ARTICLE

Cite This: Int. J. Immun and Microbiol. 2021, 1, 15-21

Sciforce

International Journal of Immunology and Microbiology

Journal homepage: www.sciforce.org

Prediction of multi-epitopic domains of a putative oral vaccine against hepatitis C virus

Olatunde Micheal Adeoti,[1a,b,c]* Aanuoluwa Victoria Akinsooto,[2a] Elisabeth Olajumoke Adedokun[3a] and Opeyemi Joy Olaoye[4a]

[a]*The Oke-Ogun Polytechnic, Saki Oyo State Nigeria, Department of Science Laboratory Technology, Microbiology Option*

[b]*Ladoke Akintola University of Technology, Department of Pure & Applied Biology, Ogbomoso*

[c]*Cellular Parasitology Unit, Department of Zoology, University of Ibadan, Nigeria*

ARTICLE INFO

Article history:
Received 20210328
Received in revised form 20210410
Accepted 20210420
Available online 20210429

Keywords:
Immunoinformatics;
Multi-Epitopes;
HCV;
Allergenicity.

ABSTRACT

For vaccine development, triggering an immune response is desired. Designing and assessing vaccine candidates for the appropriate immune response is critical for their success. Hepatitis C virus is the major cause of liver disease. Anti HCV vaccines if designed is rational decision to reinforce specific T-cell as a crucial aspect of effective antiviral treatment. This study explored the use of bioinformatics tools by retrieval of twenty (20) HCV proteins which were selected for vaccine design. These were retrieved from UniProt server based on their antigenicity, virulence, subcellular localization, essentiality non-homology and other physical parameters, including, TM helices, and relative molecular mass. BLASTp revealed 80% non-identity with Homo sapiens genes. The Epitopes obtained include: Q3S781_9HEPC$_{52-71}$, POLG_HCVBK$_{442-461}$, POLG_HCVJA$_{2-21}$, POLG_HCVJ1$_{77-95}$, POLG_HCVCO$_{445-464}$, POLG_HCVR6$_{1107-1126}$, POLG_HCVJP$_{47-66}$, POLG_HCVTW$_{664-683}$, POLG_HCVTR$_{446-465}$, LTOR5_HUMAN$_{23-42}$, POLG_HCVT5$_{100-119}$, POLG_HCVJT$_{77-96}$, HOIL1_HUMAN$_{169-188}$, POLG_HCVJ4$_{644-663}$, POLG_HCVJ8$_{47-66}$, TFB2M_HUMAN$_{49-68}$, RSF1_HUMAN$_{138-157}$, A8DGK3_9HEPC$_{77-96}$, A8DHN1_9HEPC$_{54-73}$, and A8DFL0_9HEPC$_{2-21}$. An antigenicity score of 0.6004 was obtained with the use of VaxiJen server. The allergenicity prediction showed that the vaccine is not allergenic with the use of AllerTOP v.2.0 and AlgPred servers. The molecular weights and theoretical pI of protein were 45.1 kDa and 10.24 kDa respectively. A potentially suitable vaccine candidate with multivariant regions and immunogenic which could be antagonistic to HCV was designed.

*Corresponding author. Tel.: +234-806-728-5175; e-mail: txy23m@yahoo.com

Introduction

Hepatitis C is a disease caused by hepatitis C virus (HCV): the virus can cause both acute and chronic hepatitis, ranging in severity from mild illness lasting a couple of weeks to a significant live long illness. Hepatitis C is that the major explanation for cancer of the liver. Hepatitis C virus (HCV) may be a RNA virus known to infect human and chimpanzees, causing similar disease in these 2 species. HCV is most frequently transmitted parenterally but is additionally transmitted vertically and sexually (MMWR Recomm Rep. 1998). It also requires less exposure that HIV to cause infection (Te HS, Jensen DM. 2010). HCV may be a leading explanation for chronic disease within the world [Williams et al., 2006]. The World Health Organization (WHO) estimates that 170 million people are infected with HCV globally and 3 - 4 million new infections occur annually (Madhava *et al*. 2002). Despite decades of research, there's still no effective vaccine available for HCV thanks to high genetic heterogeneity for the HCV RNA (Manns et al. 2017). Despite its high prevalence and highly infectious nature, HCV remains under-diagnosed and underreported in Africa (with the exception of Egypt). Most of the available data on HCV in Africa are old and outdated. We

therefore began to review available medical literature on HCV in Africa with a view to determining the prevalence, disease burden and customary transmission modes. Additionally, we draw attention to diagnosis, treatment and prevention of HCV.

