

ОТЧЁТ ПО КАЧЕСТВУ РАСШИФРОВКИ СТРУКТУРЫ 4R5F, ПОЛУЧЕННОЙ МЕТОДОМ РСА

МГУ им. М. В. Ломоносова Факультет
биоинженерии и биоинформатики 2016г.

Худякова Ксения Андреевна
khudyakovaks@gmail.com

Аннотация

В отчёте рассмотрены индикаторы качества модели в целом структуры цистеиновой десульфуразы IscS из организма *Archaeoglobus fulgidus* с идентификатором 4R5F, а также индикаторы качества для пяти маргинальных аминокислотных остатков.

Введение

ISC (iron-sulfur cluster) – железо-серный кластер. Это класс простетических групп, вовлеченных во многие реакции в организме. Для образования этих кластеров важен класс ферментов, известных как пиридоксаль-5'-фосфат-зависимые десульфуразы, использующие L-цистеин в качестве субстрата, и производящие персульфид. [1] Персульфид участвует в формировании железо-серных кластеров. В архее *Archaeoglobus fulgidus* была обнаружена OPC, напоминающая цистеиновую десульфуразу класса IscS, но в которой ключевой остаток Lys199 был заменён. Авторы [1] хотели понять, каковы функции этой OPC. В эксперименте разрешена структура с мутацией D199K (рис. 1). По результатам экспериментов авторы предполагают, что описанная OPC не имеет энзиматической активности, а её предполагаемой функцией может быть предоставление цистеина для неэнзиматически растущего Fe-S кластера.



Рис. 1 Структура с идентификатором 4R5F. [2]

Результаты и обсуждение

Общая информация о модели

Белок состоит из одной цепи длиной в 382 аминокислоты. Содержит пиридоксаль фосфат-связывающий сайт.

Структура была разрешена в 2014 году авторами Adrien Pagnier, Yvain Nicolet, Juan C. Fontecilla-Camps. Авторами было получено 30913 рефлексов, что составляет 97.8% полноты. Тестирование модели проводилось на 5% рефлексов. Для решения фазовой проблемы использовался метод молекулярного замещения. Разрешение структуры 1.90 Å, диапазон разрешений структурных факторов 46.49 - 1.90 Å.

Табл. 1 Параметры кристаллографической ячейки

Длина (Å) Угол (°)

a = 70.62 α = 90

b = 102.89 β = 90

c = 108.59 γ = 90

Значения индикаторов качества модели в целом

Значение R-фактора модели равно 0.1531, R-free 0.2008. R-фактор меньше 0.25 – это хороший показатель, R-free больше 0.2 – формально это плохой показатель. Но этот порог превышен очень незначительно.

Разница между значениями равна 4.7% - это меньше 10%, значит модель не переоптимизирована. Значение RSR равно 0.188, что больше 0.1, характерного для хорошей модели.

Для построения карты Рамачандрана (рис. 2) был использован сервер MolProbity [4]. Все аминокислотные остатки находятся в разрешенной области карты, выбросов нет. Помимо карты, сервер выдал таблицу с некоторыми характеристиками качества модели (табл.2).

Жёлтым в таблице обозначены те параметры, которые не удовлетворяют показателям качества хорошей модели. Тем не менее, существует ещё красное обозначение, которое совсем не удовлетворяет, но для 4R5F ничего не выделено красным. Если сравнить с последней колонкой, референсными значениями, видно, что отклонение не очень большое. Так что в целом структура очень хороша.

Сервер EDS [3] выдал значения Z-score для различных аминокислотных остатков (рис. 3). Остатки, для которых RSR Z-score превышает 2.0 (не вписываются в электронную плотность):

- 34 VAL 2.41
- 91 ILE 3.96
- 114 VAL 4.16
- 163 VAL 2.65
- 378 PRO 3.58
- 331 VAL 3.42
- 338 LYS 3.3
- 342 ALA 3.48

Интересно, что в PDB Validation report значения RSRZ другие, но маргинальные остатки те же (табл. 3).

