Immuno-proteomics approach for synthetic vaccine development form *Haemophilus influenzae*

Gomase V.S.[1], Kapoor R.A.[2], Ladak S.S.[2]
[1]School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413512, India
[2]Padmashree Dr. D.Y. Patil University, Navi Mumbai, 400614, India

Abstract- *Haemophilus influenzae* causes systemic infections leading to be an obligatory step in infection. 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 *Haemophilus influenzae* are important determinant for protection of host form pathogenic infection. In this assay, we used PSSM and SVM algorithms for antigen design and predicted the binding affinity of antigen protein having 575 amino acids, which shows 567 nonamers. Binding ability prediction of antigen peptides to major histocompatibility complex (MHC) class I & II molecules is important in vaccine development from *Haemophilus influenzae*.
Keywords- *antigen protein, 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

Haemophilus influenzae is a human-adapted commensal and pathogen that can cause mucosal infections such as sinusitis, otitis media, and bronchitis. Certain strains also cause bacteremia and meningitis. [1, 2]. *Haemophilus influenzae* 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 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 infection, a delay in symptom development, and low bacterial accumulation. Antigen protein from *Haemophilus influenzae* is necessary for new paradigm of synthetic vaccine development and target validation [3-5].

II. Methodology

In this research work antigenic epitopes of antigen protein from *Haemophilus influenzae* 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 protein 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 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 575 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 bacterial 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, 5). 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

A antigen protein from *Haemophilus influenzae* 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 epitopes of antigenic peptides to MHC class molecules are important in vaccine development from *Haemophilus influenzae*.

V. References

[1] D.A. Daines, L.A. Cohn, H.N. Coleman, K.S. Kim, A.L.Smith. *Haemophilus influenzae* Rd KW20 has virulence properties. Seattle Biomedical Research Institute, Four Nickerson Street, Suite 200, WA 98109, USA.

[2] P.W. Whitby, T.W. Seale, D.J. Morton, T.M. VanWagoner, T.L. Stull. Characterization of the *Haemophilus influenzae* tehB gene and its role in virulence. Microbiology. 2010 Apr;156(Pt 4):1188-200.

