In-silico characterization of pollen-specific protein Bnm1 from *Arabidopsis thaliana L*

Rokade S.S., Tidke J. A.and Chikhale N.J.*
Department of Botany, Sant Gadge Baba Amravati University, Amravati – 444 602 (M.S.) India.
*Shri. Shivaji Agricultural College, Amravati– 444 602 (M.S.) India.
jaikirantidke@rediffmail.com, rokadesim@rediffmail.com

Abstract- Pollen allergies (Hey-fever) are the most common and wide spread and released from the plants for the purpose of fertilization and can be carried for miles by the wind. Unfortunately when there is too much pollen in the air, it causes problems to the sensitive persons, causing some ailments and allergic disorders. Therefore, it is prime need to characterize the allergic pollen proteins. In the present investigations, pollen-specific protein Bnm1 from Arabidopsis thaliana was screened in-silico for its allergic and antigenic characters. Pollen-specific protein Bnm1 having 187 amino acids residues, which shows five antigenic determinants. Motif map shows more receptors on B-cell than on T-cell. MHC-class 1 receptors are also found in motif map. Protein statistics is also carried out with respect to several parameters with the help of Peptool 2.0. Predicted results are support to elicit, the pollen-specific protein Bnm1 is allergic.
Keywords: Pollen allergy, pollen protein, antigenicity.

INTRODUCTION
Pollen grains on their exterior coat carry different proteins involved in compatibility/ incompatibility. Although the pollen proteins are having crucial role in the process of plant reproduction, some of these proteins have allergic role when they become in contact with human body causing some allergic ailments or hey fever. Air may be greatly influenced by local sources of pollen grains emissions [1-4]. People exposed to it for several hours during daytime may develop adverse effect [5]. The onset of symptoms in pollen allergy suffers may be related to the environment to which people are largely exposed. Thus looking towards the significant role of the allergenic patients from pollen antigens in the diagnosis and therapy of allergic patients there is need of identification and characterization of potential pollen protein allergens for estimation of relative potency as a first step towards risk assessment [6]. Therefore it has been proposed to undertake such an important aspect of pollen allergenic studies.

MATERIALS AND METHODS
Prediction of hydrophobicity
The Pollen-specific protein Bnm1 protein was scanned for Hydrophobicity and Parker hydrophilicity index [7] with window size seven. Hydrophobicity (or hydrophilicity) plots are designed to display the distribution of polar and apolar residues along a protein sequence. Most commonly, this analysis has the goal of predicting membrane-spanning segments (highly hydrophobic) or regions that are likely exposed on the surface of proteins (hydrophilic domains) and therefore found to useful to identify potentially antigenic segments. Scale of hydrophobicity have been developed, which were derived from experimental studies on partitioning of peptides in apolar and polar solvents.

Prediction of antigenic sites
The Pollen-specific protein Bnm1 protein sequence of Arabidopsis thaliana L. was analyzed and characterized to study the antigenicity and MHC class peptide binding, which allows potential drug targets to identify active sites against allergenic reactions. Antigenic epitopes are determined by using the method of Kolaskar and Tongaonkar [8]. Predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes. Prediction of antigenicity program predicts those segments from within the SRK protein sequence that are likely to be antigenic by eliciting an antibody response. B-cell epitopes are predicted according to the method given by Saha and Raghava [9].

Secondary alignment
Secondary structure prediction was determined by using Chou and Fasman [10] and Garnier [11] method. These methods are based on information theory. The outputs of these programs are alpha-helix, beta-sheet and coil and gives probable value for each secondary structure. The predicted structure is one of the highest probably compatible with experimental structure.

Protein statistics
In-silico protein statistics was carried out with respect to several parameters such as atomic weight, average molecular weight, hydrophobic and hydrophilic amino acids percentage, number of basic and acidic amino acids, linear charge density and solubility of protein, amino acid frequency and hydrophobicity.

