Available online at www.elixirpublishers.com (Elixir International Journal)

Bio Physics

Elixir Bio. Phys. 39 (2011) 5021-5025

ISSN: 2229-712X

Computer aided epitope prediction for glycoprotein-B in human cytomegalovirus

Pramod Kumar Yadav[1] and Jayant Rana[2]

[1]Department of Computational Biology & Bio Informatics, Jsbb, Shiats, Allahabad-211007, India

[2]Jacob School of Biotechnology & Bio Engineering, Shiats, Allahabad-211007, India.

ARTICLE INFO

Article history:

Received: 2 August 2011;

Received in revised form:

23 September 2011;

Accepted: 30 September 2011;

Keywords

Glycoprotein-B,
Human cytomegalovirus,
Epitope,
Modeling,
Docking.

ABSTRACT

Human cytomegalovirus (HCMV), a ubiquitous and opportunistic pathogen, is a member of the Herpesviridae family of viruses. Infection with HCMV is generally asymptomatic, but naïve or immunosuppressed individuals, such as neonates, AIDS patients, and transplant recipients, often manifest serious disease. In the present work, MHC class-I, MHC class-II and B-cell epitopes for the envelope glycoprotein B (gB) of HCMV were predicted using the ProPred1, MHC2Pred and ABCpred servers respectively. The 3D structures of predicted epitopes were modeled using the HHpred server. In order to find the most relevant epitopes among all predicted T-cell epitopes, protein-protein docking was carried out for MHC-I and MHC-II receptors respectively. The energy score for every docked complex was calculated using the Hex 6.0 program. The lower energy score reveals higher binding affinity towards the receptor. It was found that the epitopes 'YLFKRMIDL' and 'KYGDVVGVN' possess highest binding affinity for MHC-I and MHC-II receptors respectively. For B-cell, the peptide 'HVTSSEAVSHRANETI' was highest ranked epitope. These predicted epitopes might be promising candidates for vaccine design against HCMV.

Introduction

Human cytomegaloviruss (HCMV), a ubiquitous, opportunistic pathogen, is a member of the *Herpesviridae* family of viruses and is alternatively known as human herpesvirus-5 (HHV-5).[1] Following infection, the virus resides in endothelial cells, macrophages, or granulocyte stem cells [2] and may cause reinfection if the host is rendered immunosuppressed. Primary HCMV infection of persons with intact immune systems often results in a self-limiting asymptomatic disease, while HCMV is a significant human pathogen for immunocompromised individuals that is often manifested as severe and debilitating sequelae. [3]

Clinically significant CMV disease frequently develops in patients immunocompromised by HIV, solid-organ transplantation, and bone-marrow transplantation. Additionally, congenital transmission from a mother with acute infection during pregnancy is a significant cause of neurological abnormalities and deafness in newborns. Symptomatic disease in immunocompromised individuals can affect almost every organ of the body, resulting in fever of unknown origin, pneumonia, hepatitis, encephalitis, myelitis, colitis, uveitis, retinitis, and neuropathy. [4] Infection of immunocompetent persons has had limited consequences in the vast majority of cases, indicating the importance of a functional immune response in the control of HCMV infections. [5]

Glycoprotein B (gB) is the most abundant constituent of the viral envelope and is a potent immunogenic HCMV protein. [6, 7] gB plays key roles in the process of CMV entry into host cells. This process is a multistep cascade beginning with attachment of the virus to the cell surface and ending with fusion of the virus envelope with the cell plasma membrane. Attachment of CMV to host cells is mediated by an initial interaction between viral gB and/or gM and cell surface heparan sulphate proteoglycans. [8] The antibody binding sites on gB have been studied by several laboratories, and both conformationally dependent and linear binding sites have been described. [9] Extensive strain- specific virus neutralization has been observed in the vast majority of studies that have employed against gB and gH in the neutralization of different clinical HCMV isolates, and some of the B-cell epitopes involved have been characterized. [10, 11, 12, 13] The aim of present work was to predict the T-cell and B-cell epitopes for envelope glycoprotein B (gB) of the human cytomegalovirus (strain AD169) and to perform the comparative docking study of all predicted epitopes (small peptides) with their respective 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. [14]

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. For the development of MHC binder prediction method, an elegant machine learning technique SVM has been used.[15]

Tele: +91-532-2434046
E-mail addresses: pkyadav.bioinformatics@gmail.com

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 artificialneural 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). [16]

HHPred server: HHsearch is a program for protein sequence searching that is free for non-commercial use. [17] 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 [18] (http://toolkit.tuebingen.mpg.de/hhpred).