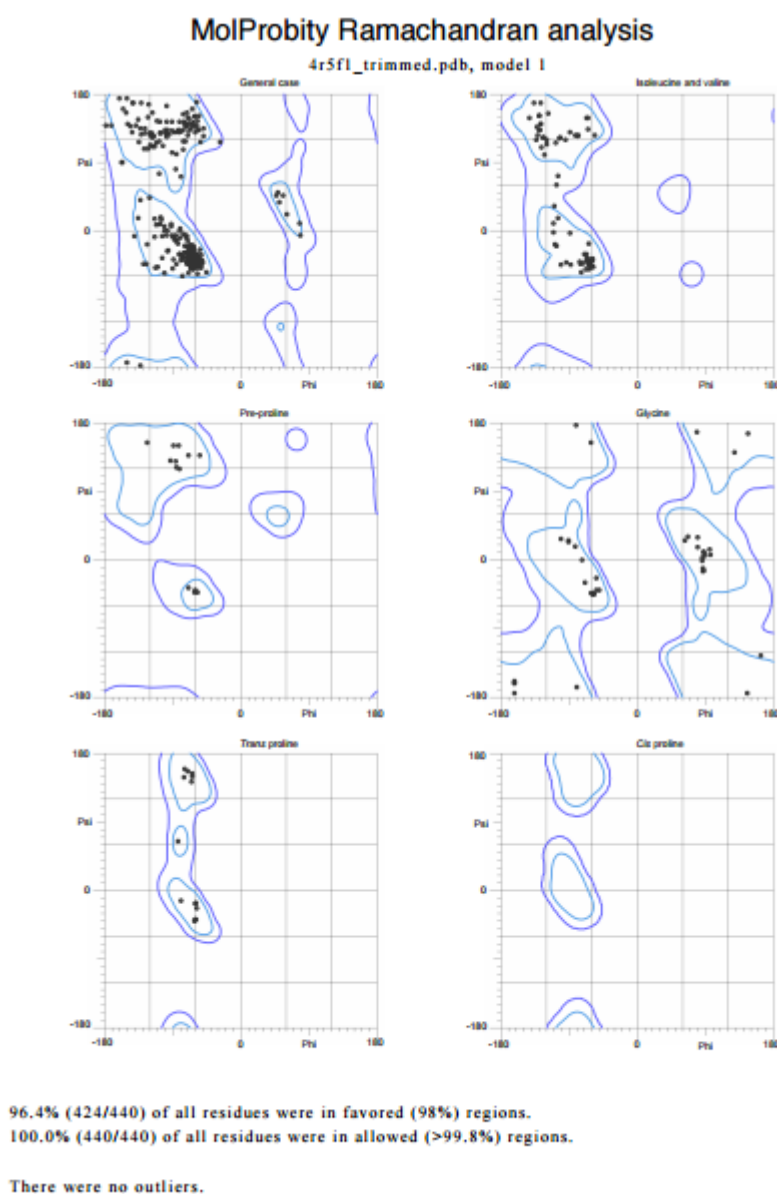


Рис. 2 Карта Рамачандрана для структуры 4R5F.

Табл. 2 Критерии качества модели, выданные сервисом MolProbity. Желтым – значения, не удовлетворяющие условиям хорошей модели. Последняя колонка – референсные значения.

Protein Geometry	Poor rotamers	4	1.22%	Goal: <0.3%
	Favored rotamers	315	95.74%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	353	96.71%	Goal: >98%
	C β deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 3166	0.00%	Goal: 0%
	Bad angles:	0 / 4329	0.00%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 14	0.00%	Expected: ≤ 1 per chain, or $\leq 5\%$

In the two column results, the left column gives the raw count, right column gives the percentage.