[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.
8mer_H2_Db	269	SGD	GQILGTTI	QSG	783.91	19.236	36.64 %
8mer_H2_Db	131	KLS	PSQKSRYY	ETL	1010.13	18.202	34.67 %
8mer_H2_Db	31	NLL	GSNFTQTL	QKD	848.9	14.378	27.39 %
8mer_H2_Db	471	ASA	TNTNTDFI	AQM	906.93	14.134	26.92 %
8mer_H2_Db	408	LAG	TDANIRYY	NLP	997.08	13.29	25.32 %
8mer_H2_Db	265	LLP	LSGDGQIL	GTT	783.88	11.876	22.62 %
8mer_H2_Db	564	TWY	QYQDGAIV	PVA	874.95	11.634	22.16 %
8mer_H2_Db	183	ALL	RSANTGVI	NNA	798.89	11.115	21.17 %
8mer_H2_Db	393	LGQ	RVGNAFNV	RWQ	857.96	10.816	20.60 %
8mer_H2_Db	252	LNF	QQTNVSQI	GLL	898.96	9.935	18.93 %
8mer_H2_Db	387	AMP	QNDLGQRV	GNA	910.98	9.766	18.60 %
8mer_H2_Db	226	QAL	QSWKNAYP	NHA	952.07	9.603	18.29 %
9mer_H2_Db	333	ADP	AQIQGMDVL	ALN	956.12	20.206	40.12 %
9mer_H2_Db	470	SAS	ATNTNTDFI	AQM	978.01	19.925	39.56 %
9mer_H2_Db	320	VGP	LLKQNLDVI	LAD	1037.26	16.685	33.13 %
9mer_H2_Db	230	SWK	NAYPNHAAA	TLF	909.96	15.621	31.02 %
9mer_H2_Db	251	LLN	FQQTNVSQI	GLL	1046.14	15.538	30.85 %
9mer_H2_Db	21	LLS	MALAGCSNL	LGS	861.04	12.713	25.24 %
9mer_H2_Db	344	LAL	NATPNSRAI	PQL	925.01	12.288	24.40 %
9mer_H2_Db	199	EGN	AALGGWLTL	IKA	860.05	9.962	19.78 %
9mer_H2_Db	506	LAK	STGGEYQLM	RLY	967.06	9.675	19.21 %
9mer_H2_Db	143	TLA	IVAENRKDM	IEA	1057.23	9.238	18.34 %
9mer_H2_Db	232	KNA	YPNHAAATL	FPK	939.04	8.855	17.58 %
9mer_H2_Db	425	TYF	VQENNSNTT	ALY	987.96	8.697	17.27 %
10mer_H2_Db	230	SWK	NAYPNHAAAT	LFP	1011.06	18.096	30.75 %
10mer_H2_Db	453	YLT	NIVPNLAIYA	SSR	1069.27	14.738	25.04 %
10mer_H2_Db	89	LRE	LGELNDAQKL	DRA	1082.22	14.207	24.14 %
10mer_H2_Db	188	ANT	GVINNASDEG	NAA	956.96	13.819	23.48 %
10mer_H2_Db	132	LSP	SQKSRYYETL	AIV	1256.39	13.447	22.85 %
10mer_H2_Db	187	SAN	TGVINNASDE	GNA	1001.01	13.288	22.58 %
10mer_H2_Db	412	DAN	IRYYNLPADV	TYF	1205.39	12.929	21.97 %
10mer_H2_Db	442	ASP	TELAEMKGYL	TNI	1136.33	12.675	21.53 %
10mer_H2_Db	395	QRV	GNAFNVRWQQ	LAG	1178.3	11.528	19.59 %
10mer_H2_Db	478	TDF	IAQMNGVQFS	DIP	1076.23	11.428	19.42 %
10mer_H2_Db	206	GWL	TLIKAYNDYI	RQP	1195.38	10.686	18.16 %
10mer_H2_Db	433	SNT	TALYAVASPT	ELA	975.11	10.34	17.57 %
11mer_H2_Db	187	SAN	TGVINNASDEG	NAA	1058.06	16.174	20.35 %
11mer_H2_Db	208	LTL	IKAYNDYIRQP	VQL	1362.56	14.205	17.87 %
11mer_H2_Db	88	LLR	ELGELNDAQKL	DRA	1211.34	10.745	13.52 %
11mer_H2_Db	395	QRV	GNAFNVRWQQL	AGT	1291.46	10.505	13.21 %
11mer_H2_Db	350	PNS	RAIPQLCYYGL	SPE	1278.55	9.623	12.11 %
11mer_H2_Db	407	QLA	GTDANIRYYNL	PAD	1281.39	9.164	11.53 %
11mer_H2_Db	468	RAS	ASATNTNTDFI	AQM	1136.17	8.8	11.07 %
11mer_H2_Db	57	GQT	QELEDQQTYKL	LAA	1376.49	8.411	10.58 %
11mer_H2_Db	193	INN	ASDEGNAALGG	WLT	942.94	7.104	8.94 %
11mer_H2_Db	43	KDA	NASSEFYINKL	GQT	1267.41	7.086	8.91 %
11mer_H2_Db	230	SWK	NAYPNHAAATL	FPK	1124.22	6.648	8.36 %
11mer_H2_Db	47	ASS	EFYINKLGQTQ	ELE	1322.48	5.265	6.62 %

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

ALLELE	Sequence	Residue No	Peptide Score
I-Ab	KTLVGPLLK	314	1.164
I-Ab	VILADPAQI	327	1.100
I-Ab	GYLTNIVPN	449	0.970
I-Ab	RIEMDKNLT	158	0.887
I-Ad	LGQRVGNAF	390	0.776
I-Ad	GNAALGGWL	197	0.712
I-Ad	DAWLLINQF	522	0.629
I-Ad	CSNLLGSNF	26	0.627
I-Ag7	YPNHAAATL	232	1.979
I-Ag7	EGNAALGGW	196	1.702
I-Ag7	YKLLAARVL	65	1.700
I-Ag7	GTDANIRYY	407	1.699
RT1.B	KVEQSAALL	78	1.078
RT1.B	QTQELEDQQ	55	1.034
RT1.B	KLGQTQELE	52	0.940
RT1.B	NFQQTNVSQ	250	0.908

