирует нерастворимый глюкан с $\alpha(1-3)$ гликозидной связью. GTF-SI формирует тетрамер в кристалле, но по результатам геле-фильтрация GTF-SI предпочтительно является мономером в растворе. Структура единиц, входящих в тетрамер почти идентичная.

На рисунке 3 представлена доменная организации структуры GTF-SI из *S. mutans*. Для дальнейшей работы была выбрана структура 3AIC - комплекс GTF-SI с акарбозой.

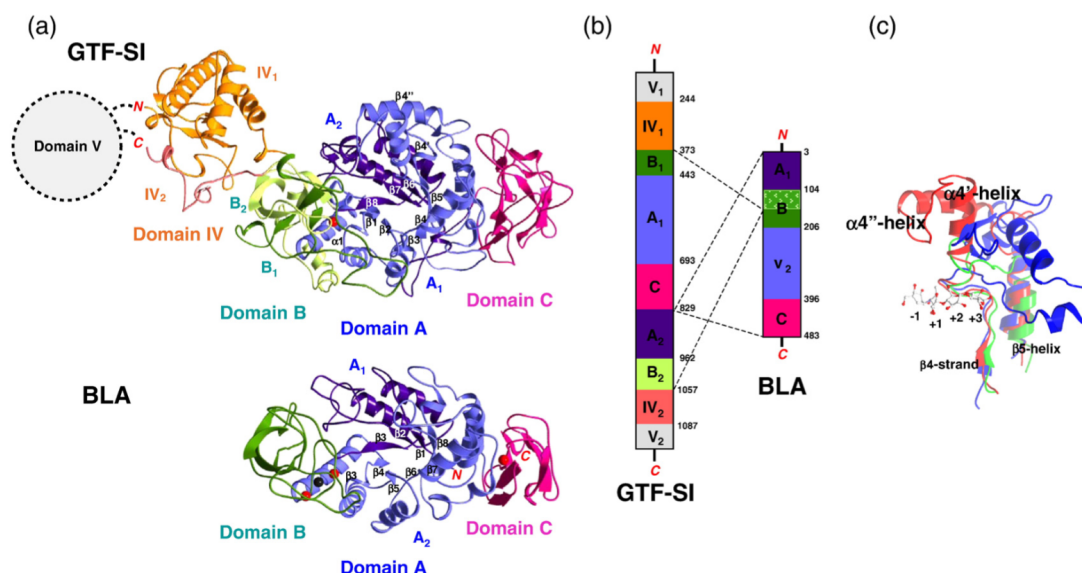


Рис. 3. Структура GTF-SI из *S. mutans*. [2] (a) Ribbon диаграмма of GTF-SI и *Bacillus licheniformis* α -амилаза (BLA). В GTF-SI, домены IV, A и B состоят из цепей IV₁ (оранжевая) и IV₂ (лососевый), цепей A₁ (синяя) и A₂ (фиолетовая) и цепей B₁ (зеленая) и B₂ (желто-зеленый), соответственно. В BLA, домен B состоит из единичной полипептидной цепи, в то время как домен A состоит из двух цепей. Домены C (магента) в обоих ферментах состоят из единичной полипептидной цепи. β -тяжи в TIM-баррели нумерованы так как они встречаются в аминокислотной последовательности. Связанный кальция (красный) и натрий (черный) ионы приведены в виде сфер. (b) Круговая перестановка доменов между GTF-SI и BLA. (c) Структура альфа-спиралей $\alpha 4'$ и $\alpha 4''$ GTF-SI. Акарбоза в активном сайте показана в модели проволочного каркаса (wireframe). GTF-SI, BLA и амилосукразы окрашены в красный цвет, зеленый и синий, соответственно.

Результаты и обсуждения

Общая информация о модели и индикаторы её качества в целом и по отдельным аминокислотным остаткам

Модель структуры ЗАИС была получена в 2011 году.[2] Общая информация о модели ЗАИС систематизирована в таблице 1.

