

Python для структур и не только

Задание 1. Prody и B-факторы, часть первая.

Для выполнения задания был взят белок метаногена *Methanosarcina mazei* (PDB ID 6AAO). Используя Prody, для структуры этого белка были найдены остатки, средний B-факторы для которых максимален и минимален. Максимальным средним B-фактором обладает остаток ASP 279 (цепь A) - 82.505.

Разброс B-фактора для атомов этого остатка: [75.18, 79.36, 79.61, 80.22, 82.96, 86.56, 87.86, 88.29]

B-фактор отражает подвижность атома в кристаллической структуре. Остаток ASP-279 находится на периферии белковой молекулы, на петле (рис. 1A), контактирует с растворителем. Разброс значений B-фактора его атомов довольно велик, значения увеличиваются в направлении от остова, поскольку возрастает степень свободы.

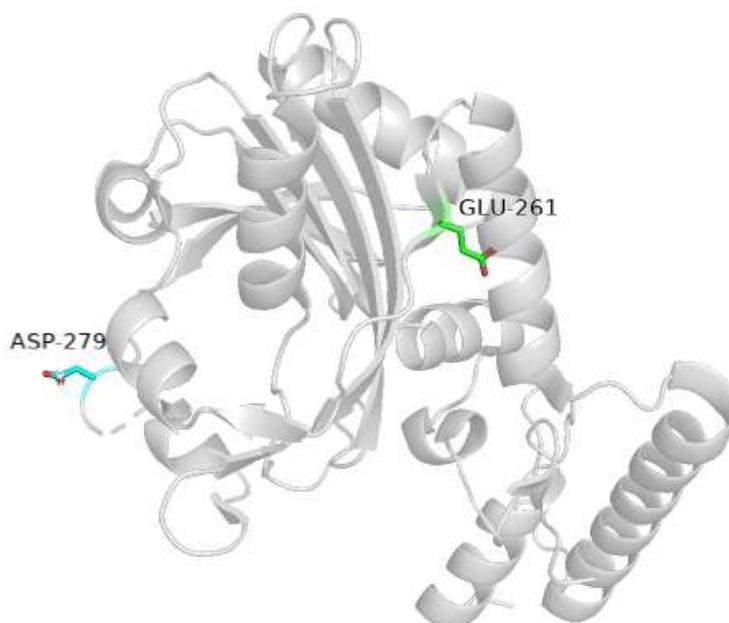


Рис. 1А. Остатки ASP-279 и GLU-261 в структуре белка с PDB ID 6AAO.

Остатком с наименьшим средним значением B-фактора 15.037 оказался GLU-261 (цепь A). Значения B-фактора для его атомов находятся в меньшем диапазоне, поскольку меньше степень свободы:

[14.14, 13.75, 13.98, 15.73, 14.48, 15.15, 15.0, 17.11, 16.0]

Подвижность этого остатка ограничена двумя взаимодействиями: солевым мостиком с LYS-263 цепи A (рис. 1Б) и солевым мостиком с аргинином молекулы-соседа (рис. 1В), оказавшейся рядом в кристалле.