OBSERVATIONS AND RESULTS
The pollen specific protein Bnm1 protein is made up of 187 amino acids and its total atomic weight is 20055.857 daltons. The In-silico studies have shown the pollen specific protein Bnm1 protein comprising average molecular weight of the amino acids is 107.251. It contains 48.1283 % hydrophilic amino acids and 51.8717% hydrophobic amino acids. The protein contains 17 basic amino acids and 25 acidic amino acids. The total linear charge density of the protein is 0.235294. The solubility of protein is 1.48957. Structural characters of the protein observed in-silico "Table (1)". Hydrophobicity information is useful in identifying coil regions, exposed loops, interior domains, B-cell antigenic determinants and membrane spanning regions with in the sequence. The maximum value of hydrophobicity for pollen specific protein Bnm1 protein is 3.00 "Fig. (1)". The Parker hydrophilicity index [7] shows average value 1.670, minimum value -4.057 (in negative) and maximum value 8.129 "Fig. (2)". The antigenic peptide prediction program found the antigenic determinants by finding the area of greatest local hydrophilicity. The Kolaskar-Tongaonkar [8] antigenicity scale shows that pollen specific protein Bnm1 protein is highly antigenic nature and has five antigenic determinants "Fig. (3)". Segment from amino acid no. 4 to 29 shows antigenic propensity more than 1.26. The average antigenic propensity of this protein is 1.0273 "Table (2)". The predicted B cell epitopes are ranked according to their score obtained by trained recurrent neural network. Higher score of the peptide means the higher probability to be as epitope "Table (3)". All the peptides shown here are above the 0.51 threshold value. Karplus-Schulz [12] flexibility displayed in graph of the main chain mobility within a protein based on sequence information alone with window size seven "Fig. (4)". It shows average value 0.986, minimum value 0.907   and maximum value 1.074. Motif map of SRK protein shows antigenic motifs which are MHC class 1 related and also have functional relationship with B-cell membrane and T-cell membrane "Fig. (5)". The Pollen-specific protein Bnm1 protein structure is predicted by using Chau – Fasman [10] and Garnier [11]. Each residue is assigned values for α-helix, □-sheet and extended coil using window size seven residues. Using this information parameter shows the structural characters of the SRK protein as α-helix 29.9%, □-sheet 28.3% and extended coil 41.8% "Fig. (6)". The beta staircase shows how the amino acids are positioned in relation to one another. The hydrophobic amino acids are seen on outside of the beta staircase "Fig. (7)". The helical wheel assumes a periodicity of 3.6 residues per helical turn. Individual residues represented as a color circles are placed successively at each node of the helix. Multiple turns of the helix spiral outwards from helix center. The inter connectivity bars indicate the residue arrangement along the helix "Fig. (8)". Serine, cystein, valine, tyrosine, phenylalanine, methionine, glycine and alanine are placed on outer side of the helix. Amino acid frequency and percentage amino acid weights are also determined in-silico by using peptool "Fig. (9 and 10)". Leucine is 12%, each in total residue of amino acid in this protein, followed by serine and glycine. Histidine is represents 2% amino acid frequency. By weight leucine and arginine the dominating amino acids in SRK protein. Cystein and histidine residues are representing 2% by weight in this protein "Fig. (10)".

CONCLUSION
Allergenic reactions to the pollen protein of Arabidopsis thaliana are causing human health hazard causing some allergic dermatitis, hay-fever and respiratory problems. Small peptide nonamers or fragments from Arabidopsis thaliana pollen involve multiple antigenic components to

direct and empower the immune system to protect the host from allergic infections. Knowledge of epitopes may be used in the design of vaccines and diagnostics tests. It is therefore of interest to develop improved methods for predicting epitops "Table (2)". The nonamers shows high antigenic response because of presence of beta sheets regions. The antigenic determinants site 4-YSGCFFVVTLAVLSQFLLAPASALVNR-30 shows high antigenic nature and form beta sheets in secondary structure "Fig. (6)" and also showing hydrophobic characteristics. These small peptide fragments of antigen can induce immune response against whole antigen. This theme is may be implemented in designing subunit and synthetic peptide vaccines. The antigenicity analysis method allows potential drug targets to identify active sites, which forms antibodies against Arabidopsis thaliana L. pollen allergy. The results are further confirmed by studying the protein statistics, B-cell epitopes "Table (3)" and antigenic motifs of SRK protein, which are found to contain antigenic sites. Antigenic epitopes of major SRK protein are important antigenic determinants against the allergic reactions.