Таблица 1. Общая информация о модели ЗАИС

Разрешение	3.11 Å
Диапазон разрешения	61.77 - 3.11 Å
Число измеренных рефлексов	238566
Число использованных рефлексов	226540
Метод решения фазовой проблемы	SAD Single-wavelength anomalous diffraction (аномальная дисперсия)
Кристаллографическая группа	P2 ₁ 2 ₁ 2

Для решения фазовой проблемы использовался метод аномальной дифракции (SAD - Single-wavelength anomalous diffraction), в структуре присутствуют аномально рассеивающие атомы селенометионина. В таблице 2 представлена информация о кристаллографической ячейке.

Таблица 2. Параметры кристаллографической ячейки

Длина (Å)	Угол (°)
a = 295.52	$\alpha = 90$
b = 214.41	$\beta = 90$
c = 220.66	$\gamma = 90$

Значения индикаторов качества модели представлены в таблице 3. Значения существенно отклоняющиеся от пороговых значений по данным с сервиса MolProbity отмечены salmon цветом.

Таблица 3. Значения индикаторов качества модели в целом

	Значение	Процент	Пороговый уровень
R-фактор	0.211		
R_free	0.244		
R-free - R	0.033		
RSR-Z = (RSR - <RSR>) / sigma	<RSR-Z> = -0.21, Sigma = 0.61		
Полные маргиналы по карте Рамачандрана	97	1.47 %	<0.05%
Число остатков в предпочитаемой области карты Рамачандрана	6012	91.16 %	>98%
Clash score число недопустимых наложений (> 0.4 Å) атомов на 1000	23.88		91st percentile* (N=59, 2.86Å - 9999Å)
Число остатков с маргинальными по отклонению от ротаметров боковыми цепями	437	7.81 %	<0.3%
Cis пролины	8 / 236	3.39 %	≤5%
Валентные углы, существенно отклоняющиеся от теории	132 / 72482	0.18 %	<0.1%

	Значение	Процент	Пороговый уровень
Ковалентные связи, существенно отклоняющиеся от теории	74 / 53370	0.14 %	0 %

Для данной структуры R-фактор равен 0.211. R_{free} равняется 0.244, а (R-free - R) равняется 0.033. Хорошие значения: $R_{free} < 20\%$, в нашем случае это значение немного превышает. $(R_{free} - R) > 10\%$ могло бы настораживать в отношении переоптимизации (overfitting), но в нашей случае у нас значение $< 10\%$, значит у нас нет переоптимизации модели.

Поиск маргинальных остатков

Real Space R (RSR) характеризует насколько модель группы атомов (или даже отдельного атома) соответствует “экспериментальной” электронной плотности. Сумма берется по узлам пространственной решетки в окружении группы атомов. Хорошие значения: $RSR < 10\%$ Плохие: $> 20\%$. Гистограмма значения RSR для всех остатков цепи А структуры 3AIC приведена на рисунке 4.

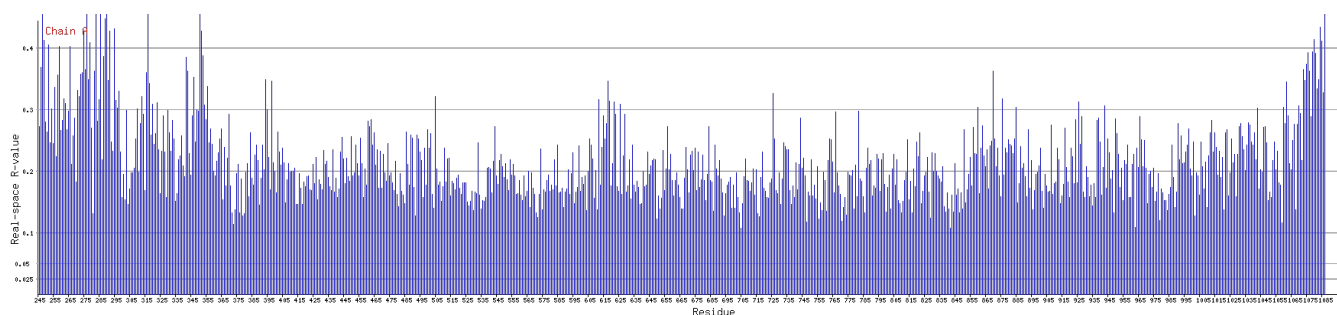


Рис.4. Пространственный R-фактор (RSR) для всех остатков цепи А структуры 3AIC (сервер EDS)

Z-score - относительный R-фактор остатков/молекул воды. Z-score вычисляется относительно всех остатков такого же типа из структур с примерно таким же разрешением. Значения z-score для цепи А приведены на рисунке 5.