Recently many studies are done introducing effective vaccines but so far no approved vaccine for HCV infection has been introduced. Generally, there are several approaches to style vaccines against microbial infection including living, inactivated, subunit, toxoids, conjugate, DNA, recombinant vector and peptide vaccines. Fundamental information about the microbe like life cycle, virulence factor(s) and host cell receptor(s) also as practical considerations, like the world where vaccine is employed and prevalence play key roles in select the sort of vaccine (Kallerup, Rie S., and Camilla Foged, 2015) The technological advances within the fields of genomics, proteomics, human immunology, and structural biology have provided the molecular information for the invention and prediction by bioinformatics tools of novel antigens, epitopes, and style of vaccines against pathogenic bacteria, like meningococcus B (Serruto et al.,2012; O'Ryan et al.,2013; Rappuoli et al., 2016).

Conventional vaccines are often characterized by an isolate, inactivate/attenuated, inject archetype of development (Ranjibar et al., 2015).The main target on vaccine design and development has changed to the assembly of peptides composed of multiple epitopes (multi epitope vaccines), supported linear arrangements, as a completely unique alternative. Additionally, epitope-based vaccines have demonstrated various advantages, including safety, the chance to rationally engineer the epitopes for increased potency, breadth, and antigenicity, and therefore the possibility to focus large repertories of immune responses on conserved epitope sequences (Livingston et al., 2002; Oyarzu´ and Kobe, 2015). In the present study, an in silico attempt towards prediction of hypervariable vaccine against *Hepatitis C*. Epitopes are enriched with charged with water-loving domains, its loops are depleted in helix to β-helix. Predicting epitopes is vital to understanding of the basis of immunological differentiation between self and non-self as well as mechanisms of bio-recognition. Much of these multi-epitopic regions are desired for its multi functionality and hyper variability in developing good vaccine candidates.

Abbreviations

CELLO: CELlular Localization; HCV: Hepatitis C- Virus; BLAST: Basic Local Alignment Search Tools

Material and methods

Selection and Retrieving sequence and antigenic evaluation of Protein

Hepatitis C proteins were selected based on reported allergenicity, antigenicity, Virulence and their relatedness to the mechanisms of adhesion. The entire twenty protein sequences were retrieved from UNIPROT reference sequence database in FASTA format. Virulence factors were predicted, gene essentiality as related to essential genes of HCV were downloaded according to Ning et al., 2014, Sayers et al., 2018). For the virus proteins; subcellular localization prediction, CELLO v2.5 and CELLO2GO web server were used (Yu *et al.*, 2006, 2014). The database of GepTop was used to evaluate gene essentiality (Chen *et al.*, 2017).

PATRIC3.5.16 databases was used for study of the virulence role of proteins (Wattam *et al.*, 2017, Garg and Gupta, 2008) and VirulentPred database for SVM based prediction method for virulent proteins in bacterial pathogens was adapted for HCV. Proteins were screened by Vaxign and BLASTp alignment were used to detect sequence homology to *Homo sapiens*. Blasting against the protein sequences was to detect and eliminate proteins that can self-react to human leucocytes or otherwise. Auto-immunity of the selected sequences should be guarded against; a good vaccine candidate must not have similarity/homology with human genes. Also, transmembrane (TM) helices prediction was done by using TMHMM v2.0 server (Krogh *et al.*, 2001). Compute pI/Mw tool was used to calculate the estimated isoelectric point and molecular weight of all amino acid sequences (Wilkins *et al*., 1999).

Phylogenetic Evolution

Construction of Phylogenetic tree and sequence similarity/dissimilarity matrices of the retrieved sequence of the capsid protein of *Hepatitis C* was created using Mega X software. 20 of the protein tree was constructed using maximum likelihood parameter in the software (fig 2). Bootstrap values was detected based on 100 replications.

T cell and B Cell Epitope Prediction

There are several antibodies which are of animal origin especially mouse, monkey or synthetic repertoires that are vital in the diagnosis and treatment of diseases. Designing and engineering of such is of vital importance to wade off infections by non- or infectious agents. This will go a long way to overcome laboratory and facility limitations as well as its economic viability. (Baran et al., 2017). To identify MHC-I binding epitopes, NetMHC 4.0 server was used (Andreatta and Nielsen, 2016). In the alternative, NetCTL 1.2 server could be used to predict CTL epitopes by integrating predictions of proteasomal cleavage, TAP transport efficiency, and MCH class 1 binding (Larsen et al., 2007). Fifty-one human leukocyte antigen (HLA) alleles (HLA-A, -B, -C, and -E) and six murine alleles (H-2) were evaluated. Predictions were calculated for nine-mers epitopes with a threshold for strong binders of 0.5% and a threshold for weak binders of 2%. For MHC-II binding epitopes, NetMHCII 2.3 server; predictions were obtained for 20HLA- DR alleles, 20 HLA-DQ, 9 HLA-DP, and 7 mouse H2 class II alleles using a threshold of -99.9, threshold for the strong binder of 5%, and threshold for the weak binder of 20%. Linear B cell epitopes of 20-mers were predicted utilizing ABCpred with a threshold of 0.7. The second was BCPred server which was applied with a specificity threshold of 75%. For BepiPred server, only amino acids with score >1.0 were considered for the downstream analysis (Jespersen *et al.*, 2017).

