Innovare Academic Sciences

International Journal of Pharmacy and Pharmaceutical Sciences
ISSN- 0975-1491 Vol 6, Issue 9, 2014

Original Article

IN SILICO PEPTIDE BASED VACCINE DESIGN AGAINST NON-STRUCTURAL PROTEIN 5 OF HEPATITIS C VIRUS

VIKAS KAUSHIK, GAUTAM CHAUHAN, JOGINDER SINGH*

School of Biotechnology and Biosciences, Lovely Professional University, Phagwara, Punjab, India.
Email: joginder.15005@lpu.co.in

Received: 04 May 2014 Revised and Accepted: 01 Sep 2014

ABSTRACT

Objective: Hepatitis C virus (HCV) is the cause of hepatitis C in human. A hepatitis C infection does not show any noticeable symptoms in the very early stages of the infection, but a chronic infection can ultimately lead to cirrhosis. The chronic condition results in liver failure or cancer. Protein-Protein interactions play a vital role in the pathogenesis of any pathogen. Protein-Protein interactions maps designed and created in this research provide accurate and valuable resource for better understanding of the pathogenicity pathways of Hepatitis C Virus. The objective of the study was to predict the epitope against non-structural protein NS (5a) of Hepatitis C Virus which could be used as suitable vaccine candidate against Hepatitis C virus infections.

Methods: A specific protein-protein interaction is selected on the basis of its significance in the pathway leading to replication of Hepatitis C genome i. e. Interaction between Hepatitis C Nonstructural protein 5A (NS5A) with sh3 domain of Fyn tyrosine protein kinase. Epitopes was predicted and screened by using various bioinformatics tools. Each of the predicted structure was docked with MHC Class I and class II molecules using PatchDock and FireDock.

Results: The MVGLNSYRI epitope was selected on the basis of half life of dissociation and binding score. The average score of half-life of disassociation ($t_{1/2}$) for MVGLNSYRI was 20 hrs, which is the greater than the other epitopes. Structure based modeling of epitopes was done and further the energy was optimized. Then after that the binding score was calculated, which was again best in case of MVGLNSYRI.

Conclusion: These findings conclude that the designed protein-protein interaction maps and predicted epitopes can be of great use in the wet laboratory formulations of vaccines against Hepatitis C Virus.

Keywords: Bioinformatics, Protein-protein interaction, HCV, Vaccine, 3D structure.

INTRODUCTION

Hepatitis C virus (HCV) is the cause of hepatitis C in humans. A hepatitis C infection does not show any noticeable symptoms in the very early stages of the infection, but a chronic infection can ultimately lead to cirrhosis. The chronic condition results in liver failure or cancer. Hepatitis C spread mostly through blood to blood contact, unsterilized equipment used for intravenous drug transfer. The number of people affected by hepatitis C infection worldwide is approximately 200 million per annum which is responsible for hundreds of thousands of deaths each year.

Hepatitis C virus infects only humans and chimpanzees. The knowledge about hepatitis C infection was first reported in the year 1989[1]. It belongs to the genus Hepacivirus and is a member of family flaviviridae [2]. It is an enveloped positive sense single stranded RNA virus with approximately 55-65 nm. Despite the discovery of HCV over 15 years ago, knowledge of the HCV lifecycle has been limited by inability to grow the virus in cell culture, as well as by the lack of small-animal models of HCV infection [3].

HCV has an RNA genome which consists of 9600 nucleotide bases. It consists of UTRs at 5' and 3' ends of the RNA. The genome contains a non-coding region (5'-3') and a coding region. HCV genome comes with a high genetic variability caused due to the mutations that happens frequently during the viral replication. These mutations vary in different genomic regions of HCV [4]. Structural proteins of HCV are core protein (E1 and E2) which are primarily present in the coding region of the genome and non-structural proteins (NS2, NS3, NS4, NS4A, NS4B, NS5, NS5A, NS5B) [2].

The proteins encoded by Hepatitis C genome (structural and non-structural proteins) are primarily responsible for the pathogenesis of Hepatitis C virus. The Hepatitis C viral proteins interact with the proteins present in the human host cells and generate certain signaling pathways which in turn regulate the replication of Hepatitis C genome and its survival in human cells. The no-structural NS1 protein is a hydrophobic transmembrane protein which basically form hydrophilic pores and there by regulating the permeability of the membrane for the propagation of viral assembly and release of viral particles to increase infectivity, NS2 protein works as to attract the envelope proteins to the assembly site and favours viral assembly, NS3 has helicase as well as protease activity. It plays a central role in the process of viral replication. It has its major role in unwinding of viral RNA alone and in complex with NS4A, NS4A functions as cofactor for proper working of NS3 protein for increased enzymatic activities. NS4A also forms complex with NS4B and NS5A to facilitate viral replication on the endoplasmic reticulum membrane, NS4B is associated with NS5B so as to modify its polymerase activity, thus showing its role in carcinogenesis. NS4B favours viral replication as it is involved in the formation of the membranous structure which acts as a platform for viral replication to happen, NS5A is necessary for the viral replication as well as viral assembly, NS5B works as RNA dependent RNA polymerase [5].