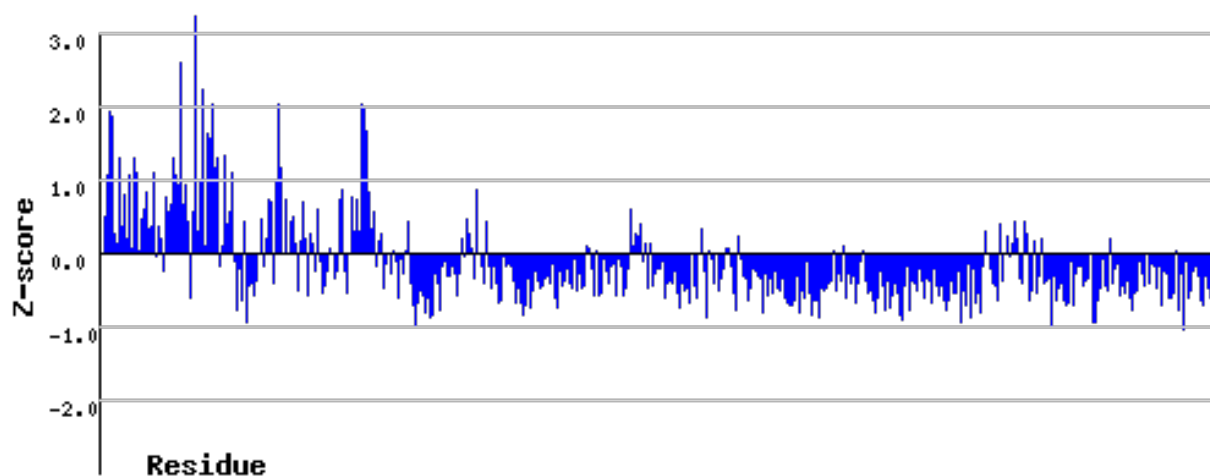


Рис.5. Значения z-score для цепи А. По оси Y отложены значения z-score, по оси X отложены номера остатков в цепи А. Значения, которые превышают 2.0 считаются маргинальными.

Высокие положительные значения $Z > 2$ свидетельствуют о том, что остаток плохо вписан в электронную плотность по сравнению с другими структурами с тем же разрешением, а значит является маргинальным. Выборочный список маргинальных остатков приведен в таблице 4.

Таблица 4. Список маргинальных остатков

Маргинальный остаток	RSR	z-score	Плохой угол
Chain A, 277 ILE	0.571	2.6	
Chain A, 283 THR	0.547	3.23	omega
Chain A, 286 GLN	0.502	2.22	
Chain A, 290 LYS	0.556	2.04	omega
Chain A, 317 LEU	0.557	2.04	
Chain A, 1087 PRO	0.792		

Для построения карты Рамачандрана на рисунках 4 и 5 были использованы серверы EDS и MolProbity соответственно. На обеих картах Рамачандрана, показывающих положение пар торсионных углов ϕ/ψ полипептидной цепи, есть остатки, которые лежат вне допустимой области.

