

PROTEIN SECONDARY STRUCTURE PREDICTION AND CALCULATION OF PHYSICOCHEMICAL PROPERTIES

Introduction :

To predict the secondary structure of a protein using chou-fausman method and to display the physicochemical properties such as molecular weight, GRAVY score, extension coefficient, amino acid composition.

Input :

Protein sequence.

Methods :

mol_weight(seq)
ext_coefficient(seq)
gravy (seq)
composition (seq)
predict(seq)

Variables:

h-helix
bs-b-strand
bt-b-turn
mol_weight
ext_coeff-extension coefficient
gravy(grand average of hydrophathy)
compo(amino acid composition)

Functions:

Predict(seq)-Predict the secondary structure of the entered protein sequence.
Composition(seq)-Gives the amino acid composition of the entered protein sequence.
Gravy(seq)-calculates the Grand average of hydrophathy scores.
Ext_coefficient(seq)-calculates the extension coefficient.
mol_weight(seq)-calculates the molecular weight of the protein sequence.

Output:

Predicted secondary structure elements-helix,b-strand,coil

Physicochemical properties:

Molecular weight.
Extension coefficient.
GRAVY scores.
Amino acid composition .

Pseudocode:

1. Input the protein sequence.
2. Create 3 dictionaries corresponding to the propensities of amino acids for forming alpha helix, beta sheet and turns.
3. Input the sequence as argument to the function predict(seq).
4. Convert the string to a list.
5. Scan the selected residues:
6. Iterating in a for loop,
7. Iteration for helix,
8. Extract residues with window size : 6
9. Check for 4 out of 6 contiguous residues have propensity(helix)>100.
10. Calculate overall helix propensity value (helix_value)
11. Iteration for beta sheet/turn
12. Extract residues with window size : 5
13. Check for 3 out of 5 contiguous residues have propensity(sheet)>100.
14. Calculate overall beta sheet propensity value (sheet_value)
15. Check for 3 out of 5 contiguous residues have propensity(turn)>100.
16. Calculate overall turn propensity value (turn_value)
17. Check for the maximum value within helix_value, sheet_value and turn_value.
18. Assign the secondary structure element based on the maximum value.
19. Call the function mol_weight(seq):
20. for i in (0,len(seq)):
21. add the molecular weights of all the amino acids from the respective dictionary.
22. Display the overall molecular weight.
23. Call the function ext_coefficient(seq):
24. for i in (0,len(seq)):
25. add up the contribution values of the 3 amino acids (Trp, Phe, Tyr) according to their presence.
26. Display the calculated extinction coefficient.
27. Call the function gravy (seq):
28. for i in (0,len(seq)):
29. add the hydropathy values of all the amino acids from the respective dictionary and divide by the length of the residues.
30. Display the obtained average hydropathy value for the protein.
31. Call the function composition (seq):
32. for i in (0,len(seq)):
33. count the number of each amino acids and calculate its percentage over the entire sequence and assign it to a dictionary.
34. Display the dictionary corresponding to the amino acid composition.

Submitted by
Hindheeya vijayakumar
Mathangi
Sivakama meenakshi

