Research Article

IN SILICO EPITOPE PREDICTION FOR GLYCOPROTEIN D IN HUMAN *HERPES SIMPLEX* VIRUS-1

Pramod Kumar Yadav[1,*], Raghuvir Singh[2], Prashant A. Jain[1], Satendra Singh[1], B. Gautam[1] and R. Farmer[1]
[1]Department of Computational Biology & Bioinformatics, JSBB, SHIATS, Allahabad-211007, India.
[2]Department of Molecular & Cellular Engineering, JSBB, SHIATS, Allahabad-211007, India.
*Corresponding author's E-mail: pkyadav.bioinformatics@gmail.com

Accepted on: 25-01-2011; **Finalized on:** 10-04-2011.

ABSTRACT

Herpes is an infection that is caused by a herpes simplex virus (HSV). There are two types of herpes simplex virus (HSV) infections. HSV type 1 (HSV-1) more commonly causes cold sores or fever blisters around the mouth. HSV type 2 (HSV-2) more commonly causes genital herpes. Entry of HSV into the host cell involves interactions of several glycoproteins on the surface of the enveloped virus, with receptors on the surface of the host cell. Studies have shown that vaccination of mice with the secreted HSV 1 gD resulted in the induction of antibodies able to neutralize both HSV 1 and HSV 2 *in vitro*. In the present work T-cell and B-cell epitopes were predicted for glycoprotein D of HSV type-1. MHC-I and MHC-II epitopes were predicted by ProPred-I and MHC2Pred servers respectively and B-cell epitopes were predicted by ABCPred server. The three-dimensional structures of top ranked epitopes (small peptides) were modelled using the HHpred server. In order to find the most relevant epitopes among the MHC-I and MHC-II predicted epitopes; protein-protein docking studies were carried out. These predicted epitopes (small peptides) might be promising candidates for the vaccine design against HSV-1.

Keywords: Herpes simplex virus, Glycoprotein D, Epitopes, Modelling, Docking.

INTRODUCTION

Herpes simplex virus 1 and 2 (HSV-1 and HSV-2) are two species of the herpes virus family, Herpesviridae, which cause infections in humans. Eight members of herpes virus infect humans to cause a variety of illnesses including cold sores, chickenpox or varicella, shingles or herpes zoster (VZV), cytomegalovirus (CMV), and various cancers, and can cause brain inflammation (encephalitis). All viruses in the herpes family produce life-long infections. They are also called *Human Herpes Virus 1* and *2* (*HHV-1* and *HHV-2*) and are neurotropic and neuroinvasive viruses; they enter and hide in the human nervous system, accounting for their durability in the human body. HSV-1 is commonly associated with herpes outbreaks of the face known as cold sores or fever blisters, whereas HSV-2 is more often associated with genital herpes.[1, 2]

As one of the largest and most complex viruses, the infectious HSV-1 virion has a highly characteristic structure consisting of four compartments: envelope, tegument, capsid, and core[3]. The core consists of the double-stranded DNA (dsDNA) genome of 152,000 bp, which is packaged into the preformed icosahedral capsid within the nucleus of the infected cell. The DNA has been reported to adopt a liquid crystalline organization within the capsid, but its precise arrangement is not known[4]. The capsid is surrounded by a proteinaceous layer of variable thickness, called the tegument, and the entire structure is bounded by the viral envelope, a spherical lipid bilayer containing 12 or more different glycoproteins[3].

