ARTICLE

In silico approach for designing a pan-proteomic and pan-genotypic hepatitis C virus multi-epitope subunit vaccine

Nyzar Mabeth O. Odchimar[1], Ella Mae Joy S. Sira[1], Edward C. Banico[1], Alea Maurice Simbulan[1], and Fredmoore L. Orosco*[1,2,3]

[1]Virology and Vaccine Research Program - Industrial Technology Development Institute, Department of Science and Technology, Taguig City, Metro Manila, 1631, Philippines
[2]S&T Fellows Program, Department of Science and Technology, Taguig, Metro Manila, 1631, Philippines
[3]Department of Biology, College of Arts and Sciences, University of the Philippines – Manila, Metro Manila, 1000, Philippines

ABSTRACT

Hepatitis C Virus (HCV) is the most infectious hepatitis virus and the leading cause of liver-related mortality. The lack of a vaccine, combined with HCV's hypervariability, contributes to this global health concern; thus, there is an urgent need for a potent, cross-protective vaccine. With the aid of computational vaccinology, we aimed to design and evaluate multi-epitope vaccines based on highly conserved epitopes from multiple proteins across 8 recognized HCV genotypes and 47 subtypes. The designed HCV vaccine constructs (VC) are composed of antigenic, non-allergenic, non-toxic, non-human homologues, and conserved epitopes within the structural and non-structural HCV proteins, linked with proper adjuvants and linkers. The final HCV VC had an instability index of 33.69. Tertiary structure validation of the modelled HCV VC revealed a high-quality protein. Molecular docking, molecular dynamic simulations, and Molecular Mechanics/Poisson–Boltzmann Surface Area (MM/PBSA) and General Born Surface Area (MM/GBSA) calculations demonstrated a strong and stable binding interaction between the HCV VC and immune receptors (TLR2, 3, and 4). The final HCV VC-TLR3 is the best docking complex (-226.61 kcal/mol) with a highly stable binding affinity in the entire 200 ns molecular dynamics simulation. Notably, immune simulation revealed that HCV VC elicited a high antibody level and stronger IFN-ɣ and IL-2 levels compared to reference studies. This study designed a novel multi-epitope HCV VC that could provide a stronger immune response and cross-protection against hypervariable HCV that requires further in vitro validation.

*Corresponding author
Email Address: florosco@up.edu.ph; orosco.fredmoore@gmail.com
Date received: 14 August 2024
Date revised: 24 February 2025
Date accepted: 10 April 2025
DOI: https://doi.org/10.54645/2025181RAI-34

INTRODUCTION

Hepatitis C virus (HCV) is recognized as the most contagious hepatitis virus (Keikha et al., 2020) and stands as a leading cause of liver-related mortality (Thimme, 2021). Globally, it affects approximately 71 million individuals, contributing to an estimated 400,000 deaths annually (World Health Organization (WHO), n.d.). HCV infection often results in chronic liver diseases and can progress to hepatocellular carcinoma (HCC), the most prevalent type of liver cancer (Shayeghpour et al., 2021).

KEYWORDS

conserved epitopes, Hepatitis C virus, hypervariability, multi-epitope subunit vaccine, vaccine design, immunoinformatics

Recent therapies for HCV with direct-acting antivirals (DAAs) such as sofosbuvir (a NS5B inhibitor) (Nakamura et al., 2016), and PEG-IFN with ribavirin, have shown significant success in curing HCV infection (Rinaldi et al., 2019). However, the high cost, adverse side effects, and difficulties in timely diagnosis due to late symptom manifestation make it challenging for people to access treatment. Even after successful HCV eradication, evidence indicates that HCC can still progress (Luna-Cuadros et al., 2022), necessitating the development of a prophylactic vaccine to prevent HCV infections and address the increasing incidence of HCC. However, there are currently no available vaccines for HCV infection.

HCV, a member of the *Flaviviridae* family, is a minute virus (45-65 nm in diameter) containing only 9.6 kb single-strand positive-sense RNA that encodes for a single polyprotein (Manns et al., 2017). This polyprotein is divided into two groups of proteins: (a) structural (i.e., core protein, envelope glycoproteins 1 (E1) and 2 (E2), p7 ion channel) and (b) non-structural proteins (i.e., NS2, NS3, NS4A, NS4B, NS5A, and NS5B). The structural proteins functions for HCV attachment for host entry during infection, whereas non-structural proteins is required for their replication (Manns et al., 2017).

HCV is a highly variable virus. They are classified into 8 genotypes that are 30-35% genetically distinct from each other and 86 subtypes that differ by 15-25% at the nucleotide level (Borgia et al., 2018; Petruzziello et al., 2016). This profound genetic heterogeneity of HCV is accounted to the lack of proofreading ability of NS5A protein (Irekeola et al., 2021). This characteristic complicates the application of live attenuated and mRNA vaccines due to the lack of specificity and their tendency to reactivate or residual virulence (Li et al., 2020). In June 2020, the National Institutes of Health (NIH) allocated research funding of US$8 million for the development of an HCV vaccine (Echeverría et al., 2021). However, the progress in HCV vaccine evaluation has been limited. Although several clinical trials investigating the use of modified adenoviruses as carriers for HCV genetic material, with the goal of expressing HCV proteins, between 2007 and 2016 (ClinicalTrials.gov Identifiers: NCT01094873, NCT01070407, NCT01701336, and NCT01296451), these trials were either not continued due to the ineffective reduction of HCV viral load or no results were disclosed (Echeverría et al., 2021). To explore alternative vaccine strategies, Chen et al. (2020) developed an experimental HCV vaccine containing glycoproteins E1/E2 [ClinicalTrials.gov ID: NCT00500747] and tested it on healthy volunteers (n = 509), C57BL/6J mice (n = 28), and rhesus macaques (n = 4). The glycoproteins E1/E2 are known to stimulate the secretion of neutralizing antibodies; however, findings from this phase I clinical trial revealed that the immunized participants and study organisms did not produce sufficient antibodies against diverse HCV isolates. These results led to the design of an HCV vaccine aimed at eliciting T-cell responses targeting HCV structural proteins and non-structural proteins NS3-NS5, which advanced to a phase II trial [ClinicalTrials.gov ID: NCT01436357] (National Institute of Allergy and Infectious Diseases (NIAID), 2019; Page et al., 2021). Unfortunately, this study did not prevent HCV infection or inhibited chronic progression of HCV infection. Reports claimed that CD8+ T cell induction was successful. However, potential mismatch of induced HCV-specific response and infecting HCV might have occurred, implicating concerns in quasi-species across subjects (Page et al., 2021). Furthermore, primary HCV isolates exhibit a limited ability to replicate in tissue culture (Adugna, 2023; Bailey et al., 2019). The absence of an effective tissue culture system for HCV replication limits research on its mechanisms for evading the host immune response and obstructs progress in creating potent HCV vaccines. To address these concerns, a multi-epitope subunit vaccine is proposed. Multi-epitope vaccines is a collection of immunogenic peptides that can readily activate both cellular (MHC I and II molecule attachment) and humoral immune response (B cell receptor recognition) to elicit immune response (Aziz et al., 2022; Orosco, 2024). In addition, multi-epitope vaccines have been preferred over conventional vaccine because they are highly immunogenic properties, safer, and can be evaluated using cell-based assays, eliminating concerns associated with the use of live HCV (Chauhan et al., 2019; Kar et al., 2020). Particularly, a multi-epitope subunit vaccine containing conserved epitopes that undergo fewer mutations should be considered countering the problem of variability and optimize cross-efficacy across HCV genotypes (Simbulan et al., 2024). As suggested by Semmo and Klenerman (Semmo & Klenerman, 2007), a highly effective prophylactic vaccine against HCV should contain numerous specific epitopes to induce a stronger response and a more specific vaccine, thereby reducing the risk of immune escape and increasing vaccine efficacy. Clinically used multi-epitope vaccines like UB-612 for COVID-19 (Wang et al., 2024) and RTS,S/AS01 (Laurens, 2020) for malaria exemplify this approach. Thus far, no additional information on HCV vaccine clinical trials, particularly on the HCV vaccine utilizing conserved proteins, has been reported. However, the identification of relevant epitopes still poses a significant challenge. Fortunately, the rise of computational vaccinology has provided a new and effective approach for epitope prediction and vaccine design, offering a promising pathway.

