Gene Ther Mol Biol Vol 12, 83-86, 2008

Prediction of MHC binder for fragment based viral peptide vaccines from *cabbage leaf curl virus*

Research Article

Virendra S Gomase[1,*] Karbhari V Kale[2]

[1]Department of Bioinformatics, Padmashree Dr. D.Y. Patil University, Navi Mumbai, 400614, India
[2]Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, 431004 (MS), India

***Correspondence**: Virendra S. Gomase, Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, 431004 (MS), India; Mobile: +91-9226960668; e-mail: virusgene1@yahoo.co.in

Key words: Antigen, 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)

Received: 6 June 2008; Revised: 9 June 2008
Accepted: 10 June 2008; electronically published: July 2008

Summary

Cabbage leaf curl viral peptides are most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. The MHC peptide binding of pathogenicity proteins is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding of pathogenicity proteins is a log-transformed value related to the IC50 values in nM units. We describe an improved method for predicting linear epitopes (Table 1). The region of maximal hydrophilicity is likely to be an antigenic site, having hydrophobic characteristics, because terminal regions of pathogenicity protein is solvent accessible and unstructured, antibodies against those regions are also likely to recognize the native protein. It was shown that a pathogenicity protein is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility. Predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

I. Introduction

Cabbage leaf curl viral 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 plant infected with a mild strain of virus is protected against a more severe strain of the same virus. The phenotype of the resistant transgenic plants includes fewer centers of initial virus infection, a delay in symptom development, and low virus accumulation. Pathogenicity protein from *Cabbage leaf curl virus* is 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 (Gomase et al 2008).

II. Methodology

Antigenic epitopes of pathogenicity protein from *Cabbage leaf curl virus* is determined using the Gomase in 2007, Hopp and Woods, Welling, Parker and Protrusion Index (Thornton) antigenicity methods (Gomase and Kale, 2007; Gomase et al, 2007a,b). The major histocompatibility complex (MHC) peptide binding of pathogenicity proteins is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding of pathogenicity proteins 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 (Reche et al, 2002; Buus et al, 2003; Nielsen et al, 2003; Bhasin and Raghava 2005; Gomase et al, 2006, 2007c). In addition, we predict those MHC ligands from whose C-terminal end is likely to be the result of proteosomal cleavage.

III. Results and Interpretations

RankPep server predicts binding of peptides to a number of different alleles using Position Specific Scoring Matrix. A pathogenicity protein sequence is 295 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 pathogenicity protein sequence are as 11mer_H2_Db, 10mer_H2_Db, 9mer_H2_Db, 8mer_H2_Db and also peptide binders to MHCII molecules of pathogenicity protein sequence as I_Ab.p, I_Ag7.p, I_Ad.p, analysis found antigenic epitopes region in putative pathogenicity protein (**Table 1**). We also found the SVM based MHCII-IAb peptide regions 109-FSLKDPIPW, 153-PFRAPTVKI, 194-IGLTGPGPI, 139GKLKLSTAK, (optimal score is 0.952); MHCII-IAd peptide regions 248-GDSASQAGL, 223-TESEVENAL, 262-TITMSVAQL, 10-NAFNYIESH, (optimal score is 0.804); MHCII-IAg7 peptide regions 3-SQLANAPNA, 174-SHVDYGRWE, 36-PSTAAQFTA, 248-GDSASQAGL, (optimal score is 1.744); and MHCII-RT1.B peptide regions 249-DSASQAGLQ, 39-AAQFTARLN, 21-EYQLSHDLT, 37-STAAQFTAR, (optimal score is 1.361) which represented predicted binders from viral pathogenicity 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 pathogenicity protein is solvent accessible and unstructured, antibodies against those regions are also likely to recognize the native protein. It was shown that a pathogenicity protein is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility. Predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

Table 1. MHC ligands from whose C-terminal end are proteosomal cleavage sites.

