Course Work: BIO213 IQB
Assignment No.2

Question 1

Part A

The code file is submitted.

Part B

-HHHHHHHHHHHSSSSSSSSSSSSSSHHHHHHHHHHSSSSHHHHHHHHHHHHH--HHHHHHHHHHSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS----HSSSSSSSSSSSSSSSSSSSS--SSSSSSSSS---HHHHHH---------

Question 2

Category I: Chou and Fasman and GOR method

Chou-Fasman:

- It uses a statistical approach based on the observed frequencies of each amino acid in known protein structures to determine their propensity to form a specific secondary structure element.
- Considers only the propensity of each residue to be in a helix, strand, or turn; conformational values for coils are not considered.
- It uses a scoring table listing the relative propensities of each amino acid to be in an α-helix, β-strand, or β-turn.
- Predicts secondary structure elements by scanning through a sequence with a certain window size to find regions with a stretch of contiguous residues, each having a secondary structure element score.
- It uses a window size of six residues for α-helices and five residues for β-strands.
- Suppose alpha helix and beta strand predictions overlap in a specific region. In that case, a prediction is made based on the criterion that if the helix's propensity score is greater than a strand's propensity score, it is declared a helix or vice versa.
- It is relatively straightforward to implement, with a simple scoring system that can be easily calculated and used for prediction.

GOR:

- It is based on information theory and uses short-range interactions of neighboring residues to predict the conformational state for each residue.
- Considers the propensity of each residue to be in one of four conformational states: helix, strand, turn, or coil.
- It uses a set of matrices that describe the probabilities of each residue in a particular conformational state given its neighboring residues.
- Examines a window of every seventeen residues and sums up propensity scores for all residues for each of the four states resulting in four summed values.
- The highest-scored state defines the conformational state for the center residue in the window (ninth position).
- It is more accurate than the Chou-Fasman method because it considers the neighboring effect of residues.
- It requires more complex calculations and matrices, making it more challenging to implement but potentially more accurate.

Category II: DSSP, P-curve, and Stride

STRIDE:

- It uses empirically derived hydrogen bond energy and phi-psi torsion angle criteria to assign secondary structure.
- Elongates helices and can veto short helices with unfavorable phi-psi angles.
- Categories include helices and sheets but do not distinguish between parallel and anti-parallel sheets.
- Fixed internal parameters for alpha-helices and beta-sheets optimized to mirror visual assignments made by crystallographers.
- Assignments agreed better with expert assignments than DSSP for the dataset used to optimize the free parameters.

DSSP:

- It uses the backbone hydrogen bonding patterns and residue properties to assign a secondary structure.

- Categories include helices, sheets (divided into parallel and anti-parallel), and coil.
- It does not elongate helices, but it allows for bulges and kinks.
- Assigns the shortest alpha-helix containing at least two consecutive i ® i+4 hydrogen bonds.
- Defines sheets as stretches of at least two adjacent strands with hydrogen bonds that form a hydrogen bonding pattern typical of beta sheets.
- Assigns the 'B' symbol to a residue involved in a hydrogen bond to another residue but not forming part of a sheet.

P-Curve:

- It uses differential geometry to calculate a helicoidal axis and match structural motifs to assign a secondary structure.
- Parameters in the motif include the radius of the helicoidal system, along with tilting, rolling, and twisting measures describing geometrical differences between two peptide planes.
- Assigns secondary structure by matching known motifs based on helicoidal parameters.
- Uses C_a-coordinates to perform parameter analysis.
- Allows for more degrees of freedom when matching a P-Curve motif than other assignment methods.
- It may fit better to curved beta-strands than other methods.