

## Оценка качества расшифровки структуры аденилаткиназы из *Bacillus subtilis* (PDBid = 3DL0)

### Введение

Аденилат киназа (АК) (EC 2.7.4.3) – очень распространенный в природе фермент, катализирующий следующую реакцию:



Существует несколько PDB структур для данного фермента, здесь рассмотрена структура 3DL0 (она же рассматривалась и в предыдущих работах).

### Общая информация о структуре

3DL0 – структура димера (2 цепи по 216 аминокислотных остатков) мутированной АК из бактерии *Bacillus subtilis* в комплексе с ионами цинка и магния, а также молекулой бисаденозинпентафосфата.

Хотя структура была получена в 2008, а сайте PDB указана ссылка на статью 2014 года за авторством Moon, S., Bannen, R.M., Rutkoski, T.J., Phillips, G.N., Bae, E [1]. Эта группа занимается оптимизацией структуры белков для получения термостойких форм. В данной статье описывается эксперимент по инженерии аденилаткиназ рода *Bacillus*, в котором есть как мезофильные так и психрофильные и термофильные организмы. Авторы использовали метод минимизации локальной структурной энтропии (LSE) для получения более устойчивых к высокой температуре аденилаткиназ. В частности, заменой некоторых аминокислотных остатков в АК из *B. subtilis* (мезофил) на аналогичные из гомолога у *B. glolispurus* (психрофил) были получены белки AKIse1, AKIse2 и AKIse3, которые, как ни странно, оказались более термостойкими чем нативный мезофильный белок.

Структура 3DL0 – это структура варианта AKIse3, который содержит 18 замен (K23D, M103Y, Q114E, T169S, A180Q, D187S, V193A, L208V, V210D, K23D, M103Y, Q114E, T169S, A180Q, D187S, V193A, L208V, V210D).

В таблице 1 приведены некоторые параметры структуры, взятые с сайта PDB.

Таблица 1. Общая информация о структуре 3DL0.

Параметр	Значение
Разрешение	1,58 Å
R-фактор	0,200
Свободный R-фактор (R-free)	0,239
Число рефлексов	59906
Полнота данных	98%
Число рефлексов для построения R-free	3038
Параметры ячейки	<b>a:</b> 34,003Å <b>b:</b> 76,688Å <b>c:</b> 77,993Å <b>α:</b> 90° <b>β:</b> 95,3° <b>γ:</b> 90°

## Оценка качества расшифровки структуры

Судить о качестве расшифровки структуры можно по нескольким показателям, в числе которых R-фактор (показатель близости экспериментальных данных к модели), свободный R-фактор (показатели степени оптимизации), а также число маргинальных аминокислотных остатков (то есть остатков с нехарактерными показателями углов и длин связей, RSR, нехарактерным окружением, etc).

В записи на сайте PDB есть краткая оценка данных параметров (рис. 1). Она строится на статистике по базе данных PDB, сравнивая качество данной модели с другими, и в качестве меры качества выдает перцентиль, соответствующий данной структуре из распределения этого показателя среди всех структур с примерно таким же разрешением или вообще всех структур.

