Biologicals xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biologicals

journal homepage: www.elsevier.com/locate/biologicals

Identification of T cell and B cell epitopes against Indian HCV-genotype-3a for vaccine development- An *in silico* analysis

Varun Chauhan, Mini P. Singh*, Radha K. Ratho

Department of Virology, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, Punjab 160012, India

ARTICLE INFO

Keywords:
Docking
Epitopes
Hepatitis C virus
Human leucocyte antigen
In silico

ABSTRACT

Hepatitis C virus (HCV) infects almost 150 million people and is a leading cause of liver disease worldwide. It has been classified into seven genotypes; the most common genotype affecting Indian population is genotype 3 (60–70%). Currently there is no vaccine for any genotype of HCV. In order to develop peptide based vaccine against HCV, it is important to identify the conservancy in the circulating genotypes, along with the Human Leucocyte Antigen (HLA) alleles in the target population. The present study aims to identify conserved CD4 and CD8 T cells and B cell epitopes against Indian HCV-genotype-3a using an *in silico* analysis. In the present study, 28 promiscuous CD4 T cell epitopes and some CD8 epitopes were identified. The NS4 region was predicted to be the most antigenic with maximum number of conserved and promiscuous CD4 T cell epitopes and CD8 T cell epitopes having strong and intermediate affinity towards a number of HLA alleles prevalent in Indian population. Additionally, some linear B cell epitopes were also identified, which could generate neutralizing antibodies. In order to ascertain the binding pattern of the identified epitopes with HLA alleles, molecular docking analysis was carried out. The authors suggest further experimental validation to investigate the immunogenicity of the identified epitopes.

1. Introduction

The Hepatitis C virus (HCV) affects nearly 3% of the world's population and is second to Human Immunodeficiency virus (HIV) in terms of morbidity and mortality among the emerging infections. The previously used interferon therapy for hepatitis C treatment was expensive, toxic, and required prolonged duration of therapy [1]. Recently the Direct Acting Antivirals (DAAs), have come up which can be given orally, require shorter duration of treatment and are known to achieve better sustained virological response (SVR). Recently the new HCV NS5A replication complex inhibitors of HCV, Ledipasvir and Daclatasvir along with Sofosbuvir (NS5B inhibitor) are usually given to patients infected with HCV genotype 1 and 3 [2,3]. However, resistance to these drugs too has been recently reported by several research groups [4–9]. Thus reinfection remains possible, making control of this leading cause of chronic liver disease difficult. The major hallmark of persistent HCV infection is the lack of B cells capable of generating neutralizing antibodies and CD4 T cells that proliferate in response to its antigenic stimulation [10–14]. There are several studies suggesting that neutralizing antibodies are crucial for protection against HCV infection [15,16].

The recent advances in the field of bioinformatics have contributed to the development of rationally designed peptide based vaccines. Several vaccine targets have been proposed so far for HCV, which have not proved to be promising due to lack of either ethnicity or specificity factors [17]. Genetic heterogeneity of HCV is the major obstacle in vaccine development [18]. The CD4 (HLA Class II) T cell epitopes are the key players of immune defense mechanism known for activation of protective B cell and CD8 T cell responses in host. The HLA molecules too are polymorphic, varying in different population. Hence, it is relevant to identify conservancy in the circulating HCV genotypes along with the HLA alleles prevalent in the target population [19]. Two peptide based vaccines are already running under clinical trials, IC41 has completed a randomized double-blind phase II study in patients with chronic HCV infection who had either relapsed or failed to respond to previous PEG-IFN/ribavirin therapy [20]. Another vaccine, composed of peptides from HCV core region (C35–44) with an emulsified incomplete Freund adjuvant ISA51, has shown to be well tolerated in HCV-infected patients in a phase I clinical trial [21]. Recently in phase II study, the researchers reported the peptide-specific IgG responses in HCV positive hepatocellular carcinoma patients and proposes further clinical trials for this vaccine [22]. These studies indicate that the peptide based vaccines could play an important role against HCV infections.