Selection and validation of Predicted Epitopes

Epitopes were selected based on the following criteria: (1) 20-mer epitopes, (2) epitopes matching on all algorithms, if possible, and (3) potential to bind with the maximum number of MHC-I and MHC-II alleles. For selection, sequences were aligned and overlapped using Clustal Omega server. Also, predicted epitopes were searched in the IEDB database (www.immuneepitope.org) to find out the already discovered experimental epitopes

Vaccine Design

Vaxign Server (Xian and He, 2013) was used for the vaccine design which shows the protein accession, gene symbol, localization probability, adhesion probability, transmembrane helixes and the protein length.

Protein Prediction and validation of secondary and tertiary structures

The secondary structure of the multi-epitope antigen was predicted using PSIPREDv3.3 (McGuffin *et al.,* 2000). The three-dimensional (3D) structure modeling was performed using Swiss-Model server (Yang and Zhang, 2015). Jmol was used for visualizing 3D structures of proteins. For refinement of 3D model structure, Galaxy Refine and Galaxy Loop were applied (Park *et al.*, 2011).

The best model was validated by the ProSA web (Wiederstein and Sippl, 2007) and ERRAT (Colovos and Yeates, 1993). The residue-by-residue stereochemical qualities of models were validated by Ramachandran plot obtained from PROCHECK server (Laskowski *et al.*, 2012). The best-refined model was selected.

Antigenicity, Allergenicity, Solubility, and physicochemical predictions of Vaccine

For antigenicity prediction, VaxiJen server was used. For allergenicity evaluation, AllerTOP v.2.0 and AlgPred servers were used. For solubility prediction, SOLpro server was used. Finally, ProtParam allowed the computation of various physical and chemical parameters (Wilkins *et al.*, 1999).

ORCID ID: https://orcid.org/0000-0003-1069-861X

Results and Discussion

Protein Selection and Evaluation

Both 9HEPC and HCVJP possessed an inner membrane location, HCVBK, HCVJA, HCVJ1, HCVCO, HCVR6, HCVTW, HCVTR, HCVT5, HCVJT, HOIL1, HCVJ4, A8DGK3, A8DHN1 and A8DFL0 were located in periplasmic region while LTOR5, TFB2M and RSF1 were of cytoplasmic location. Only HCVJ8 was predicted as outer membrane protein as shown in Table1. By convention and traditions, surface and extracellularly located proteins are good to developing a vaccine that is aiming toward prevention of viral infections and diseases (Dwivedi et al., 2016). For genes essentiality designed by GepTop, almost half of the twenty selected proteins were qualified to possess essential genes: HCVJ1, LTOR5, HCVT5, HOIL1, HCVJ8, TFB2M, RSF1 and A8DGK3. These essential genes are needed for survivability of all organisms and are critical for its existence. These essential genes are of particular importance as result of their theoretical and practical relevance for studying the viability performance of a biological system and identifying effective therapeutic targets in pathogens (Chen et al., 2017). Homology analysis of the 20 prioritized proteins using Vaxign showed about 80% non-relatedness. The prediction of the topology of proteins by TMHMM showed that LTOR5_HUMAN, POLG_HCVJ4, TFB2M_HUMAN and RSF1_HUMAN had 0, POLG_HCVT5 had 9, POLG_HCVJA, POLG_HCVJ1, POLG_HCVCO, POLG_HCVJT, POLG_HCVJ4 POLG_HCVJ8 and A8DFL0_9HEPC had 14, Q3S781_9HEPC, POLG_HCVR6, POLG_HCVJP, POLG_HCVTW, POLG_HCVTR and A8DHN1_9HEPC had 12 TM helix, POLG_HCVBK and A8DGK3_9HEPC had 11 TMH located at 2-3 amino acid position.

