

PART-II

1) Sequence: TAGAGC

Parts: TA, AG, GA, AG, GC

Score: -7, -9, -9, -9, -12

Sum: 46

Average:  $46/5 = 9.2$

$\therefore$  Avg. base stacking energy = 9.2

2) Sequence: LEADSEQ  $\Rightarrow$  L=125 E=145 A=85 D=130  
S=100 E=145 Q=140

Mol. wt =  $\Sigma$  mw. of AA - (n-1) mw. of  $H_2O$

$$= 125 + 145 + 85 + 130 + 100 + 145 + 140 - 18 \times 6$$

$$= 870 - 108$$

$$= 762 \text{ Daltons}$$

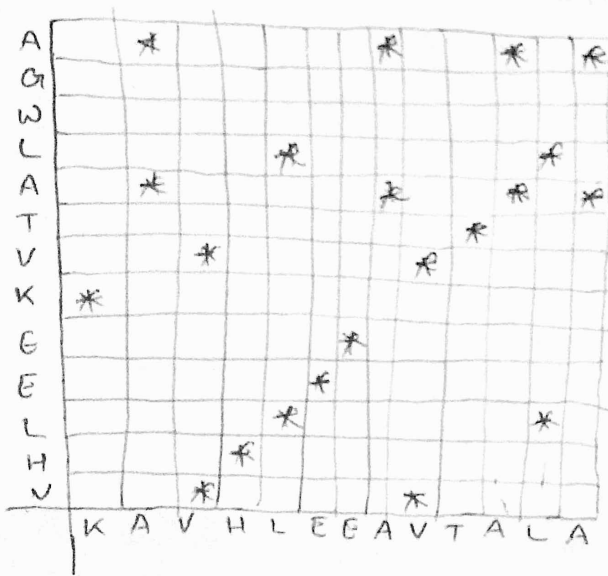
3) Sequence identity:

$$\begin{array}{ccccccc} M & L & I & T & L & A & I & W & K & V \\ | & | & & | & & & & & & \\ M & L & A & S & L & D & T & F & R & A \end{array} = \frac{3}{10} \times 100 = 30\%$$

Sequence Similarity:

$$\begin{array}{ccccccc} M & L & I & T & L & A & I & W & K & V \\ | & | & \times & \times & & & & & & \\ M & L & A & S & L & D & T & F & R & A \end{array} = \frac{5}{10} \times 100 = 50\%$$

4)



- from this dotplot, we can say that the sequences are aligned like the following:

```

      V H L E E K V T A L W G A
      | | | | | | | | | |
K A V H L E E A V T A L A
  
```

- Highest continuous line = 5 AA starting from first 'V' in 2<sup>nd</sup> sequence.

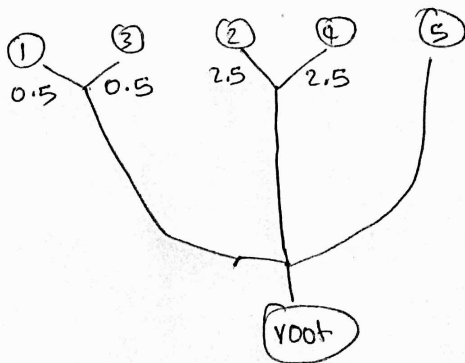
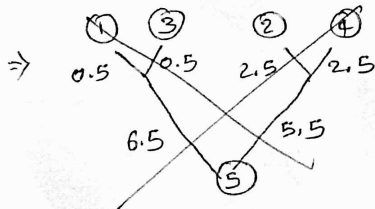
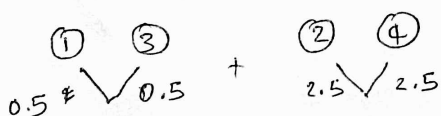
# Part- III

- 2) Sequences:
- 1: GTGCTGCACG
  - 2: AGCTGCAAGC
  - 3: GTGCTGCACCT
  - 4: CTCCTGCAGAA
  - 5: GAATCACATA

2b) Weight matrix:

	1	2	3	4	5
2	9				
3	1	9			
4	9	5	9		
5	7	8	7	8	

2a) Scaled tree:



2d) 1. GTGCTG CACG

$$M = 2$$

V. GAATCACATA

$$MM = -1$$

$$GP = -3$$

		G	T	G	C	T	G	C	A	C	G
G	0	0	0	0	0	0	0	0	0	0	0
A	0	0	1	0	1	0	0	1	2	0	0
T	0	0	0	0	0	0	0	0	3	1	0
C	0	0	0	0	0	0	0	0	0	2	0
A	0	0	0	0	0	0	1	0	0	4	1
T	0	0	0	0	0	2	0	0	2	1	6
C	0	0	0	0	0	0	1	0	0	4	3
A	0	0	2	0	0	2	0	0	1	3	2
T	0	0	0	1	0	0	1	0	2	0	2



3b) ii) Contact order:

$$CO = \frac{\sum \Delta S_{ij}}{L \cdot N}$$

$$= \frac{(20.02 - 15.94) + (26.30 - 20.02)^{\times}}{20 \times 100}$$

$$= \frac{4.08 + 6.28}{2000} = 0.00518^{\times}$$

iii) LR order:

$$LRO = \frac{\sum n_{ij}}{N}$$

$$= 6/20 = 0.3$$

i) Surrounding hydrophobicity.

$$H_p = \sum n_{ij} h_i$$

$$= 19 \times (1 + 4 + 5 + 1 + 5 + 1 + 5 + 1 + 1 + 1 + 1 + 5 + 5 + 1 + \cancel{1} + 5 + 1 + 5 + 2)$$

$$= 950$$

$$4b) 100 - 15 = 85$$

$$20 - 65$$

9 - Match.

$$TP = 9 \quad TN = 6$$

$$FP = 11 \quad FN = 5$$

$$\text{Sensitivity} = 9 / (9 + 5) = 0.643$$

$$\left[ TP / (TP + FN) \right]$$

$$\text{Specificity} = 6 / (6 + 11) = 0.353$$

$$\left[ TN / (TN + FP) \right]$$

$$\text{Accuracy} = \frac{15}{(15 + 16)} = 0.484$$

$$\left[ (TP + TN) / (TP + TN + FP + FN) \right]$$