Drug Designing: Open Access
ISSN: 2169-0138

Gomase and Kale, Drug Des 2015, 4:1
DOI: 10.4172/2169-0138.1000119

Research Article Open Access

Information of Surface Accessibility of the Peptide Fragments of Coat Protein from *Alfalfa mosaic virus* (AMV) at the Physicochemical and Immunochemical Levels

Gomase VS* and Kale KV

Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, 431004, Maharashtra, India

Abstract

Alfalfa mosaic virus (AMV) infects over 600 plant species in 70 families (experimental and natural hosts). Coat proteins are important materials for making structure-function correlations with biologically active peptides on the physicochemical and immunochemical levels and also are good models for observing the evolutional changes in protein molecules. Antigenic peptides at position 1-MSSSQKKAGGKAGKPTKRSQN-21; 151-PTHAGMQNQNF-161; 22-YAALRKAQLPKPPALKVPVVKPT-44 of *Alfalfa mosaic virus* coat protein are most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. In this research, we used PSSM and SVM algorithms for the prediction of MHC class I & II binding peptide, antigenicity, Solvent accessibility, polar and nonpolar residue to analyse the regions that are likely exposed on the surface of proteins which are potentially antigenic that allows potential drug targets to identify active sites against infection as well as to design synthetic peptide vaccine.

Keywords: *Alfalfa mosaic virus;* Antigenic peptides; MHC-Binders; TapPred; PSSM; SVM; Nonamers

Introduction

Alfalfa mosaic virus with bacilliform particles of different lengths, the largest usually c. 60 nm; in which four species of single-stranded RNA of messenger polarity are separately packaged. The three largest RNA species comprise the genome; the fourth is a sub-genomic messenger for the coat protein. The three genome RNA species and either the fourth RNA or the coat protein are needed for infectivity. *Alfalfa mosaic virus* (AMV) infects over 600 plant species in 70 families (experimental and natural hosts). Some hosts are potato (*Solanum tuberosum*), pea (*Pisum sativum*), tomato (*Lycopersicon esculentum*), tobacco (*Nicotiana tabacum*), bluebeard (*Caryopteris incana*). The virus is readily sap-transmissible, is seed-transmissible in some hosts and is transmitted in the non-persistent manner by aphids to a very wide range of host plants [1]. AMV infection symptoms are vary from wilting, mottles, white flecks, ringspots, malformation like dwarfing, mosaics to necrosis depending on the virus strain, host variety, stage of growth at infection and environmental conditions. Symptoms of infection can persist or disappear quickly. *Alfalfa mosaic virus* can be detected in each part of the host plant. The virions are mainly found in the cytoplasm and chloroplast of the infected plant as inclusion bodies.

Alfalfa mosaic virus (AMV) and ilarvirus RNAs are infectious only in the presence of the viral coat protein. To understand the coat protein's function is important for defining viral replication mechanisms. *In vitro* replication experiments shows the conformational switch model states that AMV coat protein blocks minus-strand RNA synthesis, while another research states that coat protein present in an inoculum is required to permit minus-strand synthesis [2,3]. *Alfalfa mosaic virus* causes various mosaics, mottles and malformations in lucerne (alfalfa; *Medicago sativa*) but is often symptomless in this host, especially during summer, and is most prevalent in old crops. AMV causes calico and tuber necrosis in potato. Various symptoms in tobacco and garden lupin, yellow fleck in *Caryopteris incana,* white mottle in *Philadelphus* sp., mosaic in *Malva parviflora* and *Viburnum opulus.* It causes of mosaic in red and white clover, celery, celeriac and lettuce, of yellow mosaic in cowpea (*Vigna unguiculata*), mung bean (*V. radiata*), bean (*Phaseolus vulgaris*), and chilli pepper (*Capsicum annuum*), of necrosis and stunting in pea (*Pisum sativum*), of severe necrosis in tomato, and of wilting in chickpea (*Cicer arietinum*). It found naturally in many wild and cultivated species [4-6].

Coat proteins are important biologically active peptides on the physicochemical and immunochemical levels and also are good models for observing the evolutional changes in protein molecules. Antigenic peptides from *Alfalfa mosaic virus* are most suitable for the development of subunit vaccine because a single toxin subunit can generate sufficient immune response. Major histocompatibility complex (MHC) molecules are cell surface proteins that binds to the peptides derived from host or antigenic proteins, and present them at the cell surface for recognition by cells. Cell recognition is a fundamental mechanism of the immune system by which the host identifies and responds to foreign antigens [7,8]. There are two types of MHC molecule and are extremely polymorphic. MHC class I molecules present peptides from proteins synthesized within the cell, whereas, MHC class II molecule present peptides derived from endocytosed extracellular proteins. Identification of MHC-binding peptides and epitopes helps improve our understanding of specificity of immune responses [9-13].

Virus transmission

The virus is transmitted by thrips, which have a wide range of hosts. The virus survives in these hosts and acts as a source of inoculums for the vector. The thrips are carried by wind. The population of vectors

***Corresponding author:** Gomase VS, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, 431004, MS, India, Tel: 91-9987770696; E-mail: gomase.viren@gmail.com

Received December 09, 2014; **Accepted** January 07, 2015; **Published** January 11, 2015

Citation: Gomase VS, Kale KV (2015) Information of Surface Accessibility of the Peptide Fragments of Coat Protein from *Alfalfa mosaic virus* (AMV) at the Physicochemical and Immunochemical Levels. Drug Des 4: 119. doi:10.4172/2169-0138.1000119

increases rapidly from January-March and August-September Kharif and hence the crop suffers a heavy loss in both the seasons. A prolonged dry spell favours the multiplication of thrips and spread of the virus.

Strategy

The phenotype of the resistant transgenic plants includes fewer centres of initial virus infection, a delay in symptom development, and low virus accumulation. Protoplasts from virus resistant transgenic plants are also resistant, suggesting that the protection is largely operational at the cellular level. Transgenic plants expressing nucleocapsid protein are protected against infection by virus particles but are susceptible to viral RNA, indicating that the protection may primarily involve an inhibition of virus uncoating. This approach is based on the phenomenon of cross-protection [14] hereby a plant infected with a mild strain of virus is protected against a more severe strain of the same virus. Proteins of soybean mosaic virus are necessary for its production in or on all food commodities. An exemption from the requirement of a tolerance is established for residues of the biological plant pesticide.

MHC class binding peptides

The new paradigm in vaccine design is emerging, following essential discoveries in immunology and development of new MHC Class-I binding peptides prediction tools [15]. MHC molecules are cell surface glycoproteins, which take active part in host immune reactions. The involvement of MHC class-I in response to almost all antigens and the variable length of interacting peptides make the study of MHC Class I molecules very interesting. MHC molecules have been well characterized in terms of their role in immune reactions. They bind to some of the peptide fragments generated after proteolytic cleavage of antigen [16]. This binding acts like red flags for antigen specific and to generate immune response against the parent antigen. So a small fragment of antigen can induce immune response against whole antigen. Coat protien peptides are most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. MHCpeptide complexes will be translocated on the surface of antigen presenting cells (APCs). This theme is implemented in designing subunit and synthetic peptide vaccines [22]. One of the important problems in subunit vaccine design is to search antigenic regions in an antigen [17] that can stimulate T cells called T-cell epitopes. In literature, fortunately, a large amount of data about such peptides is available. Pastly and presently, a number of databases have been developed to provide comprehensive information related to T-cell epitopes [18-21].

