

Факультет биоинженерии и биоинформатики,
МГУ им. М.В. Ломоносова

Анализ качества структуры MIF4G
домена DAP5
Homo sapiens
PDB ID – 4IUL

Мищенко Полина,
Гр.401

Москва, 2019г

Содержание

Аннотация.....	3
Введение.....	4
Результаты.....	5
Краткая сводка о PDB структуре.....	5
Основные показатели качества структуры.....	6
R-фактор (R-free).....	6
Clashscore.....	7
Карта Рамачандрана	7
Ротамеры	7
RSRZ	9
Анализ маргинальных остатков	10
Улучшение параметров модели с помощью PDB-REDO	14
Выводы	16
Список литературы.....	17

Аннотация

В работе анализируется качество структуры DAP5M с PDB ID 4IUL.

Оценка качества была проведена по основным критериям, приведенным на сайте PDB.

Были найдены маргинальные остатки и рассмотрены некоторые из них.

Введение

DAP5 – белок, который участвует в инициации трансляции и является членом факторов инициации трансляции eIF4G, которые собираются на 5'-Сар структуре мРНК в качестве преинициаторного комплекса. Как именно он участвует в инициации трансляции пока не известно. Возможно он обеспечивает независимую от Сар трансляцию посредством механизма, включающего рекрутирование рибосомы непосредственно в специфические мРНК, которые содержат внутренний сайт входа в рибосому (IRES) в их 5'-нетранслируемой области (UTR). DAP5 стимулирует IRES-зависимую трансляцию клеточных мРНК в лизатах ретикулоцитов кролика (RRLs), но полностью не зависит от сар-зависимой трансляции.

Центральный домен MIF4G рекрутирует РНК-хеликазу eIF4A. Связывающие IRES детерминанты DAP5 могут также находиться в пределах его домена MIF4G. Домен MIF4G DAP5 имеет идентичность последовательности 39% с eIF4GI и 43% с eIF4GII. Пока нет структурной информации о связывании MIF4G с IRES. Кристаллическая структура дрожжевого комплекса eIF4G-eIF4A показала, что выпуклая поверхность MIF4G вступает в контакт с eIF4A как в его N-концевых, так и в C-концевых областях. Как и человек, дрожжи также обладают двумя изоформами eIF4G. Кристаллическая структура содержит изоформу I, но для простоты в тексте упоминается eIF4G. Эта структура выявила третий сайт взаимодействия: консервативный остаток триптофана (Trp 579 в дрожжевом eIF4GI, также известном как TIF4631; Trp 734 в человеческом eIF4GI; Trp 733 в человеческом eIF4GII) на N-конце домена MIF4G, прикрепленного к eIF4A C-конца домена. Мутация этого остатка в дрожжевых eIF4G ослабляет его взаимодействие с eIF4A и приводит к потере способности стимулировать активность eIF4A-АТФазы *in vitro*. Этот остаток триптофана и большинство остатков в домене MIF4G, которые вступают в прямой контакт с eIF4A, сохраняются в DAP5.

Результаты

Краткая сводка о PDB структуре

Структура 4IUL была загружена 23 Марта 2013 года на PDB, статья с названием «Structural Analysis of the DAP5 MIF4G Domain and Its Interaction with eIF4A» была опубликована в 2013 году. Авторы исследования: Virgili, G., Frank, F., Feoktistova, K., Sawicki, M., Sonenberg, N., Fraser, C.S., Nagar, B.

Структура DAP5M с PDB ID 4IUL имеет следующие характеристики (Таблица 1).

Метод решения фазовой проблемы	Молекулярное замещение	
Число измеренных рефлексов	28017	
Полнота набора структурных факторов	99.8%(87.5)	
Диапазон разрешений структурных факторов	2.3Å-30Å	
Разрешение структуры	2.30 Å	
Параметры кристаллографической ячейки	a = 167.696	$\alpha = 90.00$
	b = 56.696	$\beta = 112.05$
	c = 74.374	$\gamma = 90.00$
Кристаллографическая группа	C2	
Некристаллографические симметрии	отсутствуют	

Таблица 1 Основные характеристики структуры

Основные показатели качества структуры

Основными показателями качества, которые приводятся на сайте PDB, являются R-фактор и связанный с ним R-free, критерий перекрывания Ван-дер-Ваальсовых радиусов атомов, то есть Clashscore, карта Рамачандрана (процент маргинальных остатков), процент маргинальных остатков по положению боковой цепи, названный Sidechain outliers (ротамеры), и RSRZ outliers, доля маргинальных остатков по RSR. Значения этих показателей показаны на рисунке 1.

