

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА

ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

**ОТЧЕТ О КАЧЕСТВЕ РАСШИФРОВКИ СТРУКТУРЫ БЕЛКА 1BL8 МЕТОДОМ
РЕНТЕГОСТРУКТУРНОГО АНАЛИЗА**

студентки 4 курса
Панкевич Евгения Вадимовна

Москва
2015

Оглавление

Аннотация	3
Введение	3
Результаты и обсуждение	3
MolProbity.....	4
PDB	7
EDS.....	7
PDB_redo	9
Примеры и анализ маргинальных остатков.....	10
Выводы	11
Литература и использованные сервисы	12

Аннотация

В данной работе представлен анализ модели 1b18 (PDB-ID) структуры калиевого канала *Streptomyces lividans*, полученной с помощью рентгеноструктурного анализа. Была сделана попытка наиболее подробным образом рассмотреть все индикаторы качества модели с помощью различных ресурсов. Были сделаны соответствующие выводы о качестве модели 1b1I. Целью данного отчета является поиск информации о модели и выявление недочетов модели с помощью различных биоинформатических ресурсов.

Введение

Расшифровка структуры калиевого канала 1b18 из организма *Streptomyces lividans* приводится в статье Doyle и коллег, 1998 года, опубликованной в журнале "Science" [1], которая процитирована 5708 раз (на момент 2015 года, ноябрь). Всего до этого и начиная с 1995 года было опубликовано 13 моделей структур калиевых каналов, на момент 2015 года их порядка 130. Мембранные каналы, в частности калиевые каналы, очень консервативны, поэтому последовательность белка 1b18 очень близка к последовательностям калиевых каналов других организмов, особенно в области поры.

Белок 1b18 является интегральным белком, состоящим из четырех субъединиц, объединяющихся в конусовидную структуру с порой в центре. Наиболее узкая часть поры имеет длину 12 Å (в то время как ширина мембраны клетки - около 70 нм), остальная часть поры изнутри выстлана гидрофобными остатками. Пора остается открытой из-за структурных ограничений, это необходимо для координирования K^+ , однако, не пропускает меньшие ионы Na^+ . Селективный фильтр содержит 2 иона K^+ на расстоянии примерно 7,5 Å друг от друга, что создает электростатические силы отталкивания необходимые для преодоления силы притяжения между ионами и селективным фильтром. В работе [1] описано, как архитектура фильтра обеспечивает селективный перенос K^+ . Для объяснения используется полученная модель 1b18.

Результаты и обсуждение

Здесь представлены результаты использования различных сервисов для оценки качества структуры и сбора информации о модели белка 1b18. Результаты охватывают всевозможные индикаторы качества модели в целом с выделением отдельных примеров.

MolProbity

Вначале сервис MolProbity выдает общую информацию о модели 1bl8 (название, количество цепей и т.д). Разрешение структуры - 3.20 Å.

Использование сервиса MolProbity для получения информации о модели 1bl8 после добавления водорода (с параметрами по умолчанию: добавить электронную плотность водородов, детектировать все возможные инверсии боковых цепей Asn, Gln, His) для выявления недопустимых наложений атомов дает информативную таблицу, представленную ниже.

Таблица 1. Данные сервиса MolProbity для 1bl8

All-Atom Contacts	Clashscore, all atoms:	58.4		31 st percentile* (N=41, 2.95Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	80	30.77%	Goal: <0.3%
	Favored rotamers	124	47.69%	Goal: >98%
	Ramachandran outliers	32	8.42%	Goal: <0.05%
	Ramachandran favored	244	64.21%	Goal: >98%
	MolProbity score [^]	4.26		6 th percentile* (N=1544, 3.20Å ± 0.25Å)
	Cβ deviations >0.25Å	0	0.00%	Goal: 0
	Bad bonds:	0 / 2896	0.00%	Goal: 0%
	Bad angles:	0 / 4004	0.00%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 12	0.00%	Expected: ≤1 per chain, or ≤5%
Low-resolution Criteria	CaBLAM outliers	24	6.45%	Goal: <1.0%

В первой графе (Clashscore, allatoms) представлено число недопустимых наложений атомов на 1000. В данном случае – это 58.4, то есть 5,84 процента остатков. Перцентиль 31 процентов говорит о том, у 37% структур с примерно таким же разрешением Clashscore

хуже данного, а хуже – значит больше. Если учесть, что лучшее значение перцентили – это 100%, то можно заключить, что значение Clashscore для структуры 1b18 неплохое, по крайней мере, выше среднего.