Several studies have explored *in silico*-designed multi-epitope vaccines against highly mutating infectious agents such as HIV-1 (Sher et al., 2023), Chikungunya virus (Mahmoodi et al., 2023), influenza A virus (Tarrahimofrad et al., 2021), and SARS-CoV-2 (Ysrafil et al., 2022). There are reported studies on HCV multi-epitope subunit vaccines. However, mostly employ screening of one HCV genotype (Ahmad et al., 2021; Ikram et al., 2014), or used pan-genotypic screening with HCV 1-6 genotypes (Donnison et al., 2020, 2022; Keck et al., 2018). Although recent studies have demonstrated the use of HCV multi-epitope subunit vaccines using conserved epitopes from both structural and non-structural proteins, and several genotypes (Behmard et al., 2022), no study has explored pan-proteomic and pan-genotypic approach, covering the entire 8 genotypes of HCV. Moreover, none of these HCV study, excluding Ahmad et al. (Ahmad et al., 2021), have an immune simulation analysis, a crucial component in vaccine research.

The present study is the first to use an *in silico* approach to explore the design of an HCV multi-epitope vaccine containing conserved epitopes from 8 genotypes and 47 subtypes. Through molecular docking, molecular dynamics simulation, and Molecular Mechanics/Poisson–Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/General Born Surface Area (MM/GBSA) analyses we explored the ability of the HCV multi-epitope vaccine to trigger an immune response by inducing TLRs (TLR2, TLR3, and TLR4). Further, we have shown that HCV multi-epitope vaccine elicited strong immune responses based on immune simulations. By exploring and targeting multiple conserved epitopes across all HCV genotypes, we designed an *in silico* broad-spectrum vaccine that may provide cross-protection against diverse HCV genotypes and address the increasing incidence of HCV infection and HCV-associated HCC.

MATERIALS AND METHODS

HCV Proteome Retrieval and Phylogenetic Analysis

Non-recombinant HCV proteomes were retrieved from GenBank in NCBI (https://www.ncbi.nlm.nih.gov/) and UniProt (https://www.uniprot.org/). Redundant, aberrant proteome sequences, and proteomes with <3000 amino acids were removed to ensure data quality. A phylogenetic tree was created to validate the genotypic clusters using MEGA 11 software. The flowchart outlining the methodology for this study is presented in Figure 1.

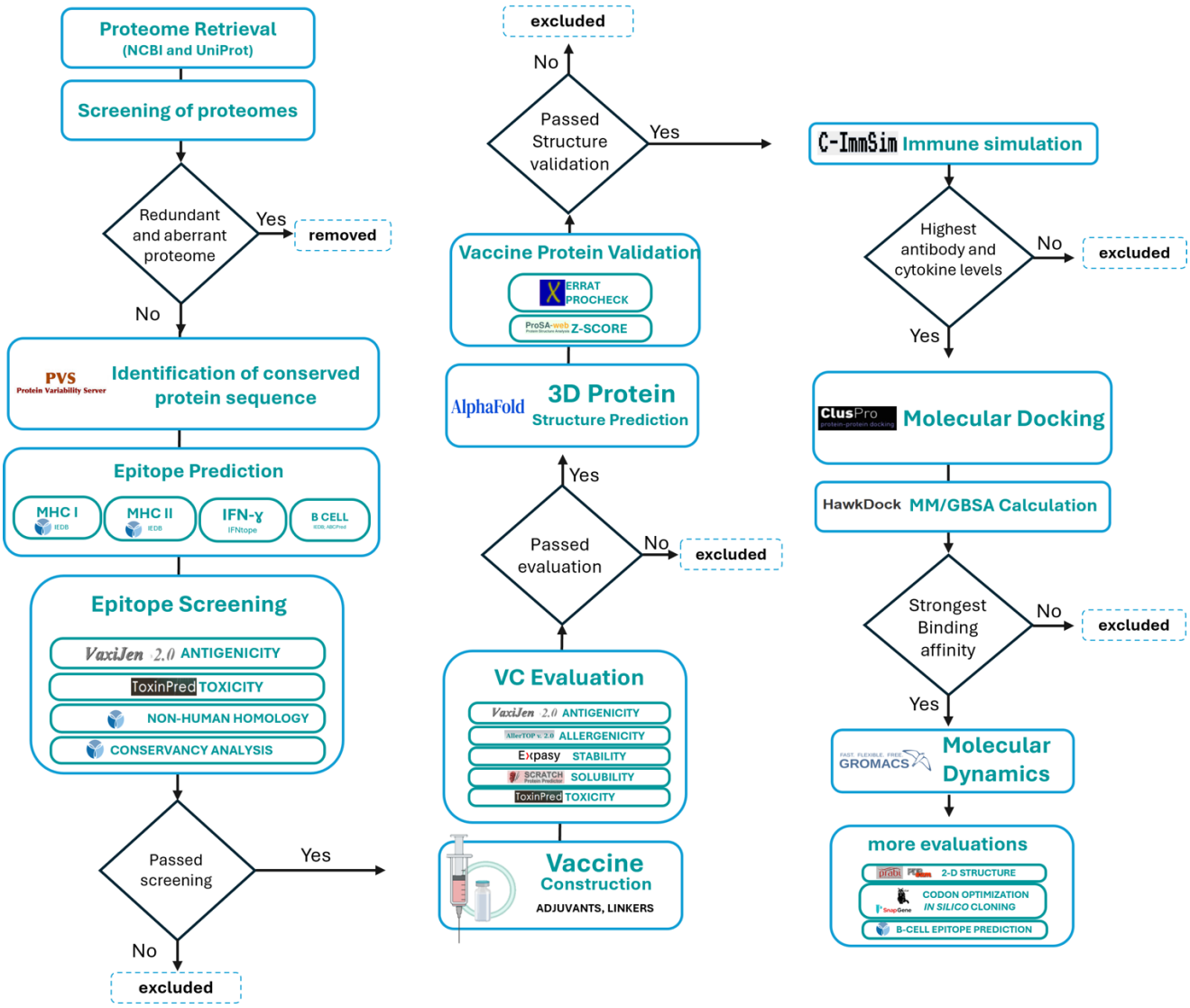

Figure 1: A diagram illustrating the in silico design process for an HCV multi-epitope subunit vaccine.

