Machine Intelligence Approach for Development of MHC Binders and Fragment Based Peptide Vaccines from *Tetrahymena pyriformis*

V.S. Gomase and N.R. Chitlange

School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413531, MS, India

Abstract: Tetrahymena pyriformis are free-living ciliate protozoa that Guppies are easily susceptible to Tetrahymena. Peptide fragments of protein antigen can be used to select nonamers for use in rational vaccine design and to increase the understanding of roles of the immune system in parasitic diseases. Analysis shows MHC class II binding peptides of protein antigen from *Tetrahymena pyriformis* are important determinant for protection of host form parasitic infection. In this assay, we used PSSM and SVM algorithms for antigen design and predicted the binding affinity of protein antigen having 389 amino acids, which shows 381 nonamers. Binding ability prediction of antigenic peptides to major histocompatibility complex (MHC) class I and II molecules is important in vaccine development from *Tetrahymena pyriformis*.

INTRODUCTION

Tetrahymena pyriformis are free-living ciliate protozoa that can switch from commensalistic to pathogenic modes of survival. They are common in fresh-water. Guppies do succumb to the disease quite easily and it can cause big problems where guppies are produced commercially. [1, 2]. *Tetrahymena pyriformis* bacterial peptides are most suitable for subunit vaccine development because with single epitope, the immune response can be strongly generated. This approach is based on the phenomenon of cross-protection, whereby a host infected with a mild strain of pathogen is protected against a more severe strain of the same pathogen. The phenotype of the resistant transgenic hosts includes fewer centers of initial pathogenic infection, a delay in symptoms development and low pathogenic accumulation. Protein antigen from *Tetrahymena pyriformis* is necessary for new paradigm of synthetic vaccine development and target validation [3-5].

METHODOLOGY

In this research work antigenic epitopes of protein antigen from *Tetrahymena pyriformis* are determined using the Gomase in 2007, Hopp and Woods, Welling, Parker and Protrusion Index (Thornton) antigenicity [6-8].

The major histocompatibility complex (MHC) peptide binding of protein antigen is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding of protein antigen is a log-transformed value related to the IC50 values in nM units. RANKPEP 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 is used for prediction of promiscuous MHC class II binding peptides. SVM has been trained on the binary input of single amino acid sequence [9-14]. In addition, we predict those MHC ligands from whose C-terminal end is likely to be the result of proteosomal cleavage [15].

RESULTS AND INTERPRETATIONS

We found binding of peptides to a number of different alleles using Position Specific Scoring Matrix. The sequence of protein antigen is 389 residues long, having antigenic MHC binding peptides. MHC molecules are cell surface glycoproteins, which take active part in host immune reactions. Both MHC class-I and MHC II

Corresponding Author: V.S. Gomase, School of Technology, S.R.T.M. University,
Sub-Centre, Latur, 413531, MS, India, Tel: +9987770696, E-mail: gomase.viren@gmail.com.