Далее, если обращаться к следующим графам таблицы, то можно извлечь следующую информацию по структуре:

- 80 остатков с маргинальными по отклонению от ротамеров боковыми цепями (Poor rotamers) Это 30.77% от всех остатков, хотя в идеале должно быть <1%; это явно плохое значение для модели;
- 32 маргинала по карте Рамачандрана, лежат вне допустимой области (Ramachandran outliers). Это 8.42% от всех остатков, тогда как при высоком разрешении таких остатков должно быть не больше 0.05%;
- 124 остатка в предпочитаемой области карты Рамачандрана (Ramachandran favored), 47.65%; в идеале таких должно быть > 98%;

По этим данным нельзя сказать о хорошем качестве модели.

Менее информативные параметры указывают на то, что в структуре 1b18 число C_beta с неприемлемым отклонением от ожидаемого положения равно 0. Хорошим значением является число ковалентных связей, существенно отклоняющихся от теории, равное 0 (в идеале оно таким и должно быть). Число валентных углов, существенно отклоняющихся от теории, также равно 0. Это говорит в пользу качества модели.

На рисунке 1 представлена карта Рамачандрана для модели 1b18 и соответствующие маргинальные остатки (выделены черным), они находятся вне выделенной области.

MolProbity Ramachandran analysis

1b18FH.pdb, model 1

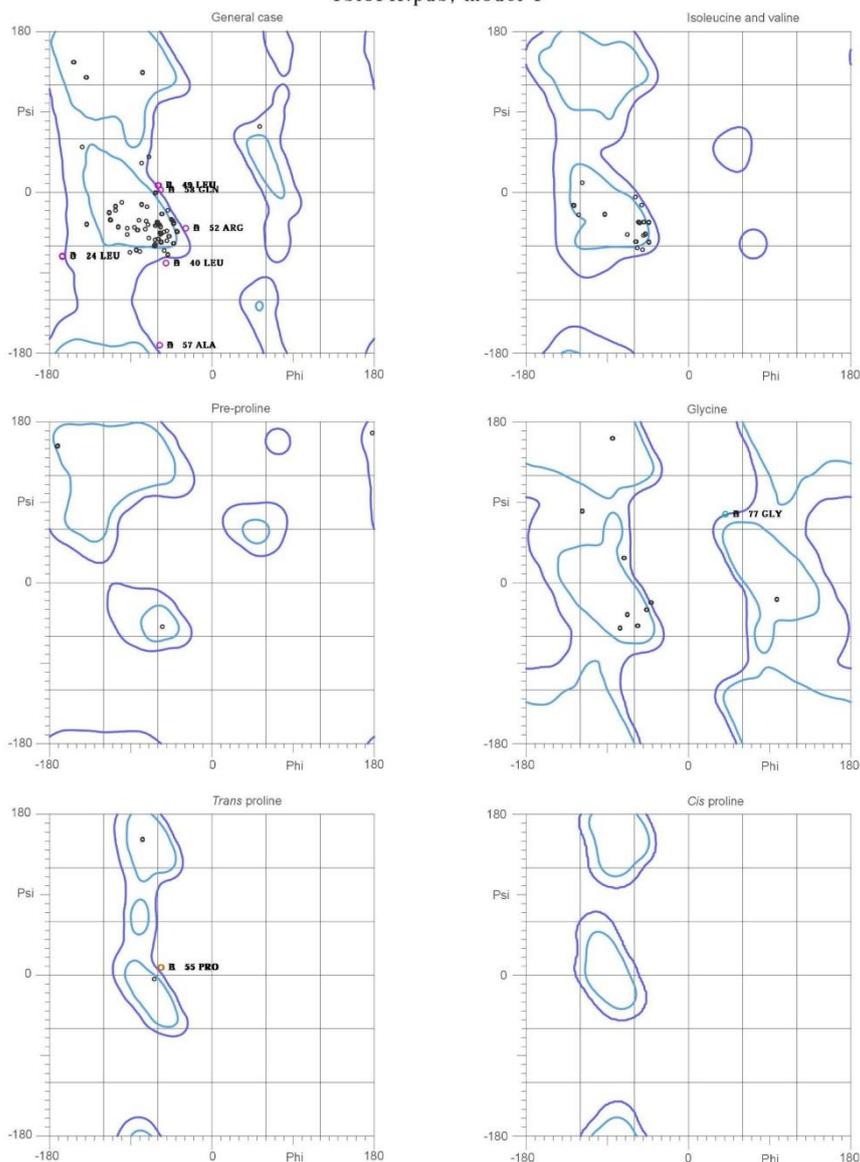


Рисунок 1. Карта Рамачандрана для модели 1b18.

Список маргинальных остатков по карте Рамачандрана:

64.2% (244/380) были в благоприятной зоне (98%).

91.6% (348/380) были в разрешенной зоне (>99.8%).

Обнаружено 32 выброса (phi, psi), для всех цепей это одни и те же остатки:

Для цепи А	Для цепи В	Для цепи С	Для цепи D
A 24 LEU (-166.7, -71.9)	B 24 LEU (-166.6, -72.0)	C 24 LEU (-166.7, -72.1)	D 24 LEU (-166.6, -71.9)
A 40 LEU (-51.3, -79.3)	B 40 LEU (-51.5, -79.2)	C 40 LEU (-51.2, -79.0)	D 40 LEU (-51.4, -79.4)
A 49 LEU (-59.9, 8.4)	B 49 LEU (-60.3, 8.4)	C 49 LEU (-59.9, 8.4)	D 49 LEU (-60.1, 8.3)