The overall pathogenicity of Hepatitis C virus depends upon the interactions between the proteins coded by its genome and their interactions with the human proteins which together result in manipulations to the originally occurring cellular processes and the pathways, thereby increasing pathogenicity. Thus in order to further known the pathogenicity pathways, various protein-protein interactions can be extracted and used in the construction of pathways. Protein–protein interaction maps provide information about the proteins which are highly involved in the viral replication and life cycle. Thus, maps provide a more easy and illustrative way of acquiring knowledge about important interactions between different proteins which are involved in the viral replication.

By studying the interaction map for Hepatitis C protein NS5A various assertions can be made supporting the fact that it is the

protein which is highly involved in the viral replication. Here the main emphasis is on the functioning and regulation of NS5A as it has key roles in various viral processes and thus it emerges as a potential target for vaccine and drug designing using bioinformatics approach.

MATERIALS AND METHODS

Extraction of Protein-Protein interactions Data

Extraction of Protein-protein interactions between different proteins present in Hepatitis C Virus and proteins present in human. For the extraction of protein-protein interaction data "HCVpro" database is utilized.

HCVpro is an online database for the protein-protein interactions between Hepatitis C Viral proteins and human proteins. Highly specific protein interactions extracted in a strain specific manner. HCVpro provide absolute interaction data containing the PubMed identifiers for each interaction.

Preparation of protein specific interaction list

The enormous raw form of protein-protein interactions was extracted from the "HCVpro" database and processed manually in a protein specific manner. Interactions of each of the Hepatitis C proteins are listed separately.

Analysis of interactions maps

Detailed analysis of protein-protein interaction maps was done by studying different interactions between Hepatitis C proteins and human proteins. A specific protein-protein interaction is selected on the basis of its significance in the pathway leading to replication of Hepatitis C genome i. e.

Interaction between Hepatitis C Nonstructural protein 5A (NS5A) with sh3 domain of Fyn tyrosine protein kinase

Analysis of protein-protein interaction

Selected protein-protein interaction is studied to find out type of interaction between NS5A and Fyn tyrosine protein kinase by mining of PubMed research articles. This interaction provided insights that NS5A can be targeted so as to develop the potential vaccine against Hepatitis C Virus.

Prediction of T-cell Epitopes for target protein

Fasta sequence of the protein NS5A is used to predict class I and Class II MHC binding peptides by using "Propred" for MHC-II [6] and "Propred-I" for MHC-I [7]. Propred is an online server for the prediction of MHC II binding peptides for 51 alleles whereas Propred-I is an online server for the prediction of MHC I binding peptides for 47 alleles.

Virtual Screening

In this step, we screened the promiscuous MHC binders separately for both the class of MHC molecules. Predicted T-cell epitopes were screened on the basis of total average score and number of alleles. We obtained 6 peptide sequences (epitopes) after virtual screening which were the binders to both MHC Class-I and MHC Class-II molecule. These were, (DEITFMVGL; WRGDGVMST; VVILDSFEP; MVGLNSYRI; DVSVLTSML; TRCSCGATI) [8].

Tertiary structure prediction for epitopes

Structure prediction is done by "PepFold" which is an online peptide structure prediction server. It uses four consecutive residues and predicts their conformations using greedy algorithm coupled with coarse grained force field. Only one conformation (model) is selected for each peptide on the basis of lowest sOPEP (OPEP-Optimized Potential for Efficient Structure Prediction) energy.

Docking using patch dock and fire dock

Each of the predicted structure was docked with MHC Class I and class II molecules using PatchDock and FireDock. PatchDock provides best binding peptide using the best binding score for each peptide. FireDock refine the PatchDock result and gives the best binding peptide using Global energy of each docking [9]. Best epitopes having the lowest global energy and binding score were selected as potential high affinity binding epitope.

Calculation of Isoelectric point (pI value) and half-life of dissociation of selected epitopes

Isoelectric point and molecular weight of the epitopes was calculated using ExPASy tool online. Lowest Pi value epitopes were selected.