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

MHC-I	POS.	N	Sequence	C	MW (Da)	Score	% OPT.
8mer_H2_Db	62	DGF	SQWDYLFY	TIP	1080.21	23.964	45.65 %
8mer_H2_Db	308	NCL	EGANCVTI	IEA	787.88	23.827	45.39 %
8mer_H2_Db	368	IVD	ISSNPFII	NKD	872.04	18.998	36.19 %
8mer_H2_Db	73	TIP	GIINFQMY	GIP	967.15	17.925	34.15 %
8mer_H2_Db	54	KYV	QHWNGDGF	SQW	918.95	14.902	28.39 %
8mer_H2_Db	216	DFY	IPQGGVFY	DFY	862.0	14.787	28.17 %
8mer_H2_Db	286	NLY	ASGSILTI	NDY	742.87	12.549	23.91 %
8mer_H2_Db	322	QGK	FENNSFSI	SLK	939.0	10.901	20.77 %
9mer_H2_Db	53	GKY	VQHWNGDGF	SQW	1018.08	21.592	42.87 %
9mer_H2_Db	294	LTI	NDYNNDENI	IRN	1092.04	21.316	42.32 %
9mer_H2_Db	307	RNC	LEGANCVTI	IEA	901.04	18.755	37.24 %
9mer_H2_Db	32	AAK	LMGRNLTFI	LSR	1046.29	18.451	36.63 %
9mer_H2_Db	268	LRS	RFLDNHFTL	YIA	1144.3	15.756	31.28 %
9mer_H2_Db	2	M	LHIPNYNLT	QYD	1066.22	15.734	31.24 %
9mer_H2_Db	173	PLM	WSFPNDDEA	LNY	1039.07	14.989	29.76 %
9mer_H2_Db	12	LTQ	YDMHNLNGF	GES	1092.19	13.152	26.11 %
10mer_H2_Db	72	YTI	PGIINFQMYG	IPF	1121.32	21.156	35.94 %
10mer_H2_Db	353	VII	AGISTFGNYI	SKI	1024.14	19.66	33.40 %
10mer_H2_Db	356	AGI	STFGNYISKI	VDI	1111.26	18.805	31.95 %
10mer_H2_Db	210	TNT	TGLDFYIPQG	GVF	1092.22	13.971	23.74 %
10mer_H2_Db	294	LTI	NDYNNDENII	RNC	1205.2	13.88	23.58 %
10mer_H2_Db	200	APV	VAQGNEVTNT	TGL	1014.04	13.396	22.76 %
10mer_H2_Db	125	SIS	QEPYAFPDYH	YVL	1248.34	13.331	22.65 %
10mer_H2_Db	62	DGF	SQWDYLFYTI	PGI	1294.47	11.829	20.10 %
11mer_H2_Db	72	YTI	PGIINFQMYGI	PFV	1234.48	34.146	42.95 %
11mer_H2_Db	141	SSS	IKTLNVRYQLL	KFY	1342.64	20.815	26.18 %
11mer_H2_Db	372	SSN	PFIINKDKQQY	DIN	1375.59	16.13	20.29 %
11mer_H2_Db	90	DDI	CGLNGNATPEL	CAR	1070.18	12.017	15.12 %
11mer_H2_Db	179	PND	DEALNYEAEFM	LGD	1313.42	11.857	14.92 %
11mer_H2_Db	293	ILT	INDYNNDENII	RNC	1318.36	11.253	14.16 %
11mer_H2_Db	355	IAG	ISTFGNYISKI	VDI	1224.42	10.852	13.65 %
11mer_H2_Db	148	NVR	YQLLKFYYHLF	VKE	1516.82	9.687	12.19 %

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

ALLELE	Sequence	Residue No	Peptide Score
I-Ab	GTIFRPLMW	165	1.104
I-Ab	ISLKNTQEQ	329	0.870
I-Ab	AKLMGRNLT	30	0.861
I-Ab	RNLTFILSR	35	0.844
I-Ad	KEQGSGTIF	160	0.942
I-Ad	AEFMLGDYL	186	0.897
I-Ad	LGSLYPFSR	107	0.677
I-Ad	AAKLMGRNL	29	0.666
I-Ag7	GESIATYEA	21	1.675
I-Ag7	LNGNATPEL	92	1.616
I-Ag7	SISQEPYAF	122	1.612
I-Ag7	IATYEAAKL	24	1.519
RT1.B	NTQEQKNTQ	333	0.976
RT1.B	ATYEAAKLM	25	0.904
RT1.B	DFYNQQRYT	224	0.877
RT1.B	TNTTGLDFY	207	0.813

molecules are involved in the response to almost all antigens. PSSM based server predicts the peptide binders to MHCI molecules of protein antigen sequence are as 11mer_H2_Db,10mer_H2_Db,9mer_H2_Db,8mer_H2_Db.