Identification of HCV Conserved Protein Sequences

To address the hypervariability problem in HCV, we used conserved protein sequences among isolates within each genotype and subtype. To do this, we generated multiple sequence alignments (MSA) using ClustalW (https://www.ebi.ac.uk/Tools/msa/clustalo/) with default parameters for each genotype. Aligned sequences were submitted to Protein Variability Server (PVS; http://imed.med.ucm.es/PVS/) with a Shannon variability threshold (*H*) of >0.5 to obtain the conserved protein sequences in each genotype. The conserved proteins in each genotype were screened separately for MHC I and II-binding epitopes, and B cell-stimulating epitope prediction.

Prediction of MHC I-binding epitopes

MHC I epitopes were generated using IEDB MHC I-binding prediction tool (http://tools.iedb.org/mhci/, accessed on 16 June 2024) for all HLA class I alleles. The IEDB default and recommended prediction method was NetMHCpan EL 4.1, which uses an artificial neural network (ANN) algorithm, a highly reliable prediction algorithm. NetMHCpan EL 4.1 incorporates peptide sequence data and peptide-MHC I molecule binding affinity based on mass spectrometry data (Reynisson et al., 2020). Predicted epitopes with 9-mer size and percentile rank score of ≤0.5 are considered possible good binders to MHC I molecules and were selected for further analysis (Nielsen et al., 2004).

Prediction of MHC II-binding and IFN-ɣ stimulating epitopes

The identification of 15-mer MHC II-binding epitopes was carried out using the IEDB MHC II binding prediction tool server (http://tools.iedb.org/mhcii/, accessed on 16 June 2024) with the default and recommended prediction method (NetMHCIIpan EL 4.1) for the reference set of HLA class II alleles (Greenbaum et al., 2011). Epitopes predicted by this method with ≤2% percentile rank were considered as "strong binders" (Soto et al., 2022) and were selected for further analysis. Furthermore, the predicted MHC II-binding epitopes were evaluated for interferon-gamma (IFN-ɣ) response. The IFN-ɣ -inducing MHC II epitopes were predicted using the IFNepitope server (http://crdd.osdd.net/raghava/ifnepitope/predict.php)

with default settings. Based on SVM method and IFN-gamma vs. non-IFN-gamma model, the MHC II epitopes exhibiting positive IFN-ɣ induction were selected (Dehghan et al., 2020).

Prediction of B cell stimulating epitopes

Two artificial neural network (ANN)-based servers were used to predict linear B-cell epitopes: (1) BepiPred 2.0 (IEDB; http://tools.iedb.org/bcell/; accessed on 21 June 2024) and (2) ABCPred (http://crdd.osdd.net/raghava/abcpred/; accessed on 21 June 2024). The identified B cell epitopes were selected based on the threshold score of ≥0.5 and ≥0.51, for BepiPred 2.0 and ABCPred, respectively. Similarly, using IEDB ElliPro server (http://tools.iedb.org/ellipro/; accessed on 21 June 2024) conformational B cell epitopes were predicted based on solvent-accessibility and flexibility of the query peptide sequence. Residue(s) with above 0.5 threshold score is considered an epitope or part of an epitope.

Evaluation of the Predicted Epitopes

The predicted MHC I, MHC II, and B cell epitopes were screened for their antigenicity property using Vaxijen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). Antigenicity prediction is entirely based on the physicochemical properties of the protein/peptide sequence (Khalid & Ashfaq, 2020). Epitopes above the threshold score (≥0.4) were regarded as antigenic and selected for further analysis. The allergenicity and toxicity properties of the predicted epitopes were also evaluated using Allertop (https://www.ddg-pharmfac.net/AllerTOP/index.html) and ToxinPred 2.0 (https://webs.iiitd.edu.in/raghava/toxinpred/index.html), respectively. Subsequently, the epitopes were screened for homology against human proteins using NCBI's BLASTp RefSeq Protein database against the human reference (Taxonomy ID: 9606) genome with E-value threshold of 0.000001. HCV epitopes with ≥30% similarity on human proteins were excluded (Haq et al., 2021). The epitopes that passed through the evaluations were selected for further analysis.

Epitope Global Conservation Analysis

Herein, we prioritized highly conserved epitopes. With this purpose, epitopes were analyzed using Epitope Conservation Analysis tool (http://tools.iedb.org/conservancy). Using the tool, the epitopes that passed the evaluations were screened against the entire HCV proteomes, covering 8 genotypes. Epitopes exhibiting ≥80% identity match were selected for vaccine designing (Bui et al., 2007).

Designing and Evaluation of Physicochemical Properties of the HCV Multi-epitope Vaccine Constructs

Multi-epitope vaccine constructs (VC) were designed by including adjuvants, linkers, and the set of antigenic, non-toxic, nonhuman homologs, and conserved MHC I, MHC II/IFN-ɣ, and B cell-stimulating epitopes. In the study, we used two different adjuvants and added them to the vaccine sequence separately: (1) human β-defensin (HBD; NCBI Accession ID: P81534) or (2) 50S ribosomal protein (L7/L12; NCBI Accession ID: WP_003974310.1), at the N-terminus of the vaccine construct and linked with EAAAK linker. There were 4 linkers: GGGS, AAY, KK, and/or GPGPG, used in different combinations to connect each epitope. Subsequently, the VCs were subjected to sequence-based physicochemical property screening. This includes Vaxijen 2.0 (≥0.4), AllerTop, and Sol-Pro (≥0.75; http://scratch.proteomics.ics.uci.edu/) for their antigenicity, allergenicity, and solubility property, respectively. In addition, ProtParam server (https://web.expasy.org/protparam/) was used to examine their molecular weight, theoretical isoelectric point (pI), estimated half-life, aliphatic index, Grand Average of Hydropathicity (GRAVY), and instability index of the VCs. The antigenic, non-allergenic, highly soluble, and stable VCs were selected for further analysis.

Prediction of HCV Multi-epitope Vaccine Construct Tertiary Protein Structure and Validation

The selected HCV VCs were subjected to 3D-protein structure modelling using AlphaFold3. The predicted 3D protein structures were subjected to refinement using GalaxyRefine (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) by correcting the chains and stabilizing the structure to improve its local and global structural quality. The quality of the HCV VC 3D protein model structures was evaluated using SAVES server (https://saves.mbi.ucla.edu/) and assessed by ERRAT and PROCHECK. ERRAT assesses the overall quality of the protein model by screening the non-bonded interactions between atoms of the residues, whereas PROCHECK determines the stereo-chemical quality by generating a Ramachandran plot of protein residues. To further validate the VC model structure, we used ProSA-Web (https://prosa.services.came.sbg.ac.at/prosa.php) which evaluates the overall quality of the predicted VC protein model by comparing their z-score with the experimentally (NMR and X-ray analysis data) determined z-scores of the native crystal structures.