Table1. The subcellular localization, gene essentiality, virulence, human homology, transmembrane helix, isoelectric point, and molecular weight predictions of hepatitis c selected proteins

Protein	Accession No	Subcellular Localization	Gene Essentiality	Human homology	TM helix	pI/MW(kDa)
Q3S781_9HEPC	tr|Q3S781|	I-1.467	N-ES	NH	12	8.69 /328461.80
POLG_HCVBK	sp|P26663|	P-1.395	N-ES	NH	11	8.48/327193.57
POLG_HCVJA	sp|P26662|	P-1.466	N-ES	NH	14	8.42/327021.39
POLG_HCVJ1	sp|Q03463|	P-1.248	ES	NH	14	8.62/327116.86
POLG_HCVCO	sp|Q9WMX2|	P-1.283	N-ES	NH	14	8.48/326906.20
POLG_HCVR6	sp|Q913V3|	P-1.303	N-ES	NH	12	8.46/326943.08
POLG_HCVJP	sp|Q9DHD6|	I-1.399	N-ES	NH	12	8.49/329985.43
POLG_HCVTW	sp|P29846|	P-1.476	N-ES	NH	12	8.43/327051.23
POLG_HCVTR	sp|Q81487|	P-1.344	N-ES	NH	12	8.47/329738.58
LTOR5_HUMAN	sp|O43504|	C-3.254	ES	H	0	4.69/9613.92
POLG_HCVT5	sp|O92529|	P-1.529	ES	NH	9	8.47/328234.84

POLG_HCVJT	sp|Q00269|	P-1.282	N-ES	NH	14	8.46/326577.54
HOIL1_HUMAN	sp|Q9BYM8|	P-2.379	ES	H	0	5.47/57571.65
POLG_HCVJ4	sp|O92972|	P-1.240	N-ES	NH	14	8.52/326766.79
POLG_HCVJ8	sp|P26661|	O-1.516	ES	NH	14	8.47/330181.85
TFB2M_HUMAN	sp|Q9H5Q4|	C-3.928	ES	H	0	9.30/45348.86
RSF1_HUMAN	sp|Q96T23|	C-1.632	ES	H	0	4.94/163820.51
A8DGK3_9HEPC	tr|A8DGK3|	P-1.319	ES	NH	11	8.53/327199.61
A8DHN1_9HEPC	tr|A8DHN1|	P-1.430	N-ES	NH	12	8.40/326511.53
A8DFL0_9HEPC	tr|A8DFL0|	P-1.490	N-ES	NH	14	8.53/327126.24

aa, amino acid; C, cytoplasmic; E, essential; ES, extracellular; N-ES, nonessential; N-H, non-homology; N-V, nonvirulent; OM, Outer membrane; P, periplasmic; TM, transmembrane; V, virulent.

Table2: Potential antigenic epitopes predicted by different servers

				B cell Epitope			T Cell Epitope	
Order	Protein	Position	Sequence	BCpred	ABCPred	BepiPred	MHC 1	MHC II
1.	Q3S781_9HEPC	52-71	TSERSQPRGRRQPIPKDRRT	0.998	0.9	0.674	3H 1M	0H 0M
2.	POLG_HCVBK	442-461	FYANSFNSSGCPERMAHCRS	0.966	0.81	0.437	49H 2M	25H 0M
3.	POLG_HCVJA	21-Feb	STNPKPQRKTKRNTNRRQD	0.984	0.83	0.49	21H 1M	3H 0M
4.	POLG_HCVJ1	77-95	AQPGYPWPLYGNEGCGWAG	0.978	0.92	0.597	21H 1M	7H 11M
5.	POLG_HCVCO	445-464	HKFNSSGCPERMASCSPIDA	0.951	0.88	0.541	10H 1M	14H 3M
6.	POLG_HCVR6	1107-1126	DLVGWQAPPGSRSLTPCTCG	0.995	0.92	0.619	30H 2M	12H 0M
7.	POLG_HCVJP	47-66	RATRKTSERSQPRGRRQPIP	0.995	0.9	0.501	33H 1M	22H 0M
8.	POLG_HCVTW	664-683	CNWTRGERCDLEDRDRSELS	0.997	0.86	0.555	24H 1M	0H 0M
9.	POLG_HCVTR	446-465	HKFNSSGCPERMSSCKPITY	0.497	0.88	0.551	31H 1M	0H 0M
10.	LTOR5_HUMAN	23-42	CTDSQGLNLGCRGTLSDEHA	0.845	0.59	0.366	6H 0M	0H 0M
11.	POLG_HCVT5	100-119	PRGSRPSWGPNDPRRRSRNL	0.997	0.88	0.582	13H 0M	0H 1M
12.	POLG_HCVJT	77-96	AQPGYPWPLYGNEGLGWAGW	0.967	0.87	0.577	12H 2M	0H 0M
13.	HOIL1_HUMAN	169-188	EPGPPKPGVPQEPGRGQPDA	1	0.94	0.612	22H 0M	5H 0M
14.	POLG_HCVJ4	644-663	CNWTRGERCNLEDRDRSELS	0.994	0.82	0.576	33H 1M	0H 0M
15.	POLG_HCVJ8	47-66	RATRKTSERSQPRGRRQPIP	0.995	0.9	0.596	34H 3M	0H 0M
16.	TFB2M_HUMAN	49-68	PQLWPEPDFRNPPRKASKAS	0.999	0.97	0.641	24H 2M	0H 0M
17.	RSF1_HUMAN	138-157	KNIINEEDADTMRLQPIGRD	0.989	0.82	0.5	37H 2M	0H 0M
18.	A8DGK3_9HEPC	77-96	AQPGYPWPLYGNEGMGWAGW	0.985	0.86	0.535	3H 1M	4H 5M
19.	A8DHN1_9HEP	54-73	ERSQPRGRRQPIPKARQSEG	0.999	0.81	0.696	13H	5H 0M

