Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

Trends in Bioinformatics 6 (1): 10-16, 2013
ISSN 1994-7941 / DOI: 10.3923/tb.2013.10.16

Screening and Structure-based Modeling of T-Cell Epitopes of Marburg Virus NP, GP and VP40: An Immunoinformatic Approach for Designing Peptide-based Vaccine

Ajay Kumar, Amisha Jain, Shraddha and Sitansu Kumar Verma
Department of Biotechnology, Institute of Biomedical Education and Research, Mangalayatan University, Aligarh, UP, 202145, India

Corresponding Author: Ajay Kumar, Department of Biotechnology, Institute of Biomedical Education and Research, Mangalayatan University, Aligarh, UP, 202145, India Tel: +91-9412883081

ABSTRACT

Marburg hemorrhagic fever (Marburg HF) is a severe, often fatal and rapidly progressive zoonotic viral disease seen in humans and non-human primates with high mortality. This study is aimed at conducting the screening, modeling and simulation studies of putative epitopes/peptides of Marburg Virus (MARV) antigenic proteins-Nucleoprotein (NP), Glycoprotein (GP) and membrane associated protein (VP40) for vaccine development. In the present study, immunoinformatic tools ProPred 1, BIMAS and SYFPETHI were used to predict the promiscuous MHC class I epitopes of Marburg NP, GP and VP40. The molecular modeling of the selected immunogenic epitopes by using Modeller and ultimately the simulation studies of epitopes for finding out the energy minimization. The first epitope/peptide DAINSGIDL at position 40-48 showed maximum binding score with B_5101 HLA allele. Second peptide EPHYSPLIL at position 106-104 also showed maximum binding score with B_5101 and the third peptide FLSFCSLFL at position 132-140 showed maximum binding score with A_0201 HLA allele. These three highest scoring nonameric epitopes predicted by ProPred 1 are also evaluated as binders by BIMAS and SYFPETHI. The present study emphasizing on the role of immunoinformatics in computational vaccinology. Vaccination is the most effective measure to treat infectious diseases, therefore, designing epitope based vaccines are easy to produce, more specific, safe and cost effective. These selected epitopes are highly potential to induce T-cell mediated immune response and are expected to be useful in the designing of either a DNA vaccine or a subunit vaccine against Marburg virus.

Key words: Marburg virus, molecular modeling, ProPred 1, T-cell epitope, vaccine designing

INTRODUCTION

Marburg virus is known to be the etiological agent of hemorrhagic fever, belongs to the family of RNA viruses known as Filoviridae and of the order Mononegavirales (Kiley *et al*., 1982; Pringle, 1991). Marburg HF, also known as Green monkey disease or African hemorrhagic fever, caused by Marburg virus, was recently recognized as a serious and recurring problem in humans. Symptoms of MARV infection include general malaise, acute fever, abdominal cramping, bleeding disorders, shock and when become severe, may include jaundice, pancreatitis, severe weight loss, delirium, shock, liver failure, massive hemorrhage and multi-organ dysfunction.

Marburg virus contains a single species, Lake Victoria Marburg virus. Six strains of Lake Victoria MARV had been recognized as of 1990; at least nine genetically distinct strains were identified from a more recent outbreak in the Democratic Republic of the Congo (DRC). Lake Victoria MARV was first discovered in Europe in 1967. Between 1967 and 1994, only six cases of Marburg were reported, all from Africa. In 1998, Marburg virus caused an epidemic disease that affected hundreds of people in the DRC. MARV is indigenous to Africa and in sizeable outbreaks; for instance in Angola from October 2004 to August 2005, an outbreak of Marburg occurred leaving 329 of 374 infected people dead and their human mortality rates have reached ~70% (Bertherat *et al*., 1999; Feldmann and Klenk, 1996; Feldmann *et al*., 1996).

Marburg viruses are enveloped, non-segmented, single stranded negative sense RNA viruses having 19 kb genome sizes which are the largest one found with negative strand RNA viruses. These are generally 80 nm in width, but vary somewhat in length which ranges from 795-828 nm. Their genome encodes the set of seven genes in the order, namely Nucleoprotein (NP), Viral Structural Protein 35 (VP35), VP40, Glycoprotein (GP), VP30, VP24 and RNA-dependent RNA polymerase (L), from the 3' to 5' end (Feldmann *et al*., 1992). Like all filoviruses, Marburg virions are filamentous particles that may appear in the shape of a shepherd's crook or in the shape of a "U" or a "6" and they may be coiled, toroid, or branched (Beer *et al*., 1999).