HSV-1 and Gylycoprotein D

Lasky *et al*. reported that vaccination of mice with the secreted HSV 1 gD resulted in the induction of antibodies able to neutralize both HSV 1 and HSV 2 *in vitro*[5]. Mice vaccinated with this antigen were protected from a lethal challenge with either HSV 1 or HSV 2. Infection of permissive cells by herpes simplex virus type 1 (HSV-1) generally requires virus attachment to the cell surface and fusion of the viral envelope with the cytoplasmic or endosomal membrane. HSV attachment involves the coordinated binding of multiple viral glycoproteins to their cognate cellular receptors. Glycoproteins B (gB) and C (gC) bind to herparan sulfate (HS) glycosaminoglycan (GAG) side chains of cell-surface proteoglycans, while glycoprotein D (gD) engages one of three HSV-1 entry receptors, HVEM (HveA), nectin-1 (Hve-C), or 3-O-sulfated HS. Receptor binding by gD is believed to set in motion the fusion process, which requires gB and the gH:gL heterodimer for completion[6]. HSV-1 membrane glycoproteins gB, gD, gH and gL are all required for virus entry The interaction between gD and cellular receptors is thought to trigger downstream events involving gB, gH and gL that result in fusion of the viral envelope with cell membranes. The domains on gD that interact with HVEM and nectin-1 have been defined by mutagenesis analysis and by resolving the crystal structure of gD or gD bound to HVEM. The resultant gD mutant protein lost the ability to interact with nectin-1, but retained the ability to bind HVEM in vitro[7]. Therefore, this gD mutant virus impaired in entry may represent a novel candidate for an attenuated live HSV-1 vaccine.

The aim of present study was to predict the T cell and B-cell epitopes for glycoprotein-D in HSV-1 using the bioinformatics tools, and to perform the protein-protein docking studies for the predicted epitopes with known receptors.

MATERIALS AND METHODS

Tools used

ProPred 1 Server: ProPred 1 is an on-line web tool for the prediction of peptide binding to MHC class-I alleles (http://www.imtech.res.in/raghava/**propred1**/). This is a matrix-based method that allows the prediction of MHC binding sites in an antigenic sequence for 47 MHC class-I alleles. The **server** represents MHC binding regions within an antigenic sequence in user-friendly formats[8].

MHC2Pred Server: The MHC2Pred is an SVM based method for prediction of promiscuous MHC class II binding peptides (http://www.imtech.res.in/raghava/mhc2pred/). The data for training has been extracted from MHCBN and JenPep database. All the peptides having IC50 value less than 500nm has been considered as binders and peptides with IC50 value greater than 500nm are considered as non-binders. The binders and non-binders for all alleles have been obtained from MHCBN and JenPep database[9]. For the development of MHC binder prediction method, an elegant machine learning technique SVM has been used[10].

ABCpred Server: ABCPred uses artificial neural networks for predicting linear B-cell epitopes. The aim of ABCpred server is to predict B cell epitope(s) in an antigen sequence, using artificial neural network (http://www.imtech.res.in/raghava/abcpred/). This is the first server developed based on recurrent neural network (machine based technique) using fixed length patterns. The target output consists of a single binary number and is 1 or 0 (epitope or non epitope)[11].

HHPred server: HHsearch is a program for protein sequence searching that is free for non-commercial use[12]. HHpred is a free protein function and protein structure prediction server based on the HHsearch method. HHpred profiles are calculated from a multiple sequence alignment of related sequences which are typically collected using the PSI-BLAST program. If a significant match with a protein of known structure (a "template") is found in the PDB database, HHpred allows to build a homology model using the MODELLER software, starting from the pairwise query-template alignment[13] (http://toolkit.tuebingen.mpg.de/hhpred).

Hex 5.1: Hex is an interactive molecular graphics program for calculating and displaying feasible docking modes of pairs of protein and DNA molecules. Hex can also calculate small-ligand/protein docking (provided the ligand is rigid), and it can superpose pairs of molecules using only knowledge of their 3D shapes. In Hex's docking calculations, each molecule is modeled using 3D parametric functions which are used to encode both surface shape and electrostatic charge and potential distributions[14].

Methodology

The sequence Glycoprotein-D of Herpes Simplex Virus-1 was retrieved from the site of Swiss-Prot (Accession No. Q69091) (http://expasy.org/sprot/). The sequence was then converted into the FASTA format. For the prediction of MHC class-I epitopes ProPred-I online prediction tool was used. The sequence was submitted to the ProPred-I server in the FASTA format. The epitopes were predicted for different alleles (HLA-A1, HLA-A2, HLA-A*0201, HLA-A*0205) of MHC class-I. Epitopes for different alleles (HLA-DR1, HLA-DR4, HLA-DR9, HLA-DR11) of MHC class-II were predicted by using the MHC2Pred online tool. The B cell epitopes were predicted using the ABCPred online tool. All the predicted epitopes were obtained in the form of small peptide sequences.

