Journal of
Data Mining in Genomics & Proteomics

Gomase and Chitlange, J Data Mining Genomics Proteomics 2012, 3:3
DOI: 10.4172/2153-0602.1000120

Research Article Open Access

Microbial Proteomics Approach for Sensitive Quantitative Predictions of MHC Binding Peptide from *Taenia ovis*

Virendra Gomase* and Chitlange NR

The Global Open University, Nagaland, India

Abstract

Taenia ovis is a tapeworm parasite with the adult stage of the parasite found in the intestines of dogs, while the intermediate or larval stage is found in the muscles of sheep, causes sheep measles. Peptide fragments of antigen protein can be used to select nonamers for use in rational vaccine design, and to increase the understanding of roles of the immune system in infectious diseases. Analysis shows MHC class II binding peptides of antigen protein from *Taenia ovis* are important determinants for protection of host from parasitic infection. In this assay, we used PSSM and SVM algorithms for antigen design and predicted the binding affinity of antigen protein having 254 amino acids, which shows 246 nonamers. Binding ability prediction of antigen peptides to Major Histocompatibility Complex (MHC) class I & II molecules is important in vaccine development against sheep measles.

Keywords: Cysticercosis; Antigen protein; Epitope; PSSM; SVM; MHC; Peptide vaccine

Abbreviations: MHC: Major Histocompatibility Complex; PSSMs: Position Specific Scoring Matrices; SVM: Support Vector Machine

Introduction

Taenia ovis are the smallest nematode parasite of sheep, are responsible for ovine cysticercosis (Sheep Measles), have an unusual life cycle, and are one of the most widespread and clinically important parasites in the world [1,2]. The small adult worms mature in the intestines of an intermediate host, such as a dog [1,2]. *Taenia ovis* antigen peptides are most suitable for subunit vaccine development, because with single epitope the immune response can be generated in a large population. This approach is based on the phenomenon of cross-protection, whereby infected with a mild strain and is protected against a more severe strain of the same. The phenotype of the resistant transgenic hosts includes fewer centers of initial infection, a delay in symptom development and low accumulation. Antigen protein from *Taenia ovis* is necessary for new paradigm of synthetic vaccine development and target validation [3-5].

Pathogen Transmission

The sheep ingests an egg. The egg hatches in the small intestine and the larval tapeworm burrows through the intestinal wall, and travels to the heart and muscles via the blood. The cysticercus develops in the cardiac and skeletal muscles, reaching the infective stage in about 46 days. When the dog eats the sheep and ingests the cysticercus, the protoscolex attaches to the small intestinal wall and the worm begins to form proglottids, and the lifecycle continues.

Methodology

In this research work, antigenic epitopes of antigen protein from *Taenia ovis* is determined using the Kyte and Doolittle [6], Bull and Breese [7], Parker et al. [8], Chothia [9], Hopp and Woods [10], Welling et al. [11], Manavalan and Ponnuswamy [12], Gomase et al. [13], hydrophobicity scale and Deleage and Roux, Chou and Fasman, Levitt (parameters) have used to predict the probability that a given sequence of amino acids would form a beta strand in antigenic epitopes [6-14]. The Major Histocompatibility Complex (MHC) peptide binding of antigen protein is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted, MHC/peptide binding of antigen protein is a log-transformed value related to the IC_{50} values in nM units. MHC2 predicts peptide binders to MHCI and MHCII molecules from protein sequences or sequence alignments, using Position Specific Scoring Matrices (PSSMs). Support Vector Machine (SVM) based method for prediction of promiscuous MHC class II binding peptides; SVM has been trained on the binary input of single amino acid sequence [15-20]. In addition, we predict those MHC ligands from whose C-terminal end is likely to be the result of proteosomal cleavage [21-25].

