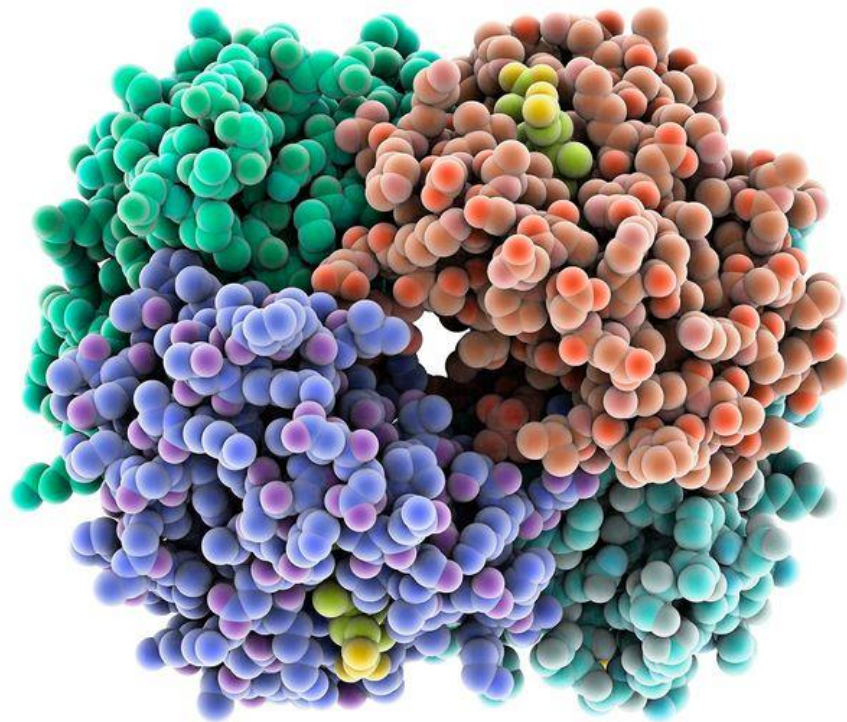


2019

Structural Bioinformatics Project Documentation



TA. Mohamed Yousef

TA. Walaa Alkady

Project Requirements:

The software application should let the user choose which Bioinformatics algorithm he wants to use. The algorithms that should be implemented are as follows:

a) Generate the theoretical spectrum of a cyclic peptide.

Input: An Amino Acid string peptide.

Output: Cyclospectrum.

b) Cyclopeptide sequencing.

Input: A protein spectrum

Output: All the linear representations of the cyclic sequence of the protein.

Project Details:

Part 1:

Circular_Spectrum:

Calculates the **circular** spectrum of a protein sequence.

Input: Peptide Sequence

Output: List of integers representing the protein's circular spectrum.

Part 2:

Helping Functions:

- **Linear_Spectrum:**

Calculates the **linear** spectrum of a protein sequence. (Same as Part 1 but with no circulation)

Input: Peptide Sequence

Output: List of integers representing the protein's linear spectrum.

- **IsConsistent:**

Checks whether a certain sub-peptide is consistent with the input spectrum by checking if its Linear Spectrum is contained within the input spectrum. (Hint: Check The built-in List functions)

Input: Sub-peptide Sequence and Theoretical Spectrum

Output: True or False

- **Initial_List:**

Create the initial list of 1-mers that will be used to extend all the sub-peptides by checking the first masses in the input spectrum.

Input: Theoretical Spectrum

Output: List of strings/chars. (Initial List of 1-mers)

Hint 1: You may use the Amino acids masses text file to create a hash table where the key is the mass and the value is the amino acid).

Hint 2: You may use the same function to predict the length of the Peptide sequence that is represented by the input spectrum by counting the values of the Initial 1-mers including repeats.

Note:

See the slides for an example of Initial List of length 4, but the length of the resulting peptide is 5, due to the repetition of an amino acid.

Main Function:

Input: Theoretical Spectrum.

Output: Peptide Sequence.

Process:

- 1) Call function **Initial_List**, Outputs: Initial_L and PeptideLength.
- 2) Copy the Initial_L to **TempList** (the list that will be extended)
- 3) Loop from 2 to PeptideLength (To Create 2-mers, 3-mers, ... etc)
 - Use another loop/s to extend each value of the TempList with each value in the Initial_L
 - Check whether the extended values/sub-peptides in the TempList are consistent using the **IsConsistent** function, delete it from the TempList if the output is false.

Note that you will keep extending until you reach the length of the Peptide (If the length is 5, the temp List wil contain 5-mers at the end of the loop)

For A full example, Check Lab 6 slides.

Project Bonus:
Implementing a good GUI using python or any language.

How to Submit the Project:

Only one Member of the team Submit a Zip file to the classroom that contains:

- Python Code file for the whole Project Named by:
(SPPProject_YourTeamNumebr.py)
- Text file (Notepad) Contains your Team Number and your Names

Notes:

- Document your code whenever is possible.
- Use meaningful names for variables and functions.
- All the members of the group should have a good understanding of all the algorithms used in the project.