Immunoproteomics approach for development of synthetic peptide vaccine from *Mycobacterium tuberculosis*

Gomase V.S.* and Chitlange N.R.
School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413531, MS, India,
Mobile- 09987770696, Mail- gomase.viren@gmail.com

Abstract- *Mycobacterium tuberculosis* causes tuberculosis leading to be an obligatory step in infection. Peptide fragments of 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 bacterial diseases. Analysis shows MHC class II binding peptides of antigen from *Mycobacterium tuberculosis* are important determinant for protection of host form tuberculosis infection. In this assay, we used PSSM and SVM algorithms for antigen design and predicted the binding affinity of antigen protein having 338 amino acids, which shows 330 nonamers. Binding ability prediction of antigen peptides to major histocompatibility complex (MHC) class I & II molecules is important in vaccine development from *Mycobacterium tuberculosis* antigen.

Keywords- *Mycobacterium tuberculosis*, antigen *n, epitope, PSSM, SVM, MHC, peptide vaccine*

Abbreviations: Goldman, Engelberg and Steitz, (GES); major histocompatibility complex, (MHC); Position Specific Scoring Matrices, (PSSMs); Support Vector Machine, (SVM)

***Correspondence**: Dr Virendra S. Gomase, School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413531, MS, India, Mobile- 09987770696, Mail- gomase.viren@gmail.com

I. Introduction

Mycobacterium tuberculosis (MTB) is a pathogenic bacteria and the causative agent of most cases of tuberculosis. The physiology of M. tuberculosis is highly aerobic and requires high levels of oxygen. Primarily a pathogen of the mammalian respiratory system, MTB infects the lungs, causing tuberculosis [1, 2]. *Mycobacterium tuberculosis* bacterial peptides are most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. This approach is based on the phenomenon of cross-protection, whereby a person infected with a mild strain of bacteria is protected against a more severe strain of the same bacteria. The phenotype of the resistant transgenic hosts includes fewer centers of initial bacterial infection, a delay in symptom development, and low bacterial accumulation. Antigen from *Mycobacterium tuberculosis* is necessary for new paradigm of synthetic vaccine development and target validation [3-5].

II. Methodology

In this research work antigenic epitopes of antigen from *Mycobacterium tuberculosis* is 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 antigen protein is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding of antigen is a log-transformed value related to the IC50 values in nM units. MHC2Pred 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 [9-14]. In addition, we predict those MHC ligands from whose C-terminal end is likely to be the result of proteosomal cleavage [15].

III. Results and Interpretations

We found binding of peptides to a number of different alleles using Position Specific Scoring Matrix. A antigen protein sequence is 338 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 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). 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 antigen protein (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 hydrophilicity 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 also likely to recognize the native protein (Fig. 1, 2). It was shown that a antigen protein 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.

IV. Conclusion

Antigen protein from *Mycobacterium tuberculosis* 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 bacterial antigen protein. These predicted of protein antigenic peptides to MHC class molecules are important in vaccine development from *Mycobacterium tuberculosis*.

V. References

[1] K.J.Ryan, C.G. Ray (editors) (2004). Sherris Medical Microbiology (4th ed.). McGraw Hill. ISBN 0-8385-8529-9.

[2] S.T.Cole, R. Brosch, J. Parkhill, et al., "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence". Nature 393 (6685): 537–44, 1998.

[3] A. Joshi, J. Tang, M. Kuzma, J. Wagner, B. Mookerjee, J. Filicko, M. Carabasi, N. Flomenberg, P. Flomenberg, "Adenovirus DNA polymerase is recognized by human CD8+ T cells". *J. Gen Virol.*, 90(Pt 1), 84-94, 2009.

[4] D. McDonald, L. Stockwin, T. Matzow, M.E. Blair Zajdel, G.E.Blair, "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, 1999.

[5] V.S.Gomase, K.V. Kale and K. Shyamkumar, "Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV)". *J. of Proteomics & Bioinformatics*, 1 (4), 188- 205, 2008.

[6] V.S.Gomase, K.V. Kale, N.J.Chikhale, S.S. Changbhale, "Prediction of MHC Binding Peptides and Epitopes from Alfalfa mosaic virus". *Curr. Drug Discov. Technol*., 4(2),117-1215, 2007.

[7] V.S.Gomase and K.V.Kale,"In silico prediction of epitopes: a new approach for fragment based viral peptide vaccines". *Int. J. of Applied Computing*, 1(1), 39-46, 2008.