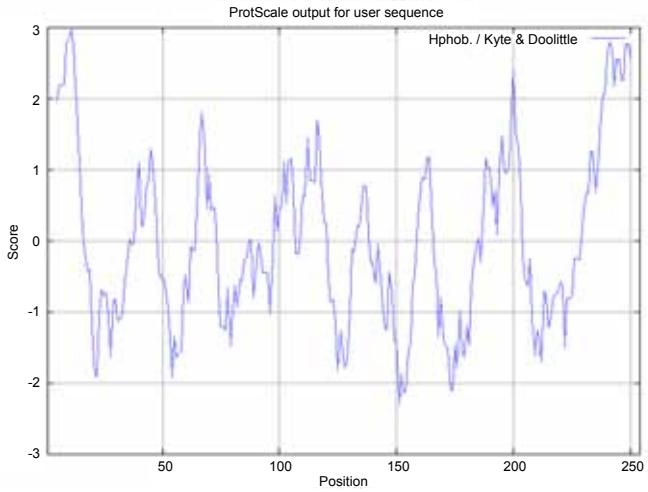

Figure 1: Hydrophobicity plot of antigen protein by Hphob/Kyte & Doolittle scale.

***Corresponding author: Virendra Gomase**, The Global Open University, Nagaland, India, E-mail: gomase.viren@gmail.com

Received July 30, 2012; **Accepted** October 25, 2012; **Published** November 02, 2012

Citation: Gomase V, Chitlange NR (2012) Microbial Proteomics Approach for Sensitive Quantitative Predictions of MHC Binding Peptide from *Taenia ovis*. J Data Mining Genomics Proteomics 3:121. doi:10.4172/2153-0602.1000121