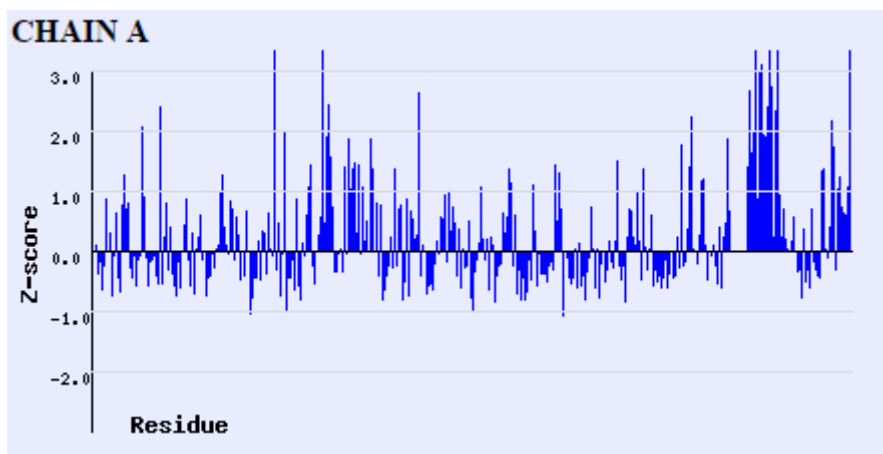


Рис. 3 Z-score для отдельных остатков.

Табл. 3 Маргинальные остатки по RSRZ согласно PDB Validation report

Mol	Chain	Res	Type	RSRZ
1	A	342	ALA	5.8
1	A	338	LYS	4.6
1	A	378	PRO	4.5
1	A	298	TYR	4.0
1	A	377	SER	3.5
1	A	343	HIS	3.1
1	A	337	LEU	3.1
1	A	299	ILE	3.0
1	A	341	GLU	2.4
1	A	333	MET	2.3
1	A	339	HIS	2.3
1	A	329	SER	2.2
1	A	373[A]	LEU	2.2
1	A	375	SER	2.0
1	A	374	ARG	2.0

Анализ отдельных маргинальных аминокислотных остатков

Табл. 4 Некоторые маргинальные остатки в структуре 4R5F

ОСТАТОК	КРИТЕРИЙ
338 LYS, 378 PRO, 342 ALA, 298 TYR, 377 SER, 343 HIS	RSRZ > 2
86 ARG, 94 SER, 220 VAL, 317 THR	Неправильная конформация боковых цепей (углы χ)

В основном выбросы есть только по критерию RSRZ, еще есть 4 (согласно MolProbity) или 2 (согласно PDB Validation report) выброса по углам χ . По остальным критериям все остатки хорошие (рис. 4). Я выбрала пять остатков с самым экстремальным отклонением: 342 ALA, 338 LYS, 378 PRO, 86 ARG, 220 VAL.

Ramachandran	Rotamer	C β deviation	Bond lengths	Bond angles	Cis Peptides
Outliers: 0 of 443	Poor rotamers: 4 of 329	Outliers: 0 of 373	Outliers: 0 of 430	Outliers: 0 of 437	Non-Trans: 0 of 394

Рис. 4 Оценка качества структуры 4R5F сервисом MolProbity.