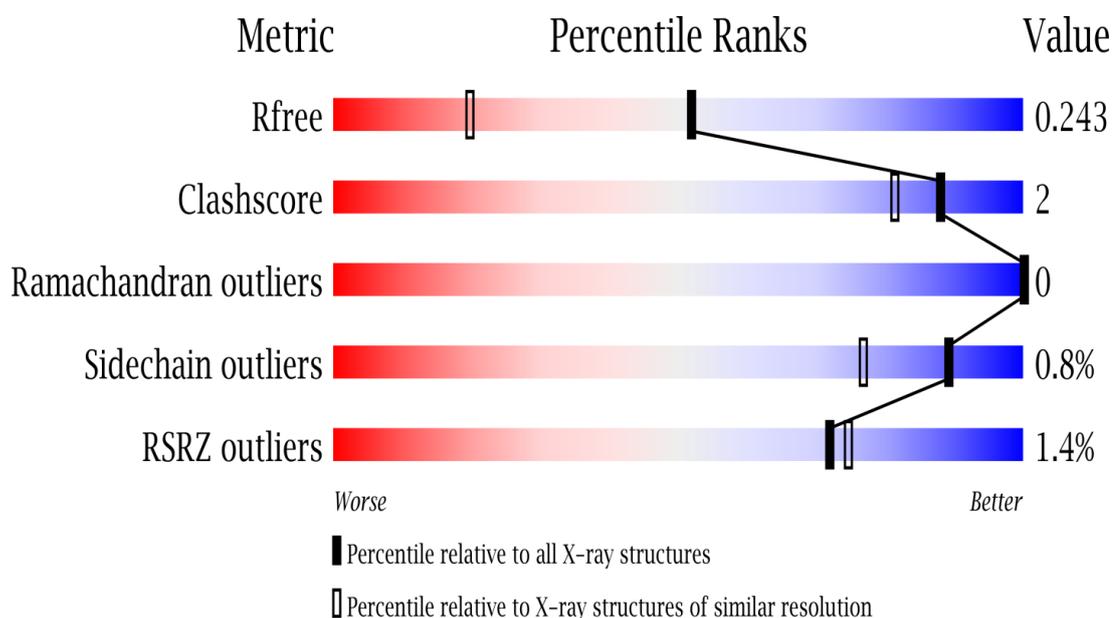


Рисунок 1. Оценка значений параметров качества для структуры 3DL0, взятая с сайта PDB.

Из рисунка 1 видно, что почти все показатели имеют хорошее значение, кроме R-free, значение которого по сравнению с другими структурами такого же разрешения плохое. Это может говорить о переоптимизации модели.

Также из PDB (через интерфейс EMBL-EBI) можно получить некоторую информацию о маргинальных остатках (подробнее см ниже), она записана в таблице 2.

Таблица 2. Маргинальные остатки структуры 3DL0.

Параметр маргинальности	Число остатков	Процент от суммарного числа остатков (%)
Угол между связями	1	0,02
Длина связи	0	0
Вписанность в карту электронной плотности	6	1,36
Карта рамачандрана	0	0
Ротамеры	4	1,07

Также с сайта PDB можно получить полный отчет об уточнении структуры [2].

Более подробную информацию (частично дублирующую информацию из PDB) о качестве структуры выдаёт сервис MolProbity. В таблице 3 представлена информация, полученная для структуры 3DL0.

Таблица 3. Информация о качестве структуры, полученная из MolProbity.

Параметр	Значение	Процент	Порог
Poor rotamers	3	0.80%	Goal: <0.3%
Favored rotamers	355	94.67%	Goal: >98%
Ramachandran outliers	0	0.00%	Goal: <0.05%
Ramachandran favored	434	99.31%	Goal: >98%
MolProbity score <sup>^</sup>	1.02		100 <sup>th</sup> percentile* (N=6761, 1.58Å ± 0.25Å)
Cβ deviations >0.25Å	0	0.00%	Goal: 0
Bad bonds:	0 / 3493	0.00%	Goal: 0%
Bad angles:	1 / 4733	0.02%	Goal: <0.1%

Таблица 4. Список маргинальных остатков.

Остаток	Критерий маргинальности
LEU47	Ротамер
MET176	Ротамер
AP5 (лиганд)	Угол связи
THR31	Ротамер

Любопытно, что на PDB указано 4 остатка с неподходящими ротамерами, а MolProbity выдает 3.

Таким образом, явно маргинальных остатков всего 3 (AP5 – лиганд), все они с плохими ротамерами, то есть с нестандартной конформацией боковой цепи. 3 маргинальных остатка на весь белок – довольно хороший показатель. Стоит отметить, что довольно распространенный тип маргинальности - нахождение торсионных углов основных цепей в запрещенных областях карты Рамачандрана - отсутствует полностью (рис. 2).

Кроме того, в структуре присутствует 6 остатков, которые плохо вписываются в электронную плотность (4 в цепи А и 2 в цепи В). У этих остатков Z-скор по RSR больше 2. На рисунке 3 показаны примеры таких остатков (структура и электронная плотность визуализированы через PyMOL). Видно, что в электронную плотность частично не вписываются боковые группы остатков.