Out of the seven structural proteins encoded by the MARV genome, three proteins-NP, GP and VP40 are found to be immunogenic. NP is antigenic and having a major role in causing pathogenicity and viral replication. GP is a transmembrane protein having a major role in viral assembly. On the other hand VP40 is the most abundant matrix protein which is necessary for the assembly and budding of new virus particles. Therefore, these 3 proteins are important in mediating and controlling the viral genome activities and its various functions.

Vaccination is the most effective of all the medical interventions to save human and animal lives and to increase production (Horzinek, 1999; Tang *et al*., 2012a). Compared to the conventional vaccines, peptide or epitope based vaccines are easy to produce, more specific, cost effective, less time consuming and also safe. It is well established that T-cells play a critical role in inducing cellular immune response against foreign antigens but they recognize antigenic fragments only when they are associated with Major Histocompatibility Complex (MHC) molecules exposed on surface of all vertebrate cells (Shekhar *et al*., 2012; Mohabatkar and Mohammadzadegan, 2007). MHC molecules are heterodimeric glycoprotein's which present a highly diverse set of peptides on the surface of a cell and induce T-cell activation, thus plays a pivotal role in regulating immune response (Viret and Janeway, 1999; Tambunan and Parikesit, 2011). They are highly polymorphic, so the binding region for one allele may not induce an immune response in a population with different alleles. This implies the need for the identification of the promiscuous viral peptides which can bind to many MHC alleles (Jiao *et al*., 2012).

Immunoinformatics approach uses computational algorithms to predict potential vaccine candidates or T-cell epitopes. Advantage of a peptide or epitope-based vaccine is the ability to deliver high doses of the potential immunogen and at a low cost (Von Hoff *et al*., 2005; Tang *et al*., 2012b). Viral protein which could act as a vaccine candidate must be surface-exposed, antigenic and responsible for pathogenicity (Cerdino-Tarraga *et al*., 2003; Verma *et al*., 2011).

MATERIALS AND METHODS

The *in silico* study was conducted at Department of Biotechnology, Mangalayatan University, Aligarh from Dec, 2011 to June, 2012.

Data collection: Nucleoprotein, glycoprotein and VP40 of Marburg virus are the major immunogenic proteins thus they have been used for immunoinformatic analysis. They are 695, 681 and 303 amino acids in length, respectively. Their amino acid sequences were retrieved from NCBI; accession code AAR85460, AAR85463 and AAR85462, respectively.

Prediction of MHC class-I binding peptides

Use of ProPred 1: The prediction of promiscuous MHC class-I binding peptides was done by using a popular immunoinformatic tool ProPred 1. It is an on-line web tool which uses matrix-based method that allows the prediction of MHC binding sites in an antigenic sequence for 47 MHC class-I alleles. It also allows the prediction of the standard proteasome and immunoproteasome cleavage sites in an antigenic sequence. The simultaneous prediction of MHC binders and proteasome cleavage sites in an antigenic sequence leads to the identification of potential T-cell epitopes. The NP, GP and VP40 protein sequences were analyzed at a threshold setting of 4%. A total of 47 MHC Class-I alleles were taken into consideration. Proteasome and immunoproteasome filters were set "on" at a threshold of 5%. Only peptides with high score were selected for analysis.

Analysis of predicted MHC class I peptides by using BIMAS and SYFPETHI: BIMAS online tool was used to analyze binding of all consensus peptides with 33 human HLA alleles that will help to identify those peptides/epitopes in the targeted proteins that binds to HLA with high affinity (Parker *et al*., 1994). The binding affinity ($T_{1/2}$) value is based on the half time (min) of dissociation of beta 2 microglobulin from HLA. SYFPETHI is also an online tool for prediction of binders for MHC class I molecules.

Structure-based modeling of T-cell epitopes: Molecular modeling and structural analysis (Chaitra *et al*., 2005) were performed for the detection of peptides binding to their respective class I MHC alleles. The modeler (Shi *et al*., 2007) was used for designing the structures of those alleles whose structures were not available in the PDB server while the CPH model server (Eswar *et al*., 2007) was used to design the structures of the predicted binding peptides. After designing the structures docking of selected alleles and peptides performed with the help of Autodock for finding out the energy minimization (Nielsen *et al*., 2010). Then PMV (Python Molecular Viewer) was used for the visualization of binding position of hydrogen bonding between the selected peptides and alleles.