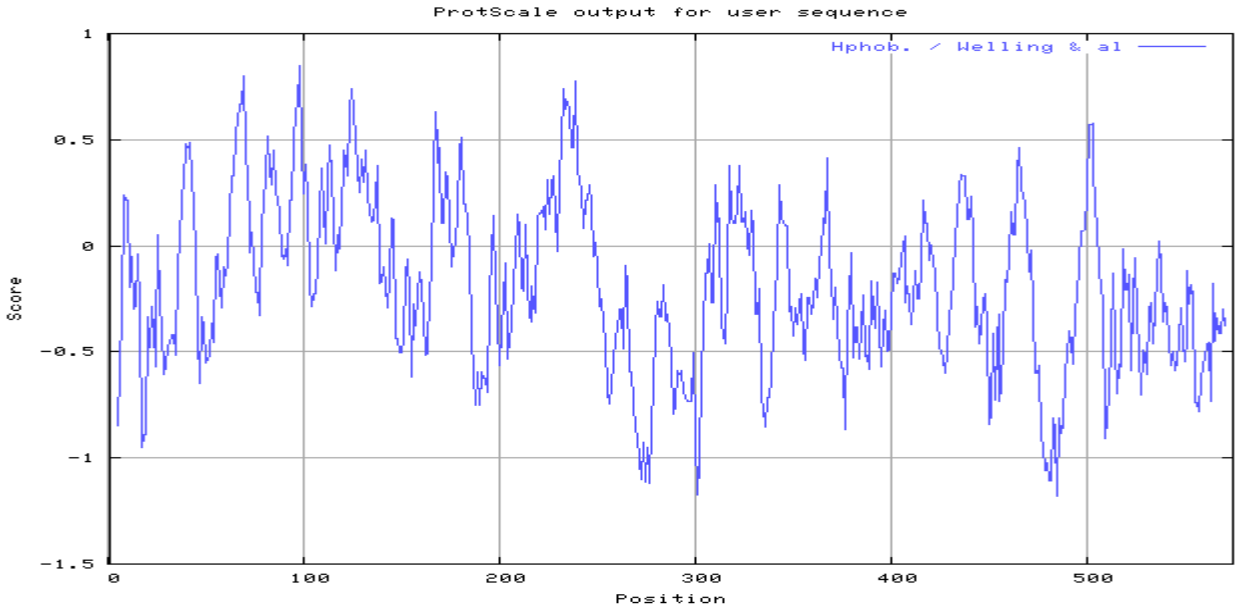

Fig. 1- Antigenicity plot of antigen protein by Welling, et al., scale

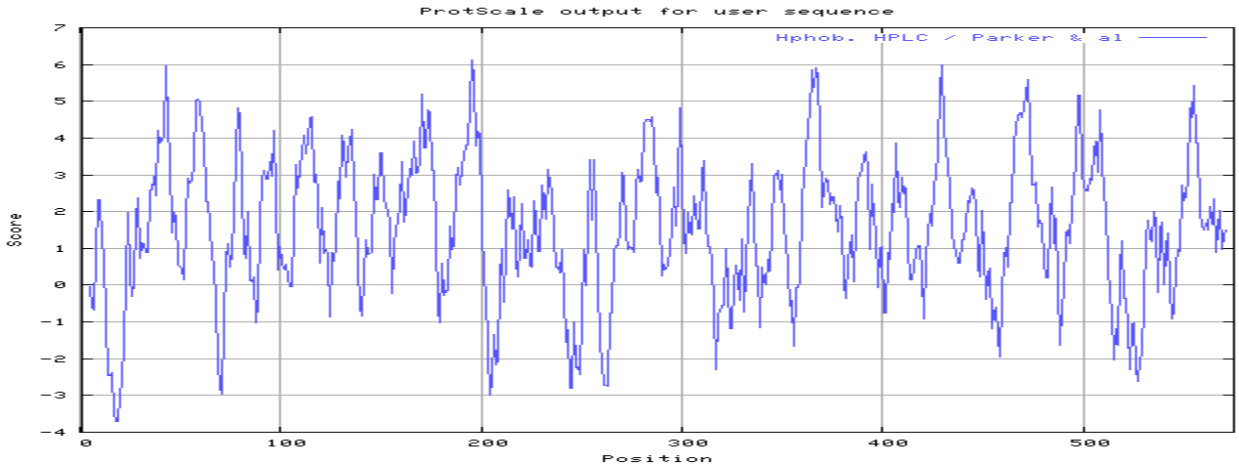

Fig. 2- Antigenicity plot of antigen protein by HPLC / Parker, et al., scale

