Global Journal of Technology and Optimization

Alam et al., Global J Technol Optim 2018, 9:2
DOI: 10.4172/2229-8711.1000224

Research Article Open Access

Epitope-Based Vaccine Design against the Outer Membrane Glycoprotein of HCMV

Saruar Alam[1], Mohammad Sayem[1], Md. Ratul Rahman[1], Zinat Sharmin[2], Mahmud Arif Pavel[3] and Md. Faruk Hossain[4*]

[1]*Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka, Bangladesh*

[2]*Department of Biology, Florida International University, Miami, Florida, USA*

[3]*The Scripps Research Institute, Jupiter, Florida, USA*

[4]*Department of Biological Sciences, St. John's University, Queens, NY, USA*

***Corresponding author:** Melese Worku, Department of Biological Sciences, St. John's University, Queens, NY, USA, Tel: +251-913 986 518; E-mail: farukbmb16@gmail.com

Received date: April 27, 2018**; Accepted date:** May 29, 2018; **Published date:** June 06, 2018

Abstract

Human cytomegalovirus is also known as (HCMV), a human beta-herpesvirus 5 which affects over 80% adults worldwide. HCMV infection is considered one of the most leading causes of morbidity and mortality during congenital infection. It can be lethal for the immunosuppressed, such as HIV-infected individuals and patients having their organ transplanted. Considering the clinical importance of HCMV, in this study we aim to design epitope based peptides to develop a vaccine against HCMV by targeting the envelope glycoprotein B of HCMV. We predicted total 22 epitopes and their variability were analysed. The analysis of population coverage reveals a high percentage of the coverage for the America and Europe with an overall 69.77% population coverage for the world. Finally, we predicted the 3D structure of epitope YAYIYTTYL and its binding profile with the human leukocyte antigen serotype (HLA-A*02:01 and HLA-A*02:03) were analyzed using Auto Dock Vina. Our study will facilitate the epitope-based vaccine designing of HCMV and could have a far-reaching benefit on the drug development for HCMV associated diseases.

Keywords Vaccine design; Glycoprotein; Human cytomegalovirus HCMV; Immunoinformatics

Introduction

HCMV, also known as herpesvirus 5 (HHV-5), belongs to the genus Cytomegalovirus of Herpesviridae family. The HCMV genome comprises of a double-stranded DNA with ~230,000 bp [1]. During its infectious cycle its complete genome remains transcriptionally active. However, a huge a part of its transcriptome has nevertheless to be expatiated [2]. HCMV is overwhelmingly cell-associated with granulocytes and macrophages in the blood [3-5].

Cytomegaloviruses (CMV) are often associated with congenital intrauterine infections [6]. Approximately, 1% of live births are affected by HCMV infection while 90% with perinatal infected neonates do not show any symptoms of the disease [7]. The CMV disease can be acquired by natural routes such as horizontal and vertical transmission [8]. The morbidity and mortality related to cytomegalovirus (CMV) contamination in immune-compromised sufferers particularly in HIV-infected sufferers and transplant recipients are well known [9,10]. CMV contamination occurs through casual contact or intimate touch with a person secreting the virus in various body fluids like urine, saliva, or other somatic fluids. CMV can also be transmitted during sexual intercourse and through breast milk. Additionally, it can occur through blood transfusions or transplanted organs and may cause mucoepidermoid carcinoma of salivary glands [3], prostate cancer, [4] and brain tumors [5] in the humans.

Since Human cytomegalovirus (HCMV) exists as a collection of numerous strains, it can re-infect patients with extra HCMV traces. Contaminations with a blend of HCMV strains are present in very common places. Considering the medical importance, research on CMV is essential to understand its pathogenesis. However, the centrality of viral genotypes to develop a vaccine for the HCMV disorder has not been examined yet. Exploring the mixed-genotypes of CMVs is important to know how the reinfections ensued from various HCMV traces and the immunization development for HCMV [11,12]. In this study, we have designed a HCMV vaccine based on the experimentally defined epitopes using bioinformatics workflow and tools (Figure 1).