Also predicts peptide binders to MHCII molecules of antigen protein sequence as I_Ab.p, I_Ad.p. Analysis found antigenic epitopes region in putative protein antigen (Table 1). We also found the SVM based MHCII-IAb peptide regions; MHCII-IAd peptide regions; MHCII-IAg7 peptide regions and MHCII- RT1.B peptide regions, which represented predicted binders from parasitic protein antigen (Table 2). The predicted binding affinity is normalized by the 1% fractil. We describe an improved method for predicting linear epitopes (Table 2). The region of maximal hydrophobicity is likely to be an antigenic site, having hydrophobic characteristics, because terminal regions of antigen protein is solvent accessible and unstructured. Antibodies against those regions are
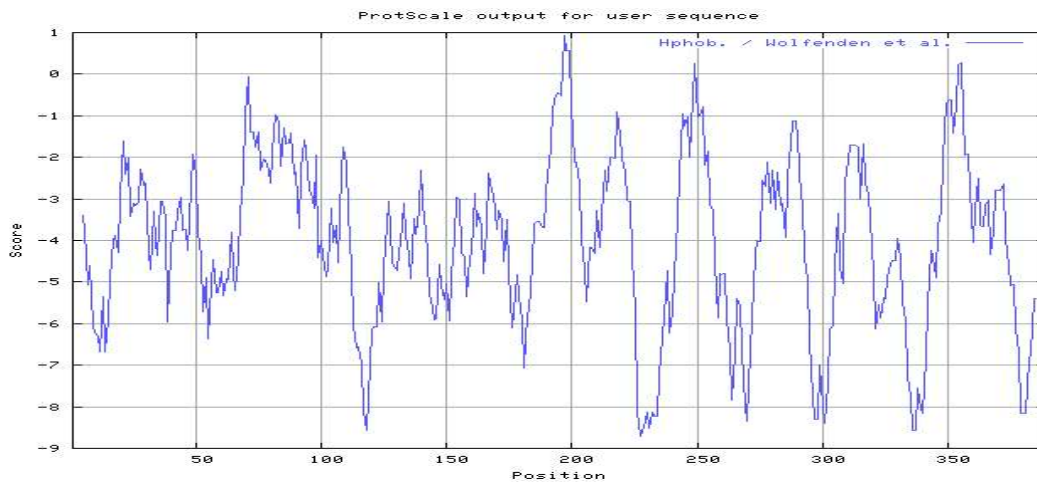
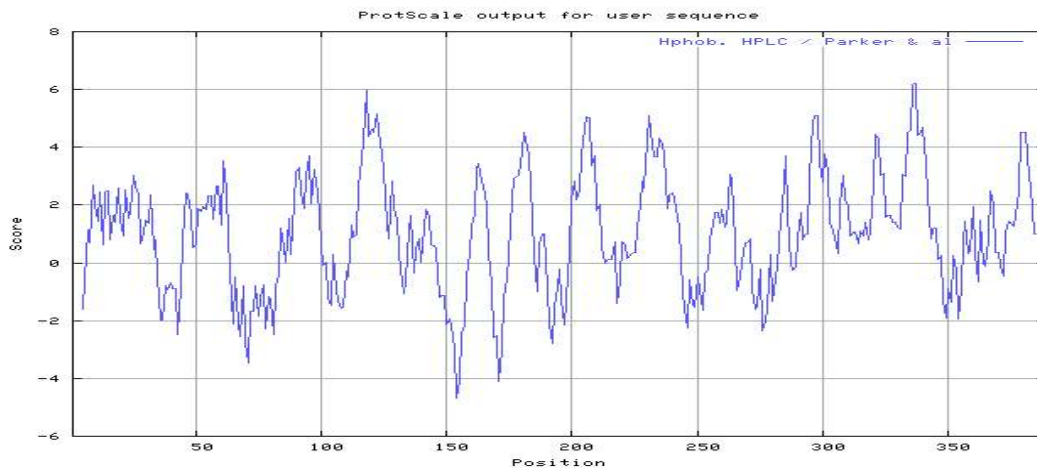Fig. 1- Antigenicity plot of antigen protein by Wolfenden, et al., scale

Fig. 2- Antigenicity plot of antigen protein by HPLC / Parker, et al., scale
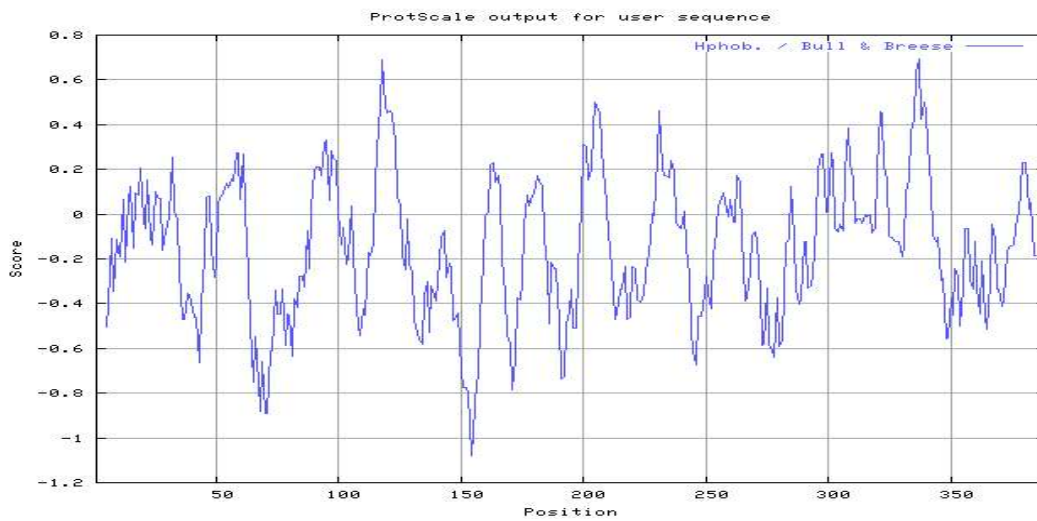