REFERENCES

[1] Solomon, W. R. (1986). Aerobiology of pollinosis. J Allergy Clin Immunol; 74: 449 – 461.
[2] Santra, S. C., S. Gupta and S. Chanda (1991). Air pollutant and aeroallergens interaction. Grana; 30: 63 – 66.
[3] Dass, A., A. Singh and A. B. Singh (1995). Impact of exposure of SO2 on pollen morphology and its protein profile. In: 29th Ann Conv Coll Aller and Appl Immunol, Calcuta, Abst. 13.
[4] Carinanos, P., J. Sanchez-Mesa, J. C. Prieto-Baena, A. Lopez, F. Guerra, C. Moreno, E. Dominguez and C. Galan (2002). Pollen allergy releted to the area of residence in the city of Cordoba, South-West Spain. J Environ Monit; 4: 734 – 738.
[5] Valenta, R., M. Duchene, C. Ebner, P. Valent, C. Sillaber, P. Deviller, F. Ferreira, M. Tejkl, H. Edelmann, D. Kraft and O. Scheiner (1992). Profilins constitute a novel family of functional plant pan-allergens. J. Exp. Med.; 175: 377 – 385.
[6] Gomase, V. S., J. A. Tidke and K. V. Kale (2008). Prediction of antigenic MHC binders of calreticulin protein from Parthenium argentatum. International Journal of Bioinformatics; 1 (1): 37 – 44.
[7] Parker, J. M., D. Guo and R. S. Hodges (1986). New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry; 23;25.
[8] Kolaskar, A. S. and P. C. Tongaonkar (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett; 276:172-174.
[9] Saha, S and Raghava G.P.S. (2006) Prediction of Continuous B-cell Epitopes in an Antigen Using Recurrent Neural Network. Proteins; 65(1):40-48.
[10] Gomase VS and Shyamkumar K. Prediction of antigenic epitopes and MHC binders of neurotoxin alpha-KTx 3.8 from Mesobuthus tamulus sindicus. African Journal of Biotechnology, 8 (23), 6658-6676, 2009.
[11] Garnier, J. and D.J. Osguthorpe and B. Robson (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol.; 120:97-120.
[12] Karplus, P. A. and G. E. Schulz (1985). Prediction of Chain Flexibility in Proteins - A tool for the Selection of Peptide Antigens. Naturwissenschafren; 72:212-3.