[8] V.S.Gomase and K.V.Kale, "Approach of proteomics system architecture in plant virus's database". *Int. J. of Applied Computin*g, 1(1), 33-38, 2008.

[9] V.S.Gomase and K.V.Kale, "Bioinformatics based sequence analysis of Nucleoplasmin like viral coat protein". Int. *J. of Information Retrieval*, 1(1), 11-15, 2008.

[10] V.S.Gomase, K.V. Kale, K. Shyamkumar and S. Shankar, "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*, 629-634, 2008.

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

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

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

[14] M. Nielsen, et al., "Reliable prediction of T-cell epitopes using neural networks with novel sequence representations". *Protein Sci*., 12, 1007-1017, 2003.

[15] M. Bhasin and G.P. Raghava, "Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences". *Nucleic Acids Research*, 33, W202-207, 2005.

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.
9mer_H2_Db	71	QFQ	SGGANSPAL	YLL	754.8	12.502	24.82 %
9mer_H2_Db	290	YNA	GGGHNGVFD	FPD	840.84	12.065	23.96 %
9mer_H2_Db	135	KAG	CQTYKWETF	LTS	1164.33	11.125	22.09 %
9mer_H2_Db	262	PSD	LGGNNLPAK	FLE	864.99	9.957	19.77 %
9mer_H2_Db	291	NAG	GGHNGVFDF	PDS	930.97	7.819	15.52 %
9mer_H2_Db	131	PAC	GKAGCQTYK	WET	937.07	6.399	12.71 %
9mer_H2_Db	254	WVY	CGNGKPSDL	GGN	871.96	5.762	11.44 %
10mer_H2_Db	294	GGH	NGVFDFPDSG	THS	1036.07	21.394	36.35 %
10mer_H2_Db	133	CGK	AGCQTYKWET	FLT	1145.28	16.436	27.92 %
10mer_H2_Db	262	PSD	LGGNNLPAKF	LEG	1012.17	13.535	23.00 %
10mer_H2_Db	71	QFQ	SGGANSPALY	LLD	917.98	11.038	18.75 %
10mer_H2_Db	312	YWG	AQLNAMKPDL	QRA	1082.28	9.195	15.62 %
10mer_H2_Db	107	YDQ	SGLSVVMPVG	GQS	927.12	9.161	15.56 %
10mer_H2_Db	72	FQS	GGANSPALYL	LDG	944.06	8.26	14.03 %
11mer_H2_Db	71	QFQ	SGGANSPALYL	LDG	1031.14	18.377	23.12 %
11mer_H2_Db	262	PSD	LGGNNLPAKFL	EGF	1125.33	10.69	13.45 %
11mer_H2_Db	130	QPA	CGKAGCQTYKW	ETF	1203.42	9.65	12.14 %
11mer_H2_Db	135	KAG	CQTYKWETFLT	SEL	1378.59	8.649	10.88 %
11mer_H2_Db	284	IKF	QDAYNAGGGHN	GVF	1085.05	6.861	8.63 %
11mer_H2_Db	289	AYN	AGGGHNGVFDF	PDS	1059.1	6.338	7.97 %
11mer_H2_Db	302	FPD	SGTHSWEYWGA	QLN	1216.3	5.72	7.20 %

Table 2- SVM based prediction of promiscuous MHC class II binding peptides from capsid protein

ALLELE	Sequence	Residue No	Peptide Score
I-Ab	GATPNTGPA	326	0.918
I-Ab	NAMKPDLQR	315	0.833
I-Ab	PGLPVEYLQ	47	0.787
I-Ab	KLIANNTRV	242	0.741
I-Ad	VVGAVGAAL	19	0.723
I-Ad	VSGLVGAVG	28	0.716
I-Ad	LQRALGATP	321	0.698
I-Ad	YAGAMSGLL	188	0.631
I-Ag7	GGTATAGAF	36	2.021
I-Ag7	MGDAGGYKA	211	1.805
I-Ag7	VVGAVGAAL	19	1.794
I-Ag7	SGGANSPAL	71	1.785
RT1.B	LNAMKPDLQ	314	1.020
RT1.B	MAASSALTL	170	0.813
RT1.B	KWETFLTSE	139	0.763
RT1.B	ANSPALYLL	74	0.711