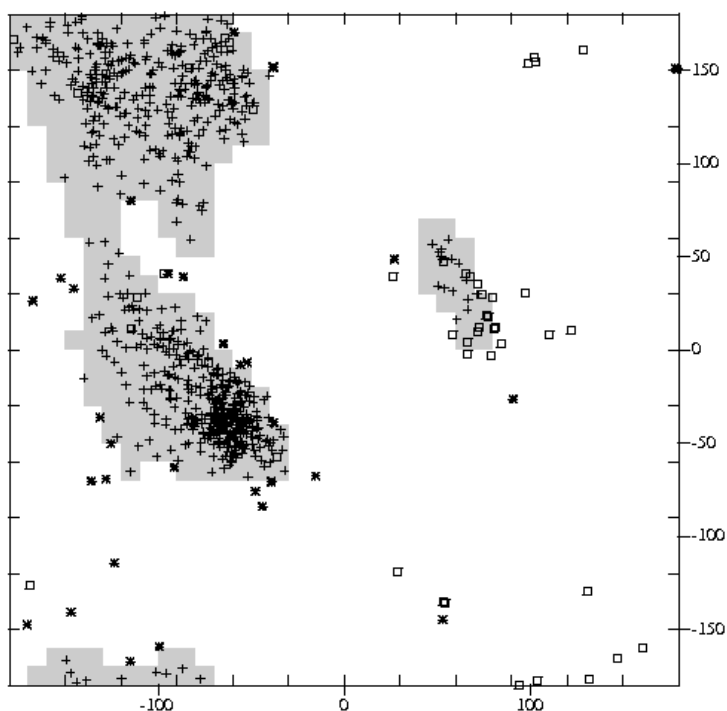


Рис. 4. Карта Рамачандрана для структуры белка 3A1C. Остатки, находящиеся в недопустимой области отмены звездочками и квадратами. Карта построена сервером EDS.