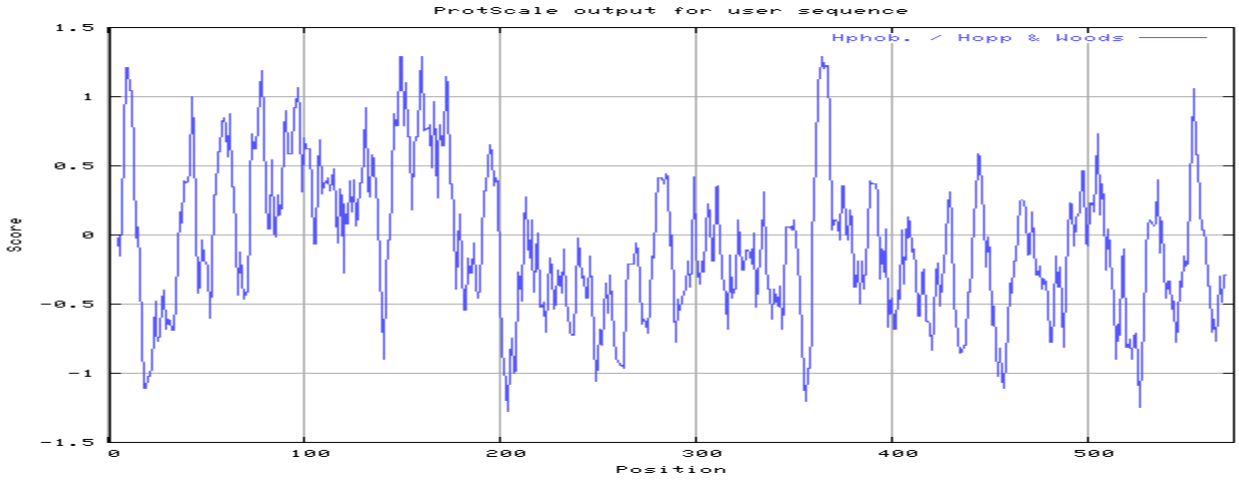

Fig. 3- Hydrophobicity plot of antigen protein by Hopp & Woods, et al., scale

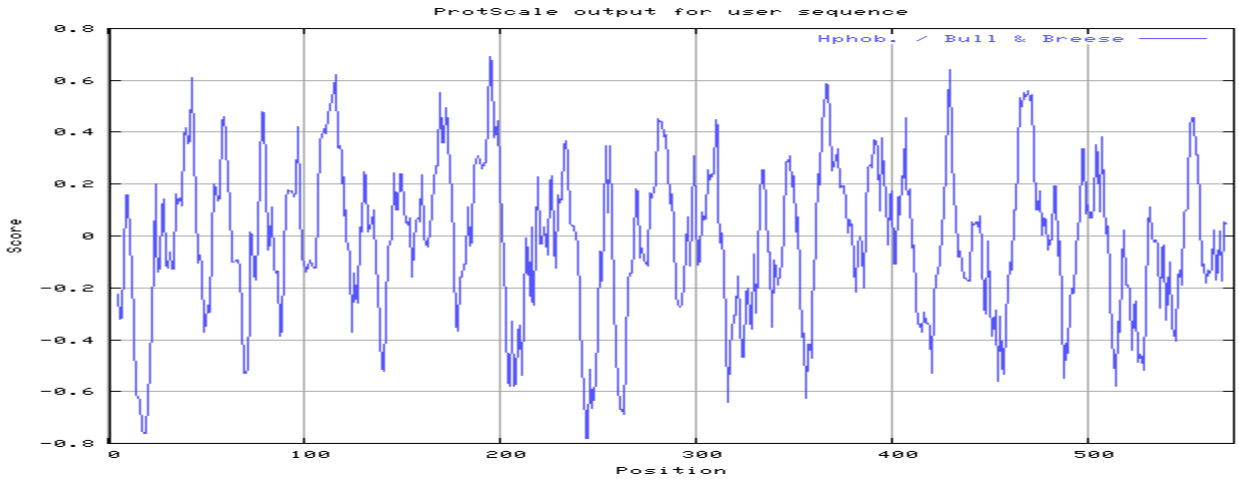

Fig. 4-Hydrophobicity plot of antigen protein by Bull & Breese scale