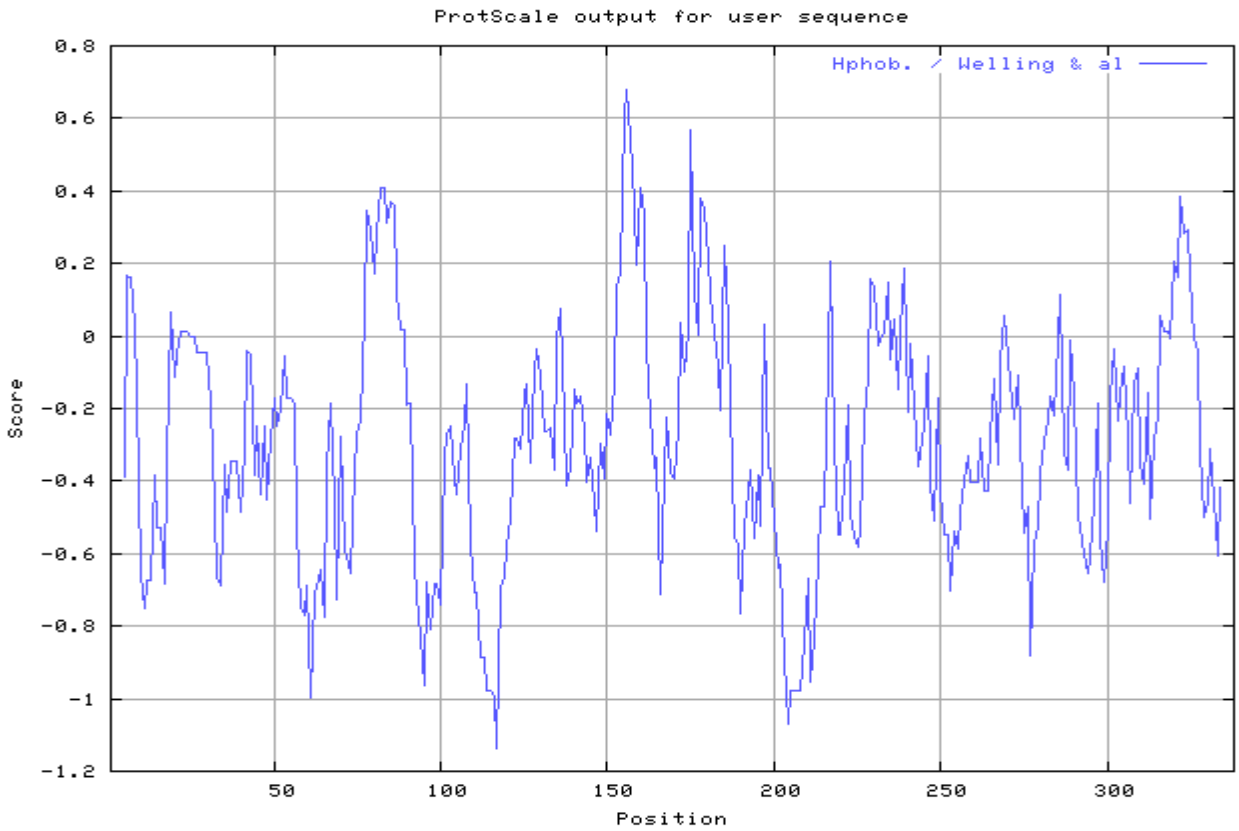

Fig. 1- Antigenicity plot of antigen protein by Welling, et al., scale

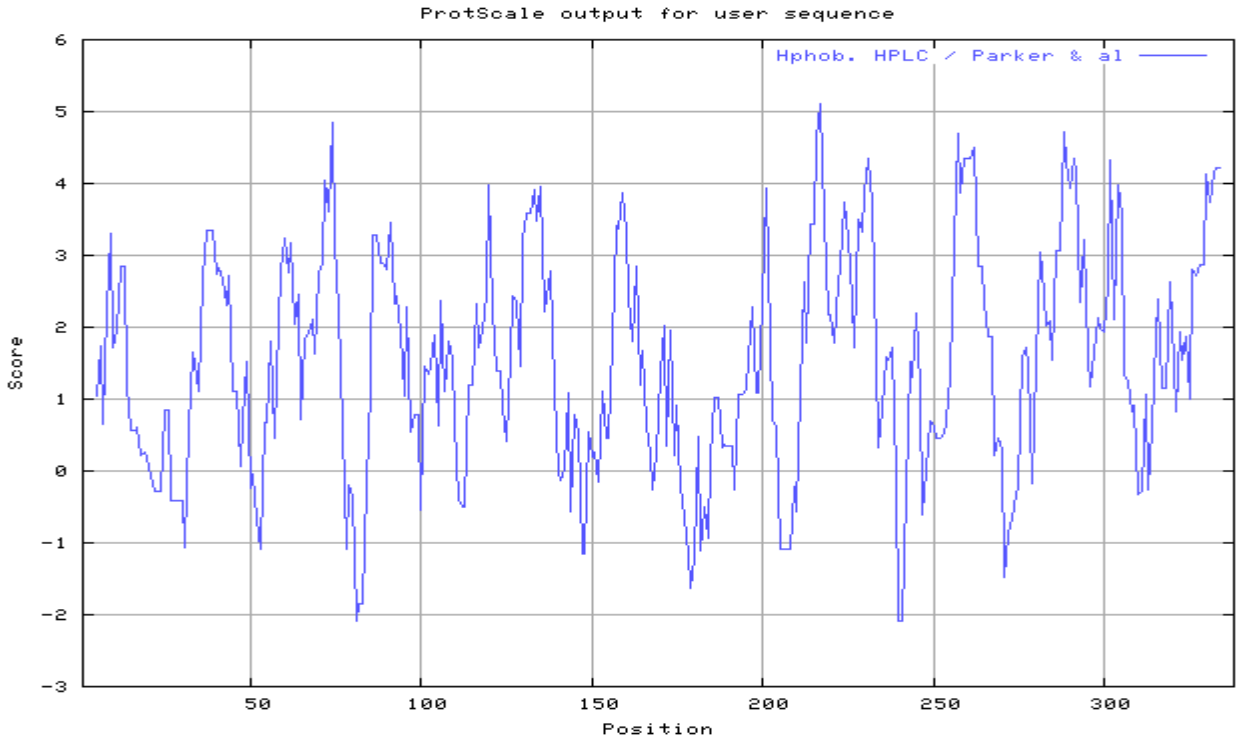