Immune Simulation of HCV VC

The immune response to the designed vaccine was determined using an online simulation tool, C-ImmSim server (https://kraken.iac.rm.cnr.it/C-IMMSIM/). Default settings were applied, except for modifications in the simulation time steps, which were set at 1, 84, and 168. The total number of simulation steps was 1000 (Singh et al., 2020).
Hence, there were three (3) injections at the interval of four weeks. The HCV VC that exhibited the highest antibody and cytokine levels was selected as the final HCV VC.

Molecular Docking and Protein-Protein Interaction Analyses

Molecular docking was used to predict the binding conformation and interactions of the final HCV VC with three Toll-like Receptors (TLRs): TLR2 (PDB ID: 3A7B), TLR3 (PDB ID: 2A0Z), and TLR4 (PDB ID: 4G8A), obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, http://rcsb.org). Docking was performed using a direct docking approach with ClusPro 2.0 server tool (https://cluspro.org/home.php). The Top 10 ClusPro cluster models, for each HCV VC-TLR (2, 3, and 4) complexes, were considered the best docked models and were further analyzed (Aiman et al., 2022). Using HawkDock (http://cadd.zju.edu.cn/hawkdock/), the binding affinity of each docked models were determined. Meanwhile, the protein-protein interactions of the final HCV VC-TLR complexes were analyzed with PDBSUM Generate tool (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html).

Molecular Dynamics Simulation and MM/P(G)BSA Calculations

Molecular dynamics simulation was performed using GROMACS 2023.3 and CHARMM36 force field parameters for the final HCV VC with TLR2, 3, and 4. The simulation was performed using SPC/E as the water model with at least 1.0 nanometer (nm) distancing the protein from the box edge. The simulation system was neutralized by adding the appropriate number of sodium (SOD) or chloride (CLA) ions. The solvated system was energy minimized in 50,000 steps using the steepest descent method iterations. Then, the system was allowed to reach an equilibrium state through the NVT ensemble by using the V-Rescale thermostat at 310 K, then through the NPT ensemble by using the Parrinello–Rahman barostat (1 atm)

(Jyotisha et al., 2022). The simulation was run for 200 ns to evaluate dynamics of final HCV VC-TLR complexes. The MD trajectory profiles of the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (rg), the number of hydrogen bonds (H-bonds), and the Solvent Accessible Surface Area (SASA) of the final HCV VC-TLR complexes were generated.

Prediction of HCV Multi-epitope Vaccine Secondary Structure

The PDBSUM Generate and Self-optimized Prediction Method with Alignment (SOPMA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) were used to examine the secondary (2D) structure of the VCs. These tools determine the 2D structure including the α-helix, β-pleated sheet, turn, and coil percentage composition.

Discontinuous B cell epitope Prediction

HCV multi-epitope vaccine contains new sets of amino acid sequence thus, re-prediction of discontinuous B-cell epitopes is necessary. Discontinuous B-cell epitopes were mapped for the validated 3D protein structure of the VCs using IEDB ElliPro: Antibody Epitope Prediction server (accessed on 2 July 2024).

Codon Optimization and In Silico Cloning of HCV Multi-epitope Vaccine Construct

Codon optimization was performed using the Java Codon Adaptation Tool (JCat) server (http://www.jcat.de/). *Escherichia coli* (strain K12) was selected as the target organism for codon optimization. In addition, parameters like *Avoid rho-independent transcription terminators*, *Avoid prokaryotic ribosome binding sites*, and *Avoid Cleavage Sites of Restriction Enzymes* were selected (Samad et al., 2022). The output of the JCat tool provided Codon Adaptation Index (CAI) value and GC content of the optimized sequence. SnapGene was used to perform *in silico* cloning of the HCV VC by integrating its optimized DNA to pET-28a, the plasmid vector, between *NdeI* and *XhoI* restriction sites (Ahmad et al., 2021).

RESULTS and DISCUSSION

HCV Proteome and Conserved Protein sequences

The study retrieved 82 non-recombinant HCV proteomes, containing ≥3000 amino acids and free of aberrant amino acids *i.e.*, X or J residues, from NCBI. Details of the 82 proteomes retrieved including their corresponding subtype designations, protein accession numbers, and sources are provided in Table S1. The neighbor-joining phylogenetic tree analysis of the 82 HCV proteomes demonstrated the presence of distinct 8 HCV genotype group (Fig. S1). The current classification of HCV, which includes 8 genotypes and several subtypes, remains comprehensive. Since the identification of HCV genotype 8 in 2018 (Borgia et al., 2018), no new HCV genotype has been recorded. Subsequently, the PVS tool identified a total of 114, 63, 135, 121, 109, 59, 78, and 103 highly conserved protein sequences across the 8 HCV genotypes, respectively, and which were then used for epitope predictions (Fig. S1).

Epitopes Prediction and Prioritization

Using the conserved protein sequences of HCV, a total of 668 9-mer (≤0.5 percentile rank scores) MHC I and 484 15-mer (≤2% percentile rank scores) MHC II-binding epitopes were predicted (Table S2). Subsequently, among the 484 MHC II epitopes, only 34 (≥1 score and positive) showed IFN-ɣ inducing epitopes. There were 1041 (≥10-mer) linear B cell stimulating epitopes with a score of ≥0.51 (Table S2). There were no discontinuous B cell stimulating epitopes identified. Following antigenicity (≥0.4 score) and toxicity assessment, and BLASTp human homology analysis, there were 11 epitopes selected.

Further, to design a cross-genotype protective HCV vaccine, the 11 initially identified epitopes were screened for IEDB epitope global conservation (EGC) analysis. Table 1 presents the 9 (out of the 11) epitopes, comprising five (5) MHC I epitopes, two (2) MHC II/IFN-ɣ stimulating epitopes, and two (2) B cell stimulating epitopes, that exhibited ≥80% protein sequence similarity signifying high conservation (Table 1) (Palatnik-de-Sousa et al., 2022; Rodrigues-da-Silva et al., 2023). Based on epitope prediction, many of the epitopes are derived from the NS3 and core proteins, indicating that these regions are highly conserved and essential to HCV across strains, making them key components for a broad-spectrum vaccine development. NS3 is an HCV protease important for initiating a series of cleaving processes in HCV, which triggers the replication cycle of the virus (Dubuisson, 2007), while core protein is vital in virus-host interaction and virion assembly (Izumida & Morita, 2024). The selected epitopes used as the basis for constructing HCV multi-epitope vaccine constructs (VC).

Table 1: The conserved MHC I, MHC II/IFN-ɣ, and B cell epitopes across HCV strains with their sequence, protein identity, epitope global conservancy percentages, and other assessment results.