A 52 ARG (-29.3, -40.9)	B 52 ARG (-29.3, -41.0)	C 52 ARG (-29.2, -40.9)	D 52 ARG (-29.4, -40.9)
A 55 PRO (-56.9, 9.5)	B 55 PRO (-57.1, 9.4)	C 55 PRO (-56.8, 9.2)	D 55 PRO (-57.0, 9.2)
A 57 ALA (-58.3, -171.7)	B 57 ALA (-58.0, -171.9)	C 57 ALA (-58.3, -171.6)	D 57 ALA (-58.1, -171.7)
A 58 GLN (-57.3, 3.6)	B 58 GLN (-57.3, 3.7)	C 58 GLN (-57.4, 3.5)	D 58 GLN (-57.4, 3.6)
A 77 GLY (41.8, 77.0)	B 77 GLY (41.9, 77.2)	C 77 GLY (41.8, 77.1)	D 77 GLY (41.8, 77.4)

PDB

На сервере PDB представлена различная информация о структуре 1bl8. Основные параметры:

- Разрешение 3.20 ангстрем (среднее, но не плохое);
- R-value 0.280 – не очень хорошее значение;
- R-free 0.290. Разница R-free - R-value = 1% – переоптимизации нет;
- Группа симметрии C 1 2 1 (моноклинной сингонии);
- 5 – Параметры элементарной ячейки кристалла: a=128.780, b=68.930, c=112.040 (в ангстремах), alpha=90.00, beta=124.63, gamma=90.00 (в градусах)

В PDB также приведен файл структурных факторов. Число рефлексов (измеренных), лежащих в файле структурных факторов – 12177.

- 10913 ("o" в колонке "_refln.status" в файле структурных факторов) использовались для оптимизации модели;
- 1265 ("f" в колонке "_refln.status" в файле структурных факторов) использовались для построения модели;
- симметрия белка: циклическая, C4.

Также благодаря PDB можно получить информацию, касающуюся белка. Так, например, можно увидеть, что в белке присутствует 4 мутации.

EDS

Если обратиться с сервису EDS, то можно дополнительно к приведенной выше найти информацию об RSR структуры 1bl8 - 0.171 (это нельзя назвать ни хорошим, ни плохим значением). Полнота данных (Completeness of data) равна 93.4%. Ранг разрешения: 9.99 - 3.20 Å. Можно найти информацию, касающуюся Z-score для отдельных остатков (среднее значение RSR и sigma считались для структур с разрешением в диапазоне 9.99 - 3.20 Å). В таблице 2 представлены данные EDS, касающиеся остатков для цепей A-D.

Таблица 2. Данные по Z-score для структуры 1bl8.