RESULTS

Analysis of MHC class-I peptides by ProPred, BIMAS and SYFPETHI: ProPred 1 was first used to predict the highest scoring epitopes of NP, GP and VP40 which binds with maximum no. of MHC Class-I alleles. After predicting and identifying the epitopes by ProPred 1, the results are evaluated by BIMAS and SYFPETHI for confirmation of predicted binders. The $T_{1/2}$ value was set at cut off $T_{1/2} \geq 100$ for peptide selection. The cut off value of SYFPETHI was set at ≥15 for peptide selection, shown in Table 1. Finally, six peptides are selected viz DAINSGIDL, EPHYSPLIL, FLSFCSLFL, LHLWGAFFL, IPAHPLRML, HPNLPPIVL with the help of scoring based algorithms of ProPred 1, $T_{1/2}$ value of BIMAS and by using SFFPETHI. These peptides have higher binding affinity which was estimated by analyzing the intra-molecular hydrogen bonds, electrostatic, van der walls and hydrophobic interactions with the protein residues in the vicinity or closed proximity.

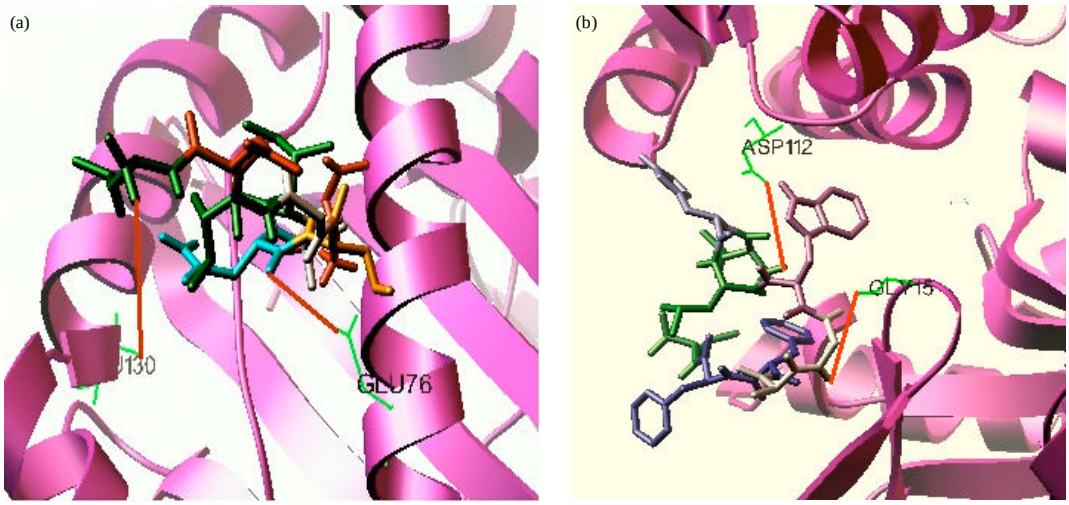

Fig. 1(a-b): Peptide (a) DAINSGIDL and (b) LHLWGAFFL derived from the NP and GP protein sequences, high binder peptides binding to the HLA class I molecule. The peptides Pink color: Very high affinity for the allele B_5101 and B_3901 modeled on to the crystal structure (1A1O and 3BUV) based on the position of the peptide, Potential hydrogen bonds Red color

Table 1: Predicted peptides from target proteins bind to different HLA class I alleles (ProPred 1, BIMAS $T_{1/2}$ = 100 and SYFPETHI value = 15)

Protein	Position	Peptide	Allele	ProPred 1	$T_{1/2}$	SYFPEITHI
	40-48	DAINSGIDL	B_5101	440	110	22
Nucleoprotein	106-114	EPHYSPLIL	B_5101	572	220	19
	132-140	FLSFCSLFL	A_0201	1794	540	21
	137-145	LHLWGAFFL	B_3901	180	180	22
Glycoprotein	440-448	REGDMFPFL	B_60	160	160	20
	508-516	NENDCDAEL	B_60	160	160	21
	130-138	IPAHPLRML	B_5101	520	157	20
VP40	198-206	HPNLPPIVL	B_5101	484	100	21

Sample peptides of high affinity binders for a few alleles whose structures are known (B_5101, A_0201 and B_3901) were modeled employing their respective structural templates (1A1O, 1AO7 and 3BUV).