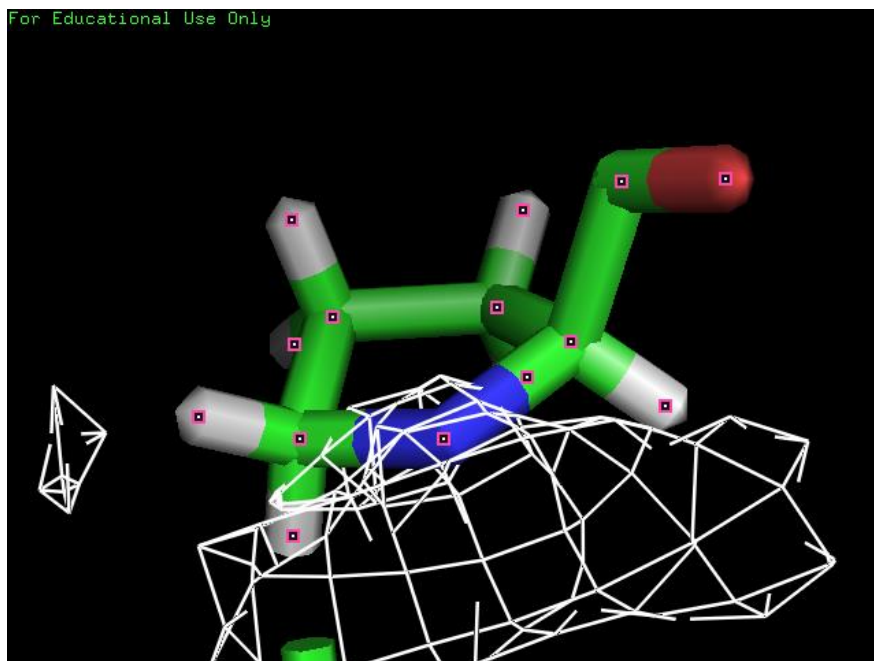


Рис. 5 378PRO

Остаток на рис.5 плохо вписан в электронную плотность (RSRZ - 4.6). Температурный фактор 130.09, что больше чем в два раза превышает средний. На рисунке видно, что электронная плотность действительно не покрывает этот остаток. Так что, по всей видимости, это ошибка расшифровки.

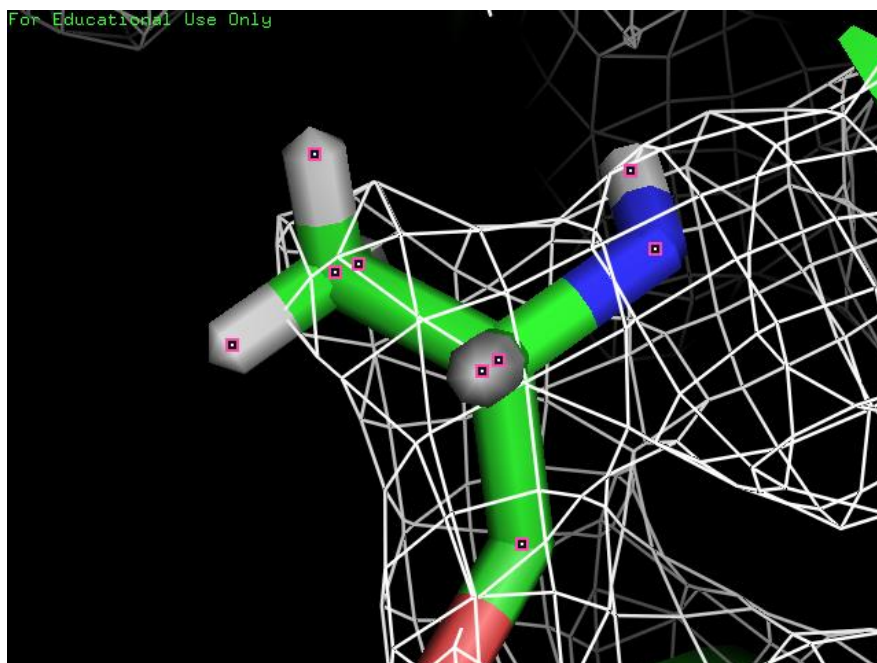


Рис. 6 342ALA

RSRZ для данного остатка – 5.8. Тем не менее, на рис.6 видно, что он прекрасно вписан в электронную плотность. Я посмотрела на RSR остатка чтобы понять, в чём же проблема. Согласно сервису EDS, RSR остатка 0.271, что больше порога в 20%. Видимо, аланины обычно вписаны в ЭП идеально, а этот чуть хуже. Следовательно, маргинальность остатка – ошибка расшифровки.

