STRUCTURAL BIOINFORMATICS LAB

Task I 2022

TA. Esraa Hamdi   TA. Nourhan Mohammed

Cyclopeptide sequencing

Main Function:

Input: Theoretical Spectrum.
Output: All the linear representations of the cyclic sequence of the protein (Peptide Sequence).

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.
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 (weight) and the value is the amino acid.

Hint 2 : To know when to stop searching for "initial (1-mer) amino acids" , you can check whether the mass you're investigating in the spectrum is contained in the dictionary you built or not.

Process:

1) Call function **Initial_List**, Outputs: Initial_L.

2) Copy the Initial_L to **TempList** (the list that will be extended)

3) Loop and only break if the generated potential sub-peptides are all inconsistent.
   - 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 are consistent using the **IsConsistent** function, keep only the consistent subpeptides to extend them in the next iteration.

4) Print the final list that contains all linear representations of the antibiotic.

General Hints :
   - Always Check the built-in functions first.
   - Document your code as much as possible. (Add comments)
   - Divide the functions amongst you, test them separately before incorporating them together.

Good Luck ☺