Methodology

Database searching

The antigenic protein sequence of coat protein of *Alfalfa mosaic virus* was retrieved from GenBank, UniProt databases are initially the most important [AAA46297] [22-28].

Prediction of antigenicity

Prediction of antigenicity program predicts those segments from coat protein that are likely to be antigenic by eliciting response. In this research work antigenic epitopes of coat protein of *Alfalfa mosaic virus* are determined by using the Gomase-Kale Method, Hopp and Woods, Welling, Parker, Bepipred , Kolaskar and Tongaonkar antigenicity methods [29-35].

Predict protein-protein binding sites

Profisis method (ISIS) is a machine learning-based method that identifies interacting residues from sequence alone. Although the method is developed using transient protein–protein interfaces from complexes of experimentally known 3D structures, it never explicitly uses 3D information. Instead, we combine predicted structural features with evolutionary information [36,37].

Prediction of MHC binding peptide

The major histocompatibility complex (MHC) peptide binding of coat protein of *Alfalfa mosaic virus* is predicted using neural networks trained on C terminals of known epitopes. Rankpep predicts peptide binders to MHC-I ligands whose C-terminal end is likely to be the result of proteosomal cleavage using Position Specific Scoring Matrices (PSSMs). Support Vector Machine (SVM) based method for prediction of promiscuous MHC class II binding peptides from protein sequence; SVM has been trained on the binary input of single amino acid sequence [38-42].

Prediction of antigenic peptides by cascade SVM based TAPPred method

In the present study, we predict cascade SVM based several TAP binders which was based on the sequence and the features of amino acids [43]. We found the MHCI binding regions, the binding affinity of coat protein of *Alfalfa mosaic virus*.

Solvent accessible regions

We also predict solvent accessible regions of proteins having highest probability that a given protein region lies on the surface of a protein Surface Accessibility, backbone or chain flexibility by Emini et al. [44] and Karplus and Schulz [45]. By using different scale we predict the hydrophobic and hydrophilic characteristics of amino acids that are rich in charged and polar residues i.e. Gomase-Kale method, Sweet et al., Kyte & Doolittle, Abraham & Leo, Bull and Breese, Guy, Miyazawa, et al., Roseman, Wolfenden et al., Wilson et al., Cowan, Chothia [46-57].

Results and Interpretations

Coat protein of *Alfalfa mosaic virus* contain a long residue with 221 amino acids [AAA46297].

MSSSQKKAGGKAGKPTKRSQNYAALRKAQLPKPPALKVPVVKPTNTILPQT-GCVWQSLGTPLSLSSFNGLGARFLYSFLKDFVGPRILEEDLIYRMVFSITPSHAGTF-CLTDDVTTEDGRAVAHGNPMQEFPHGAFHANEKFGFELVFTAPTHAGMQNQN-FKHSYAVALCLDFDAQPEGSKNPSFRFNEVWVERKAFPRAGPLRSLITVGLFDEAD-DLDRH

Prediction of antigenic peptides

Predict protein-protein binding sites are strongest predictions of the method reached over 90% accuracy in a cross-validation experiment. Our results suggest that despite the significant diversity in the nature of protein–protein interactions, they all share common basic principles and that these principles are identifiable from sequence alone (Figure 1).

In this study, we found the antigenic determinants by finding the area of greatest local hydrophilicity. The Hopp-Woods scale of hydrophilicity prediction result data found high in 165-173 [MIN: -2.633, MAX: 1.767] in a protein, assuming that the antigenic determinants would be exposed on the surface of the protein and thus would be located in hydrophilic regions (Figure 2). Welling

antigenicity plot gives value as the log of the quotient between percentage in a sample of known antigenic regions and percentage in average proteins and prediction result data found high in position 13-41[MIN: -1.248, MAX: 0.873] (Figure 3). We also study Hydrophobicity plot of HPLC / Parker Hydrophilicity prediction result data found 112-DDVTTED-118, Score: 6.357 (maximum) [Average: 1.376, Minimum: -4.571, Maximum: 6.357] (Figure 4), BepiPred predicts the location of linear B-cell epitopes Result found that 1-MSSSQKKAGGKAGKPTKRSQN-21; 151-PTHAGMQNQNF-161 (Maximum), [Average:0.227 Minimum:-1.937 Maximum:2.305 Threshold: 0.350] (Figure 5, Table 1), Kolaskar and Tongaonkar antigenicity methods (Figure 6, Table 2) predicted peptides result found i.e., 165-YAVALCL-171, Score :1.226 (maximum) [Average: 1.029 Minimum: 0.907 Maximum: 1.226 Threshold: 1.000] and the predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

Solvent accessible regions

We also predict solvent accessible regions in proteins; different measurement was performed for the prediction of antigenic activity, surface region of peptides. Emini et al. [44] (Figure 7,Table 3) predicts the highest probability i.e. found 14-KPTKRS-19, Score: 5.164 (maximum) (Average: 1.000 Minimum: 0.061 Maximum: 5.164 Threshold: 1.000), that a given protein region lies on the surface of a protein and are used to identify antigenic determinants on the surface of proteins. Karplus and Schulz (Figure 8) high score is found i.e. 2-SSSQKKA-8, Score : 1.127 (maximum) Average: 1.000 Minimum: 0.898 Maximum: 1.127 Threshold: 1.000. The Chou and Fasman scale which is commonly used to predict beta turns and position is i.e. 177-PEGSKNP-183, Score :1.334 (maximum). Predict backbone or chain flexibility on the basis of the known temperature B factors of the a-carbons. The hydrophobicity

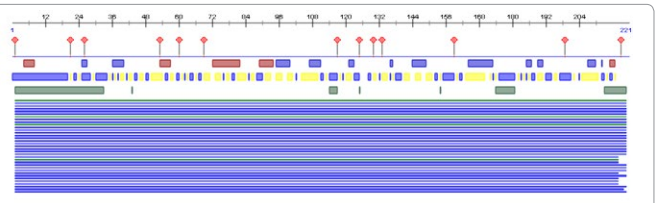

Figure 1: Predict protein-protein binding sites of coat protein of *Alfalfa mosaic virus*

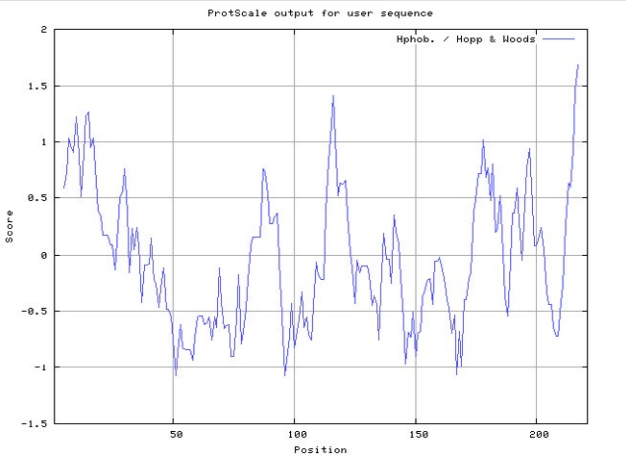

Figure 2: Hydrophobicity plot of Hopp and Woods (1981) of Coat protein of *Alfalfa mosaic virus*

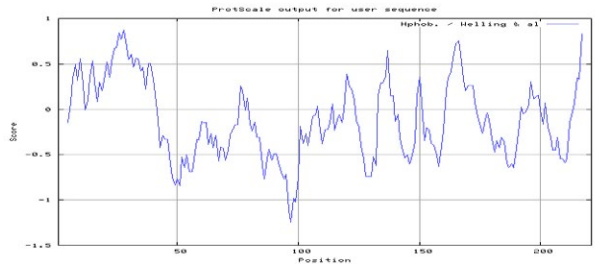

Figure 3: Hydrophobicity plot of Welling et al. (1985) of Coat protein of *Alfalfa mosaic virus*