Hex 6.0: 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. [19]

Methodology

The amino acid sequence of envelope glycoprotein B (gB) of human cytomegalovirus was retrieved from the NCBI's GenPept protein database (Accession No. AAA45926) (http://www.ncbi.nlm.nih.gov/). Subsequently, the sequence was converted into its FASTA format. For the prediction of MHC class-I epitopes, ProPred1 online prediction tool was used. The sequence was submitted to the ProPred1 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. Only those epitopes having a peptide score above the threshold value (4 %) had been selected. For the MHC class-II, epitopes were predicted for different alleles (HLA-DR1, HLA-DR4 & HLA-DR9) using the MHC2Pred online tool. Threshold 0.0 as cutoff score was set for the MHC-II epitopes and only those epitopes having a peptide score above the threshold value had been selected. For the B cell, epitopes were predicted using the ABCpred online tool. The predicted B cell epitopes were ranked according to their score obtained by trained recurrent neural network. All predicted epitopes were obtained in the form of small peptide sequences.

The three dimensional structures of all predicted epitopes (small peptides) were obligatory in order to perform the protein-protein docking. The 3D structures of all predicted small peptides were modeled by using the online 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 program. Then the 3D structures of small peptide (epitopes) were evaluated by using the SAVES server (http://nihserver.mbi.ucla.edu/SAVES/). The modeled small peptides were obtained in the PDB file format and their 3D structures were visualized by the VMD program. Protein-protein (Receptor-epitope) docking was carried out to find the epitopes with maximum binding affinity (lowest energy score) for the respective receptor. The three dimensional structures of MHC class-I (pdb id: 2X4O) and MHC class-II (pdb id: 1T5W) receptors were retrieved from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). Hex 6.0 program was used for protein-protein (receptor-epitope) docking. The receptor and the small peptide epitopes were loaded in the Hex 6.0 program. The receptor and the small peptide epitopes were docked by activating the docking option present in the tool bar of the HEX. The energy scores of all the docked complexes were calculated. The docked complexes were saved and a comparative analysis was carried out on the basis of energy score (KJ/mol).

Results & Discussion

The epitopes predicted for T-cell (MHC-I and MHC-II) and B-cell have been shown in the tabular form.

MHC class-I binding epitope prediction

Tables 1-4 show the predicted MHC class-I epitopes for various HLA alleles. These small sequence epitopes were obtained from the ProPred1 server which allows the prediction of standard proteasome and immunoproteasome cleavage sites in an antigenic sequence. It identifies the MHC binders which have cleavage sites at the C-terminus

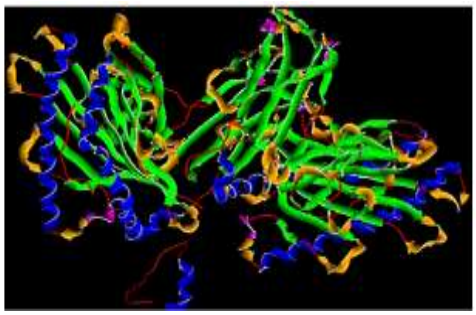

Figure 1: Docked complex of epitope 'YLFKRMIDL' with MHC-I receptor

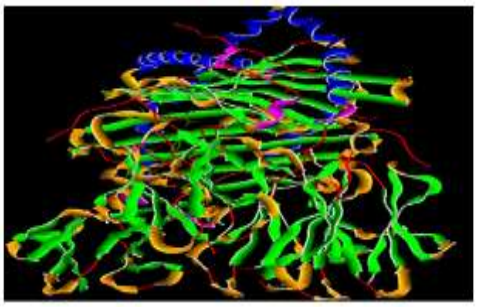

Figure 2: Docked complex of epitope 'KYGDVVGVN' with MHC-II receptor

Discussion

Human cytomegalovirus (HCMV) remains a significant pathogen in individuals with an immature or compromised immune system. In contrast, infection of immunocompetent persons has had limited consequences in the vast majority of cases, indicating the importance of a functional immune response in the control of HCMV infections.[5] One way to boost the immune response towards a specific antigen is the administration of peptides derived from this antigen which are recognized by MHC molecules. The present work was

conducted to predict B-cell and T-cell epitopes and to find out the most relevant epitope with maximum binding affinity for the envelope glycoprotein B of HCMV. The epitopes were predicted using different software in the form of small peptides, as shown in the tables 1-8. Our next task was to find the most efficient epitope among the predicted ones, which was having greatest binding affinity towards the MHC class-I, MHC class-II receptors, and which could be used for efficient vaccine design against HCMV. Epitopes of gB were docked with MHC class-I, MHC class-II receptors using the Hex 6.0 program. Earlier, similar *in silico* epitope prediction work has been carried out for Japanese Encephalitis Virus (JEV) by Kolaskar and Kulkarni. [20]