Epitope Sequence	protein source	antigenicity	allergenicity	toxicity	Human homology	epitope conservancy (%)
MHC I						
APTGSGKST	NS3	1.1795	non-allergen	non-toxic	non-human	98.78
QPRGRRQPI	CORE	0.696	non-allergen	non-toxic	non-human	90.24
KTSERSQPR	CORE	0.8428	non-allergen	non-toxic	non-human	95.12
HLIFCHSKK	NS3	0.6339	non-allergen	non-toxic	non-human	84.15
LIFCHSKKK	NS3	1.0731	non-allergen	non-toxic	non-human	84.15
MHC II						
QGYKVLVLNPSVAAT	NS3	0.4335	non-allergen	non-toxic	non-human	85.37
GGRHLIFCHSKKKCD	NS3	0.9690	non-allergen	non-toxic	non-human	84.15
B CELL						
GGRHLIFCHSKKKCDE	NS3	0.7801	non-allergen	non-toxic	non-human	82.93
VFTGLTHIDAHFLSQT	NS3	0.7588	non-allergen	non-toxic	non-human	80.49

Designing and Evaluation of Physicochemical Properties of the HCV Multi-epitope Vaccine Constructs

Twelve (12) HCV multi-epitope VCs were generated using combinations of two distinct adjuvants, EAAAK, a set of 12 conserved epitopes, and their respective linkers. Table S3 showed the physicochemical properties and other relevant evaluations for the 12 HCV VCs (HCV VC 1-12). Results showed that the HCV VC with human β-defensin (HBD) adjuvant (HCV VC 1-6) exhibited higher antigenicity (0.530-0.492) compared to the HCV VCs with L7/L12 adjuvant (0.467-

0.427). Meanwhile, only HCV VC 7-12 with L7/L12 adjuvant exhibited stability (≤40 instability index) and non-allergenic. Although among all HCV VCs with L7/L12 adjuvant, HCV VC 8 has a solubility value of 0.656 which is below the recommended value of ≥0.950 thus, were excluded for the subsequent analysis.

Prediction of HCV Multi-epitope Vaccine Construct Tertiary Protein Structure and Validation

The 3D protein model structures of HCV VC 7, 9, 10, 11, and 12 were predicted using AlphaFold3, a highly reliable tool for 3D protein structure prediction. It has gained recognition for its exceptional prediction accuracy, leading to its victory in the CASP14 competition in 2022 (Marcu et al., 2022). AlphaFold rely mainly on the query amino acid sequence and perform template-free, protein threading, and/or fragment assembly, making it well-suited for modeling novel protein sequences like those of the HCV VCs (Alican Gulsevin et al., 2022).

Following 3D structure prediction, GalaxyRefine server tool was used to refine the proteins. This tool provides 5 refined models with their respective Global Distance Test – High Accuracy (GDT-HA), root mean square deviation (RMSD), Molecular Probability (MolProb), Clash scores, Poor rotamers, and Rama-favored scores. Out of the five refined models, one model from each prediction tool was selected based on the highest GDT-HA scores and the lowest RMSD values, which are metrics used to evaluate the similarity between two protein structures. In cases where a model had the highest GDT-HA score but did not have the lowest RMSD, additional factors such as the MolProb, Clash, and poor rotamers scores were considered to ensure the selection of the most accurate and reliable model for further analysis.

Furthermore, the overall quality of the models was evaluated using ERRAT, PROCHECK, and Pro-SA (Z-score). The HCV VC with scores within the acceptable threshold: ≥95% rejection limit of ERRAT, ≥90% of residue in the favored regions, and <1% outliers in the Ramachandran plot, and Z-score within the experimental data, was selected. Table 2 presents the top 5 HCV VCs along with their ERRAT, PROCHECK, and Z-score values. The findings showed that HCV VC 9 and VC 12 did not pass the ERRAT metrics and were excluded from further analysis.

Table 2: The top HCV VC and their tertiary protein structure validation scores.

HCV VCs	ERRAT (cut-off score: ≥95.00%)	PROCHECK (cut-off score: ≥90.00%)	PRO-SA (z-score)
HCV VC 7	100.00	95.75	-2.67
HCV VC 9	94.53	96.40	-5.59
HCV VC 10	98.06	98.00	-5.07
HCV VC 11	97.00	98.00	-5.54
HCV VC 12	93.60	96.70	-4.38

Immune Simulation of HCV VC

The immune response simulation of the HCV VC 7, HCV VC 10, and HCV VC 11 were evaluated using the C-ImmSim server. Table 3 presents the levels of antibodies and relevant cytokines for HCV VC 7, HCV VC 10, and HCV VC 11, along with the levels obtained from the reference studies. As observed, HCV VC 7 has the highest level of IgM + IgG, IgG1, and IgG1 + IgG2 antibodies compared to HCV VC V10 and 11 after the 3rd injection, except for IgM antibody levels of HCV VC 10. Furthermore, HCV VC 7 demonstrated superior T-cell mediated immunity, as evident by the elevated levels of IFN-γ (420,253 ng/mL) and IL-2 (457,333 ng/mL) compared to HCV VC 10 and 11. This result also highlights the importance of IFN-γ in the immune response, underscoring its well-established role in defending the host against viral infections by activation of macrophages, enhancement antigen presentation, and promotion the development of a Th1 immune response, which is essential for controlling viral infections (A. J. Lee & Ashkar, 2018). Based on these findings, HCV VC 7 selected as the final HCV VC. Importantly, the HCV VC in this study generated higher antibodies and cytokine levels compared to the reference HCV vaccine design studies, including those by Ahmad et al. (2021), the only HCV multi-epitope vaccine study with immune response simulation analysis.

Table 3: Comparative immune simulation analysis of top HCV VCs and reference studies.

Antibody and Cytokines	Step Injection	HCV VC 7	HCV VC 10	HCV VC 11	Ahmad et al., (2021)	Singh et al., (2020)
		titer				
IgM+IgG	1st	5249	7150	5970	5533	1200
	2nd	110995	84182	102567		37800
	3rd	178456	154334	157970		57000
IgM	1st	4322	5126	4776	4201	0
	2nd	81201	63137	75104		50400
	3rd	86604	91602	79045		81800
IgG1+2	1st	1852	2968	1672		0
	2nd	30875	21855	29612		13200

	3rd	95403	64890	81194		26400
	1st	1852	2968	1672	1661	0
IgG1	2nd	30875	21855	29612		12646
	3rd	95403	64890	81194		32881
	1st	420253	418486	419911	416972	410638
IFN-g	2nd	406962	405271	405535		401064
	3rd	388608	387312	386813		397872
	1st	238000	235714	243931	195555	96774
IL-2	2nd	627667	623469	712139		519816
	3rd	457333	456122	426590		373272

In particular, the primary response of the final HCV VC was characterized by an increase in the levels of IgM + IgG, IgM, and IgG1, followed by an increase in IgG1 + IgG2 and IgG2 levels after the second and third injection simulations (Fig. 2a). Simultaneously, a decreasing level of antigen was observed. At the first injection, an increasing IFN-γ level and IL-2 were observed after each subsequent exposure (Fig. 2b). A surge in B memory cells: B cell isotypes IgM and IgG1 (Fig. 2c) and CD8 T cell populations (Fig. 2d-e) were sustained throughout the exposure period. The active and resting regulatory T-helper cell population (TR) was sustained in the entire simulation (Fig. 2f). The TH cell population was observed to increase abruptly during the second injection (Fig. 2g-h). In addition, the Th1 cell population increase from 40,360 cells per mm^3 on the 12th day to 115,150 cells per mm^3 after the 3rd injection (Fig. 2i). The activity of DCs, NK cells, and macrophages was sustained during the 350-day simulation period (Fig. 2j-l).