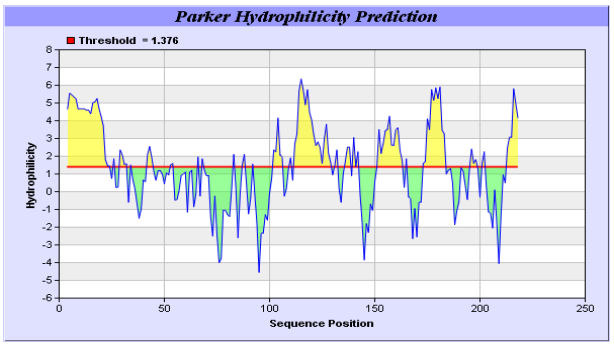

Figure 4: Hydrophobicity plot of HPLC / Parker et al. (1986) of Coat protein of *Alfalfa mosaic virus*
Average: 1.376 Minimum: -4.571 Maximum: 6.357 Threshold: 1.376

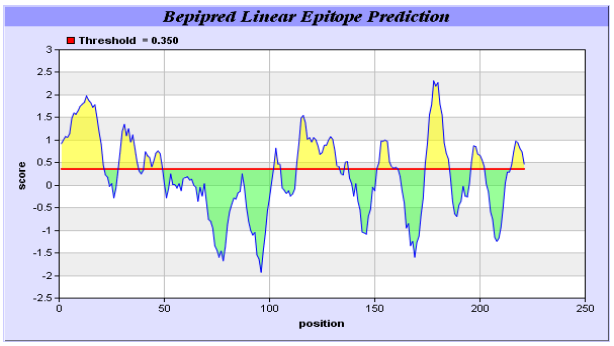

Figure 5: Bepipred Linear Epitope Prediction plot showing B-cell epitopes of Coat protein of *Alfalfa mosaic virus Average:0.227 Minimum:-1.937 Maximum:2.305 Threshold:0.350*

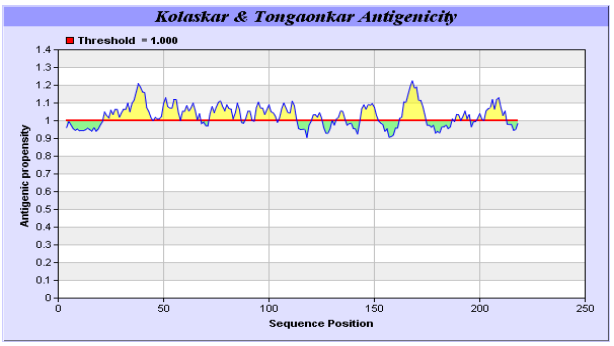

Figure 6: Kolaskar and Tongaonkar antigenicity plot for the Coat protein of *Alfalfa mosaic virus Average: 1.029 Minimum: 0.907 Maximum: 1.226 Threshold: 1.000*

and hydrophilic characteristics of amino acids is determined by using different scales that are rich in charged and polar residues i.e. Sweet et al., Kyte & Doolittle, Abraham & Leo, Bull and Breese, Guy, Miyazawa et al., Roseman, Wolfenden et al., Wilson et al., Cowan, Chothia, Chou-Fasman, Manavalan et al., [46-59] shows hydrophobicity prediction result data found high in position (Figure 8-21).

Prediction of MHC binding peptide

We found binding of peptides to a number of different alleles using position specific scoring matrix. Coat protein of *Alfalfa mosaic virus* sequence is 221 residues long, having 8mer_H2_Db, 9mer_H2_Db, 10mer_H2_Db, 11mer_H2_Db MHC I binders. MHC molecules are cell surface proteins, which actively participate in host immune reactions and involvement of MHC-I and MHC-II in response to almost all antigens. We have predicted MHC-I peptide binders of coat protein of *Alfalfa mosaic virus* was tested with on a set of 3 different alleles i.e. H2-Db (mouse) 8mer, H2-Db (mouse) 9mer, H2-Db (mouse) 10mer (Tables 4-7) and MHC-II peptide binders for I_Ab.p, I_Ad.p alleles highlighted in red represent predicted binders (Table 8). Here RANKPEP report PSSM-specific binding threshold and is obtained by scoring all the antigenic peptide sequences included in the alignment from which a profile is derived, and is defined as the score value that includes 85% of the peptides within the set [60-62]. Peptides whose score is above the binding threshold will appear highlighted in red and peptides produced by the cleavage prediction model are highlighted in violet. We also use a cascade SVM based TAPPred method which found 8 High affinity TAP Transporter peptide regions which represents predicted TAP binders residues which occur at N and C termini from coat protein of *Alfalfa mosaic virus* (Table 9).

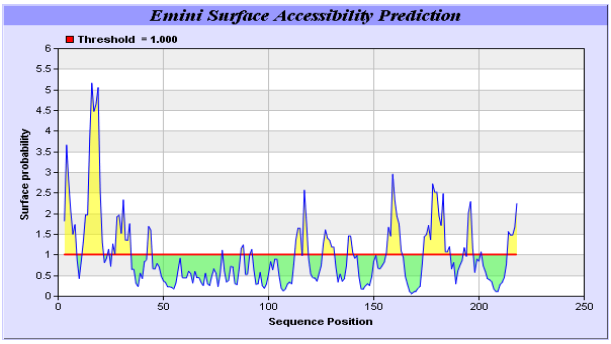

Figure 7: Emini Surface Accessibility Prediction plot of Coat protein of *Alfalfa mosaic virus Average: 1.000 Minimum: 0.061 Maximum: 5.164 Threshold: 1.000*

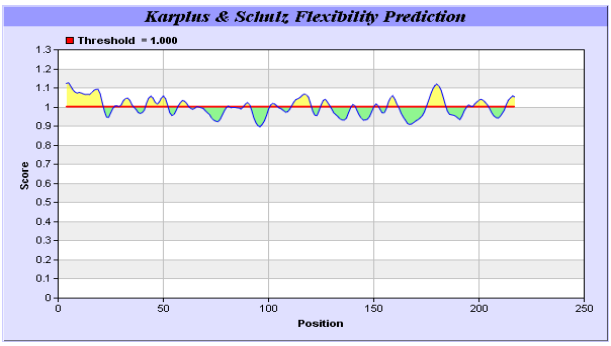

Figure 8: Karplus & Schulz Flexibility Prediction of Coat protein of *Alfalfa mosaic virus Average: 1.000 Minimum: 0.898 Maximum: 1.127 Threshold: 1.000*

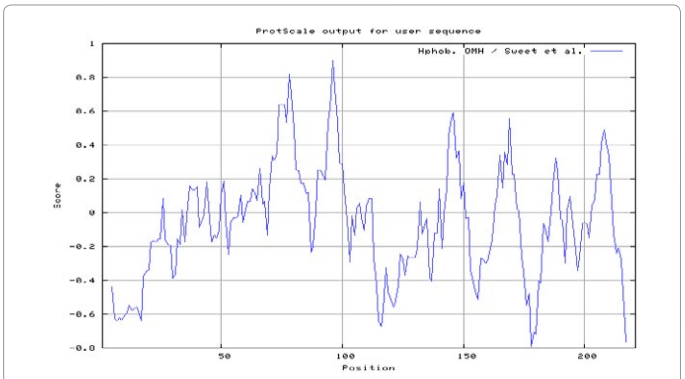

Figure 9: Hydrophobicity plot of Sweet et al. (1983) of Coat protein of *Alfalfa mosaic virus*