* Corresponding author. Room No-609, Department of Virology, Research Block A, PGIMER, Chandigarh 160012, India.
E-mail address: minipsingh@gmail.com (M.P. Singh).

https://doi.org/10.1016/j.biologicals.2018.02.003
Received 6 February 2017; Received in revised form 21 February 2018; Accepted 23 February 2018

Table 1
Physico-chemical properties and antigenicity prediction results of all the target proteins.

Protein	Molecular weight	Theoretical pI	Amino acid composition	Antigenicity score (Threshold 0.4)
Core	12963.8	12.05	Ala(A):7.4%; Arg(R):12.2%; Asn(N):3.7%; Asp(D):2.6%; Cys(C):2.1%; Gln(Q):3.7%; Glu(E):2.1%; Gly(G):13.8%; His(H):1.1%; Ile(I):4.2%; Leu(L):9.5%; Lys(K):3.7%; Met(M):0.5%; Phe(F):2.6%; Pro(P):10.6%; Ser(S):5.8%; Thr(T): 4.2%; Trp(W): 2.6%; Tyr(Y):2.1%; Val(V):5.3.%	0.3062 (non-antigenic)
E1	20974.3	6.38	Ala(A):9.4%; Arg(R):4.2%; Asn(N):3.1%; Asp(D):4.2%; Cys(C):3.1%; Gln(Q):4.2%; Glu(E):1.6%; Gly(G):8.4%; His(H):3.7%; Ile(I):4.2%; Leu(L):7.9%; Lys (K):0.5%; Met(M):5.2%; Phe(F):3.1%; Pro (P): 5.2%; Ser(S):5.8%; Thr (T):6.8%; Trp (W):4.2%; Tyr(Y):3.1%; Val(V):12.0%	0.4433 (antigenic)
E2/NS1	38787.1	8.24	Ala(A): 6.9%; Arg(R):5.7%; Asn(N):5.1%; Asp(D):4.0%; Cys(C):5.4%; Gln(Q):2.3%; Glu(E):2.9%; Gly(G):9.4%; His(H):2.9%; Ile(I):3.1%; Leu(L):8.3%; Lys (K):2.3%; Met(M):0.9%; Phe(F):6.0%; Pro(P):7.7%; Ser(S):5.7%; Thr(T):8.0%; Trp(W):3.1%; Tyr(Y):4.0%; Val(V):6.3%	0.4736 (antigenic)
NS3	15423.7	8.70	Ala(A):11.4%; Arg(R):5.4%; Asn(N):0.7%; Asp(D):3.4%; Cys(C):3.4%; Gln(Q):4.0%; Glu(E):2.7%; Gly(G):11.4%; His(H):2.0%; Ile(I):2.0%; Leu(L):9.4%; Lys(K):2.7%; Met(M):2.0%; Phe(F):2.0%; Pro(P):7.4%; Ser(S):6.7%; Thr(T):10.1%; Trp(W):1.3%; Tyr(Y):2.0%; Val(V):10.1%	0.3903 (non-antigenic)
NS4a	5751.7	4.25	Ala(A):9.3%; Arg(R):0.0%; Asn(N):0.0%; Asp(D):3.7%; Cys(C):5.6%; Gln(Q):3.7%; Glu(E):9.3%; Gly(G):11.1%; His(H):1.9%; Ile(I):3.7%; Leu(L):14.8%; Lys(K):3.7%; Met(M):1.9%; Phe(F):0.0%; Pro(P):3.7%; Ser(S):3.7%; Thr(T):1.9%; Trp(W):1.9%; Tyr(Y):5.6%; Val(V):14.8%	0.7566 (antigenic)
NS4b	20221.6	8.94	Ala(A):12.4%; Arg(R):2.6%; Asn(N):3.6%; Asp(D):1.0%; Cys(C):0.5%; Gln(Q):4.1%; Glu(E):3.1%; Gly(G):12.4%; His(H):2.1%; Ile(I):6.2%; Leu(L):12.4%; Lys(K):2.6%; Met(M):3.1%; Phe(F):4.1%; Pro(P):4.6%; Ser(S):5.7%; Thr(T):6.7%; Trp(W):2.6%; Tyr(Y):1.0%; Val(V):9.3%	0.4380 (antigenic)
NS51a	6704.7	9.44	Ala(A)8.1%; Arg(R)6.5%; Asn(N)4.8%; Asp(D)1.6%; Cys(C)6.5%; Gln(Q)1.6%; Glu(E)1.6%; Gly(G)14.5%; His(H)3.2%; Ile(I)4.8%; Leu(L)3.2%; Lys(K) 4.8%; Met(M)4.8%; Phe(F)3.2%; Pro(P)6.5%; Ser(S)4.8%; Thr(T)9.7%; Trp(W)3.2%; Tyr(Y)3.2%; Val(V)3.2%	0.1424 (non-antigenic)