	C						0M	
20.	A8DFL0_9HEPC	21-Feb	STNPKPQRKTKRNTNRRPQD	0.984	0.83	0.56	15 0M	4H 1M
H – Human; M – Murine								

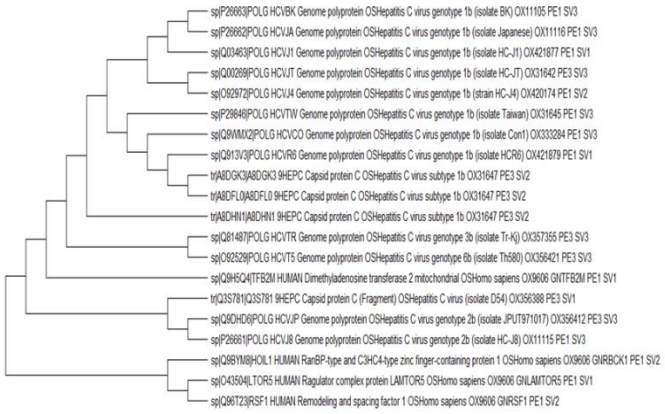
T and B cell epitopes

The prediction of T and B- cell epitopes by different bioinformatics servers for T and B cells (NetCTL 1.2 server and using MHC-I/-II alleles for human and mouse BALB/c allowed the selection of 20 epitopes based on their score, number of alleles, and agreement between the servers. Epitopes obtained using the above mentioned servers include: Q3S781_9HEPC$_{52-71}$, POLG_HCVBK$_{442-461}$, POLG_HCVJA$_{2-21}$, POLG_HCVJ1$_{77-95}$, POLG_HCVCO$_{445-464}$, POLG_HCVR6$_{1107-1126}$, POLG_HCVJP$_{47-66}$, POLG_HCVTW$_{664-683}$, POLG_HCVTR$_{446-465}$, LTOR5_HUMAN$_{23-42}$, POLG_HCVT5$_{100-119}$, POLG_HCVJT$_{77-96}$, HOIL1_HUMAN$_{169-188}$, POLG_HCVJ4$_{644-663}$, POLG_HCVJ8$_{47-66}$, TFB2M_HUMAN$_{49-68}$, RSF1_HUMAN$_{138-157}$, A8DGK3_9HEPC$_{77-96}$, A8DHN1_9HEPC$_{54-73}$, and A8DFL0_9HEPC$_{2-21}$

Protein Structure Prediction

The vaccine is composed of 600 amino acids, and prediction of secondary structure showed that it composed of 7.75% α - helices, 2.75% β sheets, and 89.5% others (random coil and b-turn), as shown in Figure 1.

Figure 1: Phylogenetic tree for the selected capsid protein of Hepatitis C created by Mega X