MHC-I	POS.	N	Sequence	C	MW Da)	Score	% OPT.
8mer_H2_Db	129	DTD	PMQNCFIW	GPV	997.23	16.483	31.40 %
8mer_H2_Db	74	TSD	LSNTKTTY	AEL	908.99	9.998	19.05 %
8mer_H2_Db	191	VDG	LVPDTLYI	VTL	915.1	9.447	18.00 %
8mer_H2_Db	40	FTW	GPVFSEFI	GLN	877.02	7.125	13.57 %
8mer_H2_Db	49	FIG	LNWNKDAF	HDA	966.09	5.692	10.84 %
8mer_H2_Db	87	LGD	GSATLDEL	TPN	786.84	5.432	10.35 %
8mer_H2_Db	2	M	ASQLCLIL	LAT	842.07	5.367	10.22 %
8mer_H2_Db	32	RHQ	SLRDIFTW	GPV	996.17	3.894	7.42 %
8mer_H2_Db	230	ATV	VTTSGSAI	VSA	716.78	3.492	6.65 %
8mer_H2_Db	16	AVL	ASDYKDTI	ERT	893.95	1.852	3.53 %
8mer_H2_Db	93	TLD	ELTPNATY	LVT	889.96	1.822	3.47 %
8mer_H2_Db	26	IER	TVARHQSL	RDI	893.01	1.469	2.80 %
8mer_H2_Db	153	QLD	PEDTHDMI	VTL	939.01	0.403	0.77 %
8mer_H2_Db	113	GNT	ILALSSTI	HTP	798.98	-0.095	-0.18 %
8mer_H2_Db	239	AIV	SAILGLLL	TCM	781.01	-0.337	-0.64 %
9mer_H2_Db	48	EFI	GLNWNKDAF	HDA	1023.14	23.089	45.84 %
9mer_H2_Db	93	TLD	ELTPNATYL	VTA	1003.12	14.322	28.44 %
10mer_H2_Db	72	VLT	SDLSNTKTTY	AEL	1111.16	21.688	36.85 %
10mer_H2_Db	92	ATL	DELTPNATYL	VTA	1118.21	16.58	28.17 %
10mer_H2_Db	125	TPA	NDTDPMQNCF	IWG	1166.24	10.033	17.05 %
10mer_H2_Db	189	VAV	DGLVPDTLYI	VTL	1087.24	10.023	17.03 %
10mer_H2_Db	202	VTL	TVLKDGRQFF	NST	1192.38	7.674	13.04 %
10mer_H2_Db	14	ATA	VLASDYKDTI	ERT	1106.24	7.079	12.03 %
10mer_H2_Db	75	SDL	SNTKTTYAEL	GDG	1109.19	5.956	10.12 %
10mer_H2_Db	105	VTA	TANISGNTIL	ALS	985.09	2.588	4.40 %
10mer_H2_Db	111	ISG	NTILALSSTI	HTP	1014.18	2.045	3.47 %
10mer_H2_Db	107	ATA	NISGNTILAL	SST	997.15	1.914	3.25 %
10mer_H2_Db	82	TTY	AELGDGSATL	DEL	914.97	1.723	2.93 %
10mer_H2_Db	233	VTT	SGSAIVSAIL	GLL	899.06	0.415	0.71 %
10mer_H2_Db	164	VTL	TAETASKPRV	ERS	1041.17	0.185	0.31 %
10mer_H2_Db	47	SEF	IGLNWNKDAF	HDA	1136.3	-0.387	-0.66 %
10mer_H2_Db	37	RDI	FTWGPVFSEF	IGL	1175.35	-0.697	-1.18 %
10mer_H2_Db	239	AIV	SAILGLLLTC	MAL	985.25	-0.957	-1.63 %
10mer_H2_Db	85	AEL	GDGSATLDEL	TPN	958.98	-1.594	-2.71 %
10mer_H2_Db	235	TSG	SAIVSAILGL	LLT	925.14	-1.977	-3.36 %
11mer_H2_Db	234	TTS	GSAIVSAILGL	LLT	982.19	9.022	11.35 %
11mer_H2_Db	227	HKE	ATVVTTSGSAI	VSA	988.09	1.554	1.95 %
11mer_H2_Db	39	IFT	WGPVFSEFIGL	NWN	1210.44	0.512	0.64 %
11mer_H2_Db	201	IVT	LTVLKDGRQFF	NST	1305.54	-0.221	-0.28 %
11mer_H2_Db	106	TAT	ANISGNTILAL	SST	1068.23	-1.84	-2.31 %
11mer_H2_Db	46	FSE	FIGLNWNKDAF	HDA	1283.48	-2.028	-2.55 %
11mer_H2_Db	188	EVA	VDGLVPDTLYI	VTL	1186.37	-2.329	-2.93 %
11mer_H2_Db	235	TSG	SAIVSAILGLL	LTC	1038.3	-2.982	-3.75 %
11mer_H2_Db	236	SGS	AIVSAILGLLL	TCM	1064.38	-3.416	-4.30 %
11mer_H2_Db	5	ASQ	LCLILLATAVL	ASD	1124.49	-3.72	-4.68 %
11mer_H2_Db	171	ASK	PRVERSESARF	TRG	1315.48	-4.096	-5.15 %

Table 1: PSSM based prediction of MHC ligands, from whose C-terminal ends are proteosomal cleavage sites.

Results and Interpretations

Binding of peptides to a number of different alleles using Position Specific Scoring Matrix have been found through this study. An antigen protein sequence is 254 residues long having antigenic MHC binding peptides. MHC molecules are cell surface glycoproteins, which take active part in host immune reactions and involvement of MHC class-I and MHC II in response to almost all antigens. PSSM based server predict the peptide binders to MHCI molecules of antigen protein sequence which are as 11mer_H2_Db, 10mer_H2_Db, 9mer_H2_Db, 8mer_H2_Db, and also peptide binders to MHCII molecules of antigen protein sequence as I_Ab.p, I_Ad.p; analysis found antigenic epitopes region in putative antigen protein (Table 1). Additionally, SVM based MHCII-IAb peptide regions were also found; MHCII-IAd peptide regions; MHCII-IAg7 peptide regions and MHCII- RT1.B peptide regions were also found, which represented predicted binders from bacterial antigen protein (Table 2). The predicted binding affinity is normalized by the 1% fractal. Through this study, an improved method for predicting linear epitopes has been described (Table 2). The region of maximal hydrophilicity is likely to be an antigenic site, having hydrophobic characteristics (Figure 1-4), because terminal regions of antigen protein is solvent accessible and unstructured; antibodies against those regions are also likely to recognize the native protein (Figure 5-7). It was shown that an antigen protein is hydrophobic in