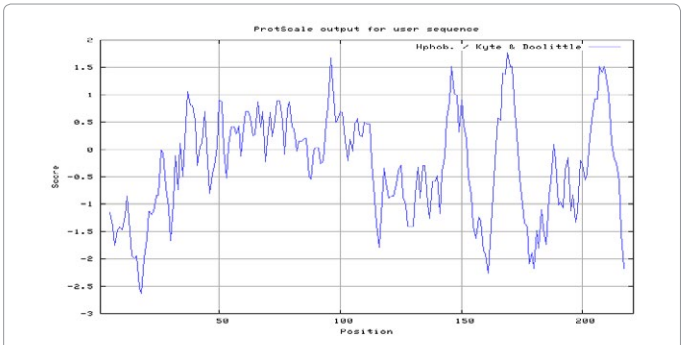

Figure 10: Kyte & Doolittle hydrophobicity plot of Coat protein of *Alfalfa mosaic virus*

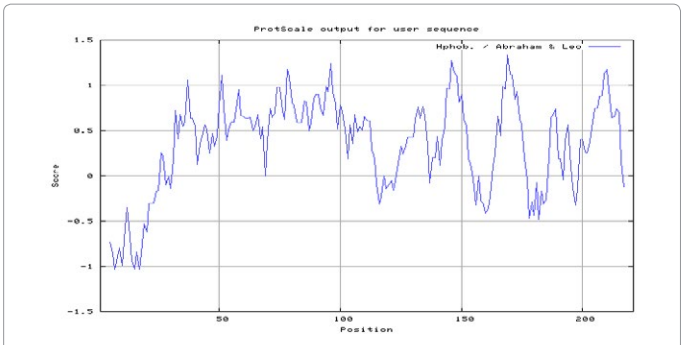

Figure 11: Abraham & Leo hydrophobicity plot of Coat protein of *Alfalfa mosaic virus*

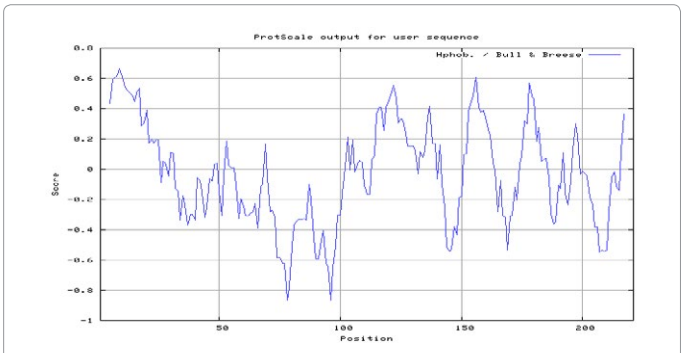

Figure 12: Bull & Breese use surface tension to measure hydrophobicity and also uses negative values to describe the hydrophobicity of Coat protein of *Alfalfa mosaic virus*

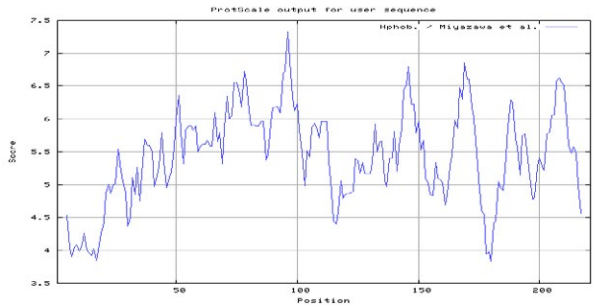

Figure 13: Hydrophobicity plot of Miyazawa et al. (1985) of Coat protein of *Alfalfa mosaic virus*

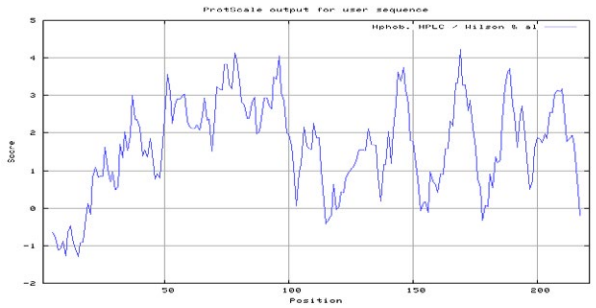

Figure 17: Hydrophobicity/HPLC plot of Wilson & al (1981) of Coat protein of *Alfalfa mosaic virus*

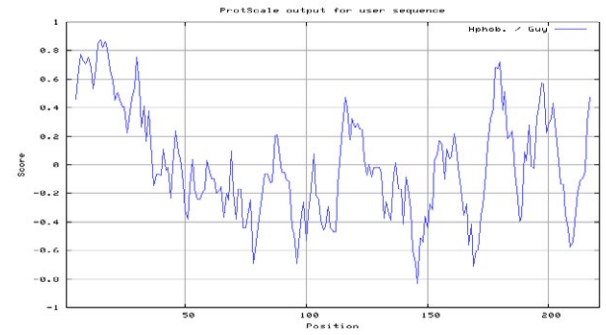

Figure 14: Hydrophobicity plot of Guy (1988) of Coat protein of *Alfalfa mosaic virus*

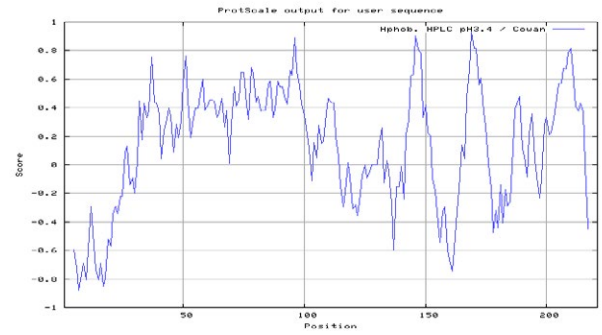

Figure 18: Hydrophobicity/HPLC pH 3.4/ plot of Cowan (1990) of Coat protein of *Alfalfa mosaic virus*

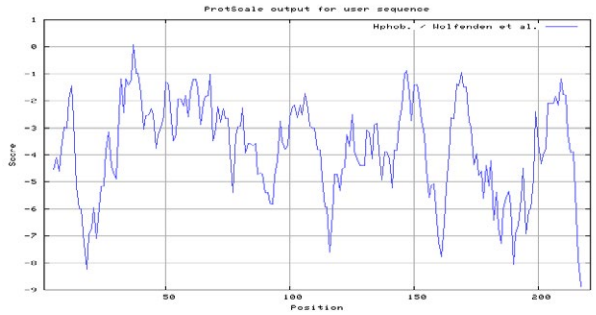

Figure 15: Hydrophobicity plot of Wolfenden et al.(1981) of Coat protein of *Alfalfa mosaic virus*

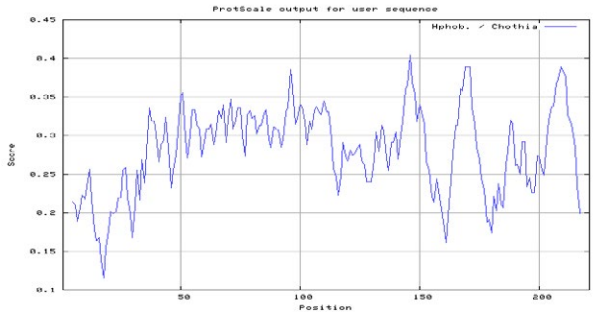

Figure 19: Hydrophobicity plot of Chothia (1976) of Coat protein of *Alfalfa mosaic virus*