The three dimensional structure of small peptides (epitopes) were required to perform docking so as to find out the most relevant epitopes for MHC class-I, MHC class-II and B cell receptors. The 3D structures of predicted peptides were modelled by using the online tool HHpred server. The small peptide sequences were submitted to the HHpred server which is a protein structure prediction server based on the HHsearch method. HHpred profiles were calculated from a multiple sequence alignment of related sequences which were collected using the PSI-BLAST program. A significant match with a protein of known structure (a "template") was found in the PDB database. Subsequently HHpred allowed to build a homology model using the MODELLER software. Then the 3D structures of small peptide (epitopes) were evaluated by using the SAVES server. The three dimensional structure of the receptors for MHC class-I (PDB ID: **1I1Y**) and for MHC class-II (PDB ID: **1KG0**) required for docking were retrieved from the Protein Data Bank (www.rcsb.org/pdb) and visualized by using the RASMOL tool.

Protein-protein docking was performed to find out the epitopes with maximum binding affinity. HEX 5.1 program was used for docking. The receptor and the small peptide (epitopes) were loaded in the HEX program. Both the receptor and the small peptide (epitope) were docked by activating the docking option present on the tool bar of the HEX window. Top ranked epitopes predicted for different MHC-I alleles (HLA-A1, HLA-A2, HLA-A*0201 and HLA-A*0205) were docked with the MHC class-1 receptor. Similarly, top ranked epitopes predicted for different MHC-II alleles (HLA-DR1, HLA-DR4, HLA-DR9 and HLA-DR11) were docked with MHC class-II receptors. Subsequently, the energy score of all docked complexes were calculated. All the docked complexes were saved and a comparative analysis was carried out on the basis of energy score (KJ/mole).

RESULTS

Result for MHC Class-I binding epitope prediction

Table 1-4 shows the predicted T-cell MHC class-I epitopes for different alleles i.e. HLA-A1, HLA-A2, HLA-**0201 and HLA-A*0205. These small sequence (epitopes) were obtained from the ProPred-I server.

ProPred-I server allows the prediction of standard proteasome and immunoproteasome cleavage sites in an antigenic sequence. It identifies the MHC binders who have cleavage sites at the C-terminus. All peptides having IC50 value less than 500nm has been considered as binders and peptides with IC50 value greater than 500nm has been considered as non-binders.

Result for MHC Class-II binding epitope peptide prediction

Table 5-8 shows the small peptides (epitopes) for MHC class-II molecules predicted by SVM based MHC2Pred prediction server. All peptides having IC50 value less than 500nm has been considered as binders and peptides with IC50 value greater than 500nm has been considered as non-binders.

Result for B -cell epitope binding peptide prediction

Table 9 shows the predicted B-cell epitopes which are in the form of small peptide sequences. B-cell epitopes were predicted using the ABCpred server.

All the predicted epitopes were ranked according to their respective scores which have been obtained by the trained recurrent neural network. Higher score of the peptides reveals the higher probability to be most reliable epitope.

Docking of the predicted epitopes with T-cell receptors

Top ranked predicted epitopes for different alleles of MHC class-I were docked with human MHC-I receptor using the Hex program. Similarly protein-protein docking was also performed for the top ranked MHC class-II predicted epitopes with the human MHC-II receptor. Their respective energy scores are depicted in the Table 10.

Table 1: MHC Class-I (HLA-A1) Binding Peptide Prediction

ALLELE:HLA-A1					
Threshold for 4 % with score: -0.693			Highest Score achievable by any peptide on ***log* scale**: 12.324		
Rank	Sequence	At Position	Real Score	***Log*** Score	% of Highest on ***log*** scale
1.	LTDPPGVRR	53	250.0000	5.5215	44.80
2.	MADPNRFRG	36	25.0000	3.2189	26.12
3.	VMEYTECSY	137	22.5000	3.1135	25.26
4.	RTQPRWNYY	155	12.5000	2.5257	20.49