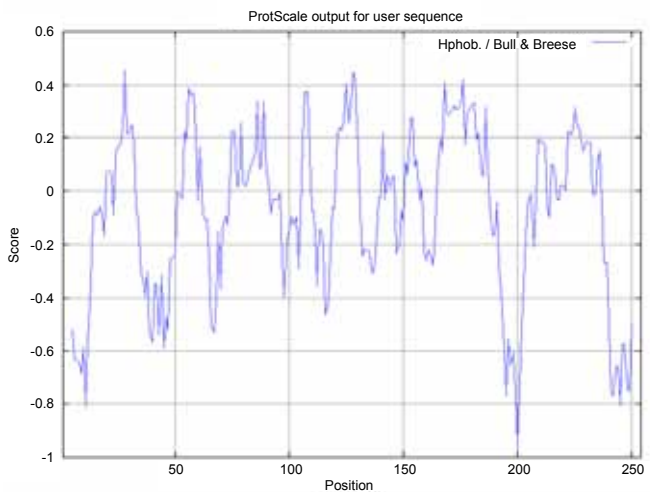

Figure 2: Hydrophobicity plot of antigen protein by Hphob/Bull & Breeze scale.

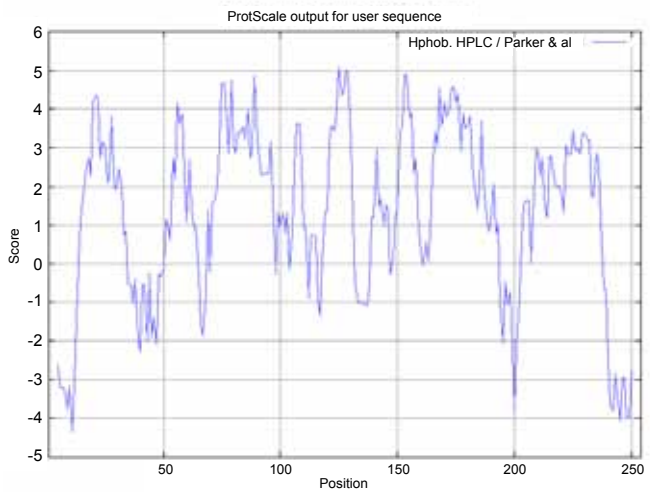

Figure 3: Hydrophobicity plot of antigen protein by Hphob. HPLC/Parker & et al., scale.

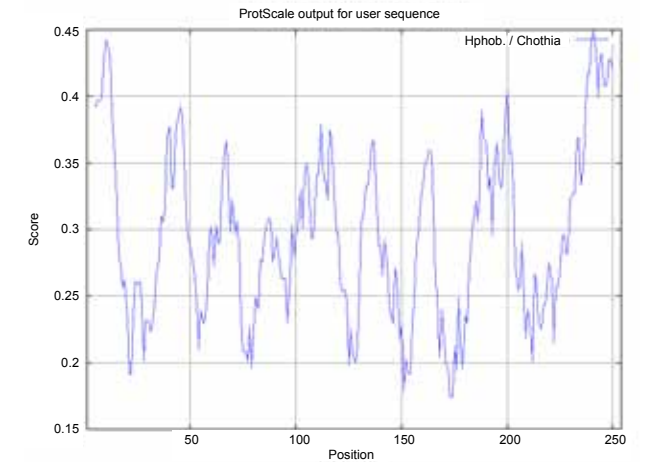

Figure 4: Hydrophobicity plot of antigen protein by Hphob./Chothia scale.