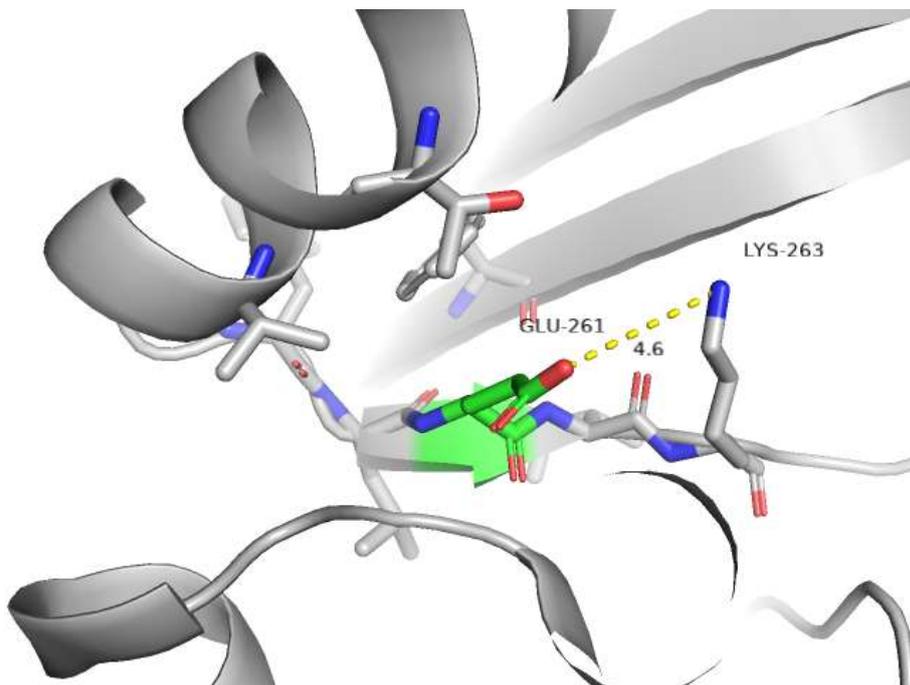


Рис.1Б. Взаимодействие остатков GLU-261 и LYS-263 цепи А.

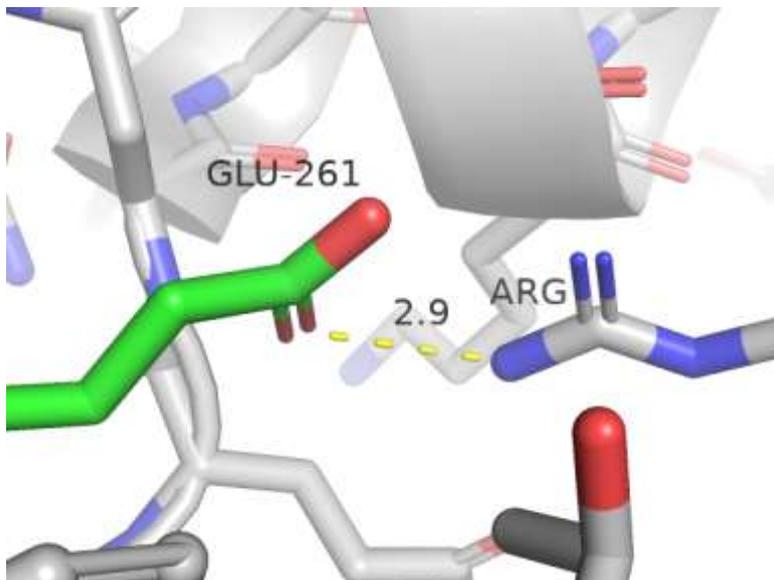


Рис. 1В. Взаимодействие остатка GLU-261 и остатка аргинина молекулы-соседа.

Задание 2. Prody и B-факторы, часть вторая.

В ходе выполнения этого задания для каждого остатка было найдены расстояние от его центра масс до центра масс белка и построен scatter-plot зависимости B-фактора остатка от расстояния до центра белка. Результат показан на графике 1. Заметен тренд к увеличению B-фактора при отдалении от центра масс. Можно выделить несколько групп с разной склонностью к соответствию тренду. Самая многочисленная группа остатков,

обведенная зеленым овалом, демонстрирует плавную зависимость. Корреляция есть, но не сильная. Это могут быть остатки, имеющие взаимодействия с остовом или даже с соседними молекулами, как в случае описанного выше GLU-261. Для таких остатков не так важно, где они находятся – близко к центру или на периферии, подвижность их все равно ограничена.

В группе, обведенной желтым овалом, промежуточная степень корреляции. Сюда также могут входить остатки, стабилизирующиеся взаимодействиями с соседними молекулами в кристалле.

В красный овал обведены остатки со средним расстоянием от центра массы белка, но при этом имеющими наибольший В-фактор. Сюда могут входить остатки на петлях, выдающихся из структуры молекулы, как в случае ранее рассмотренного ASP-279, а также остатки, смотрящие в полости белка, где вероятно взаимодействие с растворителем.