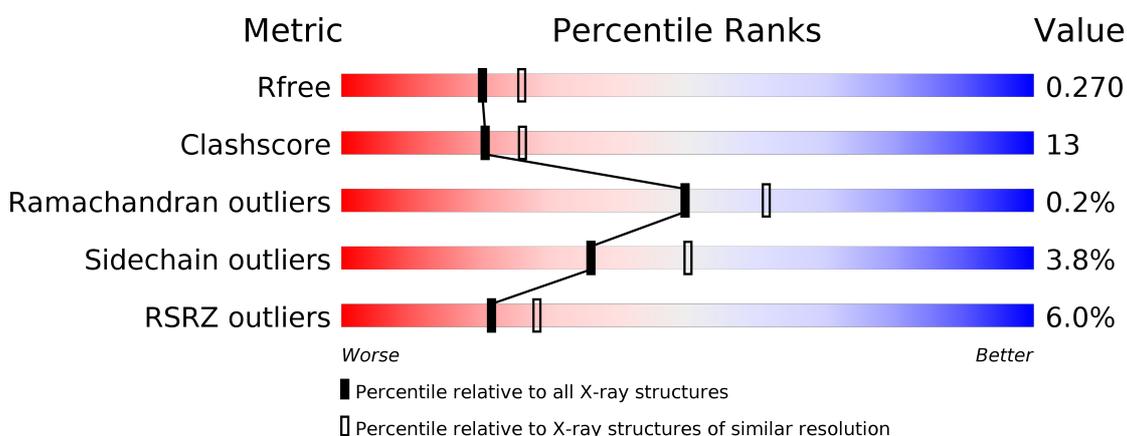


Рисунок 1. Основные показатели качества структуры

Рассмотрим подробнее каждый из этих критериев.

R-фактор (R-free)

R-фактор – мера согласованности между кристаллографической моделью и экспериментальным массивом данных. Он вычисляется по формуле: $R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$

Хорошими значениями R-фактора считают значения ниже 25%. В нашем случае – 22.2%. Это значит, что модель достаточно хорошо вписывается в экспериментальные данные.

R-free – значение R-фактора, рассчитанное для небольшого процента гармоник (~5%), не участвовавших в оптимизации модели. Значение, R-free всегда выше R-фактора. То, насколько велика эта разница, показывает насколько модель переоптимизирована. Хорошим значением для R-free так же является меньше 25%, в нашем случае - 25.6%, а разница составляет 3.4%. Это не очень хорошие показатели для модели.

Clashscore

Clashscore можно еще назвать счетчиком соударений. Для подсчета этого показателя в структуре должны быть атомы Н или их нужно добавить. Пересечение Ван-дер-Ваальсовых радиусов больше чем на 0.4\AA считается столкновением (clash). Clashscore – количество столкновений на 1000 атомов, включая Н. Clashscore нашей модели = 13. Суммарное количество столкновений в модели = 103.

Карта Рамачандрана

Карта Рамачандрана – способ визуализации двугранных углов (ψ и ϕ) аминокислот в белках. Карта имеет энергетические области 3 типов, которые соответствуют разрешенным, запрещенным и предпочитаемым комбинациям ψ и ϕ .

Число маргинальных остатков по углам ψ и $\phi = 1$ (80 LYS), что составляет 0.22%, порог для хорошей структуры $\sim <0.05\%$. Число остатков в предпочтительных позициях = 440 (95%). Порог для хорошей структуры $\sim >98\%$. Судя по этим данным, модель не является хорошей по доле маргиналов карты Рамачандрана. Однако, во всей структуре нашелся только 1 маргинал. Поэтому плохие процентные показатели можно списать на небольшое количество аминокислот в последовательности белка.