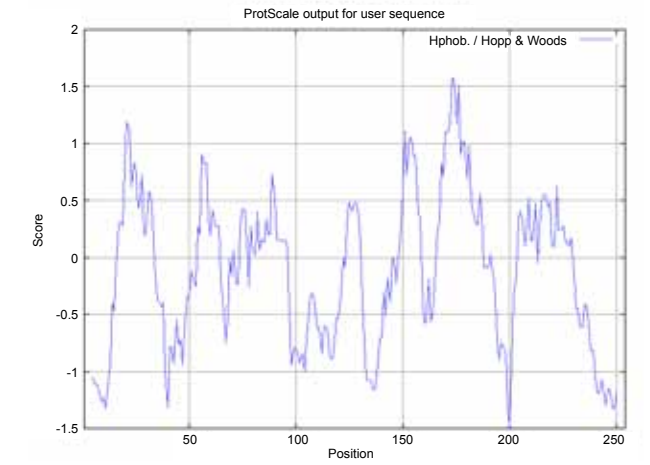

Figure 5: Hydrophobicity plot of antigen protein by Hphob./Hopp & Woods scale.

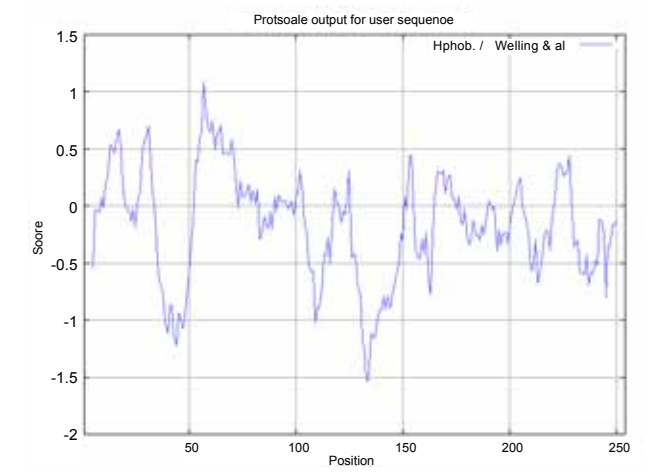

Figure 6: Hydrophobicity plot of antigen protein by Hphob./Welling scale.

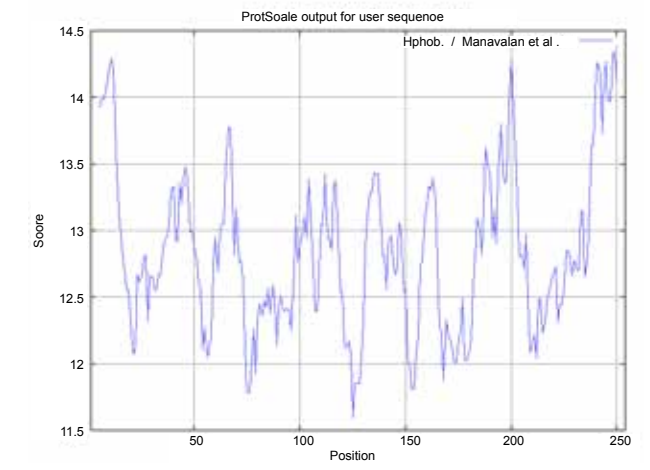

Figure 7: Hydrophobicity plot of antigen protein by Hphob./Manavalan et al. scale.

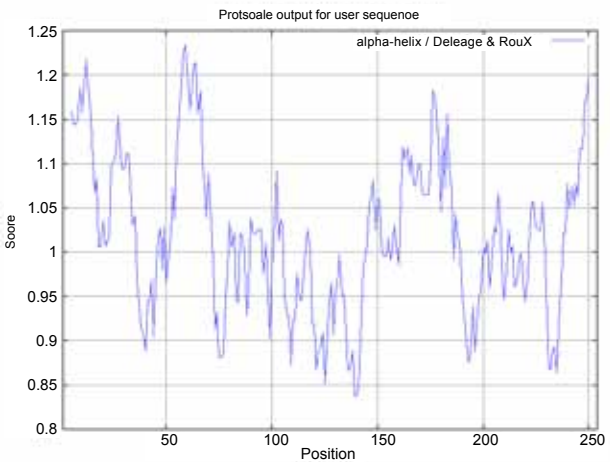

Figure 8: Antigenicity plot of antigen protein by alpha-helix/Deleage & Rouxscale.