Table 1: Various Protein-protein interactions, corresponding gene, Entrez gene IDs, Human protein names and HCVpro interactions identifiers NS5A sequence of 452 amino acids is used to predict MHC Class I and Class II binding peptides.

HCVpro ID	Molecule A	Molecule B Gene Symbols	Molecule B Gene ID	PMID	Experimental Evidence (PSI-MI Term ID)
hcv0453	NS5A	EIF2AK2	5610	9143277	Two Hybrid Test (MI: 0018)
hcv0477	NS5A	FYN	2534	14993658	Affinity Chromatography (MI: 0004)

Table 2: Docking results of Predicted Peptides with MHC molecule using PatchDock

S. No.	EPITOPES	MHC-Molecule		pI Value	Binding Score		Global Energy	Half life
1.	**DEITFMVGL**	**I**	**II**	**3.67**	**10488**	**9722**	**-1.50**	**1.1 hr.**
2.	TRCSCGATI	I		7.75	8298		215.97	7.2 hrs.
3.	DVSVLTSML	I		3.80	9198		0.84	1.1 hrs.
4.	**MVGLNSYRI**	**I**	**II**	**8.50**	**10704**	**9918**	**-45.23**	**20 hrs.**
5.	VVILDSFEP	II		3.67	8518		2.27	100 hrs.
6.	WRGDGVMST	II		5.84	7716		14.54	2.8 hrs.

Predicted structures of all 6 epitopes namely tertiary structures of MHC Class I (PDB ID: 1HSA) and MHC Class II (PDB ID: 1DLH) were retrieved from PDB (Figure 1). Also the peptide binding to the MHC class I and MHC class II molecule was observed (Figure 2). The peptides (shown in blue color) predicted a have very high affinity for the MHC-I and MHC-II modeled with the help of Gromacs server. The highest binder peptide MVGLNSYRI is derived from the HCV.

RESULTS

HCVpro database is used to retrieve protein-protein interaction data of Hepatitis C Virus in a raw form. The raw data is segregated specifically for each HCV protein in a tabular form having Various Protein- protein interactions, corresponding gene, Entrez gene IDs, Human protein names and HCVpro interactions identifiers (Table 1). Docking results obtained for predicted peptides with MHC molecule using PatchDock are shown in Table 2.

Docking result of MHC-1 and MHC-II with most valid epitopes:

We performed the docking of all the six peptide structures separately each with MHC Class-I (PDB ID: 1HSA) and MHC Class-II (PDB ID: 1DLH) molecule using an online docking tool named as Patch Dock. For the best result the docking score should be

maximum, we found that the docking of the peptide sequence (MVGLNSYRI) with MHC Class-I (PDB ID: 1HSA) showed the top most score of 10704 as compared to other peptide sequences. In second case, the docking of peptide sequence (MVGLNSYRI) with the MHC Class-II (PDB ID: 1DLH) again showed the highest score of 9918 as compared to other peptide sequence.

DISCUSSION

Biological processes are mainly controlled by various mechanisms such as signal transduction through protein-protein interactions and other protein modifications. The construction of major pathogenicity island regulatory pathways in hepatitis C virus subtype NS5A helped here in acquiring the understanding of basic regulation and mechanism of survival in host cells.

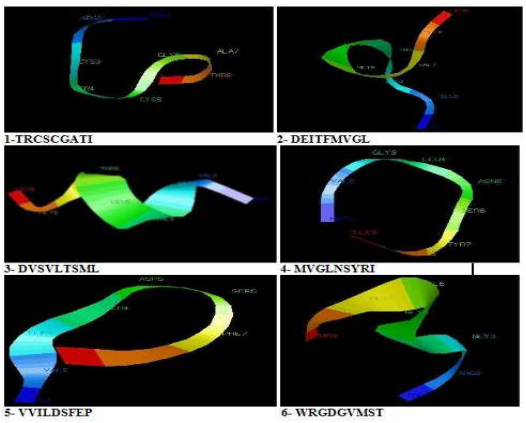

Fig. 1: Predicted structures of all 6 epitopes. Tertiary structures of MHC Class I (PDB ID: 1HSA) and MHC Class II (PDB ID: 1DLH) were retrieved from PDB.