These findings collectively demonstrate the potential of the final HCV VC to induce a strong and sustained immune response. This outcome aligns with the growing consensus on the importance of T-cell mediated immunity in achieving durable protection against HCV infection, as highlighted by Thimme et al. (2021). The inclusion of conserved NS3epitopes, a non-structural protein, within the designed HCV VC likely contributed to the strong T-cell response observed in this study, consistent with previous findings (Krishnadas et al., 2009; Laidlaw et al., 2016).Moreover, the incorporation of the core protein, a structural protein component, likely enhanced the vaccine's ability to activate both innate and adaptive immune responses, particularly the humoral response (Tzarum et al., 2020). These finding highlights the effectiveness of incorporating NS3 and core proteins in the HCV vaccine inducing a potentially strong and long-lasting immune response and can be cross-protective against several HCV strains.

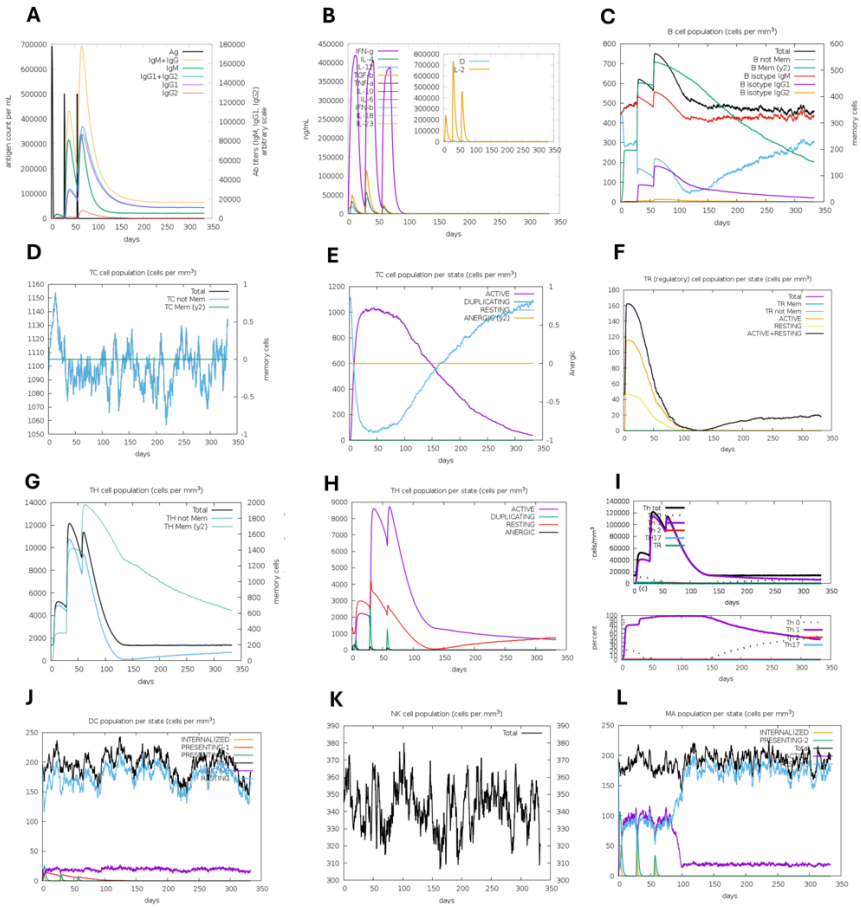

Figure 2: *In silico* immune stimulation analysis of HCV VC after three subsequent injections using C-ImmSim tool. (A) Antigen and antibody response. (B) Cytokine levels. (C) B cell population per entity state. (D) general CD8 T cell population (E) CD8 T cell population per entity state (F) regulatory Thelper (CD4) cell population. (G) Th memory T cells. (H) Th cell population per entity state. (I) Th1 cell counts. (J) NK cell population. (K) DC population. (L) macrophage cell population.

Molecular Docking and Protein-Protein Interaction Analyses

The selected final HCV VC, consisting of 266 amino acids with a molecular weight of 28149.58 kDa, was used to dock, using ClusPro, with relevant TLRs to determine possible interactions and, thereby, potential immune response stimulation. The schematic diagram showing the composition arrangement of the final HCV VC (Fig. 3a) and its 3D protein structure (Fig. 3b) are presented. The final HCV VC is comprised of a combination of 50S ribosomal protein L7/L12, AAY, GPGPG, and KK linkers between epitopes, and EAAAK linkers between the adjuvant and the set of epitopes. Apart from providing stability, L7/L12 adjuvant can also contribute to the induction of antigen-presenting cell maturation (S. J. Lee et al., 2014). In general, linkers prevent the formation of junctional epitopes or neoepitopes and are necessary for protein conformation, separation, and stability (Palatnik-de-Sousa et al., 2022). Particularly, AAY linkers are specialized cleavage sites for mammalian cell proteasomes (Ayyagari et al., 2022), while KK linkers function in fragmentation as it is targeted by the lysosome enzyme (Tarrahimofrad et al., 2021). GPGPG linker was used to link MHC II epitopes and prevent junctional epitope immunogenicity (Soltan et al., 2022). Additionally, the EAAAK linker was used to provide stiffness and prevent interactions between the adjuvant and the epitopes (Rouzbahani et al., 2022). These compositions can greatly influence the interaction with the immune receptors to elicit immune response and provide insights on the critical binding sites and interaction mechanisms that may play a role in the immune response to HCV infection. Figure 3 also showed that the docking scores and docking poses of the best docked HCV VC-TLR2, 3, and 4 complexes. The docking scores of each Cluspro model are presented in Table S4.