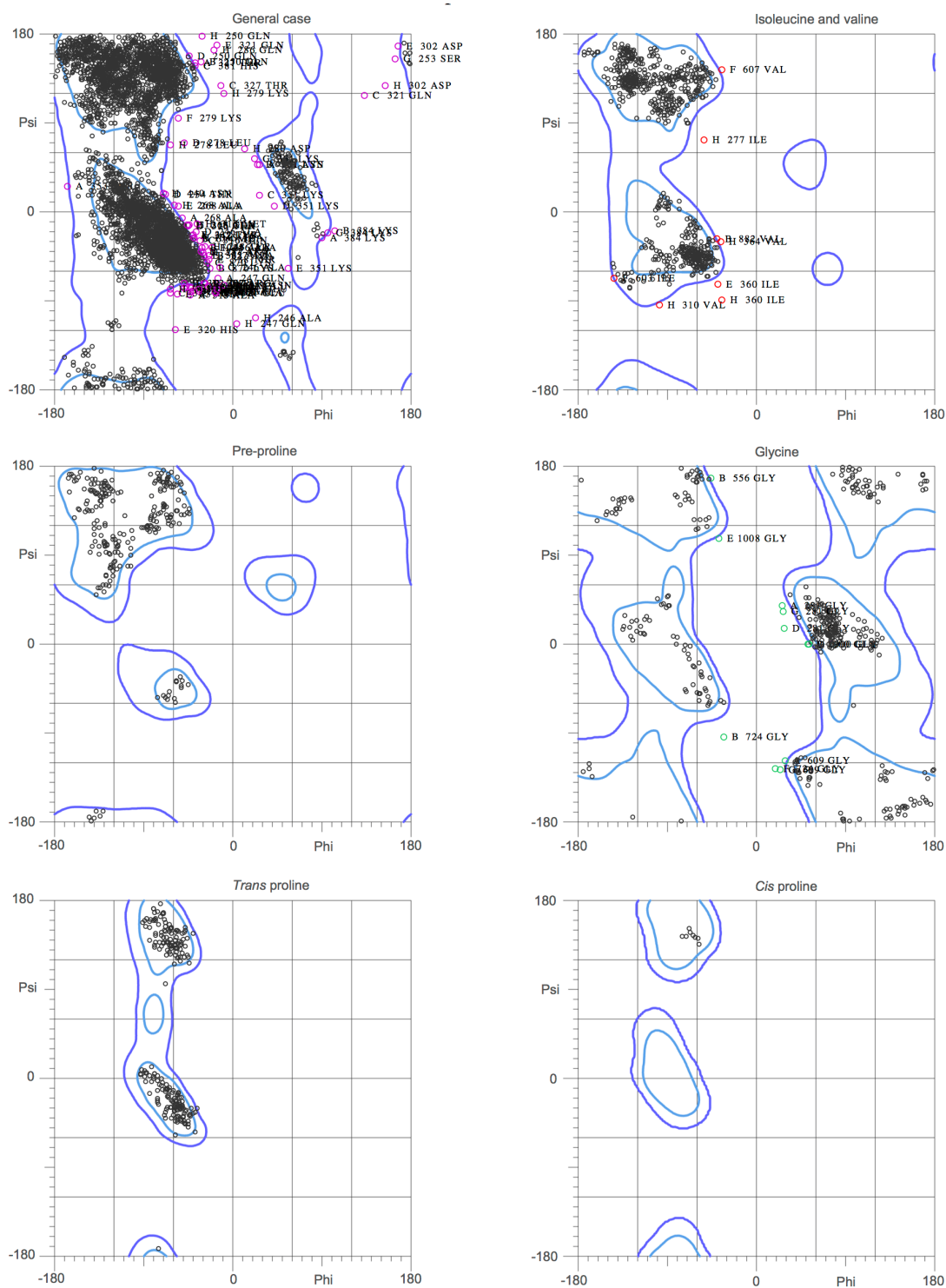


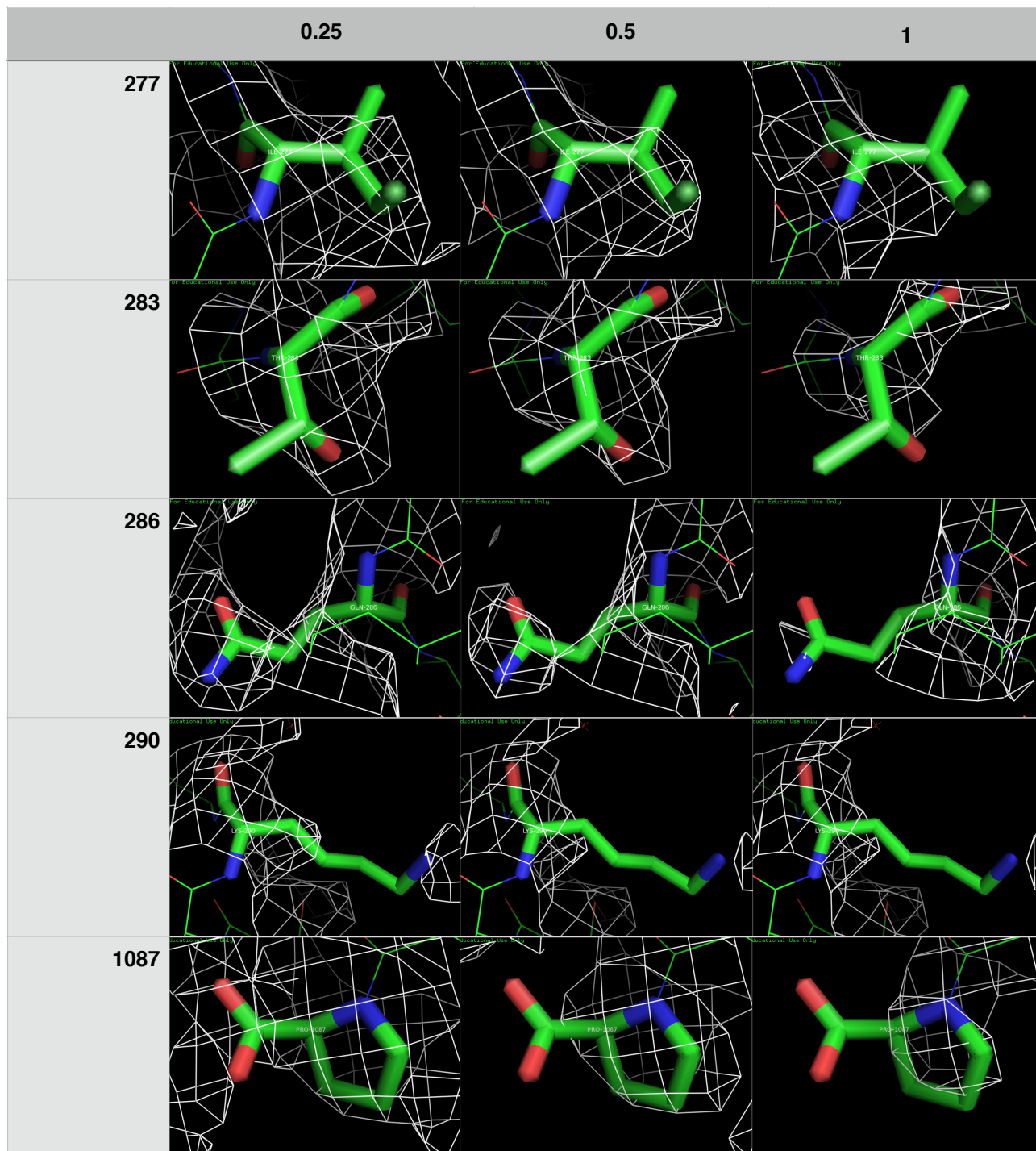
Рис.5. Карта Рамачандрана для структуры белка 3AIC. Множество пар торсионных углов ϕ/ψ для всех типов аминокислотных остатков разбито для удобства на 6 групп (общая структура, Ile и Val, Pre-proline, Gly, trans-Pro, cis-Pro). Остатки, являющиеся маргинальными по Рамачандрану обозначены цветными кружками. Остатки с допустим значениями отмечены черными кружками. Карта построена сервисом MolProbity.