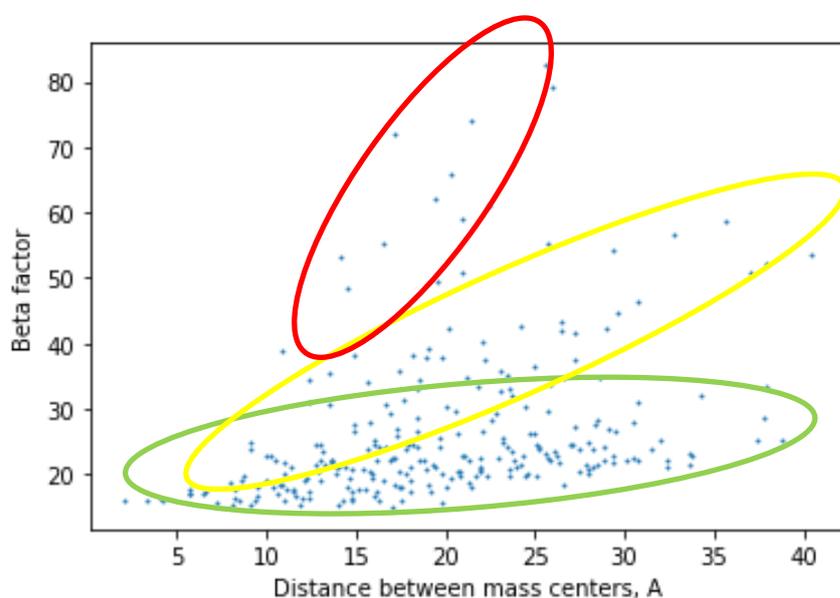


График 1. Зависимость В-факторов остатков от их расстояния до центра белка.

Задание 3. Как работает восстановление функции электронной плотности по экспериментальным данным

Задание заключалось в том, чтобы воспроизвести ход кристаллографического эксперимента и восстановить функцию ЭП.

Сначала была сгенерирована функция ЭП для двух молекул на расстоянии 3,5 Å друг от друга, состоящих каждая из 3 атомов. Для этого скрипт compile-func.py был запущен со следующими параметрами:

31,3,2,4+2.9,3.1,5+29,3,6.2+36,3,10.7+40,2.9,11.9+2,3,13.4

Затем по этой функции были рассчитаны параметры сигнала – амплитуды и фазы. По полученным данным восстанавливалась ЭП.

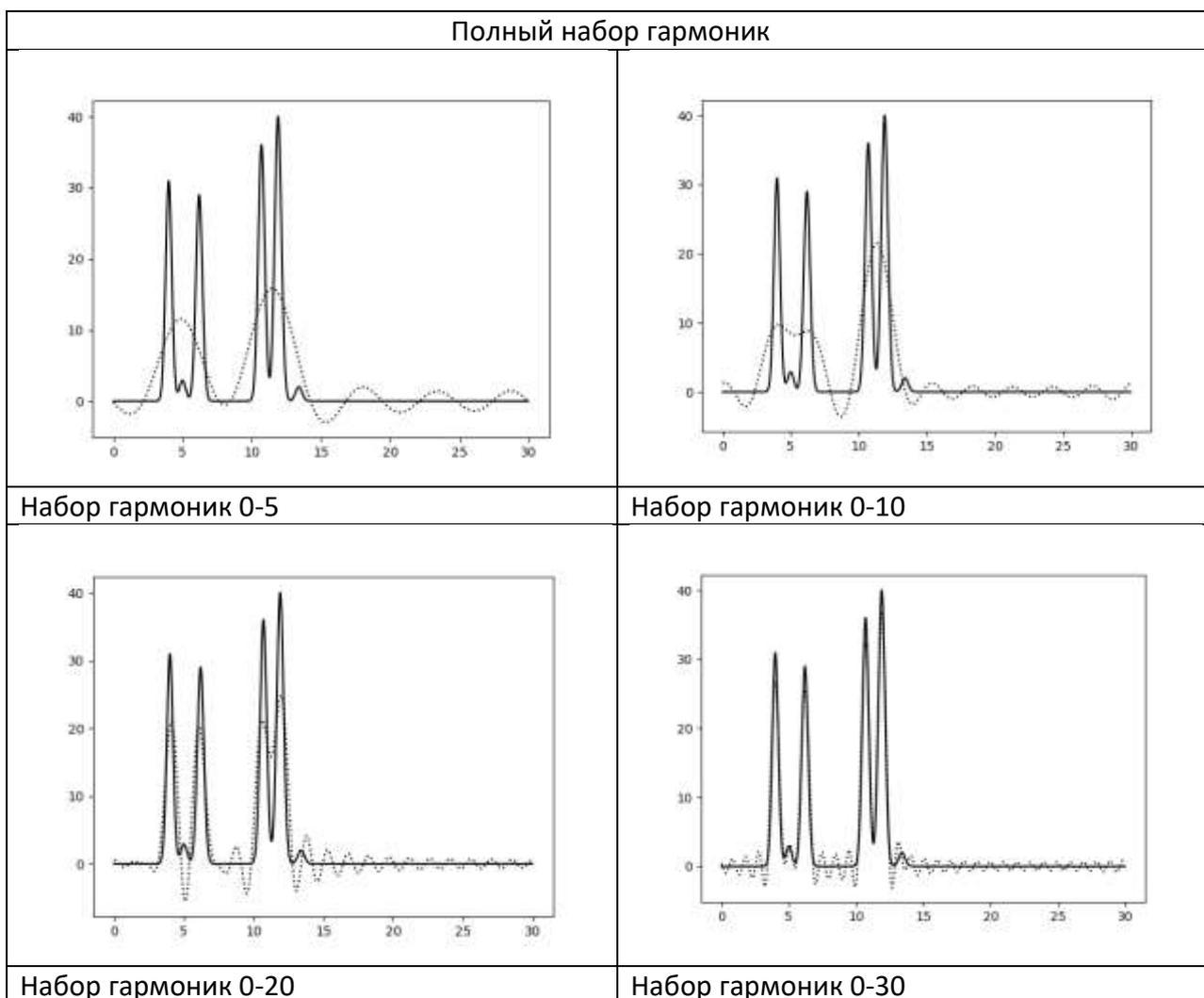
Графики приведены в таблице 1.