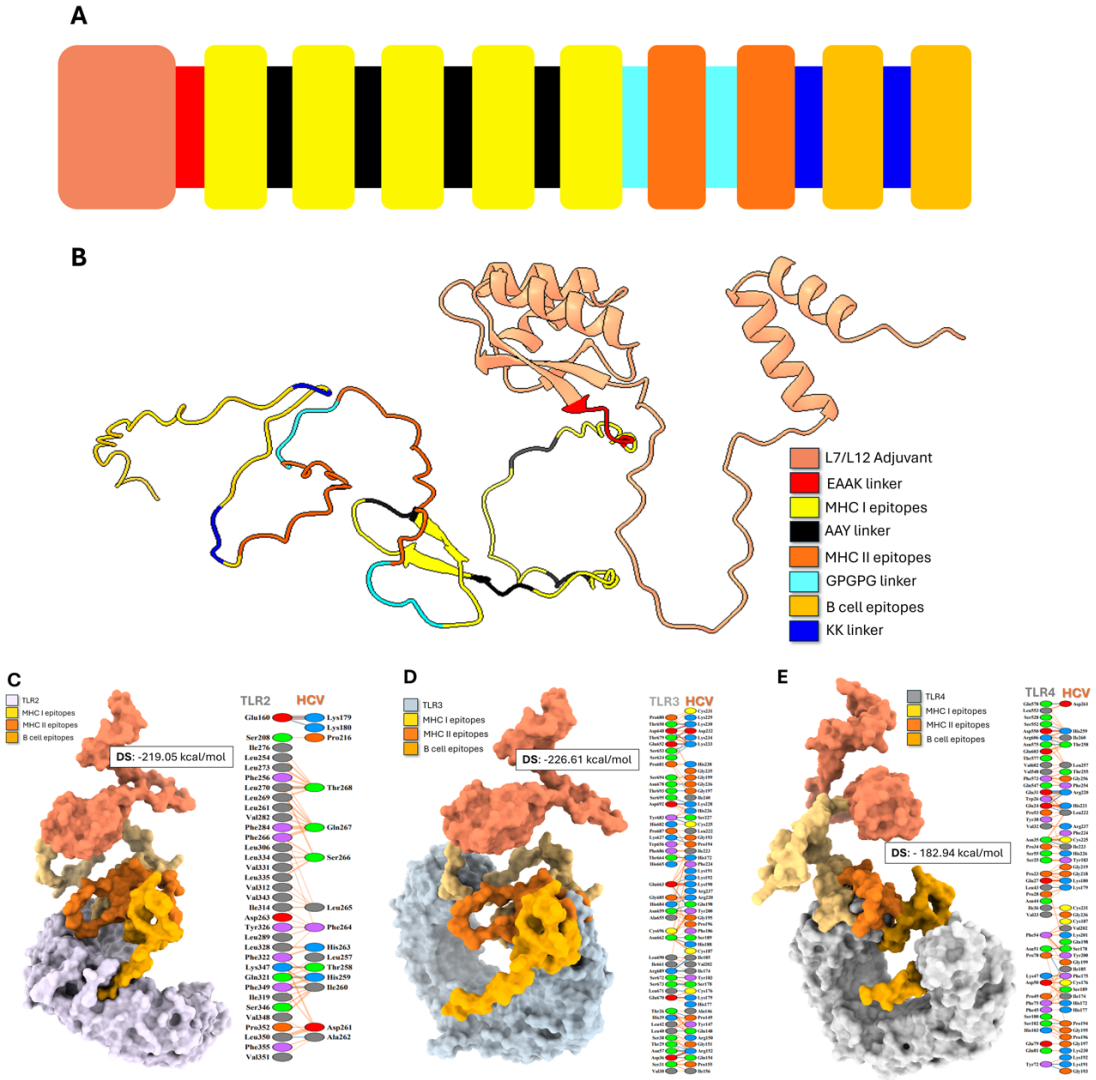

Figure 3: The final HCV multi-epitope subunit vaccine construct, its tertiary protein structure, and their intermolecular interaction analysis with TLRs. (A) The schematic diagram of the arrangement of the final HCV vaccine construct composition (B) The tertiary (3D) structure predicted by AlphaFold3 is presented in different colors for each composition. (C) Docking score and interacting residues between HCV VC-TLR2 (PDB: 3A7B) complex, (D) Docking score and interacting residues between HCV VC-TLR3 (PDB: 2A0Z) complex, and (E) Docking score and interacting residues between HCV VC-TLR4 (PDB: 4G8A) complex. Docked complex was generated by ClusPro 2.0 and intermolecular interactions was identified by PDBsum generate tool. Amino acid colors indicate properties: Negative: Red, Positive: Blue, Neutral: Green, Aromatic: Pink, Pro&Gly: Orange, Cys: Yellow, and Aliphatic: Grey.

Further, this analysis showed that the final HCV VC-TLR3 exhibited the most favorable binding with binding score of -226.61 kcal/mol. TLR3 is known to regulate innate immune responses to viral infections (Y. Chen et al., 2021), particularly to HCV (Bhattacharjee et al., 2021). As demonstrated by Sugrue et al. (2022), HCV activated TLR3 consequently triggering IFN-γ production in B cells through a T helper cell-dependent mechanism. Moreover, it was demonstrated that activation of TLR3 has been shown to inhibit HCV replication by enhancing the macrophage activity (Zhou et al., 2016). HCV VC-TLR2 exhibited a strong binding affinity with a binding energy score of -219.05 kcal/mol while HCV VC exhibited relatively weak

binding to TLR4 (-182.94 kcal/mol). The selection of TLR2 and TLR4 is based on the propensity of these TLRs towards extracellular viral components (Kayesh et al., 2022) such as core and E1/E2 glycoprotein, few components present in the final HCV VC.

Several interacting residues between the HCV VC and TLRs were observed. There were 59, 45, and 15 HCV VC residues that formed intermolecular interaction with TLR2 (Fig. 3c), TLR3 (Fig. 3d), and TLR4 (Fig. 3e), respectively. These interacting residues correspond to HCV epitopes, suggesting that they play a critical role in the recognition and binding of HCV VC by these TLRs. This interaction may potentially facilitate the activation of immune responses against HCV.

Molecular Dynamics Simulation and MM/P(G)BSA Calculations

The 200 ns molecular dynamics simulation was conducted to assess the stability and dynamics of the final HCV VC-TLR complexes. Visualization of the MD trajectories of each HCV VC-TLR complex is displayed in Figure 4. RMSD represents the conformational deviation of fitted protein based on the reference structure. At 0-10 ns, HCV VC-TLRs showed a steep increase in RMSD value indicating significant conformational changes occurring immediately after binding (Fig. 4a). From 11 ns onwards, the HCV VC-TLR2 and 3 complexes exhibited a dynamic equilibrium, indicating a generally stable binding. While HCV VC-TLR4 complex exhibited instability after 160 ns, as indicated by high RMSD values, visualization revealed that this instability was contributed to the highly moving TLR4 protein. Despite the observed high deviations, the final HCV VC protein remained in the binding pocket of the TLR4. This flexibility behavior is also aligned with the high residual fluctuations shown in Fig. 4b. The mean RMSD values were 1.42 ± 0.14 nm, 1.55 ± 0.25 nm, and 1.82 ± 0.56 nm for HCV VC-TLR2, HCV VC-TLR3, and HCV VC-TLR4 complex, respectively.

On the other hand, RMSF value gives insight into the flexibility of the residues within the protein. Figure 8b showed TLRs have relatively rigid protein, except for TLR4 which showed high fluctuations. The HCV VC residues also showed relatively higher fluctuations than the TLRs as displayed at the C terminus region (~500-900th residue). This suggests a highly flexible protein, which is consistent with the HCV VC predominantly comprising coil region structures.