Table 1- Protein Statistics of Pollen-specific protein Bnm1

| Sr. No.	Parameter	Value
1	Molecular weight (Daltans)	20055.857
2	Number of amino acids	187
3	Mean amino acid weight (Daltans)	107.251
4	Average hydrophobicity	-0.0181818
5	Ratio of hydrophilicity to hydrophobicity	1.01378
6	Percentage of hydrophilic amino acids	48.1283
7	Percentage of hydrophobic amino acids	51.8717
8	Ratio of percentage hydrophilic to Percentage hydrophobic	0.927835
9	Mean beta hydrophobic moment	0.189511
10	Mean helix hydrophobic moment	0.175531
11	Number of basic amino acids	17
12	Number of acidic amino acids	25
13	Estimated pH for protein	5.7
14	Total linear charge density	0.235294
15	Polar area of extended chain (Angs^2)	11825.0
16	Non-polar area of extended chain (Angs^2)	20281.0
17	Total area of extended chain (Angs^2)	32106.0
18	Polar ASA of folded protein (Angs^2)	2947.95
19	Non-polar ASA of folded protein (Angs^2)	5773.7
20	ASA of folded protein (Angs^2)	8721.65
21	Ratio of folded to extended area	0.293886
22	Buried polar area of folded protein (Angs^2)	7340.74
23	Buried non-polar area of folded protein (Angs^2)	12793.9
24	Buried charge area of folded protein (Angs^2)	838.942
25	Total buried surface (Angs^2)	20973.6
26	Number of buried amino acids	51
27	Packing volume (est) (Angs^3)	23940.6
28	Packing volume (act) (Angs^3)	23882.7
29	Interior volume of protein (Angs^3)	17329.4
30	Exterior volume of protein (Angs^3)	6553.25
31	Partial specific volume (ml/g)	0.72571
32	Fisher volume ratio (act)	0.378157
33	Fisher volume ratio (idealized)	0.817077
34	Protein solubility	1.48957
35	Estimated radius of folded protein (Angs)	22.1591
36	RMS end to end distance of extended chain (Angs)	143.422
37	Radius of gyration of extended chain (Angs)	58.552
38	Solv. Free energy folding (Kcal/mol)	-169.11
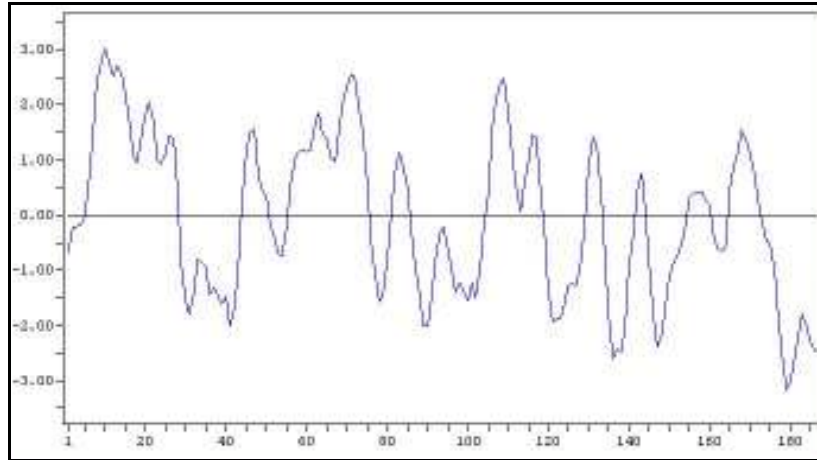
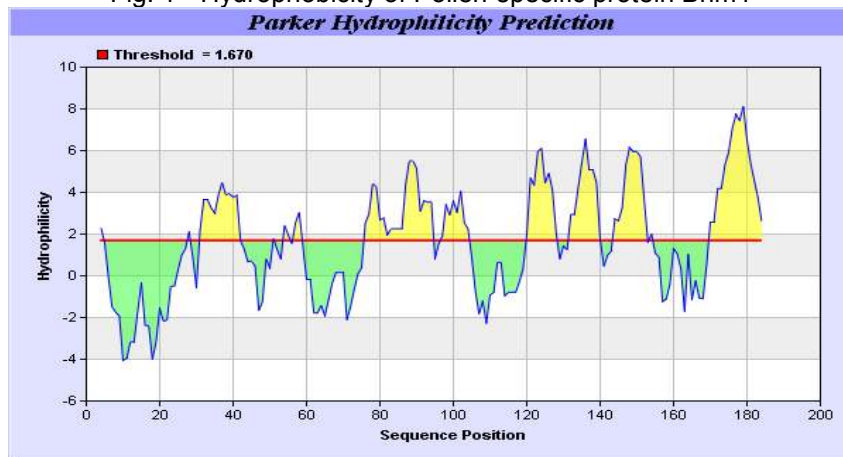Fig. 1- Hydrophobicity of Pollen-specific protein Bnm1

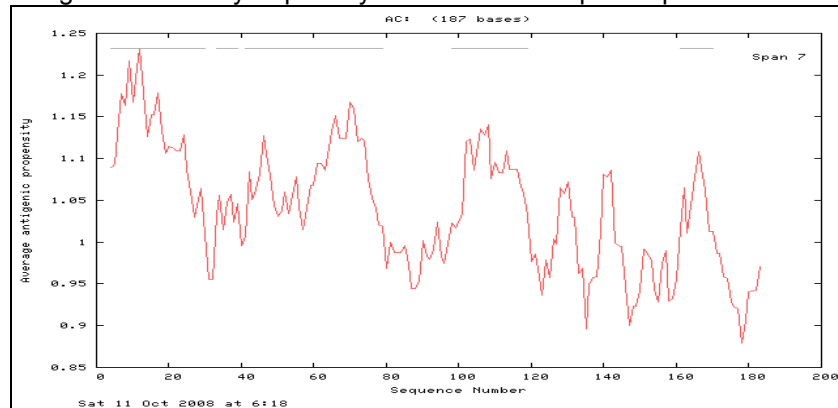Fig. 2- Parker Hydrophilicity Scale of Pollen-specific protein Bnm1

Fig. 3- Antigenic plot for sequence of Pollen-specific protein Bnm1

Table 2- Antigenic determinants in Pollen-specific protein Bnm1 protein

Sr. No.	Start Position	Sequence	End Position
1	4	YSGCFFVVTLAVLSQFLLAPASALVNR	30
2	33	IDSNCQR	39
3	41	KNKAFCIQTLTTYPPAAAATGLLPLAEAVVGLAISHCEK	79
4	98	FNECHDAYVGILASLKSALLEL	119
5	161	QMEQLLDLAA	170