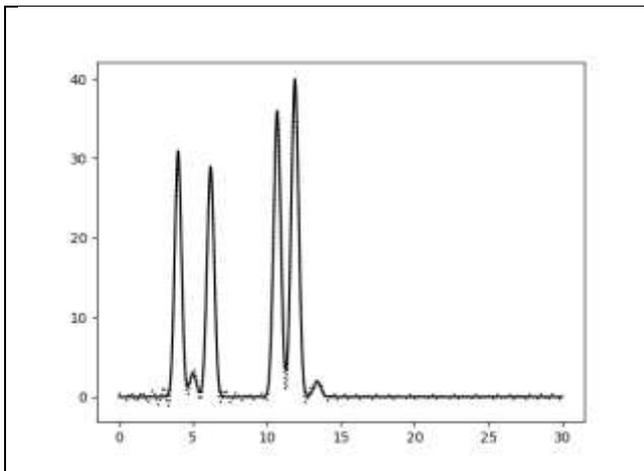
Для полного набора гармоник был определен номер первой гармоники с отличным восстановлением ЭП - 42 (на самом деле, формально отличное восстановление появилось уже на более ранних гармониках, но номер 42 показался мне первым близким к идеальному восстановлению).

Затем на наборе гармоник 0-42 было продемонстрировано влияние шума на восстановление ЭП. Качество ухудшается уже на уровне 5% шума (амплитуды и фазы одновременно), но все же остается в пределах хорошего и для 5, и для 10% шума – вероятно, потому, что изначально ЭП имеет отличное качество восстановления.

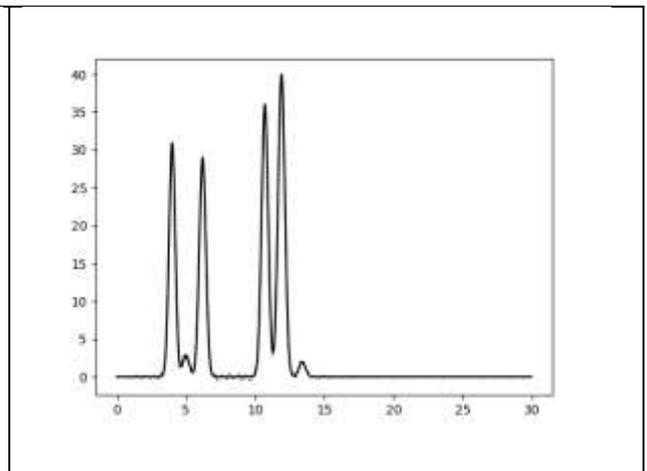
Если сравнивать влияние шума фазы и шума амплитуды, то шум фазы дает больше искажения в вид ЭП, чем амплитуда. Вероятно (и согласуется с теорией), что фазы несут больше информации для функции распределения ЭП.

Таблица 1. Графики исходной и восстановленной ЭП (пунктиром).