Simulation studies: Simulation studies of the epitope DAINSGIDL of nucleoprotein and LHLWGAFFL of glycoprotein formed stable HLA-peptide complexes with the energy minimization of -32.65 (kcal mol^{-1}) and -30.22 (kca mol^{-1}), respectively. The other peptide IPAHPLRML identified in the present study was found antigenically variable with energy minimization value of -28.12 (kcal mol^{-1}). This can possibly be targeted for designing of vaccine against Marburg virus. The structures of selected peptides are shown in Fig. 1a, b and the energies of the predicted epitopes with their binding alleles are given in Table 2.

Table 2: Conformational properties of the peptides with efficient binding energy and hydrogen bonds present on the variable regions of predicted peptide as investigated by molecular dynamics simulation using Autodock tool v3.0

Peptide/epitope	Alleles	Energy (kcal mol^{-1})	Hydrogen bonds
DAINSGIDL	B_5101	-32.65	2
EPHYSPLIL	B_5101	-28.04	2
FLSFCSLFL	A_0201	-27.76	2
LHLWGAFFL	B_3901	-30.22	2
IPAHPLRML	B_5101	-28.12	1
HPNLPPIVL	B_5101	-24.43	1

DISCUSSION

With the advent in the field of computational immunology it is now possible to drastically reduce the time for identification of putative and promiscuous antigenic peptides. The present study was undertaken with an objective to initiate the use of immunoinformatics in computational vaccinology (Brusic *et al*., 2002). These approaches are currently used for prediction of antigenic determinants in the protein sequence of NP, GP and VP40 of Marburg virus without using their cultures. The prediction of nonamer epitopes of these Marburg proteins for T cells is recognized against MHC class-I molecules. This prominent study emphasizing to find out the NP, GP and VP40 recurring peptides to allocate the suitable MHC binding peptides which serve as a better lead molecule for vaccine development. Both Propred 1 and BIMAS tools are used to find out the binding peptides and alleles along with their $T_{1/2}$ values (Lal *et al*., 2006; Nayak *et al*., 2001). Hereby main requirement is to model the prerequisite stable structure of selected peptides by accurately determining its MHC binding peptide thus paving initial step for vaccine development (Kavita *et al*., 2010; Mohabatkar and Mohammadzadegan, 2007). The predicted epitopes may be served as a useful diagnostic reagent for evaluating T-cell responses in the context of natural infection and also might be helpful for designing a subunit vaccine against Marburg virus.

CONCLUSION

The screening of putative epitopes using immunoinformatics tools thus suggests that protein NP and VP40 of Marburg virus could be used for preparation of immunological constructs. Molecular simulation and binding tests also suggest that the two nonameric epitopes DAINSGIDL and IPAHPLRML predicted and reported for the first time have considerable binding with MHC molecules and low energy minimization values providing stability to the peptide-MHC complex. These peptide construct will further undergo wet lab studies, for the development of targeted vaccine against Marburg virus strains. Using a similar approach the short listed candidate epitopes for vaccine design using other proteins can also be targeted that would reduce time and experimental expense.

ACKNOWLEDGMENTS

Authors duly acknowledge the motivation and computational facility provided by Mangalayatan University, Aligarh, U.P., India. We are grateful to Dean, I.B.M.E.R., Mangalayatan University, for providing necessary facilities and encouragement. We are also thankful to all faculty members of the Institute of Biomedical Education and Research, Mangalayatan University for their generous help and valuable suggestions throughout the study.

REFERENCES

Beer, B., R. Kurth and A. Bukreyev, 1999. Characteristics of filoviridae: Marburg and ebola viruses. Naturwissenschaften, 86: 8-17.

Bertherat, E., A. Talarmin and H. Zeller, 1999. Democratic republic of the congo: Between civil war and the Marburg virus. International committee of technical and scientific coordination of the durba epidemic. Med. Trop. (Mars), 59: 201-204.

Brusic, V., N. Petrovsky, G. Zhang and V.B. Bajic, 2002. Prediction of promiscuous peptides that bind HLA class I molecules. Immunol. Cell Biol., 80: 280-285.

Cerdino-Tarraga, A.M., A. Efstratiou, L.G. Dover, M.T.G. Holden and M. Pallen *et al.*, 2003. The complete genome sequence and analysis of *Corynebacterium diphtheria* NCTC13129. Nucl. Acids Res., 31: 6516-6523.