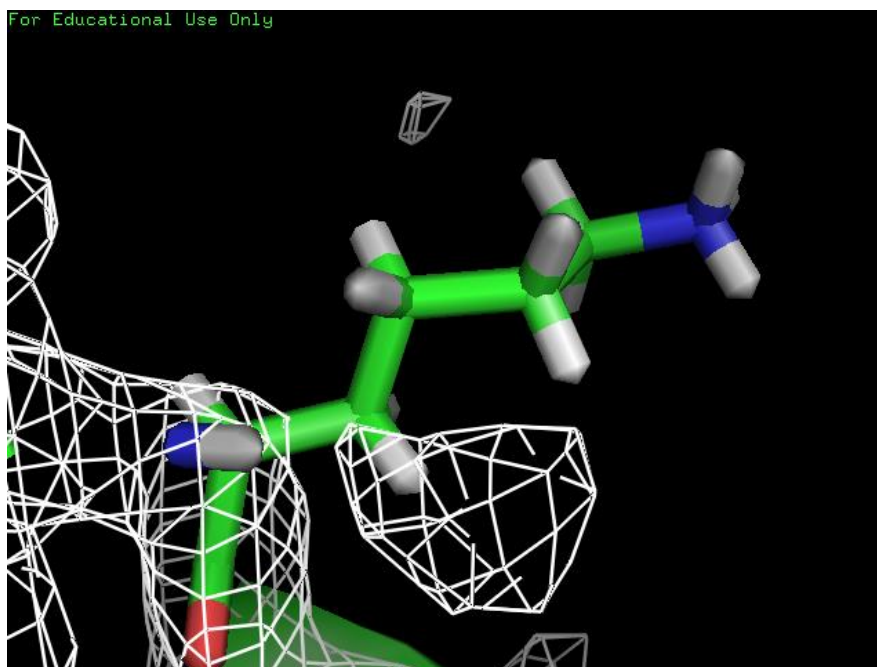


Рис. 7 338LYS

Лизин на рис. 7 совершенно не вписан в ЭП, RSR 0.377. Маргинальность остатка – ошибка расшифровки.