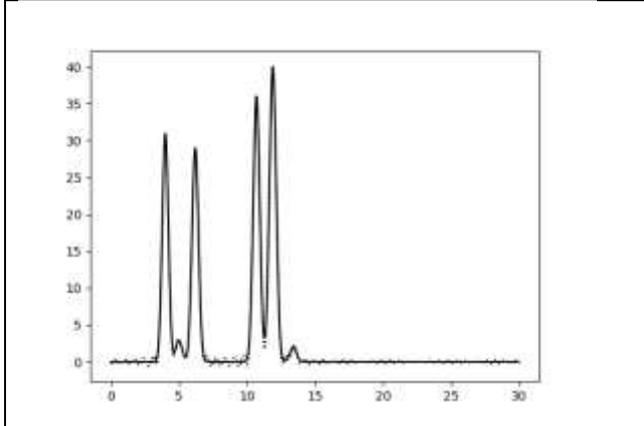




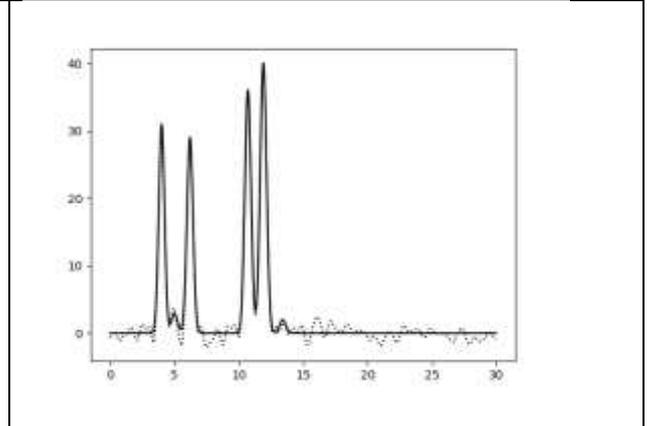
Набор гармоник 0-40



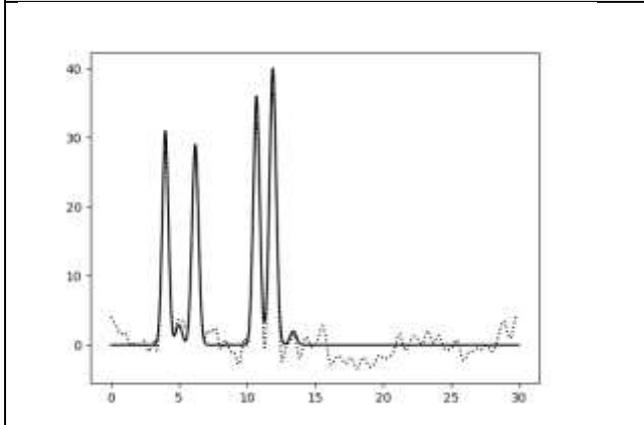
Набор гармоник 0-50



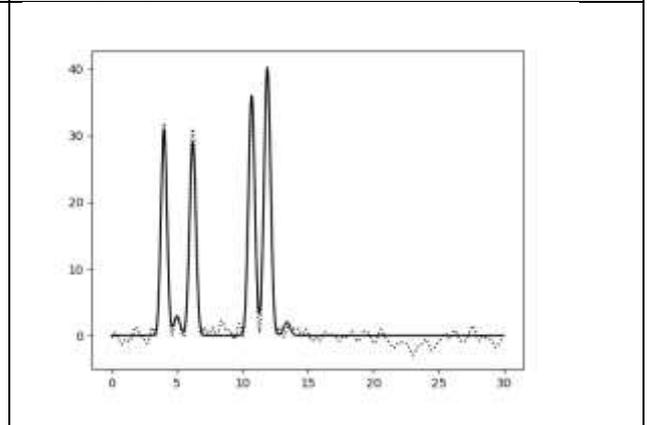
Набор гармоник 0-42



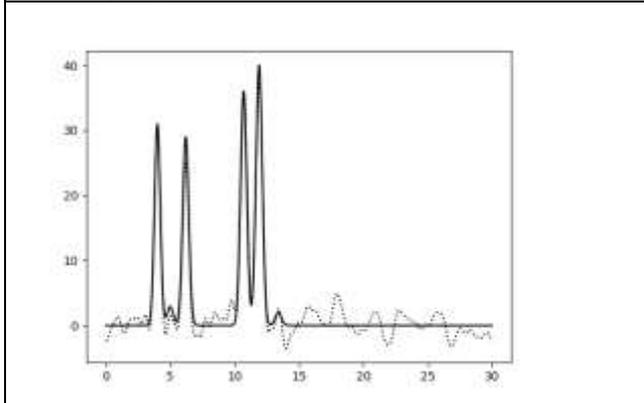
Набор гармоник 0-42, 5% шум амплитуды, 5% шум фазы