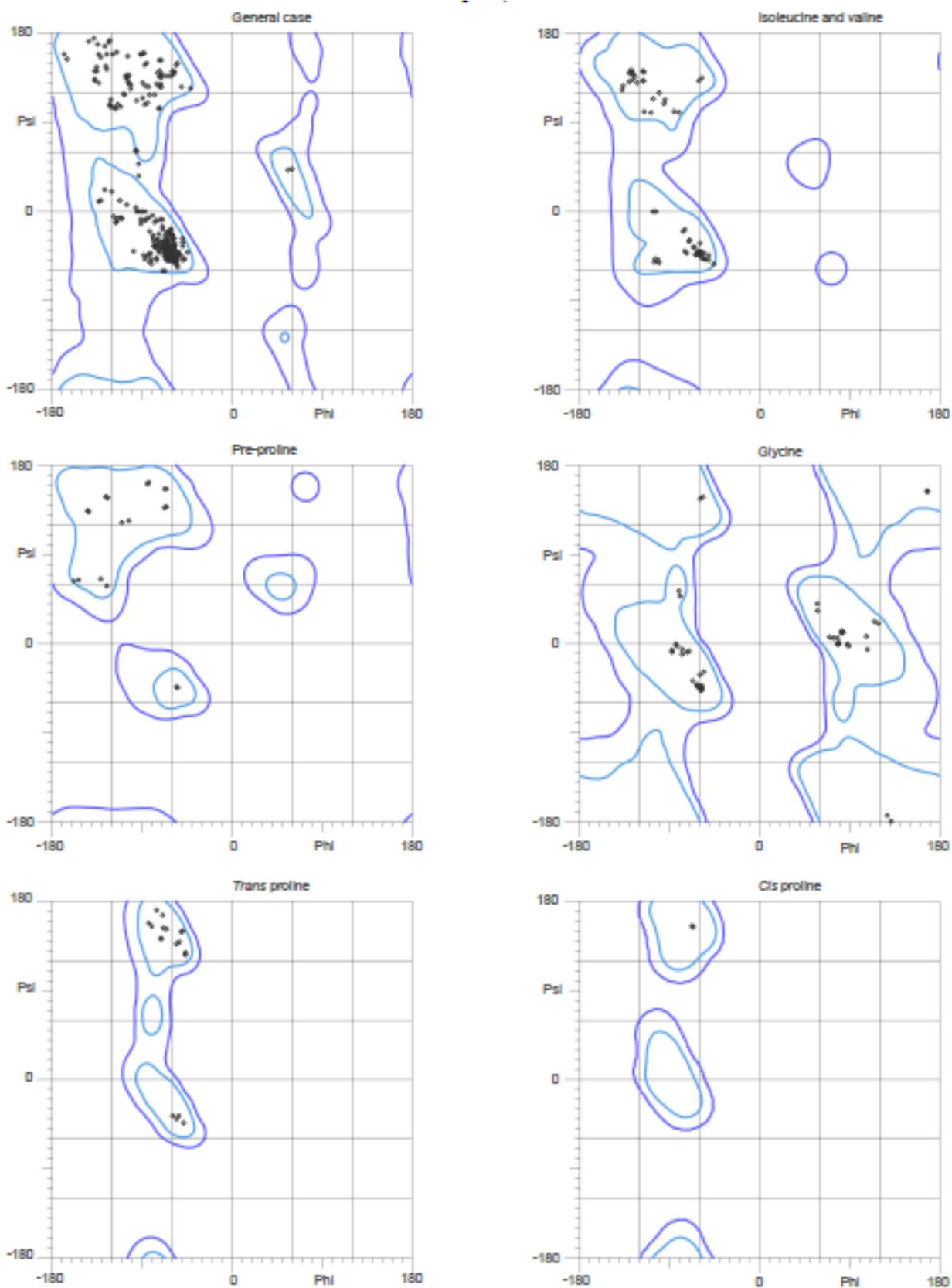


Рисунок 2. Распределение торсионных углов в структуре 3DL0 по карте Рамачандрана.