По правде говор, разобраться карте Рамачандрана для всех атомов с сервиса MolProbity достаточно не просто, так как атомы накладываются друг на друга и не видно названия остатков. Но для этого случая есть список остатков в текстовом виде, что несомненно радует.

Анализ маргинальных остатков

Будем работать с маргинальными остатками приведенными в таблице 4. Изображение электронной плотности при разных уровнях подрезки (0.25, 0.5, 1) для остатков (277 ILE, 283 THR, 286 GLN, 290 LYS, 1087 PRO) представлены в таблице 5.

Таблица 5. Изображение электронной плотности



На рисунке 6 представлены выбранные в таблице 5 остатки на фоне мономера глюкансукразы.

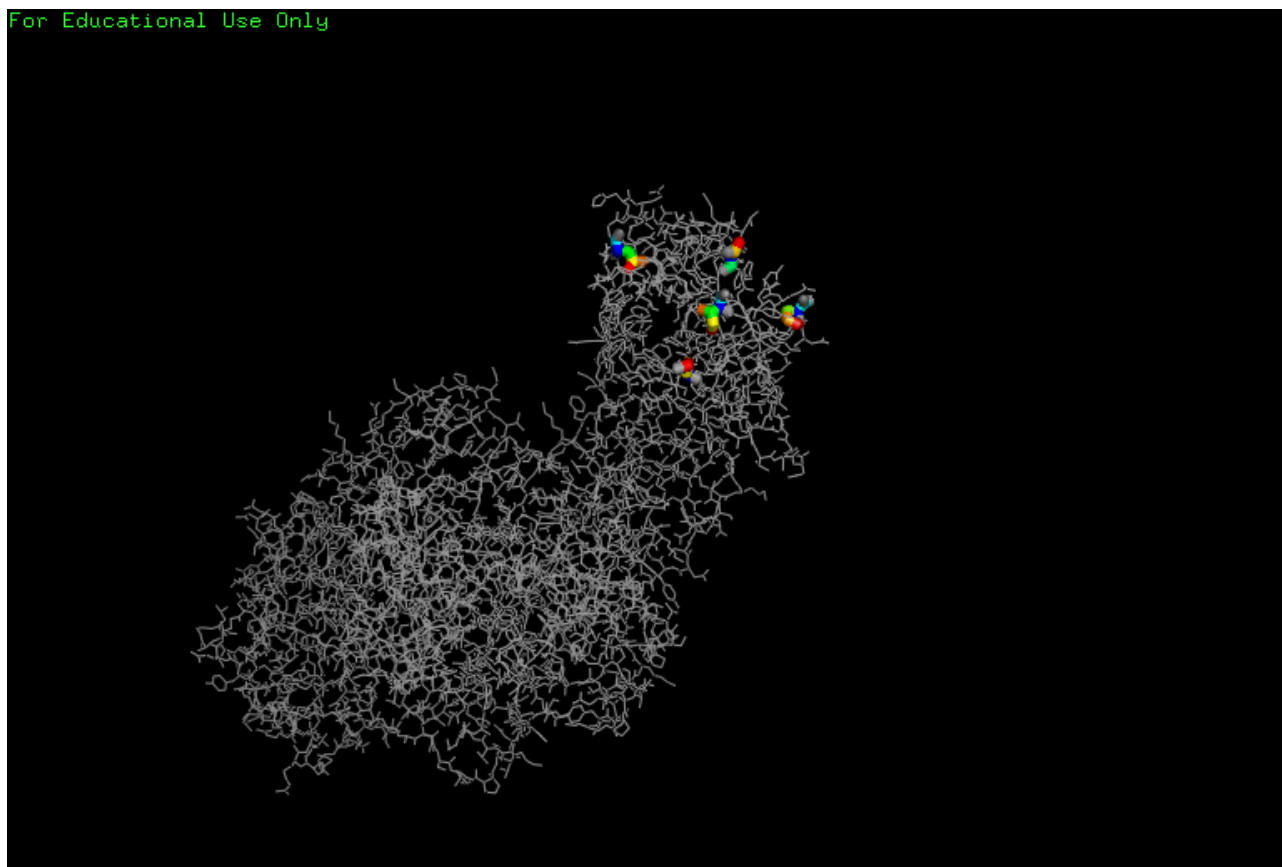
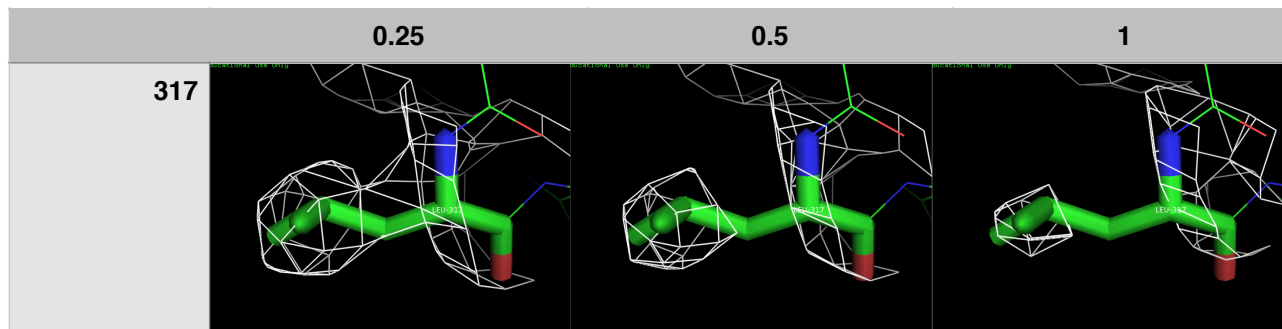


Рис.6. Маргинальные остатки цепи А структуры белка ЗАИС. Маргинальные остатки выделены разноцветными шарами на фоне общей структуры белка глюкансукразы (серая).

Для тех остатков, у которых RSR плохой, а RSR-Z – хороший, координаты атомов *resolution resolution* расшифрованы плохо, но не хуже, чем в других подобных структурах. Приблизительно такое обоснование можно привести для остатка 317, так как его RSR-Z почти подходит под хорошее значение. Изображение электронной плотности для этого остатка приведено в таблице 6. Мы видим, что в целом, остаток вписывается в электронную плотность.

Таблица 6. Изображение электронной плотности для 317 LEU



Так же из протокола качества WHAT_CHECK были получены еще данные, которые выходят за допустимые значения.