Карта Рамачандрана нашей структуры изображена на рисунке 2.

Ротамеры

Белковые боковые цепи принимают в основном определенные значения торсионных углов, называемые ротамерами, аналогично карте Рамачандрана, только для углов χ . Показатель sidechain рассчитывается как процент остатков с необычной конформацией. Доля маргиналов боковой цепи нашей структуры – 3.8% (4% по molprobity). Порог для хорошей структуры (по molprobity) - $<0.3\%$. Таким образом, можно сказать, что структура плохая по этому показателю.

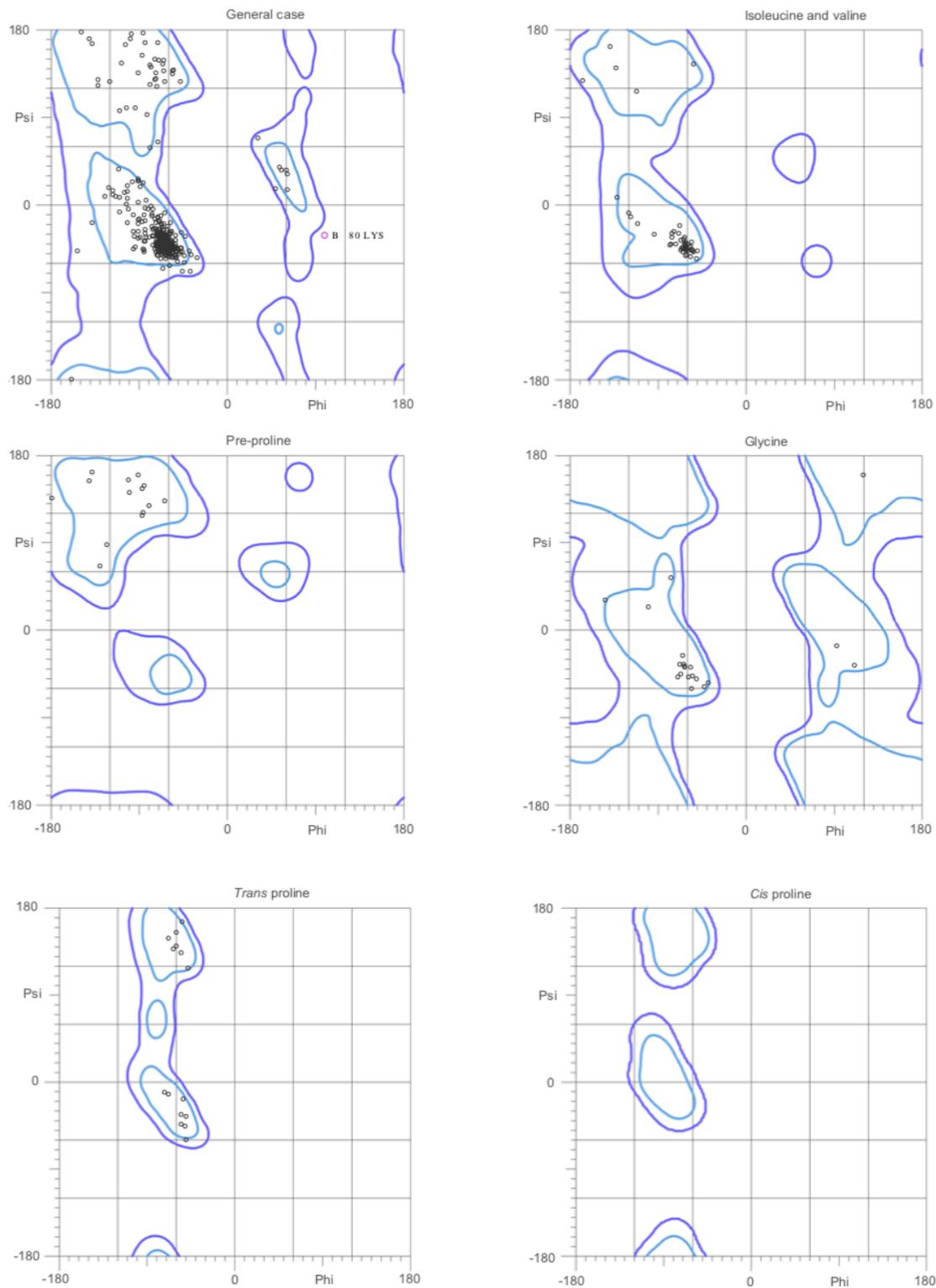


Рисунок 2. Показана карта Рамачандрана для всех остатков (верхний левый), валина и изолейцина (верхний правый), пре-пролина (центральный левый), глицина (центральный правый), транс пролина (нижний левый) и цис пролина (нижний правый). Маргинальный остаток показан красным кружком.

RSRZ

RSR (real-space R-value) – мера соответствия между частью атомной модели (один остаток) и экспериментальными данными. RSRZ (RSR Z-score) – нормализация RSR по типу остатка и разрешению. RSRZ рассчитывается только для стандартных аминокислот. Остаток считается выбросом, если его RSRZ больше 2. Показатель RSRZ показывает долю выбросов по RSRZ относительно общего количества остатков, для которых был рассчитан RSRZ.

