# Assignment 2: Main chain modelling

Yuchao He

November 24, 2024

## 1 Exercise 1

### 1.1 Analysis of CA-CA Distance Variations

In this work, I wrote a Python program named ‘cal_dist.py‘ to analyze the distances between consecutive alpha-carbon (CA) atoms. The program's results are shown in Figure 1. Based on the these results, a tolerance range of (3.7 Å, 3.9 Å) was set for subsequent analysis.

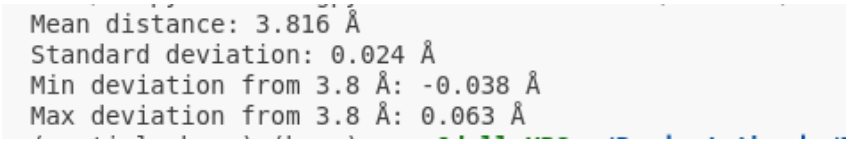


Figure 1: Calculate Distances

### 1.2 Analysis of file "test_q1.txt"

In this work, I ran the program named "task_1.py". Outputs are shown in Figure 2. Outputs here represent the sequence of the main chain.


Figure 2: Path in "test_q1.py"

### 1.3 Analysis of file "data_q1.txt"

In this work, I ran the program named "task_1.py". Outputs are shown in Figure 3. Outputs here represent the sequence of the main chain.


Figure 3: Path in "data_q1.py"

# 2 Exercise 2

## 2.1 Methods

- For each alpha-carbon atom, find its neighboring atoms within a specified distance tolerance range..
- Given that multiple neighboring atoms may exist, select the most relevant ones. In this context, we prioritize atoms whose distance deviation from the reference atom is less than 1.2 Å.
- Subsequently, algorithms such as Depth-First Search (DFS) can be employed to find paths connecting the selected atoms.

## 2.2 Analysis of file "data_q2.txt"

In this work, I ran the program named "task_2.py". Outputs are shown in Figure 4. The selection of 50 alpha-carbon atoms as the main chain meets the requirement for "correct" outputs.

```
20
9
32
15
44
62
48
73
45
19
39
63
62
63
69
83
117
160
177
203
237
256
266
301
294
243
236
227
205
223
273
262
215
181
106
184
207
129
158
169
142
86
70
72
54
27
16
3
4
19

The total number of alpha-carbon atoms in the main chain: 50
(particle_hana) (base) user@dell-XPS:~/Project_thesis/Bio-Info/Assign__2$
```

Figure 4: Path in "data_q2.py"