Table 2: MHC Class-I (HLA-A2) Binding Peptide Prediction

ALLELE:HLA-A2					
Threshold for 4 % with score: 1.553			Highest Score achievable by any peptide on ***log* scale**: 15.156		
Rank	Sequence	At Position	Real Score	***Log*** Score	% of Highest on ***log*** scale
1.	ILFVVIVGL	12	468.1870	6.1489	40.57
2.	AVILFVVIV	10	91.3433	4.5146	29.79
3.	PKAPYTSTL	269	58.4516	4.0682	26.84
4.	VLLNAPSEA	94	57.4146	4.0503	26.72

Table 3: MHC Class-I (HLA-A*0201) Binding Peptide Prediction

ALLELE: HLA-A*0201					
Threshold for 4 % with score: 1.143			Highest Score achievable by any peptide on ***log* scale**: 17.857		
Rank	Sequence	At Position	Real Score	***Log*** Score	% of Highest on ***log*** scale
1.	RLGAVILFV	7	879.8328	6.7797	37.97
2.	ILFVVIVGL	12	309.0498	5.7335	32.11
3.	ALLEDPVGT	302	113.0468	4.7278	26.48
4.	SLPITVYYA	77	62.8452	4.1407	23.19

Table 4: MHC Class-I (HLA-A*0205) Binding Peptide Prediction

ALLELE: HLA-A*0205					
Threshold for 4 % with score: 0.519			Highest Score achievable by any peptide on ***log* scale**: 10.499		
Rank	Sequence	At Position	Real Score	***Log*** Score	% of Highest on ***log*** scale
1.	ILFVVIVGL	12	71.4000	4.2683	40.65
2.	RVYHIQAGL	61	42.0000	3.7377	35.60
3.	AVILFVVIV	10	16.0000	2.7726	26.41
4.	VVIVGLHGV	15	13.6000	2.6101	24.86

Table 5: MHC class-II (HLA-DR1) Binding Peptide Prediction

ALLELE:HLA-DR1				
Threshold 0.0 as cutoff score				
Prediction Method	Rank	Sequence	Residue No.	Peptide Score
SVM	1.	TVYYAVLER	81	0.936
SVM	2.	GGAAARLGA	2	0.794
SVM	3.	YAVLERACR	84	0.703
SVM	4.	IVRGASEDV	105	0.662

Table 6: MHC class-II (HLA-DR4) Binding Peptide Prediction

ALLELE:**HLA-DR4**				
Threshold 0.0 as cutoff score				
Prediction Method	Rank	Sequence	Residue No.	Peptide Score
SVM	1	YYAVLERAC	83	1.265
SVM	2	PRFIPENQR	246	1.172
SVM	3	DWTEITQFI	197	1.044
SVM	4	LNAPSEAPQ	96	1.003

Table 7: MHC class-II (HLA-DR9) Binding Peptide Prediction

ALLELE:HLA-DR9				
Threshold 0.0 as cutoff score				
Prediction Method	Rank	Sequence	Residue No.	Peptide Score
SVM	1.	PNRFRGKDL	39	1.487
SVM	2.	QLTDPPGVR	52	1.453
SVM	3.	LFVVIVGLH	13	1.430
SVM	4.	LLEDPVGTV	303	1.376

Table 8: Result of MHC class-II (HLA-DR11) Binding Peptide Prediction

ALLELE:HLA-DR11				
Threshold 0.0 as cutoff score				
Prediction Method	Rank	Sequence	Residue No.	Peptide Score
SVM	1.	GLIAGAVGG	341	0.149
SVM	2.	SLKIAGWHG	260	0.140
SVM	3.	YAVLERACR	84	0.140
SVM	4.	GLPDPFQPP	68	0.138