Chain	Nres	<Z>	Sigma	%<-1	%<0	%>1	%>2
A	97	-0.59	0.78	15.46	91.75	2.06	0.00
B	97	-0.61	0.78	24.74	89.69	1.03	0.00
C	97	-0.55	0.78	16.49	84.54	3.09	0.00
D	97	-0.62	0.77	17.53	91.75	1.03	0.00
All	388	-0.59	0.78	18.56	89.43	1.80	0.00

Как можно видеть, в модели нет остатков с плохим z-score ($>2\%$) (таких, что координаты атомов остатков расшифрованы хуже, чем для остальных структур с подобным разрешением). Это говорит о хорошем качестве структуры.

В разделе "Significant regions" обычно указывают участки с повышенным RSR фактором, особенно подозрительные на наличие неточностей в координатах атомов. Для модели 1bl8 таковых не обнаружено (0 outliers), что также свидетельствует о хорошем качестве структуры.

EDS позволяет посмотреть на распределение значений температурного фактора для разных остатков в цепях A-D структуры 1bl8 (рисунок 2). На этом рисунке красным выделены пики, отвечающие тому самому отрезку из остатков с плохим RSR-фактором.

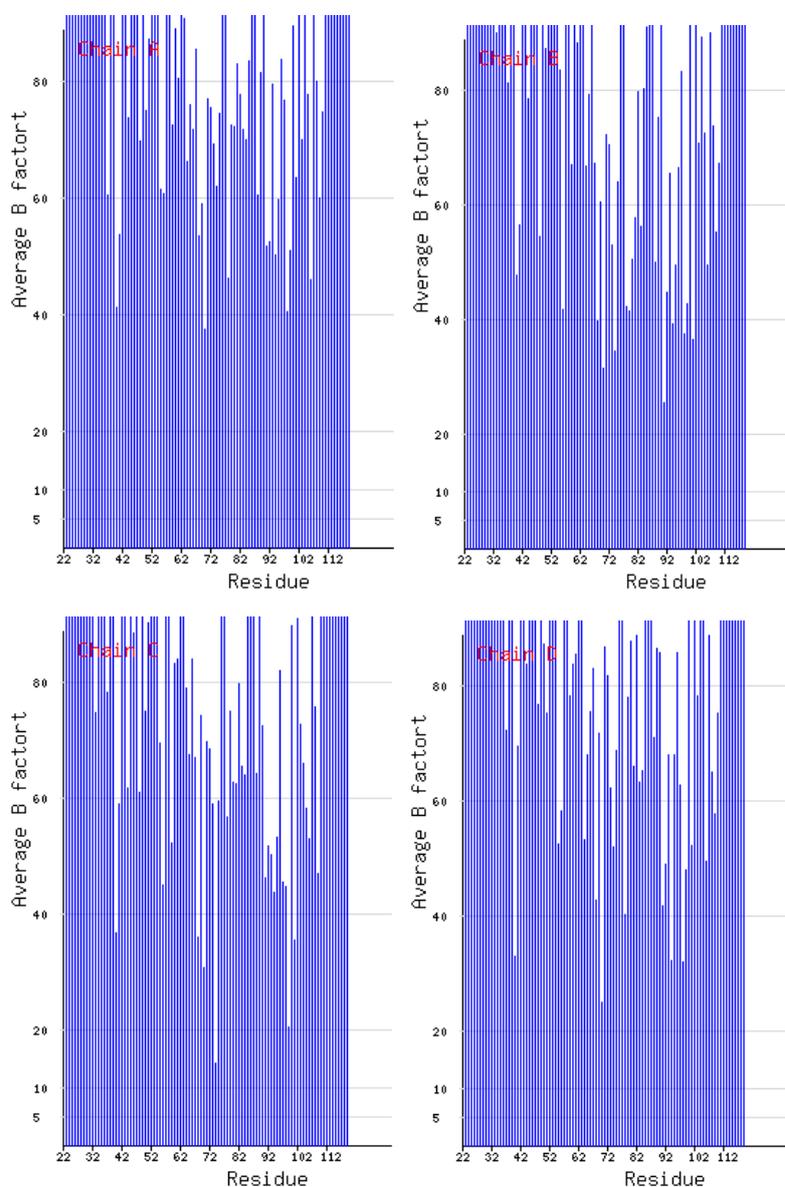


Рисунок. 2. Распределение значений температурных факторов для цепей A-D структуры 1bl8.