Fig. 2- Antigenicity plot of antigen protein by HPLC / Parker, et al., scale

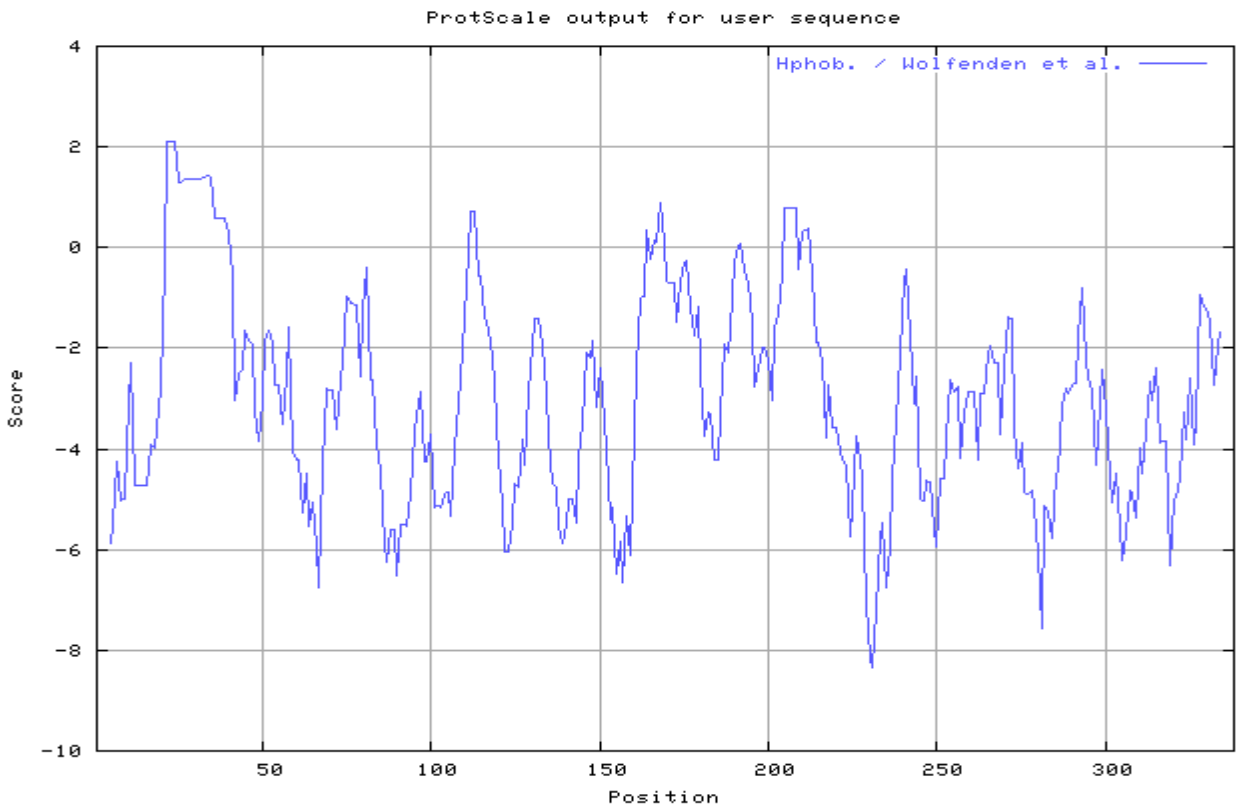

Fig. 3- Hydrophobicity plot of