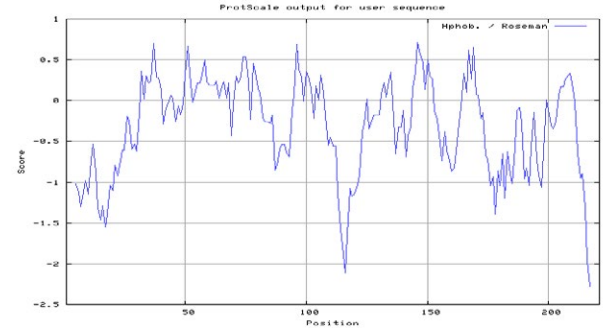

Figure 16: Hydrophobicity plot of Roseman M.A.. (1988) of Coat protein of *Alfalfa mosaic virus*

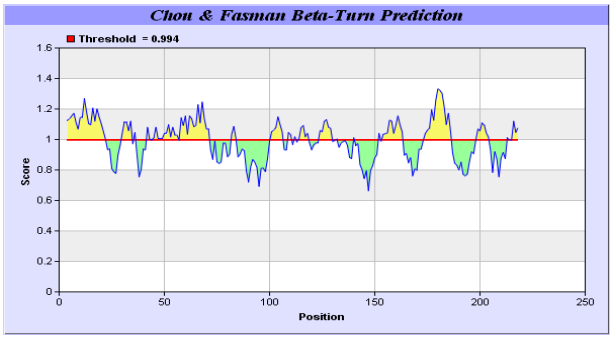

Figure20: Chou & Fasman Beta-Turn Prediction *Average: 0.994 Minimum: 0.664 Maximum: 1.334 Threshold: 0.994*

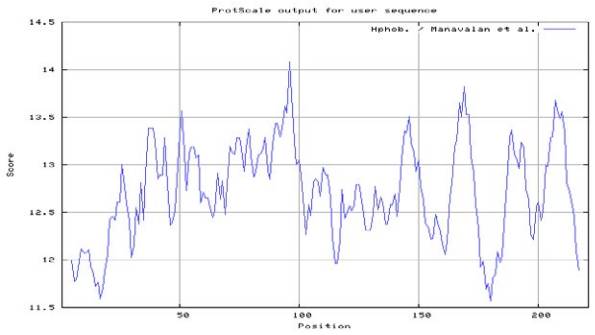

Figure 21: Hydrophobicity plot of Manavalan et al of Coat protein of *Alfalfa mosaic virus*

No.	Start Position	End Position	Peptide	Peptide Length
1	1	21	MSSSQKKAGGKAGKPTKRSQN	21
2	29	37	QLPKPPALK	9
3	41	49	VKPTNTILP	9
4	102	105	PSHA	4
5	113	133	DVTTEDGRAVAHGNPMQEFPH	21
6	136	137	FH	2
7	151	161	PTHAGMQNQNF	11
8	174	185	DAQPEGSKNPSF	12
9	196	202	AFPRAGP	7

Table 1: Bepipred linear epitope predicted epitopes

No.	Start Position	End Position	Peptide	Peptide Length
1	22	44	YAALRKAQLPKPPALKVPVVKPT	23
2	46	67	TILPQTGCVWQSLGTPLSLSSF	22
3	72	86	ARFLYSFLKDFVGPR	15
4	94	103	YRMVFSITPS	10
5	105	113	AGTFCLTDD	9
6	144	151	FELVFTAP	8
7	162	174	KHSYAVALCLDFD	13
8	189	195	EVWVERK	7
9	199	212	RAGPLRSLITVGLF	14

Table 2: Kolaskar and Tongaonkar antigenicity Predicted peptides

No.	Start Position	End Position	Peptide	Peptide Length
1	3	8	SSQKKA	6
2	12	21	AGKPTKRSQN	10
3	26	34	RKAQLPKPP	9
4	126	131	NPMQEF	6
5	156	163	MQNQNFKH	8
6	174	186	DAQPEGSKNPSFR	13

Table 3: Emini surface accessibility predicted peptides

Discussion

In this study, we found the antigenic determinants by finding the area of greatest local hydrophilicity. Hopp and Woods hydrophobicity scale is used to identify of potentially antigenic sites in proteins. Hydrophilicity Prediction result data found high in amino acid position at 165-173 [MIN: -2.633, MAX: 1.767] in a protein this scale is basically a hydrophilic index where apolar residues have been assigned negative values. The Window size of 5-7 is good for finding hydrophilic regions, greater than 0 values are consider as hydrophilic which is consider as antigenic. Welling used information on the relative occurrence of amino acids in antigenic regions to make a scale which is useful for prediction of antigenic regions and the predicted result data found high in sequence position 68-70. Welling antigenicity plot gives value as the log of the quotient between percentage in a sample of known antigenic regions and percentage in average proteins and prediction result data found high in position 13-41[MIN: -1.248, MAX: 0.873]. We also study Hydrophobicity plot of HPLC / Parker Hydrophilicity prediction result data found 112-DDVTTED-118, Score: 6.357 (maximum) [Average: 1.376, Minimum: -4.571, Maximum: 6.357]. BepiPred predicts the location of linear B-cell epitopes at position 1-MSSSQKKAGGKAGKPTKRSQN-21; 151-PTHAGMQNQNF-161 (Maximum), [Average:0.227 Minimum:-1.937 Maximum:2.305 Threshold: 0.350]. There are 3 antigenic determinant sequences is found by Kolaskar and Tongaonkar antigenicity scales the results show highest pick at position 165-YAVALCL-171, Score :1.226 (maximum) [Average: 1.029 Minimum: 0.907 Maximum: 1.226 Threshold: 1.000]. Result of determined antigenic sites on proteins has revealed that the hydrophobic residues if they occur on the surface of a protein are more likely to be a part of antigenic sites. This method can predict antigenic determinants with about 75% accuracy and also gives the information of surface accessibility and flexibility. Further this region form beta sheet which show high antigenic response than helical region of this peptide and shows highly antigenicity.

We predict solvent accessibility by using Emini et al. [44] the result found the highest probability i.e. found 5.164 Maximum in 14-KPTKRS-19, that a given protein region lies on the surface of a protein and are used to identify antigenic determinants on the surface of proteins. This algorithm also used to identify the antigenic determinants on the surface of proteins and Karplus and Schulz predict backbone or chain flexibility on the basis of the known temperature B factors of the a-carbons here we found the result with high score is i.e. 1.127 maximum in 2-SSSQKKA-8.