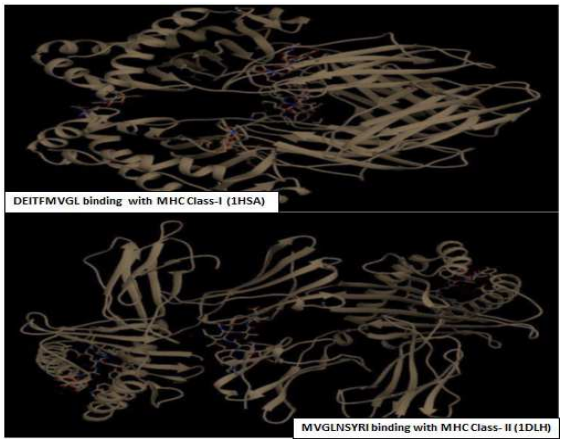

Fig. 2: The peptide binding to the MHC class I and MHC class II molecule. The peptides (shown in blue color) predicted to a have very high affinity for the MHC-I and MHC-II modeled with the help of Gromacs server. The highest binder peptide MVGLNSYRI is derived from the HCV.

This knowledge about the mechanism of survival and regulation is highly necessary for the effective and efficient identification as well as detection and treatment of this viral infection in host. Protein-Protein interactions between different human and HCV proteins provides the insights about proteins that can act as potential target so as to prevent the HCV chronic infection. .The NS5A is necessary for the viral replication as well as viral assembly. The MVGLNSYRI epitope was selected on the basis of half life of dissociation and binding score. In case of the half life of dissociation, more the time of dissociation, better the candidate epitope (Kaushik et al., 2013).

From the table, it can be inferred that, the average score of half-life of disassociation ($t_{1/2}$) for MVGLNSYRI is 20 hrs, which is the greater than the other epitopes. Structure based modeling of epitopes was done and further the energy was optimized. Then after that the binding score was calculated, which was again best in case of MVGLNSYRI. However, as shown in the hypothetical example of planning a vaccine, the discovery of suitable antigens could be a small part of the problem of producing an effective vaccine, but never the less important. The study shows that the predicted epitope can be developed as a potential vaccine candidate.

CONCLUSION

Protein-Protein interactions play a vital role in the pathogenesis of any pathogen. Proteins coded by the Hepatitis C Viral genome interact with different human proteins interfering and modulating the natural functioning of human proteins and other important factors. This promotes Hepatitis C Viral replication and production of large number of viral particles which infect other cells, causing critical antigenic condition. Protein-protein interactions maps designed and created in this research provides accurate and valuable resource for better understanding of the pathogenicity pathways of Hepatitis C Virus. Role of different proteins in the pathogenesis were interpreted using interactions present in the interaction maps. Nonstructural protein 5A (NS5A) is an important protein which is known to be highly involved in the protein-protein interactions during different regulatory pathways leading to the replication of viral genome. For the prediction of an effective vaccine against Hepatitis C Virus, non-structural protein 5A (NS5A) epitopes for MHC Class I and MHC Class II were predicted in this research. These epitopes show high affinity binding with MHC Class I and MHC Class II molecules. In contrast to these epitopes, MVGLNSYRI was found to be the high affinity binding epitope for MHC Class I and MHC Class II repectively. These findings conclude that the designed protein-protein interaction maps and predicted epitopes can be of great use in the wet laboratory formulations of vaccines against Hepatitis C Virus.

CONFLICTS OF INTEREST

Authors declare that there are no conflicts of interest.

REFERENCES

1. Houghton M. The long and winding road leading to the identification of the hepatitis C virus. J Hepat 2009;51:939-48.
2. Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, Gallegos C, *et al.* Genetic organization and diversity of the hepatitis C virus. Proc Nat Acad Sci USA 1991;88:2451-5.
3. Kolykhalov L, Rehak V, Riley LW. Structure and functions of hepatitis C virus proteins: 15 years after. Folia Microbiologica 2006;51:665-80.
4. Steinmann E, Penin F, Kallis S, Patel AH, Bartenschlager R, Pietschmann T. Hepatitis C Virus p7 protein is crucial for assembly and release of infectious virions. PloS Path 2007;3:e103.
5. Huang Y, Staschke K, De Francesco R, Tan SL. Phosphorylation of hepatitis C virus NS5A nonstructural protein: A new paradigm for phosphorylation-dependent viral RNA replication. Virol 2007;364:1-9.
6. Singh H, Raghava GPS. ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinfo 2003;19:1009-14.
7. Singh H, Raghava GPS. Propred: Prediction of HLA-DR binding sites. Bioinfo 2001;17:1236-7.
8. Kaushik V, Singh J, Kaur R, Kumar A. In Silico peptide based vaccine design against non-Structural Protein 3 of Hepatitis C virus. Conf Proc Wilkes100-ICCS 2013:172-3.
9. Schneidman Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. Patch Dock and Symm Dock: servers for rigid and symmetric docking. Nuc Acids Res 2005;33:W363-7.