Остаток, маргинал по RSRZ, скорее всего неправильно вписан. Доля таких остатков в структуре – 6% (28 остатков). Это является высоким показателем для структур данного разрешения (22 перцентиль).

Анализ маргинальных остатков

В таблице 2 указаны маргинальные остатки по различным показателям.

Показатель	Остатки
Карта Рамачандрана	LYS80
Столкновения	ARG79-LYS80, LYS80-ILE84, ASP120-GLU124, GLN243-LYS245
Ротамер	LEU167, LYS95
Инверсия боковой цепи	ASN66, GLN134, ASN86
Нестандартные аминокислоты	нет
RSRZ	LYS80 (2.7 σ), ARG241 (4 σ)
Остатки в cis-конформации	нет

Анализ остатка LYS80

LYS80 считается маргинальным остатком одновременно по нескольким показателям:

- Карта Рамачандрана (единственный маргинал);
- Столкновения;
- RSRZ

Тот факт, что он является маргиналом сразу по 3 показателям, изначально наталкивает на мысль, что остаток вписан неправильно. Однако, при детальном рассмотрении электронной плотности на рисунке 3 видно, что остаток вписан в электронную плотность не очень хорошо. Таким образом, остаток с очень высокой долей вероятности вписан не правильно и действительно является маргинальным по углам ψ и ϕ и перекрытия Ван-дер-Ваальсовых радиусов. Не возникает сомнений, что RSRZ действительно равен 2,7, что является одним из максимальных значений.

