

Assignment 11

- To find the propensity of alpha helices using the given sequences and secondary structure assignments, we need to find the total number of Amino Acids present in the given sequence and the total number of amino acids associated with the Helical structure and finally we compute the Propensity using the formula:

$$Propensity = \frac{HCNi/SNi}{nH/N}$$

Where:

HCN = Number of Helical Confirmation Amino Acids which have Helical Confirmation (i^{th} Amino Acid)

SN = Total number of Amino Acids of that Particular Amino Acid in the sequence (i^{th} Amino Acid)

nH = The total number of Amino Acids which have helical confirmation

N = The total number of Amino Acids which are in the sequence

The Code for computing the propensity is:

```
AA_all=['A','C','D','E','F','G','H','I','K','L','M','N','P','Q',
        'R','S','T','V','W','Y']
prop_dict={'A':0,'C':0,'D':0,'E':0,'F':0,'G':0,'H':0,'I':0,'K':0,
           'L':0,'M':0,'N':0,'P':0,'Q':0,'R':0,'S':0,'T':0,'V':0,'W':0,
           'Y':0}
tot_dict={'A':0,'C':0,'D':0,'E':0,'F':0,'G':0,'H':0,'I':0,'K':0,
          'L':0,'M':0,'N':0,'P':0,'Q':0,'R':0,'S':0,'T':0,'V':0,'W':0,
          'Y':0}
ans_dict={'A':0,'C':0,'D':0,'E':0,'F':0,'G':0,'H':0,'I':0,'K':0,
          'L':0,'M':0,'N':0,'P':0,'Q':0,'R':0,'S':0,'T':0,'V':0,'W':0,
          'Y':0}
```

```
wholeseq='LGASGIAAFAFGSTAILIILFNMAAEVHFDPLQFFRQFFWLGLYPKQYGM
GIPPLHDGGWWLMAGLFMTLSLGSWWIRVYSRARALGLGTHIAWNFAAAIFFVLCIGCIHPT
LVGSWSEGVPPFGIWPIDWLTAFSIRYGNFYPCWHGFSIGFAYGCGLLFAAHGATILAVAR
FGGDREIEQITDRGTAVERAALFW'
```

```
wholesecstr='XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX'
```

```
totalh=0
for i in range(len(wholeseq)):
    tot_dict[wholeseq[i]]+=1
    if wholesecstr[i]=='H':
        prop_dict[wholeseq[i]]+=1
        totalh+=1
for i in AA_all:
    ans_dict[i]=(prop_dict[i]/tot_dict[i])/(totalh/len(wholeseq))
```

```
for key, value in ans_dict.items():
    print(key, ': ', value)
```

and the output is:

```
A : 1.5510204081632653
C : 1.0204081632653061
D : 0.0
E : 0.40816326530612246
F : 1.0204081632653061
G : 1.0612244897959184
H : 0.8746355685131195
I : 1.2004801920768309
K : 0.0
L : 1.2244897959183674
M : 1.530612244897959
N : 1.3605442176870748
P : 0.22675736961451246
Q : 0.0
R : 0.6802721088435374
S : 1.3605442176870748
T : 0.7653061224489796
V : 0.5830903790087464
W : 1.1131725417439702
Y : 0.5830903790087464
```

2. We do the same procedure of question 1 in this question too:

Total number of AA in the given sequence (nH): 200

Total number of AA which have helical confirmation in the given sequence (N): 98

Denominator value of Propensity formula = $98/200 = 0.49$

Table of Propensity index of the given amino acid:

AA	HCN	SN	Ratio (numerator)	Propensity (numerator/denominator)
A	19	25	0.76	1.5510204081632653
C	2	4	0.5	1.0204081632653061
D	0	5	0.0	0.0
E	1	5	0.2	0.40816326530612246
F	10	20	0.5	1.0204081632653061
G	13	25	0.52	1.0612244897959184
H	3	7	0.42857142857142855	0.8746355685131195
I	10	17	0.5882352941176471	1.2004801920768309
K	0	1	0.0	0.0
L	12	20	0.6	1.2244897959183674
M	3	4	0.75	1.530612244897959
N	2	3	0.6666666666666666	1.3605442176870748
P	1	9	0.1111111111111111	0.22675736961451246
Q	0	4	0.0	0.0
R	3	9	0.3333333333333333	0.6802721088435374
S	6	9	0.6666666666666666	1.3605442176870748
T	3	8	0.375	0.7653061224489796

V	2	7	0.2857142857142857	0.5830903790087464
W	6	11	0.5454545454545454	1.1131725417439702
Y	2	7	0.2857142857142857	0.5830903790087464