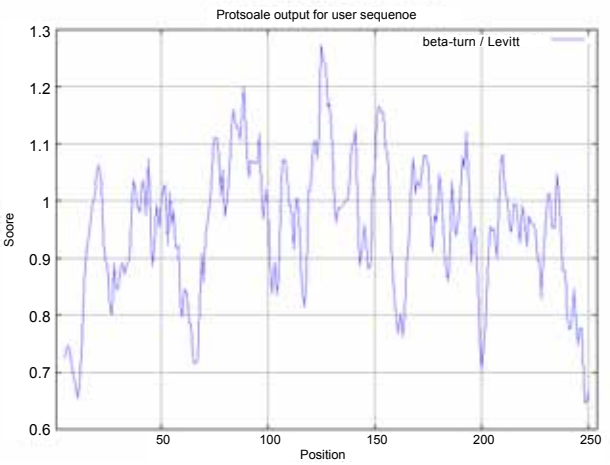

Figure 9: Antigenicity plot of antigen protein by beta-turn/Levitt scale.

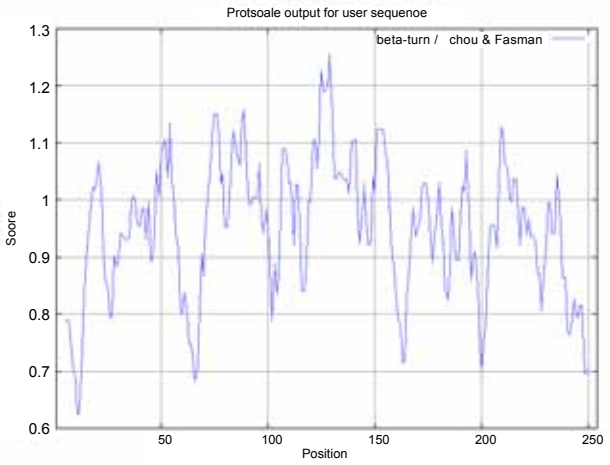

Figure 10: Antigenicity plot of antigen protein by beta-turn/Chou & Fasman scale.

MHC ALLELE	Rank	Sequence	Residue No.	Peptide Score
I-Ab	1	TASKPRVER	167	0.987
I-Ab	2	TLTAETASK	162	0.856
I-Ab	3	NATYLVTAT	97	0.822
I-Ab	4	RTLXTGHKE	218	0.762
I-Ad	1	MASQLCLIL	1	0.703
I-Ad	2	GSAIVSAIL	234	0.636
I-Ad	3	NATYLVTAT	97	0.632
I-Ad	4	LATAVLASD	10	0.622
I-Ag7	1	FHDAEHEVL	56	1.673
I-Ag7	2	TKTTYAELG	77	1.614
I-Ag7	3	NKDAFHDAE	52	1.591
I-Ag7	4	YLVTATANI	100	1.587
RT1.B	1	TKTTYAELG	77	1.136
RT1.B	2	TTSGSAIVS	231	0.961
RT1.B	3	NTKTTYAEL	76	0.890
RT1.B	4	DGSATLDEL	86	0.836

Table 2: SVM based prediction of promiscuous MHC class II binding peptides from antigen protein.

nature and contains segments of low complexity and high-predicted flexibility (Figure 8-10). Predicted antigenic fragments can bind to MHC molecule, and is the first bottlenecks in vaccine design (Figure 1-4).

Conclusion

An antigen protein from *Taenia ovis* peptide nonamers are from a set of aligned peptides known to bind to a given MHC molecule as the predictor of MHC-peptide binding. MHCII molecules bind peptides in similar yet different modes, and alignments of MHCII-ligands were obtained to be consistent with the binding mode of the peptides to their MHC class; this means the increase in affinity of MHC binding peptides may result in enhancement of immunogenicity of antigen protein. These predictions of antigen protein, antigenic peptides to MHC class molecules are important in vaccine development from *Taenia ovis*.