Table 3- Predicted B-cell epitopes in Pollen-specific protein Bnm1

Rank	Sequence	Start position	Score
1	SQFLLAPASA	17	0.74
2	SNCQRVKNKA	35	0.72
3	REYIDSNCQR	30	0.69
4	FNECHDAYVG	98	0.68
5	DATLKTQFNE	91	0.67
6	RVKNKAFCIQ	39	0.66
7	PPAAAATGLL	54	0.64
7	GILASLKSAL	107	0.64
8	CIQTLTTYPP	46	0.60
8	ASALVNREYI	24	0.60
9	SHCEKTAGFA	75	0.58
10	GLAISHCEKT	71	0.52
10	AEAVVGLAIS	66	0.52
11	LKSALLELKD	112	0.51
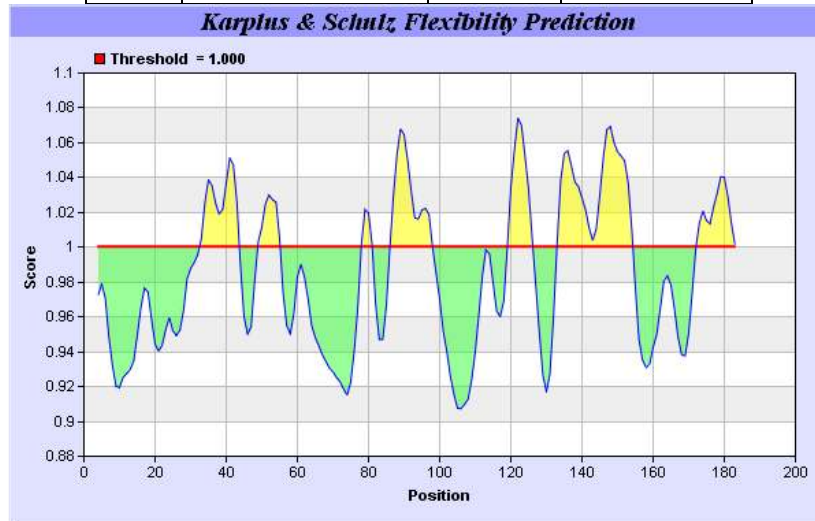
Fig. 4- Karplus-Schulz flexibility of Pollen-specific protein Bnm1 protein
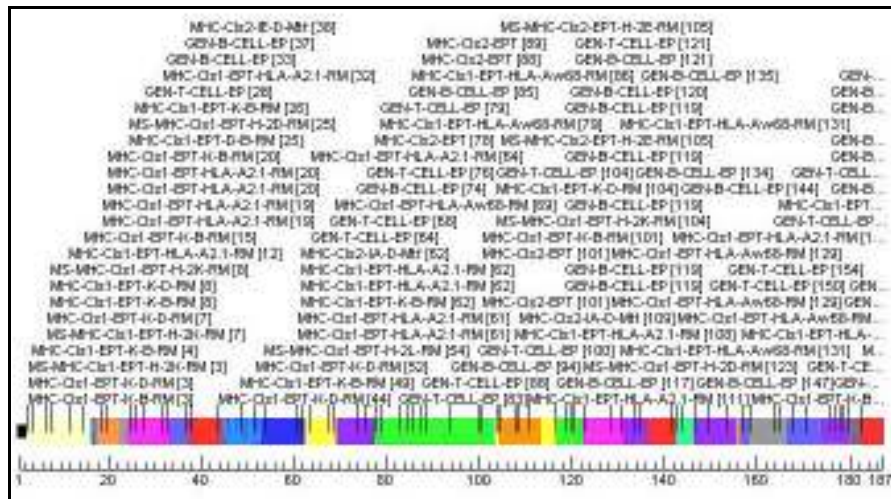
Fig. 5- Motif map of Pollen-specific protein Bnm1 protein sequence
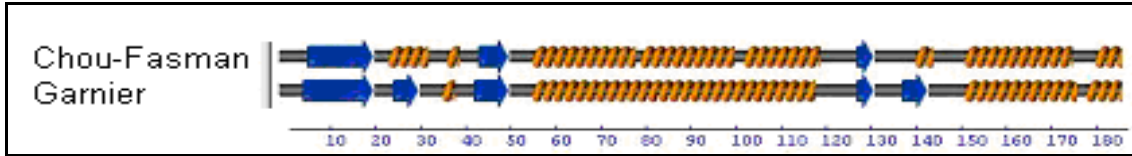
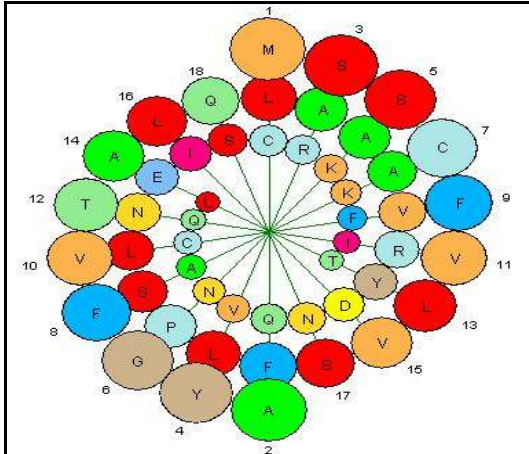Fig. 6- Structure prediction of Pollen-specific protein Bnm1