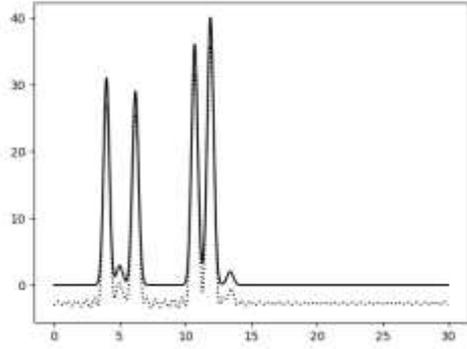
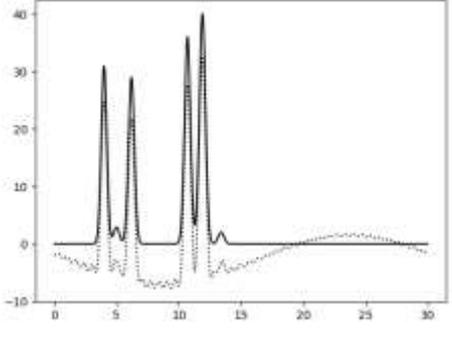
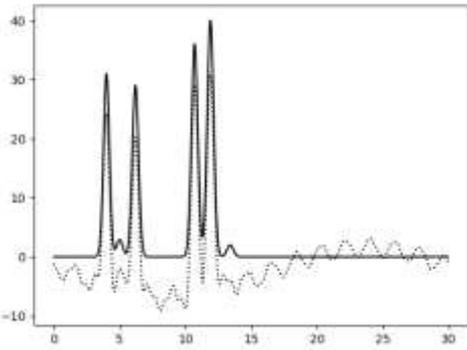
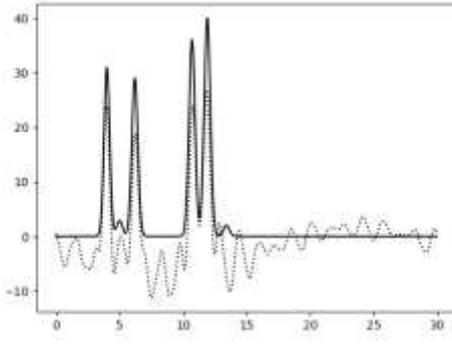


Набор гармоник 0-42, 10% шум амплитуды, 10% шум фазы



Набор гармоник 0-42, 10% шум амплитуды, 0% шум фазы



Набор гармоник 0-42, 0% шум амплитуды, 10% шум фазы		
Неполный набор гармоник		
		
Набор гармоник 1-42	Набор гармоник 2-42	
		
Набор гармоник 2-15,18-42	Набор гармоник 2-15,22-42	

Были построены также восстановления для неполного набора гармоник. Удаление первых нескольких гармоник не сильно сказывается на качестве восстановления ЭП, но вид функции распределения ЭП становится похож на синусоидальный (при удалении уже первых двух гармоник). Удаление 5-10% гармоник из середины качество заметно ухудшается, появляются пики, мешающие отличать легкие атомы. Здесь становится важным полнота. Так, для потери 15% гармоник из середины набора, полнота становится меньше приемлемого уровня 90% для разрешения, соответствующего полному набору гармоник.

Для каждого набора гармоник были определены разрешение и полнота, оценено качество восстановления. Результаты занесены в таблицу 2.

Таблица 2. Восстановление функции по коэффициентам ряда Фурье.

Набор гармоник	Разрешение (Å)	Полнота данных (%)	Шум амплитуды (% от величины F)	Шум фазы (% от величины phi)	Качество восстановления (отличное, хорошее, среднее, плохое)
Полный набор гармоник					
0-5	6 Å	100%	0	0	Плохое

0–10	3 Å	100%	0	0	Среднее
0–20	1,5 Å	100%	0	0	Среднее
0-30	1 Å	100%	0	0	Хорошее
0-40	0,75 Å	100%	0	0	Отличное
0-50	0,6 Å	100%	0	0	Отличное
0-42	0,71 Å	100%	0	0	Отличное
0-42	0,71 Å	100%	5	5	Хорошее
0-42	0,71 Å	100%	10	10	Хорошее
0-42	0,71 Å	100%	0	10	Хорошее
0-42	0,71 Å	100%	10	0	Хорошее
Неполный набор гармоник					
1–42	0,71 Å	97,7%	0	0	Отличное
2-42	0,71 Å	95,3%	0	0	Отличное
2-15,18-42	0,71 Å	90,7%	0	0	Хорошее
2-15,22-42	0,71 Å	81,4%	0	0	Среднее
2-15,22-42	1,27 Å	92,1%	0	0	Среднее