References

1. Jabbar A, Kyngdon CT, Gauci CG, Walduck AK, McCowan C, et al. (2010) Localisation of three host-protective oncospheral antigens of *Taenia ovis*. Int J Parasitol 40: 579-589.
2. Schoenian S (2009) Tapeworms: Problem or Not?
3. Dadley-Moore DL, Lightowlers MW, Rothel JS, Jackson DC (1999) Synthetic peptide antigens induce antibodies to *Taenia ovis* oncospheres. Vaccine 17: 1506-1515.
4. McDonald D, Stockwin L, Matzow T, Blair Zajdel ME, Blair GE (1999) Coxsackie and adenovirus receptor (CAR)-dependent and major histocompatibility complex (MHC) class I-independent uptake of recombinant adenoviruses into human tumour cells. Gene Ther 6: 1512-1519.
5. Gomase VS, Kale KV, Shyamkumar K (2008) Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV). J Proteomics Bioinform 1: 188-205.
6. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132.
7. 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.
8. Parker JM, Guo D, Hodges RS (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 25: 5425-5432.
9. Chothia C (1976) The nature of the accessible and buried surfaces in proteins. J Mol Biol 105: 1-12.

10. Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A 78: 3824-3828.

11. Welling GW, Weijer WJ, van der Zee R, Welling-Wester S (1985) Prediction of sequential antigenic regions in proteins. FEBS Lett 188: 215-218.

12. Manavalan P, Ponnuswamy PK (1978) Hydrophobic character of amino acid residues in globular proteins. Nature 275: 673-674.

13. 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: 117-215.

14. Gomase VS, Kale KV (2008) In silico prediction of epitopes: a new approach for fragment based viral peptide vaccines. Int J of Applied Computing 1: 39-46.

15. Gomase VS, Kale KV (2008) Approach of proteomics system architecture in plant virus's database. Int J of Applied Computing 1: 33-38.

16. Gomase VS, Kapoor RA, Ladak SS (2010) *Eimeria acervulina* analysis for binding peptides using protein profiling for target validation. International Journal of Machine Intelligence 2: 01-08.

17. Gomase VS, Kale KV, Shyamkumar K, Shankar S (2008) Computer Aided Multi Parameter Antigen Design: Impact of Synthetic Peptide Vaccines from Soybean Mosaic Virus. 1st International Conference on Emerging Trends in Engineering and Technology, ICETET '08, IEEE Xplore, Los Alamitos, California.

18. Gomase VS, Tandale JP, Patil SA, Kale KV (2006) Automatic modelling of protein 3D structure Nucleoplasmin-like viral coat protein from Cucumber mosaic virus. International Conference on Advanced Computing & Communication, ADCOM, IEEE Xplore, USA.

19. Reche PA, Glutting JP, Reinherz EL (2002) Prediction of mhc class i binding peptides using profile motifs. Hum Immun 63: 701-709.

20. Buus S, Lauemøller SL, Worning P, Kesmir C, Frimurer T, et al. (2003) Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens 62: 378-384.

21. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, et al. (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12: 1007-1017.

22. Bhasin M, Raghava GPS (2005) Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. Nucleic Acids Res 33: W202-W207.

23. Gomase VS, Kapoor RA, Ladak SS (2010) Immuno-proteomics approach for synthetic vaccine development form *Haemophilus influenzae*. Journal of Infectious Diseases Letters 1: 01-06.

24. Gomase VS, Shyamkumar K (2009) Prediction of antigenic epitopes and MHC binders of neurotoxin alpha-KTx 3.8 from Mesobuthus tamulus sindicus. African Journal of Biotechnology 8: 6658-6676.

25. Gomase VS, Kapoor RA, Ladak SS (2010) Computer intelligence approach for prediction of binding ability and fragment based peptide vaccines from *Leishmania* protozoa peptides. International Conference on Software and Computing Technology, ICSCT, Kunming, China.