We predicted solvent accessibility of coat protein of *Alfalfa mosaic virus* for delineating hydrophobic and hydrophilic characteristics of amino acids. Solvent accessibility used to identify active site of functionally important residues in membrane proteins. Solvent-accessible surface areas and backbone angles are continuously varying because proteins can move freely in a three-dimensional space. The mobility of protein segments which are located on the surface of a protein due to an entropic energy potential and which seem to correlate well with known antigenic determinants. We also found the hydrophobicity prediction result data found high in position. These scales are a hydrophilic with a polar residues assigned negative value. Because the N- and C- terminal regions of proteins are usually solvent accessible and unstructured, antibodies against those regions recognize the antigenic protein. Gomase method, B-EpiPred Server, Hopp and Woods, Welling, Parker, Kolaskar and Tongaonkar antigenicity scales were designed to predict the locations of antigenic determinants in *Alfalfa mosaic virus* (coat protein). Coat protein shows beta sheets regions, which are high antigenic response than helical region of this peptide and shows highly antigenicity. We also found the Sweet hydrophobicity, Kyte & Doolittle hydrophobicity, Abraham & Leo , Bull & Breese hydrophobicity, Guy, Miyazawa hydrophobicity, Roseman hydrophobicity, Cowan HPLC pH7.5 hydrophobicity, Rose hydrophobicity, Eisenberg hydrophobicity, Manavalan hydrophobicity, Black hydrophobicity, Fauchere hydrophobicity, Janin hydrophobicity, Rao & Argos hydrophobicity, Wolfenden hydrophobicity, Wilson HPLC hydrophobicity, Cowan HPLC pH3.4, Tanford hydrophobicity, Rf mobility hydrophobicity and Chothia hydrophobicity scales, Theses scales are essentially a hydrophilic index, with a polar residues assigned negative values.

MHC:I Allele Mouse	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
8mer_H2_Db	183	SKN	PSFRFNEV	WVE	977.1	15.274	29.10%
8mer_H2_Db	48	NTI	LPQTGCVW	QSL	862.04	12.589	23.98%
8mer_H2_Db	115	DDV	TTEDGRAV	AHG	829.86	7.115	13.55%
8mer_H2_Db	15	AGK	PTKRSQNY	AAL	975.07	6.269	11.94%
8mer_H2_Db	180	PEG	SKNPSFRF	NEV	964.1	6.076	11.57%
8mer_H2_Db	202	RAG	PLRSLITV	GLF	880.1	5.516	10.51%
8mer_H2_Db	124	AVA	HGNPMQEF	PHG	941.03	4.134	7.88%
8mer_H2_Db	57	VWQ	SLGTPLSL	SSF	768.91	3.424	6.52%
8mer_H2_Db	34	PKP	PALKVPVV	KPT	804.04	2.857	5.44%
8mer_H2_Db	18	PTK	RSQNYAAL	RKA	904	2.673	5.09%
8mer_H2_Db	103	ITP	SHAGTFCL	TDD	816.93	2.282	4.35%
8mer_H2_Db	139	FHA	NEKFGFEL	VFT	965.08	1.594	3.04%
8mer_H2_Db	166	HSY	AVALCLDF	DAQ	833.02	1.347	2.57%
8mer_H2_Db	69	SFN	GLGARFLY	SFL	878.05	1.216	2.32%
8mer_H2_Db	10	KAG	GKAGKPTK	RSQ	767.91	1.015	1.93%
8mer_H2_Db	199	AFP	RAGPLRSL	ITV	851.03	0.795	1.51%
8mer_H2_Db	29	RKA	QLPKPPAL	KVP	845.06	0.452	0.86%

Table 4: Promiscuous 8mer_H2_Db (Mouse) MHC I ligands, having C-terminal ends are proteosomal cleavage sites of coat protein of Alfalfa mosaic virus

MHC:I Allele Mouse	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
9mer_H2_Db	17	KPT	KRSQNYAAL	RKA	1032.17	10.272	20.40%
9mer_H2_Db	199	AFP	RAGPLRSLI	TVG	964.19	9.692	19.24%
9mer_H2_Db	102	SIT	PSHAGTFCL	TDD	914.05	7.3	14.49%
9mer_H2_Db	67	LSS	FNGLGARFL	YSF	976.15	6.405	12.72%
9mer_H2_Db	184	KNP	SFRFNEVWV	ERK	1142.32	3.937	7.82%
9mer_H2_Db	134	FPH	GAFHANEKF	GFE	1002.1	3.723	7.39%
9mer_H2_Db	28	LRK	AQLPKPPAL	KVP	916.14	3.591	7.13%
9mer_H2_Db	50	ILP	QTGCVWQSL	GTP	980.13	2.371	4.71%
9mer_H2_Db	123	RAV	AHGNPMQEF	PHG	1012.11	1.777	3.53%
9mer_H2_Db	71	NGL	GARFLYSFL	KDF	1055.26	1.467	2.91%
9mer_H2_Db	74	GAR	FLYSFLKDF	VGP	1161.38	1.307	2.60%
9mer_H2_Db	161	NQN	FKHSYAVAL	CLD	1017.2	0.639	1.27%

Table 5: Promiscuous 9mer_H2_Db (Mouse) MHC I ligands, having C-terminal ends are proteosomal cleavage sites of coat protein of Alfalfa mosaic virus

MHC:I Allele Mouse	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
10mer_H2_Db	27	ALR	KAQLPKPPAL	KVP	1044.31	16.664	28.31%
10mer_H2_Db	209	LIT	VGLFDEADDL	DRH	1075.15	12.405	21.08%
10mer_H2_Db	181	EGS	KNPSFRFNEV	WVE	1219.37	9.238	15.70%
10mer_H2_Db	162	QNF	KHSYAVALCL	DFD	1086.32	7.505	12.75%
10mer_H2_Db	45	KPT	NTILPQTGCV	WQS	1027.19	5.087	8.64%
10mer_H2_Db	156	HAG	MQNQNFKHSY	AVA	1278.4	3.317	5.64%
10mer_H2_Db	122	GRA	VAHGNPMQEF	PHG	1111.24	2.966	5.04%
10mer_H2_Db	17	KPT	KRSQNYAALR	KAQ	1188.36	2.696	4.58%
10mer_H2_Db	39	LKV	PVVKPTNTIL	PQT	1063.29	1.449	2.46%
10mer_H2_Db	9	KKA	GGKAGKPTKR	SQN	981.15	1.237	2.10%
10mer_H2_Db	198	KAF	PRAGPLRSLI	TVG	1061.31	1.04	1.77%
10mer_H2_Db	135	PHG	AFHANEKFGF	ELV	1149.28	0.791	1.34%
10mer_H2_Db	160	QNQ	NFKHSYAVAL	CLD	1131.3	0.78	1.33%
10mer_H2_Db	74	GAR	FLYSFLKDFV	GPR	1260.51	0.645	1.10%
10mer_H2_Db	101	FSI	TPSHAGTFCL	TDD	1015.15	0.442	0.75%
10mer_H2_Db	8	QKK	AGGKAGKPTK	RSQ	896.04	0.219	0.37%

Table 6: Promiscuous 10mer_H2_Db (Mouse) MHC I ligands, having C:terminal ends are proteosomal cleavage sites of coat protein of Alfalfa mosaic virus

MHC:I Allele Mouse	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
11mer_H2_Db	17	KPT	KRSQNYAALRK	AQL	1316.53	11.716	14.74%
11mer_H2_Db	100	VFS	ITPSHAGTFCL	TDD	1128.31	6.124	7.70%
11mer_H2_Db	64	PLS	LSSFNGLGARF	LYS	1150.31	4.796	6.03%
11mer_H2_Db	134	FPH	GAFHANEKFGF	ELV	1206.33	2.131	2.68%
11mer_H2_Db	69	SFN	GLGARFLYSFL	KDF	1225.47	0.713	0.90%
11mer_H2_Db	136	HGA	FHANEKFGFEL	VFT	1320.48	0.278	0.35%
11mer_H2_Db	161	NQN	FKHSYAVALCL	DFD	1233.5	0.067	0.08%

Table 7: Promiscuous 11mer_H2_Db (Mouse) MHC I ligands, having C:terminal ends are proteosomal cleavage sites of coat protein of Alfalfa mosaic virus