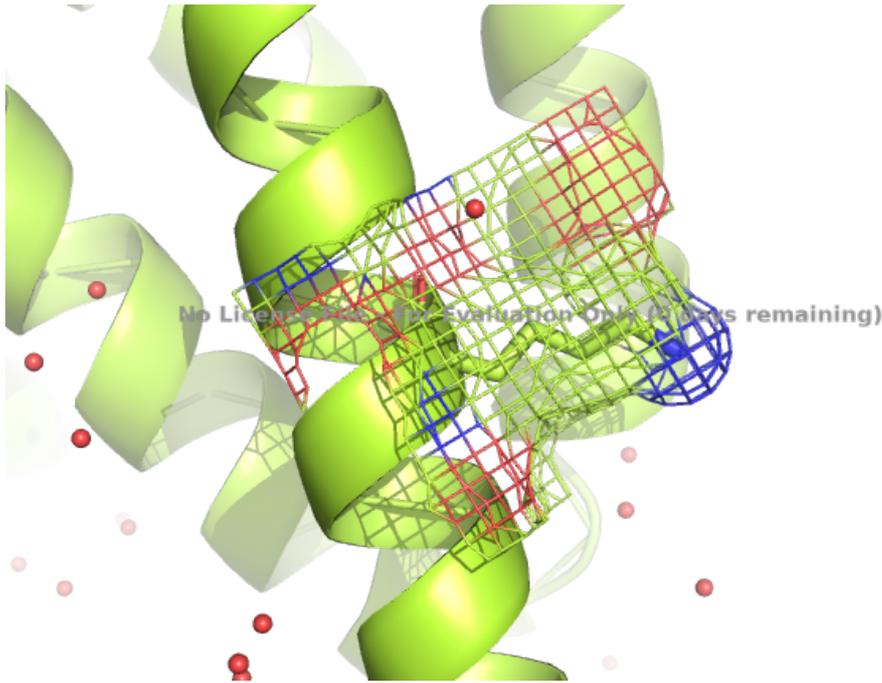


Рисунок 3. LYS80 с электронной плотностью

Анализ остатка ASN66

Остаток ASN66 считается кандидатом на инверсию боковой цепи. Посмотрим на электронную плотность на рисунке 4. Остаток очень хорошо вписывается в электронную плотность. Кроме того, белок выходит на поверхность и образует водородные связи своей боковой цепью. Таким образом, остаток не является маргинальным.

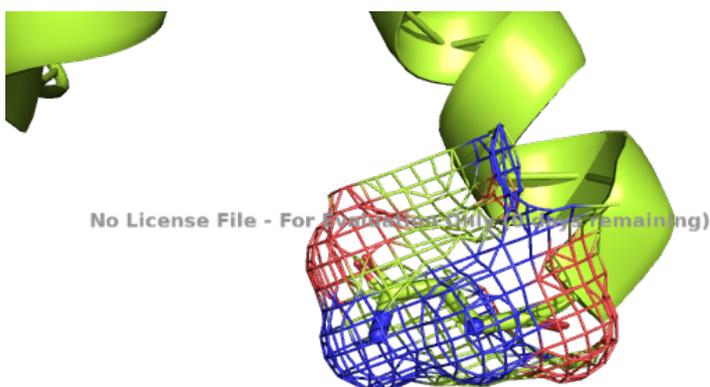


Рисунок 4. ASN66 с электронной плотностью

Анализ остатка LEU167

LEU167 считается маргиналом по боковой цепи. Посмотрим на рисунке 5 электронную плотность этого аспарагина. По очертаниям электронной плотности сложно что-то сказать, скорее всего этот остаток вписан не правильно с такими нехарактерными углами. Ошибка может быть связана с особенностями кристаллизации молекулы.