For MHC class-I epitope prediction, ProPred1 server was used. Three epitopes were predicted for different alleles of MHC class-I (HLA-A1, HLA-A2, HLA-A*0201 and HLA-A*0205). For HLA-A1, it was found that the epitope 'NSAYEYVDY' has the highest peptide score of 1.500. For HLA-A2, the epitope 'YLFKRMIDL' has the highest score of 418.024. For HLA-A*0201, the epitope 'YLFKRMIDL' has highest score of 836.253. For HLA-A*0205, the epitope 'YLFKRMIDL' has the highest score of 126.000 (Tables 1-4). Out of the three predicted epitopes for each allele of MHC class-I, top ranked epitopes were selected for protein-protein docking with the human MHC Class-I receptor (pdb id: 2X4O). After docking, the energy score of different MHC Class-I epitopes were calculated (Table 9). The peptide 'YLFKRMIDL' showed the minimum energy score of -243.02 KJ/mol which reveals highest binding affinity towards the MHC class-I receptor (Figure 1). Interestingly we found that the predicted epitope 'YLFKRMIDL' was top ranked epitope for HLA-A2, HLA-A*0201 and HLA-A*0205 alleles. Therefore, the peptide sequence 'YLFKRMIDL' might be an efficient epitope for vaccine design against HCMV. For MHC class-II epitope prediction MHC2Pred server was used. Three epitopes were predicted for different alleles of MHC class-II (HLA-DR1, HLA-DR4 and HLA-DR9). For HLA-DR1, it was found that the epitope 'YINRALAQI' has the highest peptide score of 1.064. For HLA-DR4, the epitope 'LVAFLERAD' has the highest peptide score of 1.576. For HLA-DR9, the epitope 'GITVTSGST' has the highest peptide score of 1.813 (Tables 4.5, 4.6 and 4.7). Out of the three predicted epitopes, top ranked epitope was docked with the human MHC Class-II receptor (pdb id: 1T5W). After protein-protein docking, the energy score of different MHC Class-II epitopes were calculated (Table 10). The peptide 'KYGDVVGVN' showed the lowest energy score of -597.80 KJ/mol, which reveals highest binding affinity towards the MHC class-II receptor (Figure 2). Therefore, the peptide sequence 'KYGDVVGVN' could be a promising epitope for vaccine design against HCMV. Lastly, for B cell epitope prediction ABCpred server was used. Twenty five epitopes were predicted and ranked on the basis of their score obtained by trained recurrent neural network (Table 8). Those predicted epitopes, which were having a score more than 0.90 might be most suitable predicted epitopes for the B-cell. The peptide sequence 'HVTSSEAVSHRANETI' was having highest score of 0.95 which could be an efficient epitope for B-cell.

Conclusion

The epitopes of gB of HCMV for B-cell and T-cell were predicted using ABCpred, ProPred1 and MHC2Pred servers respectively. The predicted epitopes were modelled by homology medelling method followed by protein-protein docking with the respective MHC-I and MHC-II receptors. It was found that the epitope 'YLFKRMIDL' has the least energy score (-243.02 KJ/mol) for MHC-I receptor and epitope 'KYGDVVGVN' has the least energy score (-597.80 KJ/mol) for MHC-II receptor which reveals highest binding affinity for the respective receptors. The epitope 'HVTSSEAVSHRANETI' was the highest ranked epitope for B-cell receptor. Above predicted epitopes might be promising vaccine candidates for HCMV. Further wet laboratory experiments can be carried out to verify the feasibility of above predicted epitopes.

References

[1] Ryan, K.J. and Ray, C.G. (ed). (2004). *Sherris Medical Microbiology*, 4th ed. McGraw Hill. pp 566.

[2] Plotkin, S.A. (1994). Vaccines for varicella-zoster virus and Cytomegalovirus: recent progress. *Science*. 265:1383-1385.