MHC:II Allele	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
MHC:II I_Ab	94	DLI	YRMVFSITP	SHA	1095.33	13.779	38.67%
MHC:II I_Ab	120	EDG	RAVAHGNPM	QEF	934.08	12.843	36.04%
MHC:II I_Ab	148	ELV	FTAPTHAGM	QNQ	914.04	11.612	32.59%
MHC:II I_Ab	197	RKA	FPRAGPLRS	LIT	982.17	10.747	30.16%
MHC:II I_Ab	165	KHS	YAVALCLDF	DAQ	996.2	10.614	29.79%
MHC:II I_Ab	174	LDF	DAQPEGSKN	PSF	926.94	10.54	29.58%
MHC:II I_Ab	146	GFE	LVFTAPTHA	GMQ	938.09	9.898	27.78%
MHC:II I_Ab	30	KAQ	LPKPPALKV	PVV	944.23	9.818	27.55%
MHC:II I_Ad	118	TTE	DGRAVAHGN	PMQ	877.91	12.73	23.95%
MHC:II I_Ad	96	IYR	MVFSITPSH	AGT	1000.18	10.768	20.26%
MHC:II I_Ad	161	NQN	FKHSYAVAL	CLD	1017.2	10.357	19.49%
MHC:II I_Ad	98	RMV	FSITPSHAG	TFC	897.99	9.948	18.72%
MHC:II I_Ad	68	SSF	NGLGARFLY	SFL	992.15	8.966	16.87%

Table 8: Prediction of I_Ab & I_Ad MHCII ligands all rows highlighted in red represent predicted binders.

Peptide Rank	Start Position	Sequence	Score	Predicted Affinity
1	72	ARFLYSFLK	8.259	High
2	92	LIYRMVFSI	7.788	High
3	123	AHGNPMQEF	7.263	High
4	159	QNFKHSYAV	6.942	High
5	179	GSKNPSFRF	6.507	High
6	203	LRSLITVGL	6.337	High
7	88	LEEDLIYRM	6.3	High
8	195	KAFPRAGPL	6.027	High

Table 9: cascade SVM based High affinity TAP Binders of coat protein of Alfalfa mosaic virus

In this study, we found predicted MHC-I peptide binders of toxin protein for 11mer_H2_Db, 10mer_H2_Db, 9mer_H2_Db, 8mer_H2_Db alleles and I_Ab,I_Ad, for MHC II allele was tasted. The predicted binding affinity is normalized by the 1% fractil. The MHC peptide binding is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding is a log-transformed value related to the IC50 values in nM units. These MHC binding peptides are sufficient for eliciting the desired immune response. Predicted MHC binding regions in an antigen sequence and there are directly associated with immune reactions, in analysis we found the MHCI and MHCII binding regions. We also use a cascade SVM based TAPPred method which found 8 High affinity TAP Transporter peptide regions which represents predicted TAP binders residues which occur at N and C termini from coat protein of *Alfalfa mosaic virus*. TAP is an important transporter that transports antigenic peptides. TAP binds and translocate selective antigenic peptides for binding to specific MHC molecules. The efficiency of TAP-mediated translocation of antigenic peptides is directly proportional to its TAP binding affinity. Thus, by understanding the nature of peptides, that bind to TAP with high affinity, is important steps in endogenous antigen processing. The correlation coefficient of 0.88 was obtained by using jackknife validation test. In this test, we found the MHCI and MHCII binding regions. Cell immune responses are derived by antigenic epitopes hence their identification is important for design synthetic peptide vaccine. Cell epitopes are recognized by MHCI molecules producing a strong defensive immune response against of coat protein of *Alfalfa mosaic virus*. Therefore, the prediction of peptide binding to MHCI molecules by appropriate processing of antigen peptides occurs by their binding to the relevant MHC molecules. Because, the C-terminus of MHCI-restricted epitopes results from cleavage by the proteasome

and thus, proteasome specifity is important for determing cell epitopes. Consequently, RANKPEP also focus on the prediction of conserved epitopes. C-terminus of MHCI-restricted peptides is generated by the proteasome, and thus RANKPEP also determines whether the C-terminus of the predicted MHCI-peptide binders is the result of proteasomal cleavage. Moreover, these sequences are highlighted in purple in the output results. Proteasomal cleavage predictions are carried out using three optional models obatined applying statistical language models to a set of known epitopes restricted by human MHCI molecules as indicated here.

Conclusion

From the above result and discussion it is concluded that the ability of RANKPEP to predict MHC binding peptides, and thereby potential epitopes, antigenic peptide that binds to MHC molecule are antigenic that means hydrophilic in nature. This means the increase in affinity of MHC binding peptides may result in enhancement of immunogenicity of coat protein of *Alfalfa mosaic virus.* Hence synthetic peptides will helpful in the designing of peptide vaccine. This approach can help reduce the time and cost of experimentation for determining functional properties of coat protein of *Alfalfa mosaic virus.* Overall, the results are encouraging, both the 'sites of action' and 'physiological functions' can be predicted with very high accuracies helping minimize the number of validation experiments.

Future Perspectives

This method will be applicable in immunodiagnostics, vaccine design for understanding of autoimmune susceptibility. Coat protein of *Alfalfa mosaic virus* involved multiple antigenic components and useful to protect the host from the nucleocapsid. MHC molecules are cell surface proteins, which take active part in host immune reactions and involvement of MHC class in response to almost all antigens and it give effects on specific sites. Predicted MHC binding regions acts like red flags for antigen specific and generate immune response against the parent antigen. So a small fragment of antigen can induce immune response against whole antigen. The method integrates prediction of peptide MHC class binding; proteosomal C terminal cleavage and TAP transport efficiency. This theme is implemented in designing subunit and peptide vaccines.