Видно, что значения температурных факторов для остатков из структуры 1b18 практически все превышают 40. Это достаточно плохие значения температурного фактора. Это означает, что для каждого атома пик электронной плотности очень размытый. При этом коэффициент заполнения для всех атомов равен 1.00. Исходя из этого, трудно сказать что-то окончательное по поводу всей структуры. Вероятно, были просто определены вероятные положения атомов, а различные конформации боковых цепей не рассматривались.

PDB_redo

Сервис PDB_redo позволяет оптимизировать (уточнить) любую структуру с помощью ряда инструментов через полностью автоматизированный процесс. На выходе имеются данные об исходных и новых (для структуры после оптимизации) значений R-факторов, а также протокол WHAT_CHECK анализа структуры и дополнительная информация о полученных изменениях.

Была получена оптимизированная структура 1b18. Значения R-факторов и z-score ухудшились. Помимо этого улучшились показатели для карты Рамачандрана и для ряда других факторов. Улучшились значения для ротамеров и RMS z-score для длин связей и углов между ними, исчезли некоторые неудовлетворительные доноры/акцепторы водородных связей. На рисунке 3 приведен кусок " Δ RSCC (Real Space Correlation Coefficient) plot" для каждого аминокислотного остатка (показана разница между RSCC-значениями до и после оптимизации, показаны улучшения - зеленым, а также плохие остатки - красным); полностью можно посмотреть на [PDB_redo для 1b18](#). RSCC - близкий к RSR параметр, рассчитывается по следующей формуле:

$$RSCC = \frac{\sum |\rho_{obs} - \langle \rho_{obs} \rangle| \sum |\rho_{calc} - \langle \rho_{calc} \rangle|}{(\sum |\rho_{obs} - \langle \rho_{obs} \rangle|^2 \sum |\rho_{calc} - \langle \rho_{calc} \rangle|^2)^{1/2}}$$

Видно, что многие остатки оптимизированы, осталось всего 4 плохих.

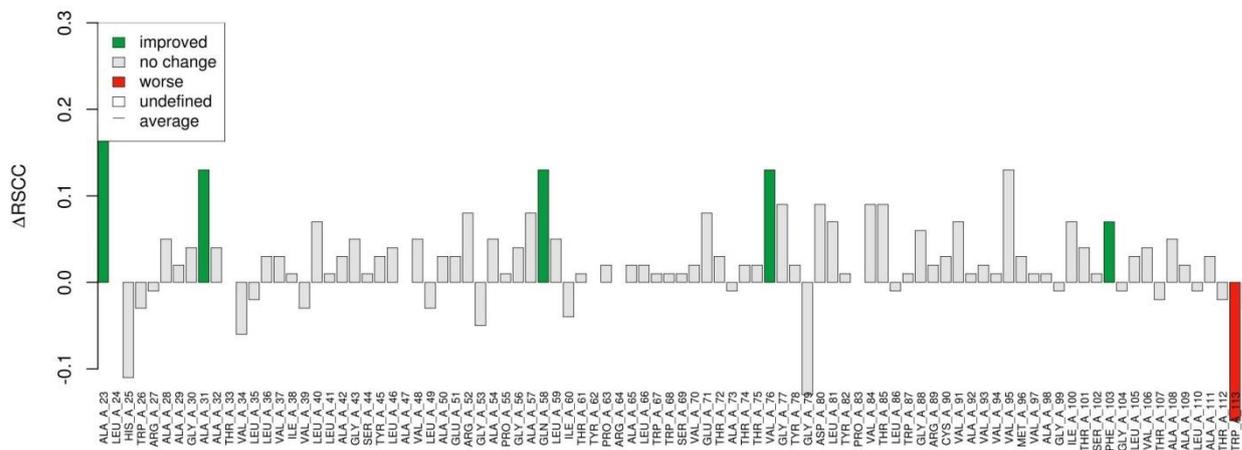


Рис. 3. Часть Δ ARSCC-графика для 1b18 на PDB_redo.