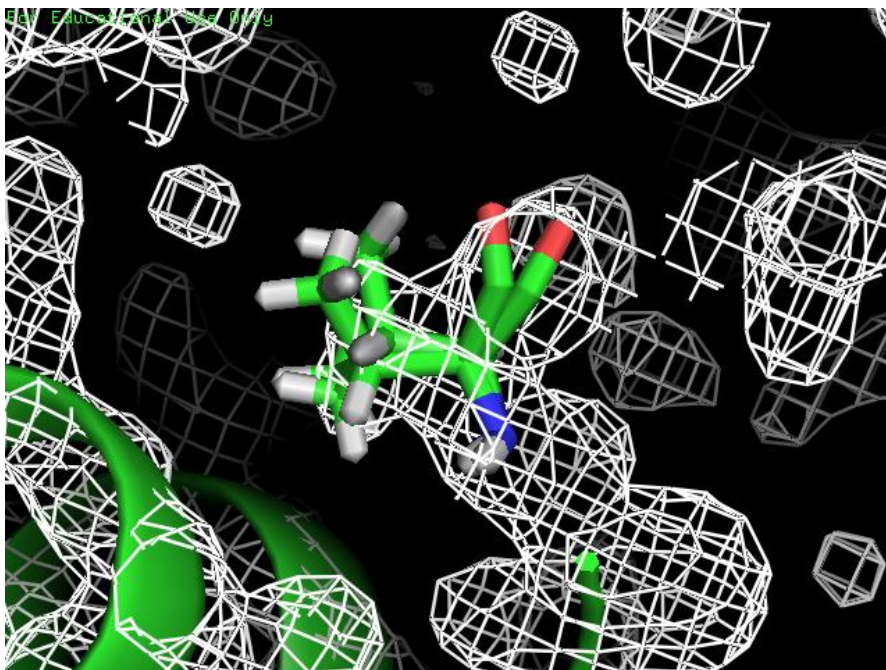


Рис. 8 220VAL

Проблема с остатком 220 на рис. 8 – угол боковой цепи ($\chi = 136.4$). Возможно, это нарушение угла связано с тем, что аминокислотный остаток получился «раздвоенным». При этом причин для этого я не нахожу – ЭП не раздваивается вместе с ним. Атом CG1 в обоих вариантах не вписывается в ЭП. Температурный фактор этого остатка почти равен среднему по модели, поэтому на температурные флуктуации раздвоенность и нарушение угла тоже списать нельзя. Видимо, маргинальность остатка – ошибка расшифровки.

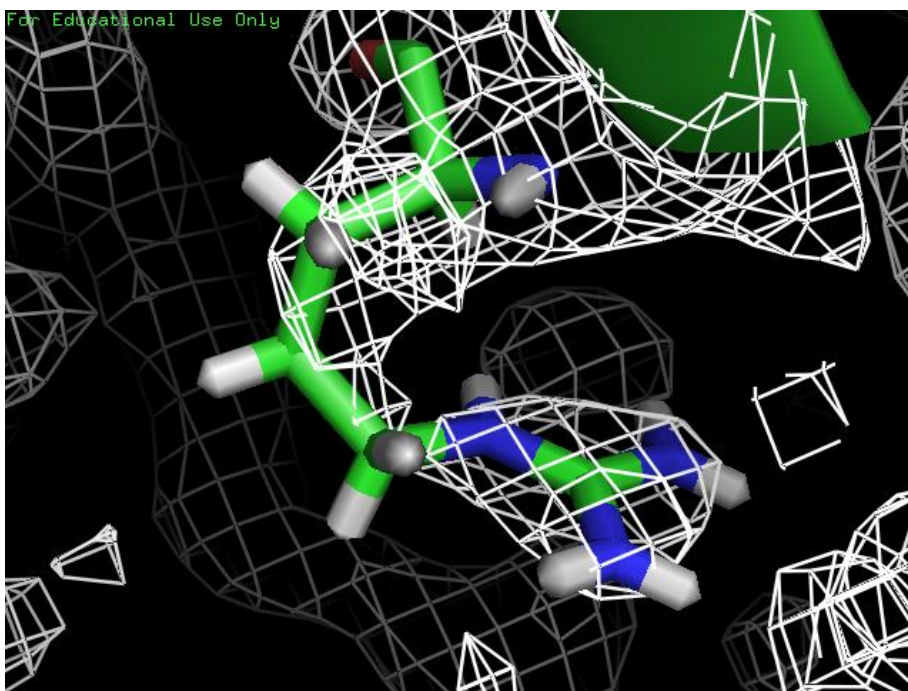


Рис. 9 86ARG

У остатка на рис.9 неправильные углы χ : 246.8, 64.2, 283.2, 157.8. При этом боковая группа не образует водородных связей, что странно. С другой стороны, азоты радикала очень

хорошо вписаны в электронную плотность. Возможно, что этот остаток не маргинал, а его изгиб и отсутствие водородных связей – его функциональная особенность.

Сравнение модели из PDB с моделью из PDB_redo

С использованием сервера PDB_REDO была построена еще одна структура (4r5f_redo) по тем же экспериментальным данным, что и исходная. В таблице 5 можно увидеть некоторые характеристики двух моделей.

Табл. 5 Сравнение моделей 4r5f и 4r5f_redo по некоторым индикатором качества. Зелёным – хороший показатель, жёлтым – средний, красным – плохой.