Fig. No.7. Beta staircase arrangement of amino acids in Pollen-specific protein Bnm1
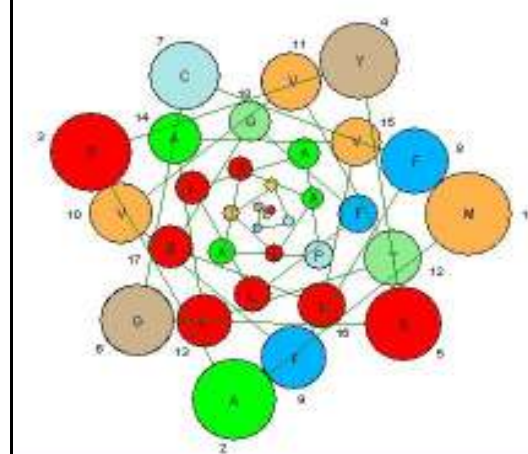

Fig. No.8. Helical wheel arrangement of amino acids in Pollen-specific protein Bnm1
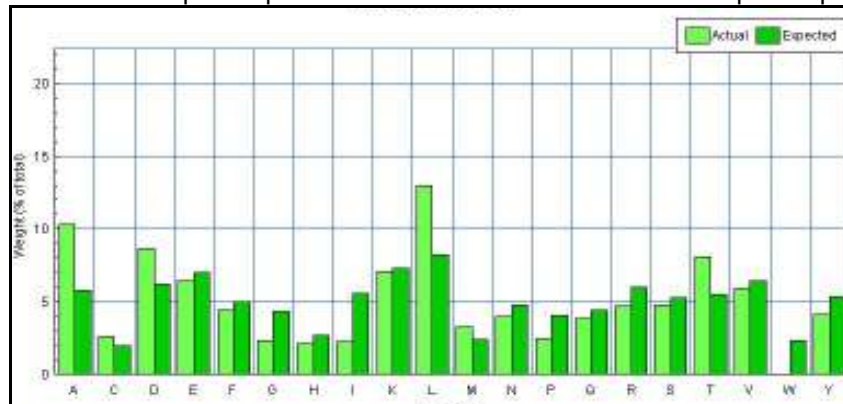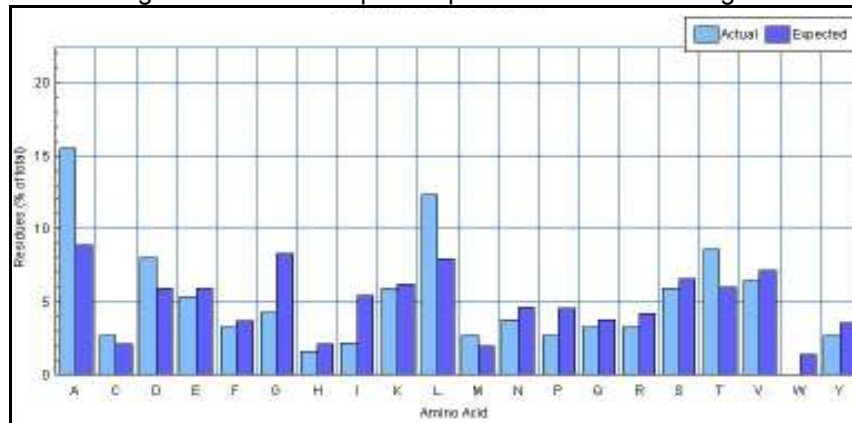

Fig. 9- Actual and expected percent amino acid weight

Fig. 10- Actual and expected amino acid frequency