Keeping these points under consideration, the following study aims to identify B cell epitopes and HLA Class II (CD4) T cell epitopes using *in silico* tools which could be utilized for vaccine development against HCV genotype 3 which is the predominant genotype in Indian population.

2. Methodology

2.1. Retrieval of amino acid sequences, molecular and structural analysis

The amino acid sequences of the poly-protein of Indian HCV genotype-3a were retrieved from NCBI database. The names and accession numbers of all the sequences analyzed in the present study were: ADE10208, AGQ17412, AGQ17413, AGQ17414, AGQ17415, AGQ17416, AFH74066, AFH74067, AFH74069, AFH74070, AFH74071, AFH74072, AFA36246, and ADV04529. The amino acid sequences obtained were further aligned using multiple sequence alignment tool Clustal Omega in order to find the conservancy among the proteins of HCV genotype 3a. Different physicochemical properties of the target proteins were also analyzed using ExPASy ProtParam tool (http://web.expasy.org/protparam/). The antigenicity determination of the target proteins was also carried out using VaxiJen v2.0 server, which is used for the prediction of subunit vaccines and protective antigens [23]. Here, the default parameter of the server was used for antigenicity determination. The threshold for antigenecity prediction was kept 0.4.

The secondary structure of the proteins was predicted using the improved self-optimized prediction method (SOPMA) software [24]. The four conformational states, including sheets, coils, helices and turns were analyzed in the protein sequence.

2.2. Identification of T and B cell epitopes

For the identification of HLA class-II T cell epitopes, the servers NetMHCIIpan 3.1, ProPred, MultiPred2, IEDB-NN and IEDB-SMM were used [25–29]. The alleles predominant in Indian population were selected for the study [30–34].

The alleles selected were: HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DRB1*08:03, HLA-DRB1*11:01, HLA-DRB1*12:02, HLA-DRB1*14:01 and HLA-DRB1*15:01. Based on the binding affinity for HLA alleles, the T cell epitopes were classified into the following 3 groups. The peptides having half-maximal inhibitory concentration less than 50 (IC_{50s} < 50 nM) were considered as high-affinity binding epitopes, whereas IC_{50s} of < 500 nM were considered as intermediate-affinity binding epitopes; and IC_{50s} of < 5000 nM were low-affinity binding epitopes. HLA Class I T cell epitopes were identified using Net MHC 4.0 server. The threshold for predicting the strong binding epitopes was kept 0.5% and for weak binding epitopes was kept 2.0% [35].

The identification of linear B cell epitopes was carried out using BCPREDS, ABCpred, BepiPred 1.0, and LBtope online web servers [36–39]. All the servers were run at the default parameters without any alteration in the prediction methods. BLASTP search was also performed against human proteome in order to find any peptides showing identity to human proteins.

2.3. Characterization of the predicted peptides

Different parameters like molecular weight, theoretical isoelectric point (pI), estimated half-life of the predicted T and B cell epitopes were calculated using ExPASy ProtParam tool (http://web.expasy.org/protparam/). The antigenicity of each predicted epitope was analyzed by VaxiJen v2.0 tool (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). The threshold for antigenicity prediction was kept 0.4. The conservancy analysis of the predicted T and B cell epitopes was carried out by IEDB Conservancy Analysis tool [40].