Fig. 3- Hydrophobicity plot of antigen protein by Bull &Breese, et al., scale
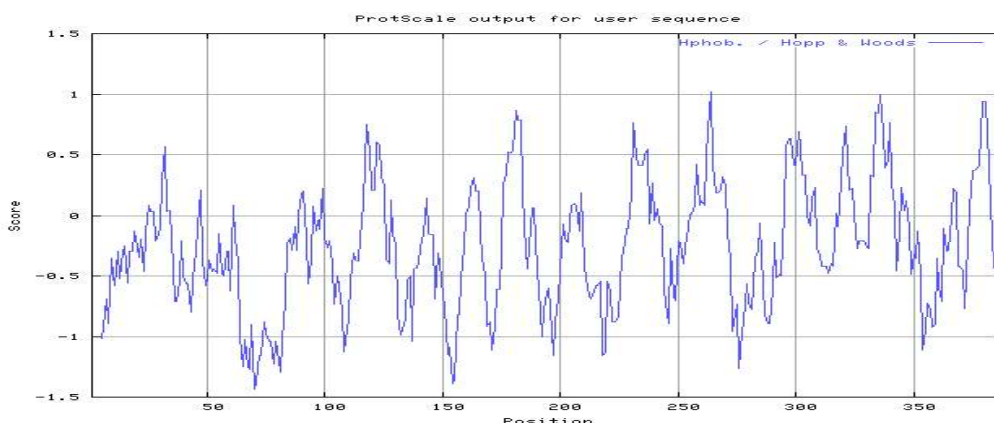
Fig. 4: Hydrophobicity plot of antigen protein by Hopp and Woods scale

also likely to recognize the native protein (Fig. 1, 2,where is fig. 5). It was shown that protein antigen is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility (Fig. 3, 4). Predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

CONCLUSION

A protein antigen from *Tetrahymena pyriformis* peptide nonamers is 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 parasitic antigen protein. These predicted bacterial protein antigenic peptides to MHC class molecules are important in vaccine development from Tetrahymena pyriformis.

REFERENCES

1. Alfred M. Elliott, 1973. Biology of Tetrahymena, Dowden, Hutchinson and Ross, ISBN 0-87933-013-9.
2. Turkewitz, A.P., E. Orias and G., Kapler, 2002. Functional Genomics: The coming of age for Tetrahymena thermophila. Trends in Genetics, 18: 35-4.
3. Joshi, A., J. Tang, M. Kuzma, J. Wagner, B. Mookerjee J. Filicko, M. Carabasi, N. Flomenberg and P. Flomenberg, 2009. Adenovirus DNA polymerase is recognized by human CD8+ T cells. J. Gen Virol., 90(Pt 1): 84-94..
4. McDonald D., L. Stockwin, T. Matzow, M.E. Blair Zajdel, G.E. Blair, 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(9):1512-9.
5. Gomase, V.S., K.V. Kale and K. Shyamkumar, 2008. Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV). J. of Proteomics and Bioinformat., 1(4): 188- 205.
6. Gomase, V.S., K.V. Kale, N.J. Chikhale and S.S. Changbhale, 2007. Prediction of MHC Binding Peptides and Epitopes from Alfalfa mosaic virus. Curr. Drug Discov. Technol., 4(2): 117-1215.
7. Gomase, V.S. and K.V. Kale, 2008. In silico prediction of epitopes: a new approach for fragment based viral peptide vaccines. Int. J. of Applied Computing, 1(1): 39-46..
8. Gomase, V.S. and K.V. Kale, 2008. Approach of proteomics system architecture in plant virus's database. Int. J. of Applied Computing, 1(1): 33-38.
9. Gomase, V.S. and K.Shyamkumar, 2009. Prediction of antigenic epitopes and MHC binders of neurotoxin alpha-KTx 3.8 from Mesobuthus tamulus sindicus. African Journal of Biotechnol., 8(23): 6658-6676.
10. Gomase, V.S., K.V. Kale, K. Shyamkumar and S. Shankar, 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, pp: 629-634.

11. Gomase, V.S., J.P. Tandale, S.A. Patil and K.V. Kale, 2006. Automatic modeling of protein 3D structure Nucleoplasmin-like viral coat protein from Cucumber mosaic virus. 14th International Conference on Advance Computing and Communication, Published by IEEE Computer Society in IEEE Xplore USA, pp: 614-615.

12. Reche, P.A., J.P. Glutting and E.L. Reinherz, 2002. Prediction of MHC Class I Binding Peptides Using Profile Motifs. Hum Immun., 63: 701-709.

13. Buus S., *et al*., 2003. Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens, 62: 378-384.

14. Nielsen, M., *et al*. 2005. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci., 12: 1007-10170

15. Bhasin, M. and G.P. Raghava, 2005. Pcleavage: An SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. Nucleic Acids Research, 33: 202-207.