Примеры остатков, в которых полярные атомы (а) не образуют водородных связей и при этом (б) не выходят на поверхность белка представлены в таблицах 7 и 8. Наличие их странно, так как доноры водородной связи, которые находятся внутри белка обычно используют все их водороды чтобы образовывать водородные связи в белке. В очень хороших структурах число перечисленных атомов, как правило, будет нулю.

Таблица 7. Полярные атомы (доноры), не образующие водородных связи и невыходящие на поверхность белка

Остаток	Номер	Атом
TYR	248	N
SER	253	N
HIS	264	N
LEU	266	N
LEU	278	N
GLN	321	N

Таблица 8. Полярные атомы (акцепторы), не образующие водородных связи и невыходящие на поверхность белка

Остаток	Номер	Атом
ASP	280	OD1
GLN	446	OE1
ASP	477	OD2
ASP	480	OD1
ASP	521	OD1
ASP	588	OD1

Молекулы растворителя (воды), не фиксированные водородными связями (в том числе с молекулами из соседних ячеек) приведены ниже в списке:

HOH (5) B O
 HOH (18) C O
 HOH (221) C O
 HOH (33) E O
 HOH (229) E O
 HOH (189) G O

Второй столбец соответствуют номеру молекулы, третий - цепи, в которой он обнаружен. Наличие таких молекул это серьезный признак того, что были проблемы очистки фермента.

Сравнение модели из PDB с моделью из PDB_redo

Было проведено сравнение модели ЗАИС из PDB с оптимизированной моделью из PDB_REDO. Значения некоторых параметров для оценки качества двух моделей приведены в таблице 9.

Таблица 9. Сравнение модели из PDB с моделью из PDB_redo

	PDB	PDB_REDO
R	0.2110	0.2093
R-free	0.2440	0.2273
R-free - R	0.033	0.018
Unsatisfied H-bond donors/ acceptors	739	803

Значения R-фактора и R_{free} улучшились, и разница их также уменьшилась. Значит, новая модель лучше подогнана под экспериментальные данные. Но при этом увеличилось число полярных атомов, не образуют водородных связей. Так же по данным PDB_redo значительно улучшились показатели для 193 остатков и при этом ухудшились для 119.

Заключение

ЗАИС - это модель среднего разрешения. В целом качество её разрешения и можно оценить как «среднее» на основании индикаторов качества (надо так же учитывать, что в основном мы рассматривали только одну цепь). Из наблюдаемых в ЗАИС аномалий можно отметить высокие показатели RSR-фактора и Z-score (RSR), а так же остатки, аномальные по длинам и углам ковалентных связей. Сервер PDB_REDO улучшил такие показатели модели, как R-фактор и R_{free} , но при этом появились маргинальные остатки по другим параметрам. Да и в целом улучшение показателей для 193 остатков и при этом ухудшение для 119 остатков говорит о посредственном улучшении модели.

Список литературы

[1]

Статья с сайта PDB <http://pdb101.rcsb.org/motm/138>

[2]

Keisuke Ito, Sohei Ito, Tatsuro Shimamura, Simone Weyand, Yasuaki Kawarasaki, Takumi, Misaka,

Keiko Abe, Takuya Kobayashi, Alexander D. Cameron and So Iwata (2011), Crystal Structure of Glucansucrase from the Dental Caries Pathogen *Streptococcus mutans*, *J. Mol. Biol.* (2011) 408, 177–186, doi:10.1016/j.jmb.2011.02.028

[3]

Monchois, V., Willemot, R. M. & Monsan, P. (1999). Glucansucrases: mechanism of action and structure– function relationships. *FEMS Microbiol. Rev.* 23, 131–151.

[4]

Ito, K., Ito, S., Shimamura, T., Kawarasaki, Y., Abe, K., Misaka, T. et al. (2010). Crystallization and preliminary X-ray analysis of a glucansucrase from the dental caries pathogen *Streptococcus mutans*. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* 66, 1086–1088.

Сервисы

EDS <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=3aic>

PDB <http://www.rcsb.org/pdb/explore/explore.do?structureId=3aic>

molprobity <http://molprobity.biochem.duke.edu/index.php>

PDB_redo www.cmbi.ru.nl/pdb_redo/ai/3aic/index.html