Table 9: Result of B-cell epitope prediction

Sr. No.	Rank	Sequence	Start position	Score
1.	1	PSIQDAATPYHPPATP	322	0.95
2.	2	ACPIRTQPRWNYYDSF	151	0.92
3.	3	YHPPATPNNMGLIAGA	331	0.91
4.	3	PITVMEYTECSYNKSL	134	0.91
5.	4	RVYHIQAGLPDPFQPP	61	0.90
6.	5	GWHGPKAPYTSTLLPP	265	0.89
7.	5	VDSIGMLPRFIPENQR	239	0.89
8.	6	RFRGKDLPVLDQLTDP	41	0.88
9.	6	PELSETPNATQPELAP	280	0.88
10.	7	PITVYYAVLERACRSV	79	0.87
11.	7	SLKMADPNRFRGKDLP	33	0.87
12.	8	VAPQIPPNWHIPSIQD	311	0.86
13.	8	PELAPEDPEDSALLED	291	0.86
14.	8	FETAGTYLRLVKINDW	183	0.86
15.	9	PSACLSPQAYQQGVTV	224	0.84

Table 10: Energy Score of Docked Complexes of T-Cell Epitopes

S. No.	Name of the Epitope (Small peptide)	Energy Score of Docked Complex (KJ/mole)
1.	MADPNRFRG	-658.94
2.	ILFVVIVGL	-668.03
3.	RLGAVILFV	**-792.42**
4.	ILFVVIVGL	-668.03
5.	TVYYAVLER	-489.48
6.	YYAVLERAC	-563.67
7.	PNRFRGKDL	**-746.57**
8.	GLIAGAVGG	-574.35

DISCUSSION

The present work was conducted to predict the epitopes for B-cell and T-cell and to find the most efficient epitope with maximum binding affinity for glycoprotein D of *Herpes simplex* virus type 1. **Awasthi** *et al.* suggested that glycoprotein D mutant virus may represent a novel candidate for HSV-1 vaccine[7]. Earlier, *in silico* epitope prediction has been carried out by **Tambunan** *et al.* for Dengue virus[15].

The epitopes were predicted in the form of small peptides using different bioinformatics tools, as shown in Table **1-9**. For MHC class-I epitope prediction, ProPred-I server was used. Four epitopes were predicted for different alleles (HLA-A1, HLA-A2, HLA-A*0201, HLA-A*0205) of MHC class-I and rank has been given on the basis of log score. For MHC class-II epitope prediction MHC2pred server was used. Four epitopes were predicted by SVM method for different alleles (HLA-DR1, HLA-DR4, HLA-DR9 and HLA-DR11) of MHC class-II and rank has been given on the basis of peptide score. For B cell epitope prediction ABCpred server was used which predicted 34 epitopes and rank has been given the basis of score. Out of 34, top 15 ranked epitopes has been shown in Table 9. In the next step, the three dimensional structure of top ranked predicted T-cell and B-cell epitopes (small peptides) were modelled using the HHpred server which uses homology modelling method.

In order to find out the most relevant epitopes among the MHC class-I and MHC class-II predicted epitopes; protein-protein docking method was carried out. Three dimensional structures of predicted epitopes of glycoprotein D for HSV-1 were docked in different orientations with MHC class-I and MHC class-II receptors respectively by using the HEX program. Subsequently, the energy score of docked complexes were calculated. Lower energy score reveals higher binding affinity towards receptors. Table 10 shows the energy score of different epitopes after the protein-protein docking. For MHC class-I receptor, epitope (small peptide) 'RLGAVILFV' showed the minimum energy score of **-792.42** KJ/mole, while for MHC class-II receptor, the epitope (small peptide) ' PNRFRGKDL' showed the minimum energy score of **-746.57** KJ/mole. In the case of B cell epitope prediction, highest ranked epitopes were 'PSIQDAATPYHPPATP' and 'ACPIRTQPRWNYYDSF'. Those above predicted epitopes might be promising vaccine candidates against *Herpes simplex* virus type-1.

CONCLUSION

T-cell and B-cell epitopes were predicted for glycoprotein D of *Herpes simplex* virus type-1. After docking studies, it was found that 'RLGAVILFV' epitope was having the least energy score (**-792.42** KJ/mole) for MHC-I receptor which reveals highest binding affinity for MHC-I. For MHC-II receptors, it was found that 'PNRFRGKDL' epitope was having the least energy score (**-746.57** KJ/mole) which reveals highest binding affinity for MHC-II receptor. The small peptides 'PSIQDAATPYHPPATP' and 'ACPIRTQPRWNYYDSF' were most probable epitopes for the B-cell. In future, above predicted epitopes (small peptides) can be synthesized in wet laboratory which might be promising candidates for the vaccine design against HSV-1.