References

1. Taschner PE, van der Kuyl AC, Neeleman L, Bol JF (1991) Replication of an incomplete *Alfalfa mosaic virus* genome in plants transformed with viral replicase genes. Virology 181: 445-450.
2. Olsthoorn RC, Mertens S, Brederode FT, Bol JF (1999) A conformational switch at the 3′ end of a plant virus RNA regulates viral replication. EMBO J 18: 4856-4864.
3. Neeleman L, Van der Vossen EA, Bol JF (1993) Infection of tobacco with *Alfalfa mosaic virus* cDNAs sheds light on the early function of the coat protein. Virology 196: 883-887.
4. Clark MF, Adams AN (1977) Characteristics of the Microplate Method of Enzyme-Linked Immunosorbent Assay for the Detection of Plant Viruses. J Gen Virol 34: 475-483.
5. Gossen BD, Martin RR (1993) Can J Plant Pathol 15: 314.
6. Jaspars EMJ, Bos L (1980) Descriptions of plant viruses. No. 229. CMI/AAB.
7. Batalia MA, Collins EJ (1997) Peptide Binding by Class 1 and Class II MHC Molecules. Biopoly 43: 281-302.
8. Flower DR (2008) "Vaccines: how they work," in Bioinformatics for Vaccinology. Wiley-Blackwell, Oxford, UK.
9. Marrack P, Scott-Browne JP, Dai S, Gapin L, Kappler JW (2008) Review Evolutionarily conserved amino acids that control TCR-MHC interaction. Annu Rev Immunol 26: 171-203.
10. Chapman HA (1998) Endosomal proteolysis and MHC class II function. Curr Opin Immunol 10: 93-102.
11. Watts C (2004) The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. Nat Immunol 5: 685-692.
12. Neefjes J, Jongsma ML, Paul P, Bakke O (2011) Review Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol 11: 823-836.
13. Laforest SM, Gehrke L (2004) Spatial determinants of the *Alfalfa mosaic virus* coat protein binding site. RNA 10: 48-58.
14. Valkonen JP, Rajamäki ML, Kekarainen T (2002) Mapping of viral genomic regions important in cross-protection between strains of a potyvirus. Mol Plant Microbe Interact 15: 683-692.
15. Cui J, Han LY, Lin HH, Tang ZQ, Jiang L (2006) MHC-BPS: MHC-binder prediction server for identifying peptides of flexible lengths from sequencederived physicochemical properties. Immunogenetics 58: 607-613.
16. Kumar M, Gromiha MM, Raghava GP (2007) Identification of DNA-binding proteins using support vector machines and evolutionary profiles. BMC Bioinformatics 8: 463.
17. Schirle M, Weinschenk T, Stevanovic S (2001) Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens. J Immunol Methods 257: 1-16.
18. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50: 213-219.
19. Blythe MJ, Doytchinova IA, Flower DR (2002) JenPep: a database of quantitative functional peptide data for immunology. Bioinformatics 18: 434-439.
20. Schonbach C, Koh JL, Flower DR, Wong L, Brusic V (2002) FIMM, a database of functional molecular immunology: update 2002. Nucleic Acids Res 30: 226-229.
21. Korber TMB, Brander C, Haynes BF, Koup R, Kuiken C, et al. (2001) Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico LA UR 02-4663.
22. Gomase VS, Kale KV, Chikhale NJ, Changbhale SS (2007) Prediction of MHC Binding Peptides and Epitopes from *Alfalfa mosaic virus*. Curr Drug Discov Technol 4: 117-121.
23. Gracy J, Argos P (1998) Automated protein sequence database classification. I. Integration of compositional similarity search, local similarity search, and multiple sequence alignment. Bioinformatics 14: 164-173.
24. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, et al. (2000) The Pfam protein families'database. Nucleic Acids Res 28: 263-266.
25. Barker WC, Bairoch A, Apweiler R, Wu CH, BoeckmannB, et al. (2000) The Protein Information Resource (PIR). Nucleic Acids Res 31: 345–347.
26. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, WheelerDL (2003) GenBank. Nucleic Acids Res. 31: 23.
27. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, et al. (2005) The Universal Protein Resource (UniProt). Nucleic Acids Res. 33: D154-159.
28. Joshi S, Neeleman L, Pleij CW, Haenni AL, Chapeville F, et al. (1984) Nonstructural *Alfalfa mosaic virus* RNA-coded proteins present in tobacco leaf tissue. Virology 139: 231-242.
29. Hoop TP, Woods KR (1978) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A 78: 3824.
30. Welling GW, Weijer WJ, van der Zee R, Welling-Wester S (1985) Prediction of sequential antigenic regions in proteins. FEBS Lett 188: 215-218.
31. Parker KC, Bednarek MA, Coligan JE (1994) Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunology 152: 163-175.
32. Larsen JE, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2: 2.
33. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276: 172-174.

34. Gomase VS, Kale KV (2008) Development of MHC class nonamers from Cowpea mosaic viral protein. Gene Therapy and Molecular Biology 12: 87-94.

35. Gomase VS, Kale KV (2008) Prediction of MHC binder for fragment based viral peptide vaccines from cabbage leaf curl virus. Gene Therapy and Molecular Biology 12: 83-86.

36. Ofran Y, Rost B (2007) ISIS: interaction sites identified from sequence. Bioinformatics 23: e13-16.

37. Peter Hönigschmid, Edda Kloppmann, Burkhard Rost (2012) "Thesis: "Improvement of DNA- and RNA-Protein Binding Prediction", Dipl.-Bioinf. (univ)

38. Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC class I binding peptides using profile motifs. Hum Immunol 63: 701-709.

39. Reche PA, Reinherz EL (2003) Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. J Mol Biol 331: 623-641.

40. Craiu A, Akopian T, Goldberg A, Rock KL (1997) Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. Proc Natl Acad Sci U S A 94: 10850-10855.

41. Pieters J1 (2000) MHC class II-restricted antigen processing and presentation. Adv Immunol 75: 159-208.

42. Gomase VS, Kale KV (2008) Antigenic epitopes of viral polyprotein: an approach for fragment based peptide vaccines from Papaya Ringspot virus. Gene Therapy and Molecular Biology 12: 31-38.

43. Bhasin M, Raghava GP (2004) Analysis and prediction of affinity of TAP binding peptides using cascade SVM. Protein Sci 13: 596-607.

44. Emini EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol 55: 836-839.

45. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins: a tool for the selection of peptide antigen. Natur wissen schaften 72: 212-213.

46. Sweet RM, Eisenberg D (1983) Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. J Mol Biol 171: 479-488.

47. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132.

48. Abraham DJ, Leo AJ (1987) Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. Proteins 2: 130-152.

49. Bull HB, Breese K (1974) Surface tension of amino acid solutions: a hydrophobicity scale of the amino acid residues. Arch Biochem Biophys 161: 665-670.

50. Miyazawa S, Jernigen RL (1985) Estimation of Effective Interresidue Contact Energies from Protein Crystal Structures: Quasi-Chemical Approximation. Macromolecules 18: 534-552.

51. Roseman MA1 (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. J Mol Biol 200: 513-522.

52. Wolfenden R, Andersson L, Cullis PM, Southgate CC (1981) Affinities of amino acid side chains for solvent water. Biochemistry 20: 849-855.

53. Wilson KJ, Honegger A, Stötzel RP, Hughes GJ (1981) The behaviour of peptides on reverse-phase supports during high-pressure liquid chromatography. Biochem J 199: 31-41.

54. Cowan R, Whittaker RG (1990) Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography. Pept Res 3: 75-80.

55. Chothia C (1976) The nature of the accessible and buried surfaces in proteins. J Mol Biol 105: 1-12.

56. Gomase VS1 (2006) Prediction of antigenic epitopes of neurotoxin Bmbktx1 from Mesobuthus martensii. Curr Drug Discov Technol 3: 225-229.

57. Gomase VS, Chitlange NR, Changbhale SS, Kale KV (2013) Prediction of Brugia malayi Antigenic Peptides: Candidates for Synthetic Vaccine Design Against Lymphatic Filariasis. Protein and Peptide Letters 20: 864-887.

58. Gomase VS, Kale KV, Shyamkumar K (2008) Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV). Journal of Proteomics & Bioinformatics 1: 188- 205.

59. Gomase VS, Kale KV, Shyamkumar K, Shankar S (2008) Computer Aided Multi Parameter Antigen Design: Impact of Synthetic Peptide Vaccines from Soybean Mosaic Virus. ICETET 2008, IEEE Computer Society in IEEE Xplore, Los Alamitos, California.

60. Reche PA, Glutting JP, Zhang H, Reinherz EL (2004) Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. Immunogenetics 56: 405-419.

61. Gomase VS, Changbhale SS, Chitlange NR, Sherkhane AS (2012) Prediction of Antigenic Epitopes from Tityus serrulatus Venom allergen 5: an aid to Antitoxic Vaccines. Journal of Toxicology Research 2: 20-24.

62. Gomase VS, Chitlange NR, Changbhale SS, Sherkhane AS, Kale KV (2012) In-Silico Approach for Prediction of Vaccine Potential Antigenic Peptides from 23-kDa Transmembrane Antigen Protein of Schistosoma haematobium. International Journal of Bioinformatics Research 4: 276-281.