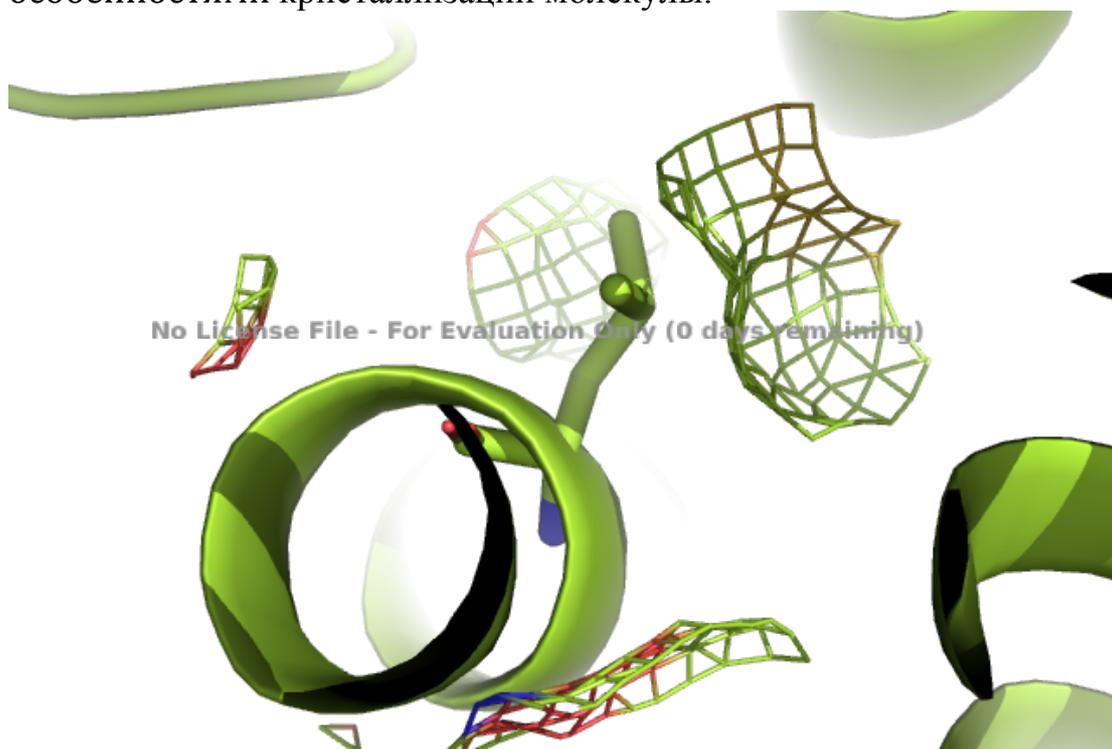


Рисунок 5. LEU167 с электронной плотностью

Анализ столкновения ASP120-GLU124

Рассмотрим остатки ASP120 и GLU124 считающиеся кандидатами на пересечение ван-дер-ваальсовых радиусов. Рассмотрим эти остатки на рисунке 6. Расстояние между ASP120.O и GLU124.HG3 составляет 2Å. Судя по электронной плотности можно сказать, что атомы вписаны правильно и расстояние между приведенными выше атомами действительно такое. Значит, между этими остатками происходит столкновение.

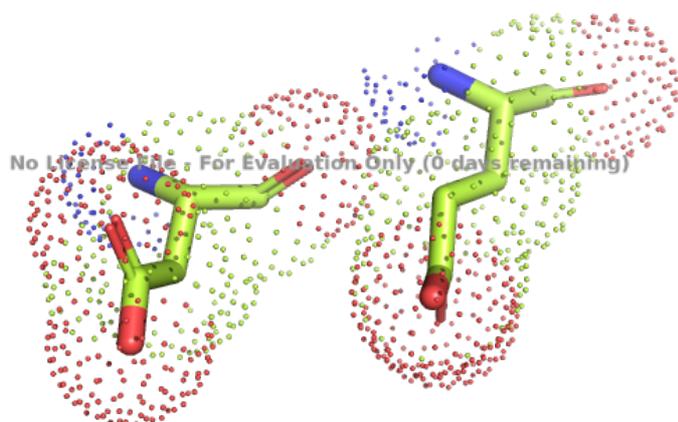


Рисунок 6. Сферами из точек окружены атомы-кандидаты на пересечение радиусов Ван-дер-Ваальса ASP120 и GLU124

Улучшение параметров модели с помощью PDB-REDO

PDB-REDO – сервис, оптимизирующий модели кристаллических структур. Алгоритм имеет полностью автоматическую систему принятия решения по уточнению, восстановлению и проверки структуры.

Сервис приводит таблицы сравнения исходной структуры и измененной. Приведем в таблице 3 ряд параметров, которые изменились после улучшения структуры. Зеленым выделены параметры, которые были улучшены.

	PDB	PDB-REDO
Crystallographic refinement		
<i>R</i>	0,2925	0,1874
<i>R-free</i>	0,3218	0,2361
<i>Bond length RMS Z-score</i>	0,298	0,598
<i>Bond angle RMS Z-score</i>	0,572	0,853
Model quality (raw scores percentiles)		
<i>Ramachandran plot appearance</i>	24	75
<i>Rotamer normality</i>	37	40
<i>Coarse packing</i>	72	84
<i>Fine packing</i>	47	69
<i>Bump severity</i>	56	82
<i>Hydrogen bond satisfaction</i>	53	82

Таблица 3. Сравнение параметров модели до и после работы сервиса PDB-REDO

На рисунке 7 приведено сравнение структур 4IUL и 4IUL_REDO относительно других структур схожего разрешения (3275 структур). По рисунку видно, что PDB_REDO делает статистически более хорошие структуры. Каждый из параметров улучшается по сравнению с PDB структурами.