Chaitra, M.G., S. Hariharaputran, N.R. Chandra, M.S. Shaila and R. Nayak, 2005. Defining putative T cell epitopes from PE and PPE families of proteins of mycobacterium tuberculosis with vaccine potential. Vaccine, 23: 1265-1272.

Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan and D. Eramian *et al.*, 2007. Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein. Sci.

Feldmann, H., E. Muhlberger, A. Randolf, C. Will, M.P. Kiley, A. Sanchez and H.D. Klenk, 1992. Marburg virus, a filovirus: Messenger RNAs, gene order and regulatory elements of the replication cycle. Virus Res., 24: 1-19.

Feldmann, H. and H.D. Klenk, 1996a. Marburg and Ebola viruses. Adv. Virus Res., 47: 1-52.

Feldmann, H., H. Bugany, F. Mahner, H.D. Klenk, D. Drenckhahn, H.J. Schnittler, 1996b. Filovirus-induced endothelial leakage triggered by infected monocytes/macrophages. J. Virol., 70: 2208-2214.

Horzinek, M.C., 1999. Vaccination: A philosophical view. Adv. Vet. Med., 41: 1-6.

Jiao, W., W. Yin, X. Li and J. Liu, 2012. Research advancement in RV novel vaccine. Asian J. Anim. Vet. Adv., 7: 1234-1244.

Kavita, S., C. Priyanka and M. Namrata, 2010. Bioinformatics approach for screening and modeling of putative T cell epitopes from Por B protein of *Neisseria meningitides* as vaccine constructs. Indian J. Biotechnol., 9: 351-359.

Kiley, M.P., E.T. Bowen and G.A. Eddy, 1982. Filoviridae: A taxonomic home for Marburg and Ebola viruses. Intervirology, 18: 24-32.

Lal, G., M.S. Shaila and R. Nayak, 2006. Activated mouse T cells downregulate, process and present their surface TCR to cognate anti-idiotypic CD4+ T cells. Immunol. Cell Biol., 84: 145-153.

Mohabatkar, H. and R. Mohammadzadegan, 2007. Computational comparison of T-cell epitopes of gp120 of Iranian HIV-1 with different subtypes of the virus. Pak. J. Biol. Sci., 10: 4295-4298.

Nayak, R., S. Mitra-Kaushik and M.S. Shaila, 2001. Perpetuation of immunological memory: A relay hypothesis. Immunology, 102: 387-395.

Nielsen, M., C. Lundegaard, O. Lund and T.N. Petersen, 2010. CPHmodels-3.0--remote homology modeling using structure-guided sequence profiles. Nucl. Acids Res., 38: W576-W581.

Parker, K.C., M.A. Bednarek and J.E. Coligan, 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J. Immunol., 152: 163-175.

Pringle, C.R., 1991. The order mononegavirales. Arch. Virol., 117: 137-140.

Shekhar, C., K. Dev, S.K. Verma and A. Kumar, 2012. In-silico: Screening and modeling of CTL binding epitopes of crimean congo hemorrhagic fever virus. Trends Bioinformatics, 5: 14-24.

Shi, Y., C. Wu, W.Y. Zhou, X.H. Mao, G. Guo and Q.M. Zou, 2007. Identification of H-2d restricted Th epitopes in Urease B subunit of Helicobacter pylori. Vaccine, 25: 2583-2590.

Tambunan, U.S.F. and A.A. Parikesit, 2011. *In silico* design of drugs and vaccines for dengue disease. Trends Bioinform., 4: 1-9.

Tang, H., X.S. Liu, Y.Z. Fang, L. Pan and Z.W. Zhang *et al.*, 2012a. Advances in studies on vaccines of foot-and-mouth disease. Asian J. Anim. Vet. Adv., 7: 1245-1254.

Tang, H., X.S. Liu, Y.Z. Fang, L. Pan and Z.W. Zhang *et al.*, 2012b. The epitopes of foot and mouth disease. Asian J. Anim. Vet. Adv., 7: 1261-1265.

Verma, S.K., S.S.P. Yadav and A. Kumar, 2011. In silico T cell antigenic determinants from proteome of H1N2 swine influenza A virus. Online J. Bioinformatics, 12: 371-378.

Viret, C. and C.A. Janeway Jr., 1999. MHC and T cell development. Rev. Immunogenet, 1: 91-104.

Von Hoff, D.D., D.B. Evans and R.H. Hruban, 2005. Pancreatic cancer. Jones and Bartlett, Sudbury, MA.