MHC-I	POS.	N	Sequence	C	MW (Da)	Score	% OPT.
11mer_H2_Db	43	AQF	TARLNRSCMKI	DHC	1274.56	103.0	48.58 %
11mer_H2_Db	63	YRQ	QVPINATGSVI	VEI	1080.24	79.0	37.26 %
11mer_H2_Db	56	IDH	CVIEYRQQVPI	NAT	1329.59	69.0	32.55 %
11mer_H2_Db	253	SAS	QAGLQRAQSTI	TMS	1154.29	59.0	27.83 %
11mer_H2_Db	260	QRA	QSTITMSVAQL	SEL	1160.34	58.0	27.36 %
10mer_H2_Db	3	MN	SQLANAPNAF	NYI	1014.11	115.0	52.04 %
10mer_H2_Db	6	SQL	ANAPNAFNYI	ESH	1076.18	108.0	48.87 %
10mer_H2_Db	254	ASQ	AGLQRAQSTI	TMS	1026.16	91.0	41.18 %
10mer_H2_Db	225	HTE	SEVENALHPY	REL	1140.23	78.0	35.29 %
10mer_H2_Db	37	QFP	STAAQFTARL	NRS	1047.18	78.0	35.29 %
9mer_H2_Db	3	MN	SQLANAPNA	FNY	866.93	76.0	52.05 %
9mer_H2_Db	265	TIT	MSVAQLSEL	VRT	959.13	72.0	49.32 %
9mer_H2_Db	6	SQL	ANAPNAFNY	IES	963.02	72.0	49.32 %
9mer_H2_Db	79	IHD	KRMTDDESL	QAS	1076.19	70.0	47.95 %
9mer_H2_Db	43	AQF	TARLNRSCM	KID	1033.23	67.0	45.89 %
9mer_H2_Db	144	LKL	STAKHSVDI	PFR	939.03	64.0	43.84 %
8mer_H2_Db	260	QRA	QSTITMSV	AQL	847.97	84.0	46.93 %
8mer_H2_Db	179	VDY	GRWERKTL	RSK	1004.19	76.0	42.46 %
8mer_H2_Db	209	PGD	SWASRSTI	GFP	865.98	73.0	40.78 %
8mer_H2_Db	220	GFP	NPHTESEV	ENA	893.91	71.0	39.66 %
8mer_H2_Db	66	QVP	INATGSVI	VEI	755.86	67.0	37.43 %
MHC-II							
I_Ab	60	VIE	YRQQVPINA	TGS	1070.22	17.225	48.34 %
I_Ab	154	DIP	FRAPTVKIH	SKQ	1050.27	12.866	36.11 %
I_Ab	120	WKL	YYRVSDTNV	HQR	1098.18	10.475	29.40 %
I_Ab	3	MN	SQLANAPNA	FNY	866.93	9.945	27.91 %
I_Ad	254	ASQ	AGLQRAQST	ITM	913.0	15.227	28.65 %
I_Ad	174	VDF	SHVDYGRWE	RKT	1107.19	13.222	24.88 %
I_Ad	247	ALD	PGDSASQAG	LQR	770.76	12.975	24.41 %
I_Ad	164	IHS	KQFSHRDVD	FSH	1113.2	9.257	17.42 %
I_Ad	226	TES	EVENALHPY	REL	1053.15	9.09	17.10 %
I_Ad	229	EVE	NALHPYREL	NLL	1094.25	8.784	16.53 %
I_Ag7	130	NVH	QRTHFAKFK	GKL	1144.34	11.452	28.02 %
I_Ag7	96	FPL	RCNIDLHYF	SSS	1162.34	10.569	25.86 %
I_Ag7	37	QFP	STAAQFTAR	LNR	934.02	10.338	25.29 %
I_Ag7	226	TES	EVENALHPY	REL	1053.15	9.943	24.33 %
I_Ag7	168	QFS	HRDVDFSHV	DYG	1093.17	9.71	23.76 %
I_Ag7	13	NAF	NYIESHRDE	YQL	1144.18	9.151	22.39 %

Table 2. MHC class II binding peptide nonamers from pathogenicity protein.

MHC ALLELE	Rank	Sequence	Residue No.	Peptide Score
I-Ab	1	FSLKDPIPW	109	0.952
I-Ab	2	PFRAPTVKI	153	0.921
I-Ab	3	IGLTGPGPI	194	0.885
I-Ab	4	GKLKLSTAK	139	0.833
I-Ad	1	GDSASQAGL	248	0.804
I-Ad	2	TESEVENAL	223	0.759
I-Ad	3	TITMSVAQL	262	0.695
I-Ad	4	NAFNYIESH	10	0.639
I-Ag7	1	SQLANAPNA	3	1.744
I-Ag7	2	SHVDYGRWE	174	1.485
I-Ag7	3	PSTAAQFTA	36	1.408
I-Ag7	4	GDSASQAGL	248	1.365
RT1.B	1	DSASQAGLQ	249	1.361
RT1.B	2	AAQFTARLN	39	1.126
RT1.B	3	EYQLSHDLT	21	0.845
RT1.B	4	STAAQFTAR	37	0.813

IV. Conclusion

A pathogenicity proteins from *Cabbage leaf curl virus* 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 viral pathogenicity protein. These predicted of pathogenicity protein antigenic peptides to MHC class molecules are important in vaccine development from Cabbage leaf curl virus

References

Bhasin M, Raghava GP (**2005**) Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. **Nucleic Acids Research** 33, W202-207.

Buus S, Lauemøller SL, Worning P, Kesmir C, Frimurer T, Corbet S, Fomsgaard A, Hilden J, Holm A, Brunak S (**2003**) Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. **Tissue Antigens** 62, 378-384.

Gomase V S, Tandale JP, Patil SA, Kale KV (**2006**) 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.

Gomase VS, Kale KV (**2007**) Binding affinity prediction of Cabbage leaf curl virus pathogenicity protein for TAP transporter, National Symposium on Genomics, Proteomics and Bioinformatics, Oamanabad, India.

Gomase VS, Kale KV, Jyotiraj A, Vasanthi R (**2007**) Identification of mhc ligands from alfalfa mosaic virus, *CTDDR-2007*. **Medicinal Chemistry Research** 15, 160.

Gomase VS, Kale KV, Dede PV, Patil SY, Patil SS (**2007**) ANN based prediction of MHC alleles in vaccine design, International Conference on Intelligent Systems & Networks IISN-2007, Jagadhri, India, 223-227.

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.

Gomase VS, Gomase SV, Tagore S, Bhiwgade DA, Baig MMV, Kale KV (**2008**) Prediction of MHC binder for fragment based viral peptide vaccines from cabbage leaf curl virus. 12th Annual International Conference on Research in Computational Molecular Biology (RECOMB08), National University of Singapore, Singapore.

Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, Brunak S, Lund O, (**2003**) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. **Protein Sci** 12, 1007-1017.

Reche PA, Glutting JP, Reinherz EL (**2002**) Prediction of MHC Class I Binding Peptides Using Profile Motifs. **Hum Immun** 63, 701-709.