antigen protein by Wolfenden, et al., scale

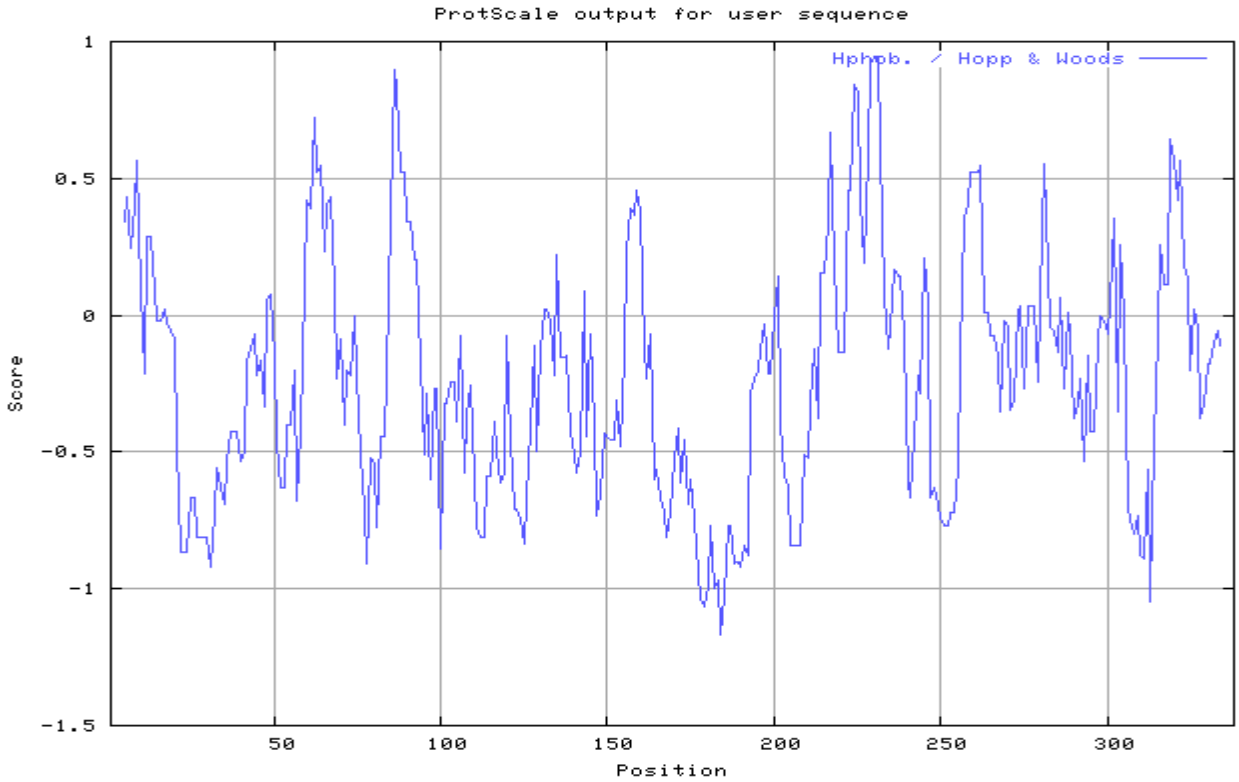

Fig. 4-Hydrophobicity plot of antigen protein by Hopp & Woods scale