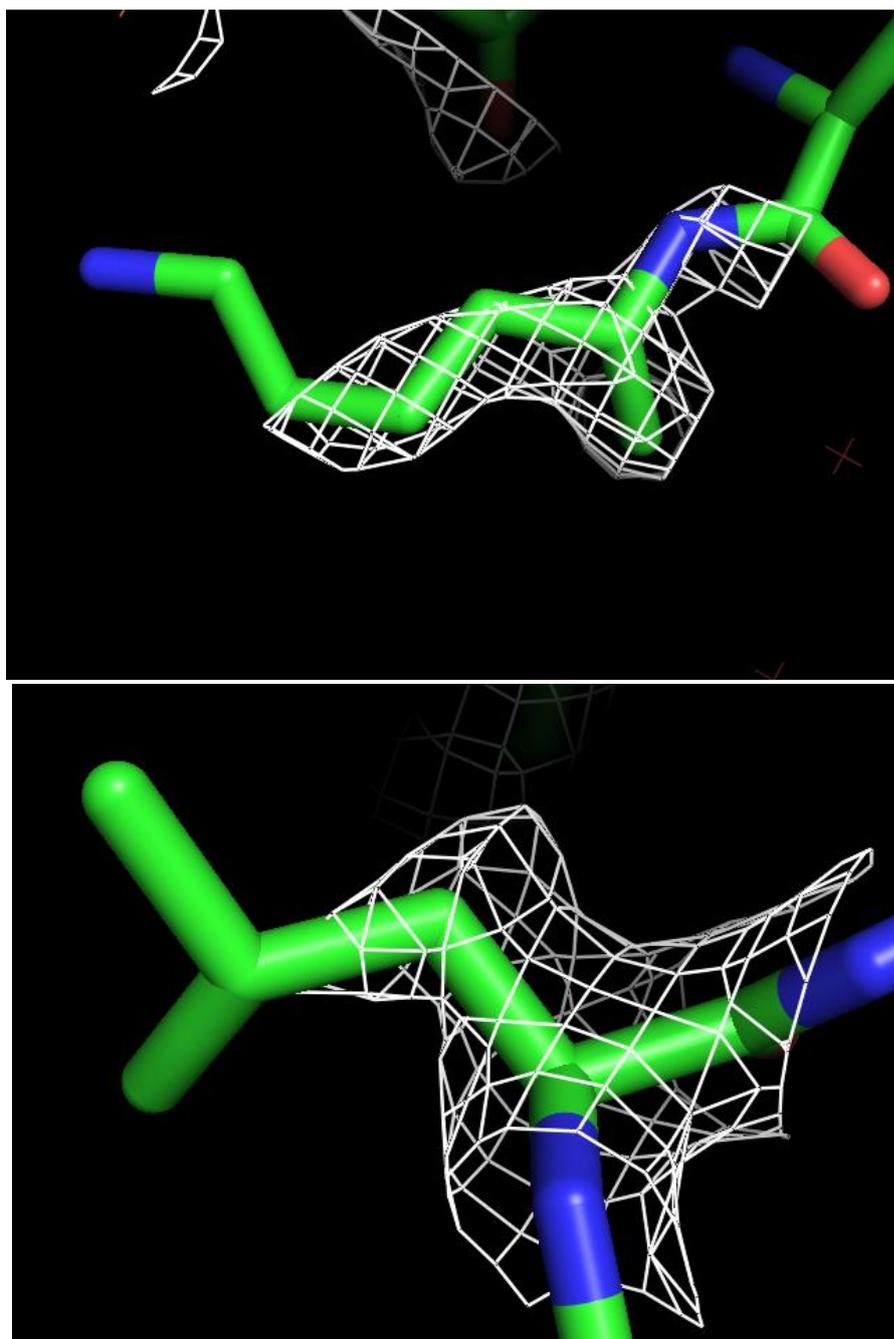


Рисунок 3. Остатки K216 (сверху) и L211 (снизу) цепи А, вокруг которых визуализирована электронная плотность с уровнем подрезки 1.

Для изучаемой структуры была построена модель PDB\_REDO. Её параметры были оценены, как и для исходной структуры, с помощью сервиса MolProbity. Результат показан в таблице 5.

Таблица 5. Информация о качестве структуры PDB\_REDO, полученная из MolProbity.

All-Atom Contacts	Clashscore, all atoms:	0.57	100 <sup>th</sup> percentile* (N=588, 1.52Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein	Poor rotamers	5	1.33% Goal: <0.3%

Geometry	Favored rotamers	358	95.47%	Goal: >98%
	Ramachandran outliers	0	0.00%	Goal: <0.05%
	Ramachandran favored	435	99.54%	Goal: >98%
	MolProbity score <sup>^</sup>	0.79		100 <sup>th</sup> percentile* (N=4870, 1.52Å ± 0.25Å)
	Cβ deviations >0.25Å	2	0.51%	Goal: 0
	Bad bonds:	1 / 3493	0.03%	Goal: 0%
	Bad angles:	19 / 4733	0.40%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	2 / 20	10.00%	Expected: ≤1 per chain, or ≤5%

Если сравнить таблицу 5 с таблицей 2, видно, что структура PDB\_REDO явно хуже. В ней больше маргинальных остатков по углам и длинам связей. С другой стороны увеличилось количество хороших ротамеров и остатков в преимущественных областях на карте Рамачандрана.

#### Вывод

3DL0 – пример весьма качественной структуры. При высоком разрешении и полноте данных она имеет хорошие показатели и в ней немного маргинальных остатков.

#### Ссылки

1. Bae, E. Effectiveness and limitations of local structural entropy optimization in the thermal stabilization of mesophilic and thermophilic adenylate kinases. 2631–2642 (2014). doi:10.1002/prot.24627
2. [http://ftp.wwpdb.org/pub/pdb/validation\\_reports/dl/3dl0/3dl0\\_full\\_validation.pdf](http://ftp.wwpdb.org/pub/pdb/validation_reports/dl/3dl0/3dl0_full_validation.pdf)
3. <http://www.ebi.ac.uk/pdbe/entry/pdb/3dl0/experiment>
4. <http://www.rcsb.org/pdb/explore/explore.do?structureId=3dl0>
5. <http://eds.bmc.uu.se/cgi-bin/eds/uusfs?pdbCode=3dl0>