K.PAVANI
BE19B023

1) molecular weight of Peptide DATARE (in Daltons)

Molecular weight $= \Sigma n(i) * w(i) - 18 * (N-1)$

D → ASP

A → Ala

T - Thr

A → Ala

R → Arg

E → Glu

$= 133 + 89 + 119 + 89 + 174 + 147 - (6 \times 18)$

$= 751 - 108$

$= 643$

2) Given sequence TATAGC

Calculate stacking energy

given energies

AA:-5, AT:-7, AC=-4, TA:-7, AG:-9, GC:-12, GA:-9, CG:-12

TATAGC

=) TA, AT, TA, AG, GC

Base stacking energy $= \dfrac{-7 + (-7) + (-7) + (-9) + (-12)}{5}$

$= \dfrac{-42}{5}$

$= -8.5$

Average base stacking energy $= -8.5$ cal/mol

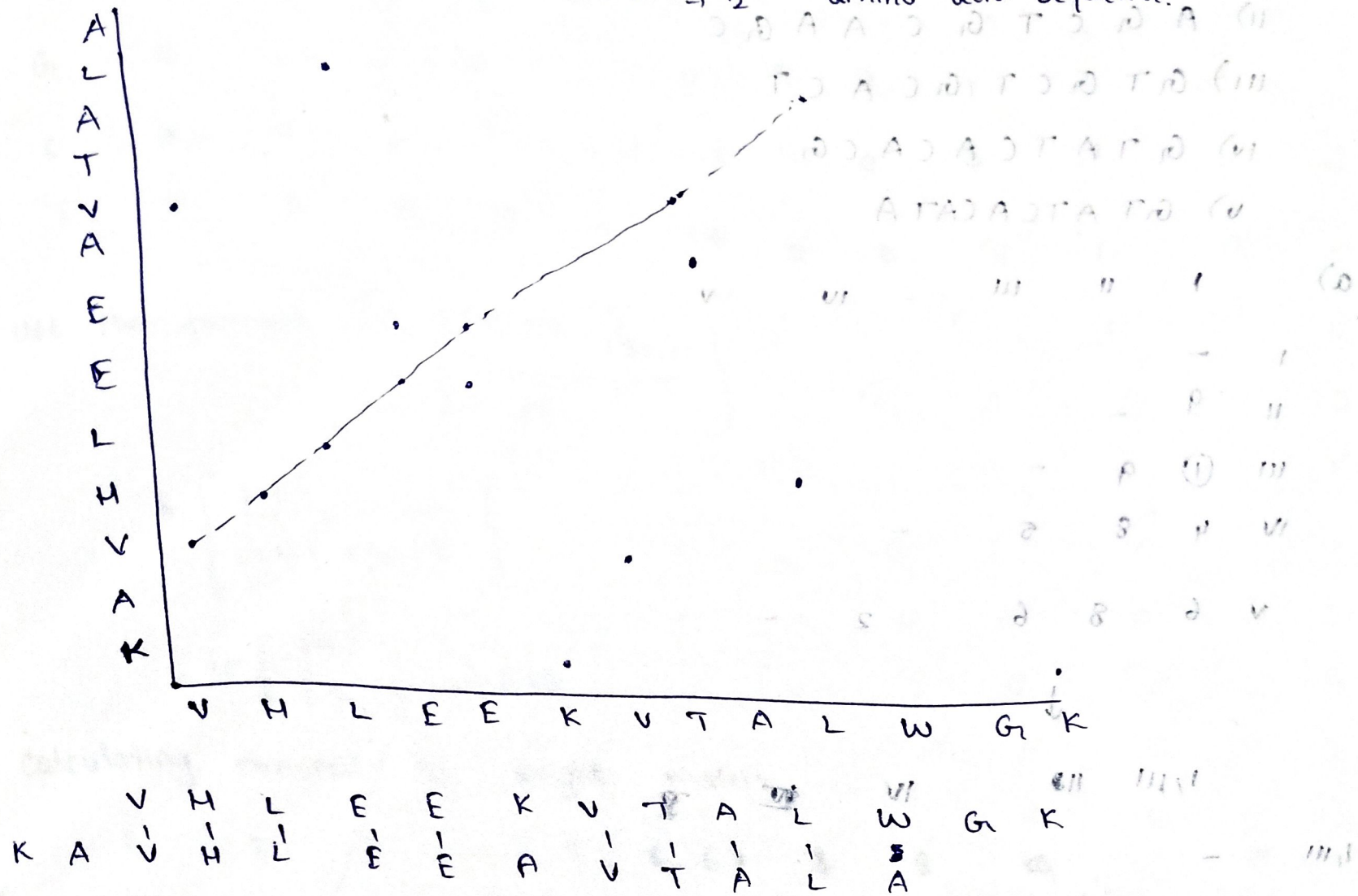3) M L I T L A I W K V
   M V A S L D T W R A

Identity sequences are M, L, W, ⊕

Sequence identity $= \dfrac{3}{10} \times 100 = 30\%$

Similarity $= \dfrac{5}{10} \times 100 = 50\%$

a) V H L E E K V T A L W G K — 1st amino acid sequence.

K A V H L E E A V T A L A → 2nd amino acid sequence.