[3] Alford, C.A. and Britt, W.J. (1990). Cytomegalovirus. In: Fields BN, Knipe DM, Chanock RM, Hirsch MS, Melnick JL, Monath TP, Roizman B (eds). *Fields virology* 2nd ed. New York: Raven Press, pp. 1981-2010.

[4] Gandhi, M.K. and Khanna, R. (2004). Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. Lancet Infect Dis., 4(12):725-38.

[5] Ho, M. (1991). Cytomegalovirus: biology and infection. *Plenum Medical Book co*. 2nd ed. New York. pp 75-78.

[6] Britt, W.J., Vugler, L., Butfiloski, E.J. and Stephens, E.B. (1990). Cell surface expression of human cytomegalovirus (HCMV) gp55–116 (gB): use of HCMV-vaccinia recombinant virus-infected cells in analysis of the human neutralizing antibody response. *J. Virol*. 64:1079–1085.

[7] Marshall, G.S., Rabalais, G.P., Stout, G.G. and Waldeyer, S.L. (1992). Antibodies to recombinant-derived glycoprotein B after natural human cytomegalovirus infection correlate with neutralizing activity. *J. Infect. Dis*. 165:381–384.

[8] Compton, T., Nowlin, D.M. and Cooper, N.R. (1993). Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology* 193:834–841.

[9] Kneiss, N., Mach, M., Fay, J. and Britt, W.J. (1991). Distribution of linear antigenic sites on glycoprotein gp55 of human cytomegalovirus. *J. Virol*. 65:138–146.

[10] Rasmussen, L.E., Nelson, R.M., Kelsall, D.C. and Merigan, T.C. (1984). Murine monoclonal antibody to a single protein neutralizes the infectivity of human cytomegalovirus. *Proc. Natl. Acad. Sci*. 81:876–880.

[11] Baboonian, C., Blake, K., Booth, J.C. and Wiblin, C.N. (1989). Complement independent neutralising monoclonal antibody with differential reactivity for strains of human cytomegalovirus. *J. Med. Virol*. 29:139–145.

[12] Meyer, H., Sundqvist, V.A., Pereira, L. and Mach, M. (1992). Glycoprotein gp116 of human cytomegalovirus contains epitopes for strain-common and strain-specific antibodies. *J Gen Virol*. 73:2375–2383.

[13] Urban, M., Britt, W. and Mach, M. (1992). The dominant linear neutralizing antibody-binding site of glycoprotein gp86 of human cytomegalovirus is strain specific. *J. Virol*. 66:1303–1311.

[14] Singh, H. and Raghava, G.P.S. (2003). ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics*. 19(8):1009–1014.

[15] Bhasin, M., Singh, H. and Raghava, G. P. S. (2003). *MHCBN*: A comprehensive database of MHC binding and non-binding peptides. *Bioinformatics* 19: 665-666.
[16] Saha, S and Raghava G.P.S. (2006). Prediction of Continuous B-cell Epitopes in an Antigen Using Recurrent Neural Network. *Proteins*. 65(1):40-48.
[17] Söding, J., Biegert, A. and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33:W244–W248.
18] Sali, A., Potterton, L5024., Yuan, F., van Vlijmen, H. and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins*. 23:318–326.
[19] Ritche, D.W. and Kemp, G.J.L. (2000). Protein Docking Using Spherical Polar Fourier Correlations. *PROTEINS: Struct. Funct. Genet*. 39:178-194.
[20] Kolaskar, A. S. and Kulkarni, U. (1999). Prediction of three-dimensional structure and mapping of conformational antigenic determinants of envelope glycoprotein of Japanese encephalitis virus. *Virology*. 261:31-42.

Table 1: MHC Class-I Binding Peptides Predicted for HLA-A1 allele

ALLELE: HLA-A1				
Analysis at 4 % threshold [Numerical Value = -0.693]				
Rank	Sequence	Peptide Position	Peptide Score	*Binder/Non-Binder*
1	NSAYEYVDY	3	1.500	*Predicted Binder*
2	AYEYVDYLF	5	0.500	*Predicted Binder*

Table 2: MHC Class-I Binding Peptides Predicted for HLA-A2 allele

ALLELE: HLA-A2				
Analysis at 4 % threshold [Numerical Value = 1.553]				
Rank	Sequence	Peptide Position	Peptide Score	*Binder/Non-Binder*
1	YLFKRMIDL	11	418.024	*Predicted Binder*
2	SAYEYVDYL	4	56.112	*Predicted Binder*
3	YVDYLFKRM	8	2.566	*Predicted Binder*