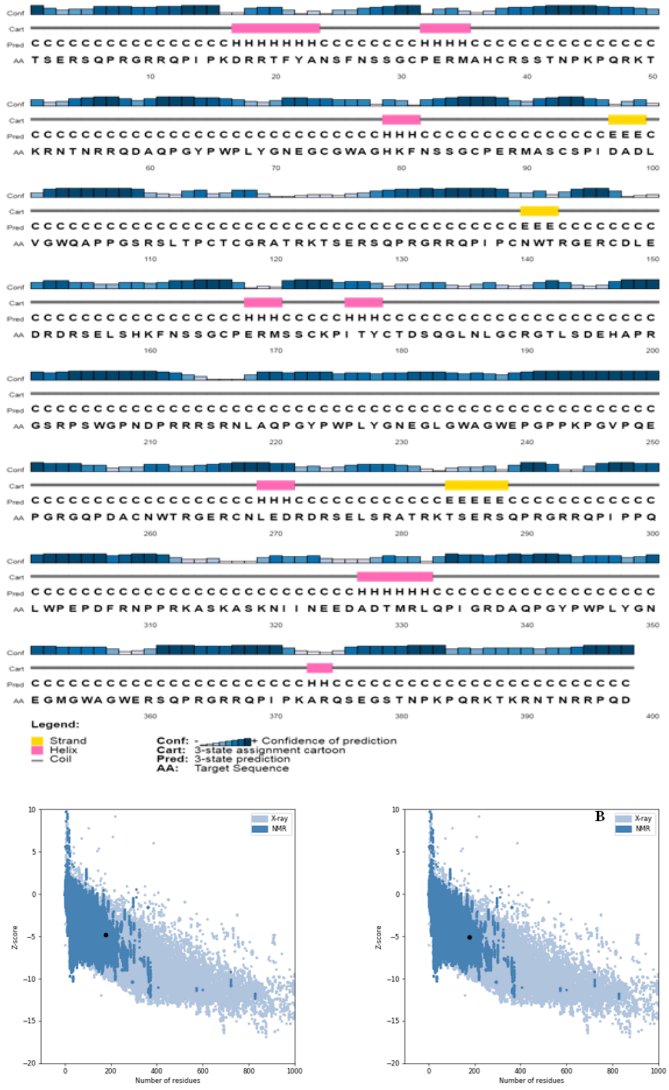
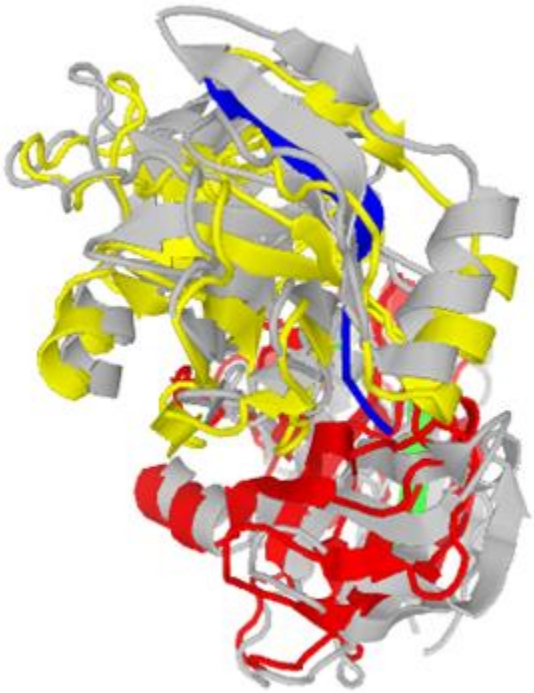


Figure 3: The z-score plot of unrefined and refined 3D structure of vaccine by ProSA-web. (a) The z-score of the starting model is -5.07, (b) The z-score of model after refinement steps is -4.77. the z-scores indicates overall model quality and is depicted as a black spot. The z-scores of all experimentally determined protein chains in current protein data bank (PDB) from NMR spectroscopy (Charcoal) and X-ray crystallography (silver). 3D, three dimensional

Five 3D models of protein vaccine were generated among which the model with the highest c-score = 2 was selected for further refinement; the c-score range is typically from 2-3, the higher the value, the higher the confidence.



Prediction of vaccine by PSIPRED. The protein vaccine consists of 7.75% a helix (H, cylinder), 2.75% b strand (E, arrow), and 89.5% coil (C, line) secondary structural elements. The bar chart represents the percentage of confidence.

The quality and potential errors in the best model were analyzed. The initial input model z-score was -5.07, which falls within those commonly observed in similar size-native proteins (Fig.4A).

ProSA-web indicated that the preliminary model requires refinement processes. Hence, the raw model was subjected to loop refinement and energy minimization using galaxy refine. After all refinement procedures, ERRAT factor was improved from 91.56 to 98.58. The z-score of the final model reached a value of -4.77 (Fig. 4B). The starting models was given (Fig.5).