(y-axis, top to bottom): A L A T V A E E L H V A K

(x-axis): V H L E E K V T A L W G K

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

1)a) Database:-

Database contains collection of information in a computer readable form

b) i) Updating the data and maintaining database.
ii) proper order and format

c) i) data should be proper and format
ii) organize the database with proper definitions
iii) presentation of results.
iv) Interlinkages with the databases
v) links for original publications.

d) DDBJ, EMBL

b) Domain knowledge
Scalability
Accessibility
Audience.
Tracking and analytics.
Providing security.
Optimization

8) I) G T G C T G C A C G

II) A G C T G C A A G C

III) G T G C T G C A C T

IV) G T A T C A C A C G

V) G T A T C A C A T A

a)

	I	II	III	IV	V
I	–				
II	9	–			
III	①	9	–		
IV	4	8	5	–	
V	6	8	6	2	–

↓

	I,III	II	IV	V
I,III	–			
II	9	–		
IV	4.5	8	–	
V	6	8	3	–

	I,III	II	IV,V
I,III	–		
II	9	–	
IV,V	5.25	8	–

↓

	(I,III),(IV,V)	II
(I,III)(IV,V)	–	
II	8.5	–

Scaled tree
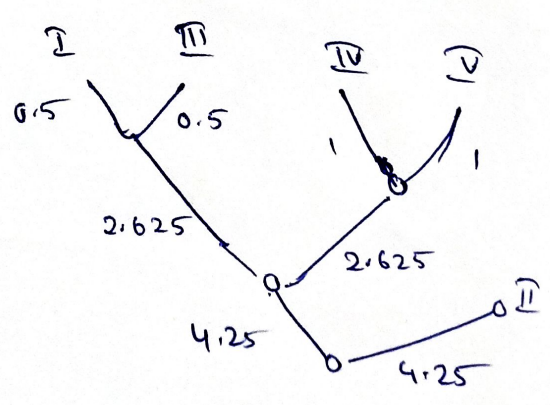
b) Freq matrix.

Freq	1	2	3	4	5	6	7	8	9	10
A	1	0	2	0	0	2	1	5	0	1
G	4	1	2	0	1	0	20	0	0	1	2
C	0	6	1	2	2	1	4	0	3	1
T	0	4	0	3	2	0	0	0	1	1

use the formula
$$\ln\left[\frac{(n_{ij}+P_i)/(N+1)}{P_i}\right]$$

$$\ln\left[\frac{1.25}{(5+1)\times 0.25}\right]$$

$$=\ln\left[\frac{1.25}{1.5}\right] \quad P = -0.18.$$

calculating elements of weight matrix.

	1	2	3	4	5	6	7	8	9	10
A	-0.182	-1.79	0.405	-1.79	-1.79	0.405	-0.182	1.252	-1.79	-0.182
G	1.04	-0.15	0.405	-1.79	-0.182	0.405	-1.79	-1.79	-0.182	0.405
C	-1.79	-1.79	-0.182	0.405	0.405	-0.182	1.04	-1.79	0.723	-0.182
T	-1.79	1.04	-1.79	0.723	0.45	-1.79	-1.79	-1.79	-0.182	-0.182

2c) At position L

Using unweighted freq

$f''(T) = \frac{2}{5} = 0.6$ $f''(A) = \frac{1}{5} = 0.2$

$f''(G) = \frac{1}{5} = 0.2$ $f''(c) = \frac{0}{5} = 0$

Using entropy based measure.

$$= \sum f(i) + \ln f(i)$$

$$= 0.6\ln 0.6 + 0.2\ln 0.2 + 0.2\ln(0.2)$$

$$= -0.3065 + (-0.3219) \times L$$

$$= -0.9503.$$

At position 8   A, A, A, G, A.

Using unweighted freq.

$f''(T) = 0$ $f''(A) = 0.8$ $f''(G) = \frac{1}{5} = 0.2$ $f''(c) = 0$

Using entropy based measure

$$= 0.8\ln 0.8 + 0.2\ln(0.2)$$

$$= -0.223 - 0.3219$$

$$= -0.5449$$

d)

	G	A	A	T	C	A	C	A	T	A
O	-3	-6	-9	-12	-15	-18	-21	-24	-27	-30
G -3	☐2	-1	-4	-7	-10	-13	-16	-19	-22	-25
T -6	-1	☐1	-2	-2	-5	-8	-11	-14	-17	-20
G -9	-4	-2	☐0	-3	-3	-6	-9	-12	-15	-18
C -12	-7	-5	-3	☐-1	-1	-4	-4	-7	-10	-13
T -15	-10	-8	-6	-1	☐-2	-2	-5	-5	-5	-8
G -18	-13	-11	-9	-4	-2	☐-3	-3	-6	-6	-6
C -21	-16	-14	-12	-7	-2	-3	☐-1	-4	-7	-7
A -24	-19	-14	-12	-10	-5	0	-3	☐1	-4	-5
C -27	-22	-19	-15	-13	-8	-3	2	-1	☐0	-3
					-16	-11	-6	-1	1	-2	☐-1
G. -30	-25	-20	-18	-16	-11	-6	-1	1	-2	☐-1

score = -1

3a)

alpha segment

A I K S W V KTI A R N L

$$a_1 = \frac{4+5+5+5}{4}$$

$$= 4.75$$

$$a_3 = \frac{1+1+1}{3} = 1$$

$$a_2 = \frac{5+5+4}{3} = 4.67$$

Power of amphipothicity $= |4.75 + 4.67 - (1+2)|$

$$= 6.42$$

beta segment   E A R I K L N A

$$b_1 = \frac{1+1+1+2}{4} = 1.25$$

$$b_2 = \frac{4+5+5+4}{4} = 4.5$$

P. O. A $= |b_1 - b_2| = 3.25$