Methodology

Collection of HCMV specific epitopes

First, immunogenic epitopes of glycoprotein B of HCMV were amassed from an online database Immune Epitope Database and Analysis Resource [IEDB] (http://www.iedb.org/) web server [13]. The data we included were narrowed by the following criteria: positive T cell assays and human host only. in addition, for MHC-I epitopes we best taken into consideration peptides consisting 9 amino acids since most regarded epitopes processed via class I HLA are 9-mers [14] and for MHC-II epitopes we only considered those with 15 residues.

Data collection and generation of multiple sequence alignment of HCMV-glycoprotein B through clustering

To build a multiple sequence alignment (MSA), complete HCMV-glycoprotein B FASTA sequences were taken from UniProtKB. Fragment sequences were excluded. We generated the multiple

sequence alignment (MSA) encompassing entire HCMV-glycoprotein B using MUSCLE [15], with default settings. MEGA 7 software package was employed for generating the MSA.

Variability analysis of HCMV-glycoprotein B

Aligned sequences for HCMV-glycoprotein B were subjected to protein Variability Server [PVS] to analyze the sequence variability of the Protein [16]. We chose Shannon entropy (H) as the variability metric [17], and the variability threshold was set at 0.5. Afterward, we picked only those conserved epitopes that precisely coincided with the generated sequences and consensus throughout their whole length.

Determination of population coverage by these epitopes

In order to trace minimal sets of epitopes (optimal epitope combinations) with a target PPC we first obtained Class I binding profiles assigning the IEDB class I HLA binding prediction tool web server. We selected a class I HLA reference set as these alleles prevalence frequently in the population [18]. For each MHC-II epitope, we predicted HLA II binding affinities to different alleles found in the human population [19], using IEDB class II HLA binding prediction server.

Only those epitopes having a predicted score of IC50<50nM were retained. Population protection coverage (PPC) calculation was performed using the IEDB Population coverage prediction tool web server for 7 populations of interest: Central America, South America, North America, Europe, Southeast Asia, West Africa, and West Indies.

HLA-epitope binding prediction

The 3D structure of MHC class I molecule; HLA-A*02:01 (PDB ID: 4U6Y) and HLA-A*02:03 (PDB ID: 3OX8) were retrieved from an online database Protein Data Bank web server and processed via the PYMOL software. Before the docking experiment, all the water molecules present in the HLA protein were removed using pymol. Additionally, B chain was removed from HLA-A*02:01 and C and F chains were removed from the HLA-A*02:03. The 3D structure of YAYIYTTYL was predicted by the PEP-FOLD server: predict de novo structure of peptide. The Vina wizard tool of a Pyrx software package (version 1.5.6) was used for docking analysis. Both the protein [HLA-A*02:01] and [HLA- A*02:03] were docked against ligand (epitope YAYIYTTYL) files were first converted into PDB format before docking study. The grid space for the center was set manually to increase the interaction of an epitope at the binding groove of HLA-A*02:01 and HLA-A*02:03.

Result and Discussion

Primary selection of Epitopes

Primary epitopes were selected from IEDB database and all the possible strains of HCMV were considered. To find out the epitopes, we applied these search criteria was positive T cell assays and human host only, epitopes of glycoprotein B of HCMV strains AD169 and 10359. HCMV strains AD169 and 10359 are the most common strains. We considered both the MHC class 1 and MHC class 2 epitopes. We found 8 epitopes for the MHC I each having 9 amino acids, while 6 were found for the MHC II each containing 15 amino acids (Table 1).

Secondary selection of Epitopes for MHCI and MHCII based on predicted ANN IC50 score

We exploited and tools to analyze the MHCI and MHCII binding prediction respectively. Throughout our analysis we found only 4 epitopes have ANN IC50 score less than 50 among previously primary selected 8 epitopes for MHC I. Following same procedure, we found only 4 epitopes for MHC II out of previously primary selected 6 epitopes (Table 2).

Final selection of epitopes from PVS result

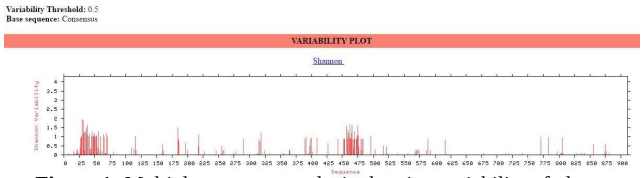

Figure 1: Multiple sequence analysis showing variability of glycoprotein.