Параметр	4r5f	4r5f_redo
R-фактор	0.1510	0.1753
R-free	0.2010	0.2180
Число маргинальных остатков по карте Рамачандрана	0	0
Число остаток в предпочитаемой области карты Рамачандрана	353(96.71%)	356(97.53%)
Неправильные углы χ	4(2.34%)	8(1.22%)
Неправильная длина связей	0	3
Отклонения угла $C\beta$	0	0

Все показатели, которые изменились, сервер ухудшил. Кроме числа остатков в предпочитаемой области карты Рамачандрана – их количество он увеличил, но это не помогло структуре перейти порог в 98%.

Сервер ухудшил соответствие с электронной плотностью для девяти остатков, но улучшил для двенадцати. 377й серин, который и так был плохо вписан в ЭП (см. табл. 4), он существенно ухудшил, а соседний 378й пролин, который тоже был плохо вписан, сервер значительно улучшил. Состояние остальных остатков, которые раньше были выделены как плохо вписанные в ЭП, PDB_REDO значительно не изменил.

В целом сервер не смог улучшить качество модели.

Заключение

Структура 4r5f расшифрована очень хорошо, что подтверждает график на рис.10. В ней практически нет маргинальных остатков. По количеству остатков с RSRZ > 2 она превышает среднее для моделей с таким же разрешением, но это единственный показатель, по которому она хуже других моделей.

Соответствие всех показателей качества стандартам, кроме RSRZ, может натолкнуть на мысль, что структура переоптимизирована. Но это не так, потому что разница между R-фактором и R-free составляет 0.05.

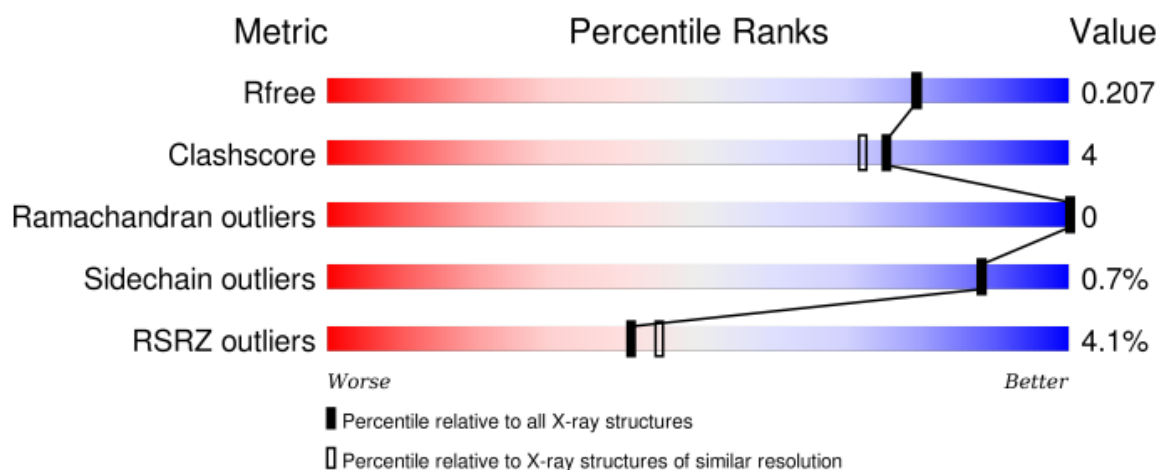


Рис. 10 Характеристики качества структуры в целом с сайта PDB

Список литературы

1. IscS from *Archaeoglobus fulgidus* has no desulfurase activity but may provide a cysteine ligand for [Fe₂S₂] cluster assembly. Adrien Pagnier, Yvain Nicolet, Juan C. Fontecilla-Camps, 2015, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research
2. <http://www.rcsb.org/pdb/explore/explore.do?structureId=4R5F>
3. <http://eds.bmc.uu.se/eds/>
4. <http://molprobity.biochem.duke.edu/>
5. <http://swift.cmbi.ru.nl/gv/pdbreport/>