REFERENCES

1. Ryan, K.J. and Ray, C.G. (editors). *Sherris Medical Microbiology (4th ed.).* McGraw Hill. 2004, pp. *555–562.*
2. Matis, J. and Kudelova, M. Early shutoff of host protein synthesis in cells infected with herpes simplex viruses. *Acta Virol., 2001, 45 (5-6): 269–277.*
3. 3. Steven, AC and Spear PG. Herpes virus capsid assembly and envelopment, Structural biology of viruses. New York, N.Y: Oxford University Press, 1997, pp: 312–351.
4. Booy, F. P., Newcomb, W. W., Trus, B. L., Brown, J. C., Baker, T. S. and Steven, A. C. Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell, 1991,* 64, pp:1007–1015.
5. Lasky, L.A., Dowbenko, D., Simonsen, C.C. and Berman, P.W. Protection of Mice from Lethal Herpes Simplex Virus Infection by Vaccination with a Secreted form of Cloned Glycoprotein D. *Bio/Technology, 1984,* 2:527 - 532.
6. Tsvitov, M., Frampton, A.R., Shah, W.A., Wendell, S.K., Ozuer, A., Kapacee, Z., Goins, W.F., Cohen, Z.B. and Glorioso, J.C. Characterization of Soluble

Glycoprotein D Mediated Herpes Simplex Virus Type 1 Infection. *Virology*, 2007, 360(2): 477–491.

7. Awasthi S., Lubinski JM, Eisenberg RJ, Cohen GH, Friedman HM. An HSV-1 gD mutant virus as an entry-impaired live virus vaccine. *Vaccine*. 2008, 26(9):1195-203.
8. Singh, H. and Raghava, G.P.S. Propred1: prediction of promiscuous MHC class-1 binding sites. *Bioinformatics,* 2003, 19 (8): 1009-1014.
9. Bhasin, M., Singh, H. and Raghava, G.P.S). MHCBN: a comprehen-sive database of MHC binding and non-binding peptides, *Bioinformatics*. 2003, 19(5): 665-666.
10. Joachims, T. Advances in Kernel Methods- Support Vector Learning, Making large-Scale SVM Learning Practical, MIT Press. 1999, 11:102-104.
11. Saha, S and Raghava, G.P.S. Prediction of Continuous B-cell Epitopes in an Antigen Using Recurrent Neural Network. Proteins, 2006, 65(1): 40-48
12. Hildebrand, A., and Remmert. Fast and accurate automatic structure prediction with Hhpred. *Proteins: Structure, Function, and Bioinformatics*, 2009, 77 (S9): 128–132.
13. Jaroszewski, L., Rychlewski, L. and Godzik, A. Improving the quality of twilight-zone alignments. *Protein Science*, 2000, 9 (8): 1487-1496.
14. Ritchie, D.W. and Kemp, G.J.L. Protein Docking Using Spherical Polar Fourier Correlations, PROTEINS: Struct. Funct. Genet., 2000, 39: 178-194.
15. Tambunan, U.S.F., A.A. Parikesit, Hendra, R.I. Taufik, F. Amelia and Syamsudin. In silico analysis of E DENV-2 and E DENV-3 protein as the backbone of dengue virus tetravalent vaccine by using homology modeling method. *OnLine J. Biol. Sci*., 2009, 9: 6-16.

About Corresponding Author: Mr. Pramod Kumar Yadav

Mr. Pramod Kumar Yadav graduated in Pharmacy from CCS University, Meerut, India and post graduated in IT (specialization in Bioinformatics) from Indian Institute of Information Technology, Allahabad, India. His main area of research includes, molecular modeling, docking, structure based drug design, molecular dynamics simulation, chemoinformatics and epitope prediction. Currently working as Assistant Professor in the Dept. of Computational Biology & Bioinformatics at the Sam Higginbotom Institute of Agriculture, Technology & Sciences - Deemed University, Allahabad, India.