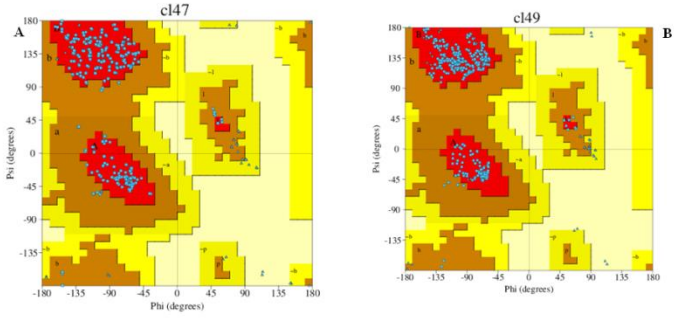


Fig. 5: Validation of vaccine 3D model using Ramachandran plots of (a) the unrefined model and (b) the refined model. The most favored (A, B, and L) and additional allowed (a, b, l and p) regions were demonstrated with charcoal and silver gray colors respectively. The generously allowed regions (-a, -b, -l and -p) are indicated in silver, and the disallowed regions are in white color. Glycine residues are shown in black triangles and other residues of protein are shown in black squares

To validate the 3D models, Ramachandran plot analysis was performed before and after refinement processes. The Ramachandran plots of the unrefined model indicated that 92.8% of residues were located in most-favored regions, 7.2% in the additional allowed region, 0.1% in generously allowed regions, and 0.3% in disallowed regions of the plot (Fig. 6A). The refined model showed that 96.7% of residues were located in most-favored regions, 3.3% in additional allowed regions, 0.0% in generously allowed regions, and only 0.0% in disallowed regions (Fig.6B). Our results indicated that the quality and stability of the final refined model were slightly improved based on Ramachandran plot predictions.

Antigenicity, Allergenicity, Solubility and physicochemical parameters of the vaccine

An antigenicity score of 0.6004 was obtained. The allergenicity prediction showed that the vaccine is not allergenic. The molecular weight and theoretical pI of protein were 45.1 kDa and 10.24, respectively. The recombinant protein vaccine solubility upon over expression in *Hepatitis C* was 0.821589. Half-life was estimated to be 7.2 hours in mammalian reticulocytes, >20 hour in yeast and >10 hour in *Hepatitis C.* The vaccine was found as unstable within instability index of 72.40.GRAVY and aliphatic index were assessed as -1.469 and 30.73, respectively.

Conclusion

In this study, we designed presumptive multi epitopes oral vaccine against *Hepatitis C Virus* based on bioinformatics approach to predict structure that could be capable of provoking cellular and humoral immune response against *Hepatitis C Virus* could be a good vaccine candidate against *Hepatitis C.* However, to make the therapeutic and prophylactic effect of our oral vaccine design valid, *in vitro* and *in vivo* immunological studies are required.

Conflicts of interest

The authors hereby declare no conflict of interests

References

1. Agostini, F.; Vendruscolo, M.; Tartaglia, G. G. Sequence based prediction of protein solubility. *J Mol Biol.* **2012,** 421, 237–241.
2. Andreatta, M.; and Nielsen, M. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics.* **2016,** 32, 511–517.
3. Chan, W. C.; Liang, P. H.; Shih, Y. P. Learning to predict expression efficacy of vectors in recombinant protein production. *BMC Bioinformatics.* **2010,** 11(Suppl 1), S21.
4. Chen, W.-H.; Lu, G.; Chen, X.; OGEE v2: An update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines. *Nucleic Acids Res.* **2017,** 45, 940–944.
5. Davis, G. D.; Elisee, C.; Mewham, D. M. New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnology Bioeng.* **1999,** 65, 382–388.
6. Dwivedi, P.; Alam, S. I.; Tomar, R. S. Secretome, surfome and immunome: Emerging approaches for the discovery of new vaccine candidates against bacterial infections. *World J. Microbiol. Biotechnol.* **2016,** 32, 155.