В целом, можно сказать, что все индикаторы качества структуры немного изменились. Программа оценивает эти изменения как улучшение структуры. Человеку трудно оценить эти изменения. Складывается ощущение, что они не значительны. Это говорит в пользу работы, проведенную авторами при расшифровке структуры 1b18.

Примеры и анализ маргинальных остатков

Исходя из всех тех индикаторов качества структуры, которые описаны выше, была построена сводная таблица маргиналов (таблица 3), взятых для примера 10 аминокислот с выраженными отклонениями. В ней приводятся номера остатков, цепи, к которой они принадлежат, и признаки, на основе которых этот остаток выделен как маргинал.

Таблица 3. Примеры маргинальных остатков

Маргинальные остатки		
остаток	Цепи	причина
58 GLN	A, B, C, D	Вне разрешенной области на карте Рамачандрана; плохой z-score и температурный фактор
24 LEU	A, B, C, D	Вне разрешенной области на карте Рамачандрана; плохой температурный фактор
70 VAL	A	Плохой z-score, очень плохой температурный фактор
114 PHE	A	Плохой RSR-фактор (и после оптимизации лучше не становится), плохой температурный фактор
25 HIS	B	Плохой RSR-фактор (и после оптимизации лучше не становится), плохой температурный фактор

На рисунках 4-5 показаны электронные плотности вокруг остатков 114 TRP A и 25 HIS B, даже на уровне подрезки 2 Å видно, что атомы вписаны в плотность плохо.

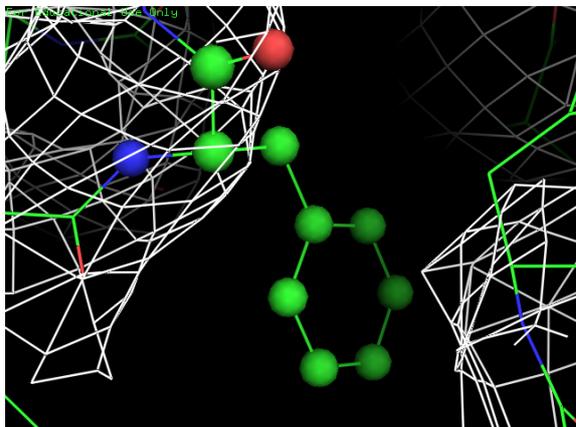


Рис. 4. 114 PHE A, уровень подрезки 2 Å.

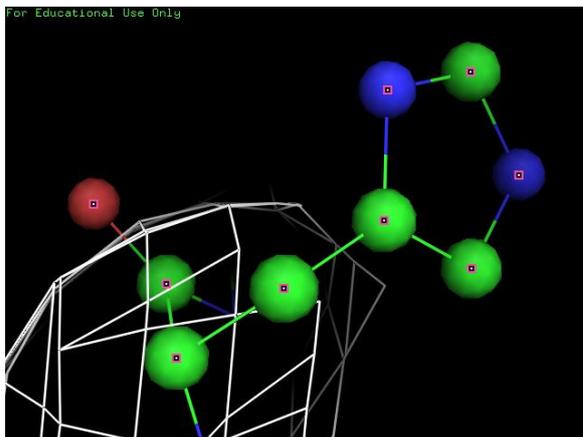


Рис. 5. 25 HIS B, уровень подрезки 2 Å.

Выводы

По результатам проведенного анализа структуры 1b18 можно сказать, что качество последней является удовлетворительным. Экспериментальные данные дали среднее разрешение (3.20 Å), структура довольно неплохо оптимизирована. Однако есть довольно большое число недостатков:

1. достаточно большое количество маргиналов на карте Рамачандрана;
2. немалое число аминокислот с неправильной конформацией боковой цепи;
3. очень плохой средний температурный фактор;

4. большое количество остатков с неудовлетворительным z-score.

Не смотря на перечисленные недостатки нельзя сказать, что работа авторов расшифровки структуры 1bl8 неудовлетворительна. Напротив, работа выполнена хорошо. Также модель трудно улучшить с помощью сервисов переоптимизации структур.

Литература и использованные сервисы

[1] D. A. Doyle, J. Morais Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon, “The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity.,” *Science*, vol. 280, no. 5360, pp. 69–77, Apr. 1998.

- сервис [MolProbity](#);
- сервис [PDB](#);
- сервер [EDS](#);
- программа [PyMol-1.7.2.1](#).