For the final selection, we performed multiple sequence analysis followed by analysis of the variability of glycoprotein B using Protein variability server. We found both the glycoprotein B from AD169 and 10359 strains were consensus under the variability threshold of 0.5 (Figure 2). Our result suggests that the amino acids between 25th to 70th positions and 450th to 485th positions of glycoprotein B of both the HCMV strain 10359 and AD169 are highly variable. Other regions are consensus and the amino acids in the window of 625th to 770th are highly conserved. We then selected only those epitopes that found conserved precisely throughout their full-length with the generated consensus sequences. Finally, we found 4 epitopes for MHCI and 2 epitopes for MHCII that satisfy these criteria (Table 3).

Population coverage calculation

Combined Population coverage of the predicted epitopes for both Class I and Class II MHC was calculated by the IEDB population coverage tool at from the IEDB analysis resource. We set 6 epitopes for calculation and combined MHC Class I and MHC Class II (Table 4). Interacting alleles were put against each epitopes as MHC Restricted Allele(s). Total 16 regions of the world as well as the overall world coverage were calculated. Of 16 regions, 7 regions have population coverage more than 60% and overall world population has 69.77% coverage (Figure 3).

HLA-epitope binding prediction

To investigate the interaction between HLA and the selected epitopes, we first determined the 3D structure of the epitopes employing PEP-FOLD an online server. In this case, the YAYIYTTYL epitope was selected for docking because of having maximum interactions with different HLAs. While processing in PEP-FOLD, linear as well as disulfide bonded cyclic peptides with 9-36 amino acids were allowed for the treatment [20]. Again virtual molecular screening was done to dock small-molecule libraries to a macromolecule in order to find lead compounds with desired biological function. PyRx, a software used for performing specific steps for considerations for data preparation, docking study, and data analysis [21]. Using the predicted structures, we employed AutoDock Vina to dock the predicted epitopes to MHC class I molecule, HLA-A*02:01 and HLA-A*02:03. Of 9 predicted possible binding models, here we presented the model with the highest score (Figure 4). Because of the higher binding energy with HLA-A*02:01 and HLA- A*02:03, the best output model for

YAYIYTTYL epitope was predicted to have a binding energy of -7.4 and -8.4 kcal/mol respectively (Figure 4).

Discussion

Glycoproteins (GP) can be the excellent targets for epitope-based vaccine design as they are associated with viral attachment and invasion to the cell [22,23]. The key role of GPs of HCMV during cellular infection makes them a suitable antigenic target for development of vaccine. However, the limited amount of information representing the HCMV worldwide population demands a computational based epitope prediction study as presented here would be efficient and cost-effective in the case of HLA class I molecules.

Traditional vaccines generally take longer time for improvement and have much less energy to distinct between infected and vaccinated cells [24]. In Contrast, peptide-based vaccines are more convenient than other vaccines, such as peptide stability, ease of synthesis, storage, circumventing infectious agents while manufacturing and other molecules can be combined with peptides to augment their immunogenicity [25]. Envelope glycoprotein B is one of the best options which is both available in all strains of HCMV and is virus-specific [23]. In this study, we have predicted HCMV epitopes based on the combination of legacy experimentation consisting of experimentally defined epitopes with immunoinformatics tools.

We employed a strategy to assemble CD8 T cell epitope vaccines and also considered the CD4 T cell epitope. We utilized the experimentally validated epitopes rather than the predicted epitopes and successfully identified epitopes that are more appropriate for epitope based vaccine design. Therefore, our approach is conducive as it saves time and resources. The HLA allele frequencies vary in a population with different genetic backgrounds. Therefore, we target the population that would be administered with the vaccine and considered the candidate epitope that particularly binds with the ubiquitous HLA molecules in that population. This interaction of epitopes can be enhanced by linking with some other molecules which could increase immunogenicity.

In summary, we predicted 6 prospective epitopes of HCMV and identified their peptide binding cores using bioinformatics tools. Further in vitro and vivo studies using these predicted epitopes may pave the way for the development of more effective vaccine against human cytomegalovirus infection.

Conflict of Interests

The authors have declared no conflict of interest with any parties which may arise from this publication.