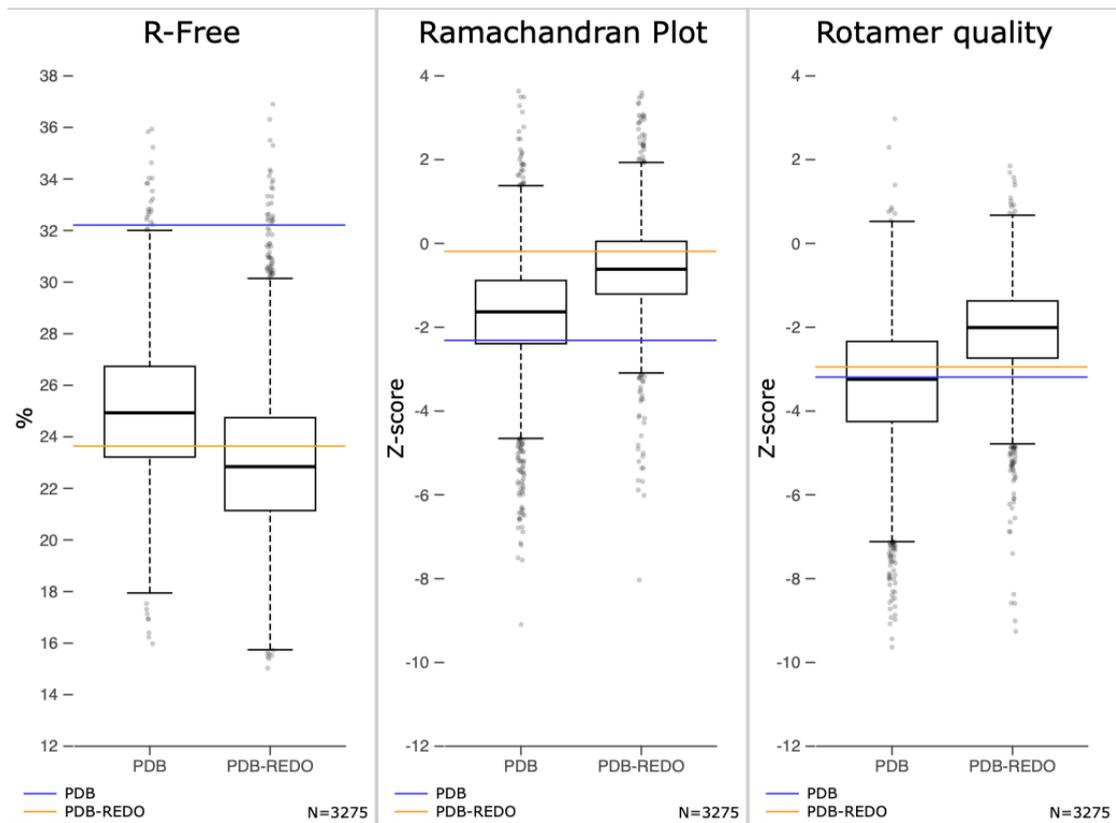


Рисунок 7. Сравнение основных параметров модели 4IUL и 4IUL_REDO относительно 3275 структур схожего разрешения. Синей полосой показаны значения в структуре PDB, а желтой – PDB-REDO. Бокс плоты показывают распределение этих параметров среди структур такого же разрешения. Для каждого показателя приведено 2 бокс плота, для PDB моделей структур и PDB-REDO моделей, после улучшения.

Выводы

Анализ качества структуры показал, что по большинству критериев структура является хорошей, а по некоторым – не очень хорошей. В структуре есть некоторое количество остатков, считающихся маргинальными. Подробное рассмотрение 5 из них показало, что только в одном из этих случаев было произведено неадекватное вписывание в электронную плотность. В остальных случаях, модель хорошо соответствовала электронной плотности. Некоторые из рассмотренных остатков имеют особенности строения, сильно отходящие от нормы для таких аминокислот: очень близкие контакты (столкновения), ротамеры и маргиналы карты Рамачандрана.

Список литературы

1. Structural analysis of the DAP5 MIF4G domain and its interaction with eIF4A PMID: 23478064
2. https://files.rcsb.org/pub/pdb/validation_reports/iu/4iul/4iul_full_validation.pdf
3. <https://www.ebi.ac.uk/pdbe/entry/pdb/4iul>
4. <https://pdb-redo.eu/db/4iul>