Table 3: MHC Class-I Binding Peptides Predicted for HLA-A*0201 allele

ALLELE: HLA-A*0201				
Analysis at 4 % threshold [Numerical Value = 1.143]				
Rank	Sequence	Peptide Position	Peptide Score	*Binder/Non-Binder*
1	YLFKRMIDL	11	836.253	*Predicted Binder*
2	SAYEYVDYL	4	24.129	*Predicted Binder*
4	YVDYLFKRM	8	1.520	*Predicted Binder*

Table 4: MHC Class-I Binding Peptides Predicted for HLA-A*0205 allel

ALLELE: HLA-A*0205				
Analysis at 4 % threshold [Numerical Value = 0.519]				
Rank	Sequence	Peptide Position	Peptide Score	*Binder/Non-Binder*
1	YLFKRMIDL	11	126.000	*Predicted Binder*
2	SAYEYVDYL	4	7.560	*Predicted Binder*
3	YVDYLFKRM	8	2.400	*Predicted Binder*

In all the tables (Table 1-4), epitopes which were having a peptide score above the threshold value (at 4%) have been selected. 4% threshold score means that there is 4% chance that our predicted binder is random peptide.

Table 5: MHC Class-II Binding Peptides Predicted for HLA-DR1 allele

ALLELE: HLA-DR1				
Threshold 0.0 as cutoff score				
Prediction method	Rank	Sequence	Residue No.	Peptide Score
SVM	1	YINRALAQI	493	1.064
SVM	2	NAAPETHRL	313	0.977
SVM	3	IYNKPIAAR	531	0.775

Table 6: MHC Class-II Binding Peptides Predicted for HLA-DR4 allele

ALLELE: HLA-DR4				
Threshold 0.0 as cutoff score				
Prediction method	Rank	Sequence	Residue No.	Peptide Score
SVM	1	LVAFLERAD	321	1.576
SVM	2	SHATSSTHN	29	1.253
SVM	3	LVAIAVVII	758	1.230

Table 7: MHC Class-II Binding Peptides Predicted for HLA-DR9 allele

ALLELE: HLA-DR9				
Threshold 0.0 as cutoff score				
Prediction method	Rank	Sequence	Residue No.	Peptide Score
SVM	1	GTTVTSGST	794	1.813
SVM	2	YSRVIGGTV	189	1.659
SVM	3	KYGDVVGVN	77	1.612

For MHC class-II epitope prediction, the epitopes which were having peptide score above the threshold value 0.0 have been selected as an epitope.

Table 8: Results of B-cell epitope prediction

S.No.	Rank	Sequence	Start position	Score
1.	1	HVTSSEAVSHRANETI	56	0.95
2.	2	TTSAQTRSVYSQHVTS	44	0.92
3.	3	CSMAQGTDLIRFERNI	94	0.91
4.	3	TSYNQTYEKYGNVSVF	406	0.91
5.	4	SSLNITHRTRRSTSDN	449	0.90
6.	4	NGTNRNASYFGENADK	281	0.90
7.	4	YSRVIGGTVFVAYHRD	189	0.90
8.	5	DTSLQAPPSYEESVYN	804	0.89
9.	5	TVDSMIALDIDPLENT	643	0.89
10.	5	GRCYSRPVVIFNFANS	571	0.89
11.	5	AQLQFTYDTLRGYINR	481	0.89
12.	5	DSVISWDIQDEKNVTC	329	0.89
13.	6	DSLDGQTGTQDKGQKP	866	0.88
14.	6	TARSKYPYHFFATSTG	256	0.88

The predicted B cell epitopes were ranked according to their score obtained by trained recurrent neural network. Higher score of the peptide reveals the higher probability to be an epitope.

Table 9: Docking energy score of MHC Class-I epitopes

S. No.	Predicted Epitope	Energy Score (KJ/mol)
1.	YLFKRMIDL	-243.02
2.	YVDYLFKRM	-187.39
3.	SAYEYVDYL	-165.70
4.	NSAYEYVDY	-141.61

Table 10: Docking energy score of MHC Class-II epitopes

S. No.	Predicted Epitope	Energy(KJ/mol)
1.	KYGDVVGVN	-597.80
2.	LVAFLERAD	-533.07
3.	GTTVTSGST	-214.67
4.	YSRVIGGTV	-203.21