7. Grimm, S. K.; Ackerman, M. E. Vaccine design: Emerging concepts and renewed optimism. *Curr. Opin. Biotechnol*. **2013,** 24, 1078–1088.
8. Idicula-Thomas, S.; Kulkarni, A. J.; Kulkarni, B. D. A support vector machine-based method for predicting the propensity of a protein to be soluble or to form inclusion body on overexpression in Escherichia coli. Bioinformatics. **2006,** 22, 278–284.
9. Jespersen, M. C.; Peters, B.; Nielsen, M. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **2017,** 45, 24–29.
10. Kallerup. Rie, S.; Camilla Foged. "Classification of vaccines." Subunit vaccine delivery. Springer New York**, 2015**. P 15-29.
11. Krogh, A.; Larsson, B.; Von Heijne, G. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001,** 305, 567–580.
12. Laskowski, R. A.; MacArthur, M. W.; Thornton, J. M. PROCHECK: Validation of protein-structure coordinates. *Int. Tables Crystallogr. F.* **2012,** 684–687.
13. Livingston, B.; Crimi, C.; Newman, M. A rational strategy to design multiepitopeimmunogens based on multiple The lymphocyte epitopes. *J. Immunol.* **2002**, 168, 5499–5506.
14. Madhava, V.; Burgess, C.; Drucker E. Epidemiology of chronic hepatitis C virus infection in sub-Saharan Africa. *Lancet Infect Dis.* **2002,** 2(5), 293-302.
15. Magnan, C. N.; Randall, A.; Baldi, P. SOLpro: accurate sequence- based prediction of protein solubility. *Bioinformatics.* **2009,** 25, 2200–2207.
16. Manns, M. P.; et al. Hepatitis C virus infection. *Nature reviews diseases primers.* **2017,** 3, 17006.
17. McGuffin, L. J.; Bryson, K.; Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics.* **2000,** 16, 404–405.
18. Naz, A.; Awan, F. M.; Obaid, A. Identification of putative vaccine candidates against Escherichia coli exploiting exoproteome and secretome: A reverse vaccinology based approach. *Infect. Genet. Evol.* **2015,** 32, 280–291.
19. Nezafat, N.; Eslami, M.; Negahdaripour, M. Designing an efficient multi-epitope oral vaccine against *Helicobacter pylori* using immunoinformatics and structural vaccinology approaches. *Mol. Biosyst.***2017,** 13, 699–713.
20. O'Ryan, M.; Stoddard, J.; Toneatto, D. A multi-component meningococcal serogroup B vaccine (4CMenB): The clinical development program. *Drugs.* **2013,** 74, 15–30.
21. Oyarzu´n, P.; Kobe, B. Recombinant and epitope based vaccines on the road to the market and implications for vaccine design and production. *Human Vaccine Immunotherapy.* **2015,** 12, 763–767.
22. Park, H.; Ko, J.; Joo, K. Refinement of protein termini in template-based modeling using conformational space annealing. *Proteins.* **2011,** 79, 2725–2734.
23. Rappuoli, R.; Bottomley, M. J.; D'Oro, U. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *J. Exp. Med.* **2016,** 213, 469–481.
24. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. Centers for Disease Control and Prevention. *MMWR Recomm Rep***. 1998,** 47(RR-19), 1-39.
25. Rosa, D. S.; Ribeiro, S. P.; Fonseca, S. G. Multiple approaches for increasing the immunogenicity of an epitope-based anti-HIV vaccine. *AIDS Res. Hum. Retroviruses.* **2015,** 31, 1077–1088.
26. Rossi, G.; Ruggiero, P.; Peppoloni, S. Therapeutic vaccination against *Helicobacter pylori* in the Beagle Dog Experimental Model: Safety, immunogenicity, and efficacy. *Infect. Immun.* **2004,** 72, 3252–3259.
27. Serruto, D.; Bottomley, M. J.; Ram, S. The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: Immunological, functional and structural characterization of the antigens. *Vaccine.* **2012,** 30, 87–97.
28. Singh, H.; Raghava, G. P. ProPred: Prediction of HLA-DR binding sites. *Bioinformatics.* 17, 1236–1237.
29. Singh, H.; Raghava, G. P. 2003. ProPred1: Prediction of promiscuous MHC Class-I binding sites. *Bioinformatics.* **2001,** 19, 1009–1014.
30. Smialowski, P.; Doose, G.; Torkler, P. PROSO II – A new method for protein solubility prediction. *FEBSJ.* **2012,** 279, 2192–2200.
31. Sy, T.; Jamal, M. M. Epidemiology of hepatitis C virus (HCV) infection. Int J Med Sci. **2006**, 3(2), 41-46.
32. Te. H. S.; Jensen, D. M. Epidemiology of hepatitis B and C viruses: a global overview. Clin Liver Dis. **2010,** 14(1), 1-21
33. United Nations Children's Fund (UNICEF). Committing to child survival: promise renewed progress. **2015.**
34. WHO Years of life lost due to premature mortality-trends and causes. World Health Statistics. **2014,** pp 45–49.
35. WHO. Key Facts about Escherichia coli: World Health Progress Report. **2018.**
36. Wilkins, M. R.; Gasteiger, E.; Bairoch, A. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999,** 112, 531–552.
37. Williams, R. Global challenges in liver disease. Hepatology. **2006,** 44(3), 521-526
38. Xiang, Z.; He, Y. Genome-wide prediction of vaccine targets for human herpes simplex viruses using Vaxign reverse vaccinology. *BMC Bioinformatics*. **2013,** 14(Suppl 4), S2.
39. Yu, C. S.; Chen, Y. C.; Lu, C. H. Prediction of protein subcellular localization. *Proteins.* **2006,** 64, 643–651.