Acknowledgement

The authors are grateful to Md. Kamrul Hasan for his suggestions to this work.

References

1. Schottstedt V (2010) Human Cytomegalovirus (HCMV)-Revised, Transfus. Med Hemother 37: 365-375.
2. Balázs Z, Tombácz D, Szűcs A, Snyder M, Boldogkői Z, et al. (2017) Long-read sequencing of the human cytomegalovirus transcriptome with the Pacific Biosciences RSII platform. Sci Data 4: 170-194.
3. Melnick M, Sedghizadeh PP, Allen CM, Jaskoll T (2012) Human cytomegalovirus and mucoepidermoid carcinoma of salivary glands: Cell-specific localization of active viral and oncogenic signaling proteins is confirmatory of a causal relationship. Exp Mol Pathol 92: 118-125.
4. Geder L, Sanford EJ, Rohner TJ, Rapp F (2012) Cytomegalovirus and cancer of the prostate: in vitro transformation of human cells. Cancer Treat Rep 61: 139-146.
5. Nauclér CS, Johnsen JI (2015) Cytomegalovirus in human brain tumors: Role in pathogenesis and potential treatment options. World J Exp Med 5: 1-8.
6. Hamprecht K, Maschmann J, Jahn G, Poets CF, Goelz R, et al. (2008) Cytomegalovirus transmission to preterm infants during lactation. J Clin Virol 41: 198-205.
7. Śmiechura M, Strużycka M, Konopka W (2014) Congenital and acquired cytomegalovirus infection and hearing evaluation in children. Otolaryngol Pol 68: 303-307.
8. Forbes BA (1989) Acquisition of cytomegalovirus infection: an update. Clin Microbiol Rev 2: 204-216.
9. Rafailidis PI, Mourtzoukou EG, Varbobitis S, Falagas ME (2008) Severe cytomegalovirus infection in apparently immunocompetent patients: a systematic review. Virol J 5: 1-47.
10. Taylor GH (2003) Cytomegalovirus. Am Fam Physician 67: 519-524.
11. Gasteiger E (2005) Protein identification and analysis tools on the ExPASy Server," in the Proteomics Protocols Handbook, Totowa. NJ: Humana Press 4: 571-607.
12. Wilkinson GWG (2015) Human cytomegalovirus: taking the strain. Med Microbiol Immunol 204: 273-284.
13. Schubert B, Lund O, Nielsen M (2013) Evaluation of peptide selection approaches for epitope-based vaccine design. Tissue Antigens 82: 243-251.
14. Guerra RES, Gomez RN, Alonso ODG, Mendoza S (2015) An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. J Biomed Inform 53: 405-414.
15. Edgar RC (2004) Muscle: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792-1797.
16. Boronat MG, Diez CM, Reinherz EL, Reche PA (2008) PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. Nucleic Acids Res 36: 35-41.
17. Stewart JJ (1997) A Shannon entropy analysis of immunoglobulin and T cell receptor. Mol Immunol 34: 1067-1082.
18. Weiskopf D (2013) Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci 110: 2046-2053.
19. Greenbaum J, Sidney JJ, Chung C, Brander B, Sette A, et al. (2011) Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. Immunogenetics 63: 325-335.
20. Thevenet P, Shen T, Maupetit J, Guyon R, Derreumax P, et al. (2012) PEP-fold: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 40: 288-293.
21. Dallakyan S, Olson AJ (2015) Small-Molecule Library Screening by Docking with PyRx, in methods in molecular biology (Clifton, N.J.) 1263: 243-250.
22. Banerjee N, Mukhopadhyay S (2016) Viral glycoproteins: biological role and application in diagnosis. Virus Disease 27: 1-11.
23. Isaacson MK, Compton T (2009) Human Cytomegalovirus Glycoprotein B is required for virus entry and cell-to-cell spread but not for virion attachment, assembly, or egress. J Virol 83: 3891-3903.
24. Sallusto F, Lanzavecchia A, Araki K, Ahmed R (2010) From Vaccines to Memory and Back. Immunity 33: 451-463.
25. Skwarczynski M, Toth I (2016) Peptide-based synthetic vaccines. Chemical Science 7: 842-854.