Similarly, the total radius of gyration (Rg), which denotes the compactness and thereby the stability of the protein structure or complex, showed that the final HCV VCin complex with TLR2 and TLR3 was relatively lesser compared to the final HCV VC-TLR4 complex (Fig. 4c). This suggests greater compactness and stability of the HCV VC-TLR2 and HCV VC-TLR3 complexes throughout the entire simulation. In addition, the H-bond analysis showed that the average number of H-bonds formed in the interface of final HCV VC-TLR2, 3, and 4 complexes were 3.69 ± 0.06, 8.44 ± 2.35, and 8.62 ± 7.46, respectively (Fig. 4d). The average number of hydrogen bonds formed at the HCV VC-TLR2 interface was low, indicating a relatively weaker interaction compared to the other complexes. In contrast, the HCV VC-TLR3 and HCV VC-TLR4 complexes formed significantly more hydrogen bonds indicating stronger and more stable interactions between HCV VC and TLR3, as well as TLR4. The larger standard deviation observed in the TLR4 complex (7.46) compared to the TLR3 complex (2.35) indicates greater variability in the hydrogen bonding interactions during the simulation. These results imply that TLR2 forms fewer and more consistent hydrogen bonds with HCV VC, while TLR3 and TLR4 form more numerous and potentially more dynamic hydrogen bonds, contributing to the overall stability and variability in the interaction of these complexes. Lastly, the SASA analysis was performed to assess the molecular surface area exposed to the solvent. As presented in Figure 4e, larger SASA correlates with increased protein-protein interactions. The HCV VC-TLR4 complex had a larger SASA (456 nm^2 ± 19.99), indicating a more exposed and less compact interface compared to the HCV VC-TLR2 (446 nm^2 ± 21. 85) and HCV VC-TLR3 (471 nm^2 ± 16.70) complexes. These results collectively imply that TLR2 forms fewer and more consistent hydrogen bonds with HCV VC, leading to a more exposed interaction interface, while TLR3 and TLR4 form numerous and potentially more dynamic hydrogen bonds, resulting in a more compact and stable interaction interface, contributing to the overall stability of these complexes throughout the simulation.

Meanwhile, the MM/P(G)BSA calculations align with the observations from the hydrogen bond and SASA analyses (Fig. 4f). The HCV VC-TLR3 complex exhibited the highest total binding energy of -171 kcal/mol (GBSA: -160 kcal/mol) compared to TLR2 with -116 kcal/mol (GBSA: 113 kcal/mol) and TLR4 with -108 kcal/mol (GBSA: -105 kcal/mol) binding energy scores. This consistency across the analysis highlights the reliability of these findings and underscores the varying strengths of the HCV interactions with different TLRs.

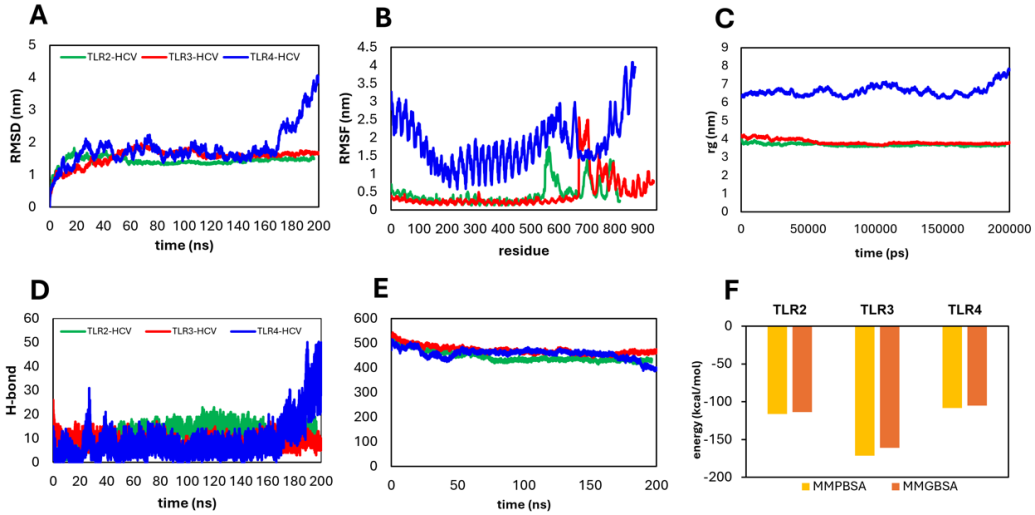

Figure 4: Molecular dynamics simulation and MM/P(G)BSA analyses of HCV VC-TLR (2, 3, and 4) complexes. (A) RMSD values, (B) RMSF values, (C) Rg of the C-α backbone-atoms of the HCV VC-TLR complexes, and the (D) Number of hydrogen bonds formed, (E) SASA values between HCV VC and the TLRs in the entire 200 ns simulation, and the (F) MM/P(G)BSA calculations within the last 100 ns simulation.

Prediction of HCV Multi-epitope Vaccine Secondary Structure and B-cell epitopes

The predicted secondary structure of the final HCV VC consisted of 37.31% α-helix, 13.43% extended strand, and 49.25% random coils (Fig. S2). The unique structure and folding of this new protein may generate novel conformational B-cell epitopes, leading to additional predictions. The ABCpred server predicted 28 linear epitopes on final HCV VC (Table S5), highlighting potential targets for B-cell recognition. Additionally, the Ellipro server identified 11 discontinuous epitopes with a score higher than 0.6 (Table S6), suggesting strong candidate regions for epitope-based vaccine design.

Codon Optimization and In Silico Cloning of HCV Multi-epitope Vaccine Construct

Java Codon Adaptation Tool (JCAT) produced an optimized codon sequence of the HCV vaccine, consisting of 813 nucleotides. The Codon Adaptation Index (CAI) of the optimized sequences was 0.94, and the GC content was 54.10%. CAI values higher than 0.50 are considered acceptable while a GC% range of 40-60% is considered good (Palatnik-de-Sousa et al., 2022). This codon optimization and cloning is important in consideration for potential expression of the vaccine construct using *E. coli* K12 strain. The cloned plasmid had a final length of 6,102 bp (Fig. 5).

CONCLUSION

This study in silico designed a promising cross-protective multi-epitope subunit vaccine against Hepatitis C Virus strains. By analyzing the conserved proteins across eight (8) HCV genotypes and 47 subtypes using immunoinformatic tools, a set of antigenic, non-allergenic, non-toxic, non-human homologous, and highly conserved epitopes was identified and incorporated into a stable HCV vaccine construct. Moreover, molecular docking and molecular dynamics simulations provided evidence of the strong affinity and binding stability of the vaccine to TLRs. Overall, this study represents a significant step in the development of a potential HCV vaccine that could offer strong cross-protection against hypervariable HCV that affects millions of individuals in developing countries. The integration of computational structural biology, and immune simulation analyses provides a solid foundation for further *in vitro* and *in vivo* experimental validations of this promising HCV multi-epitope subunit vaccine.

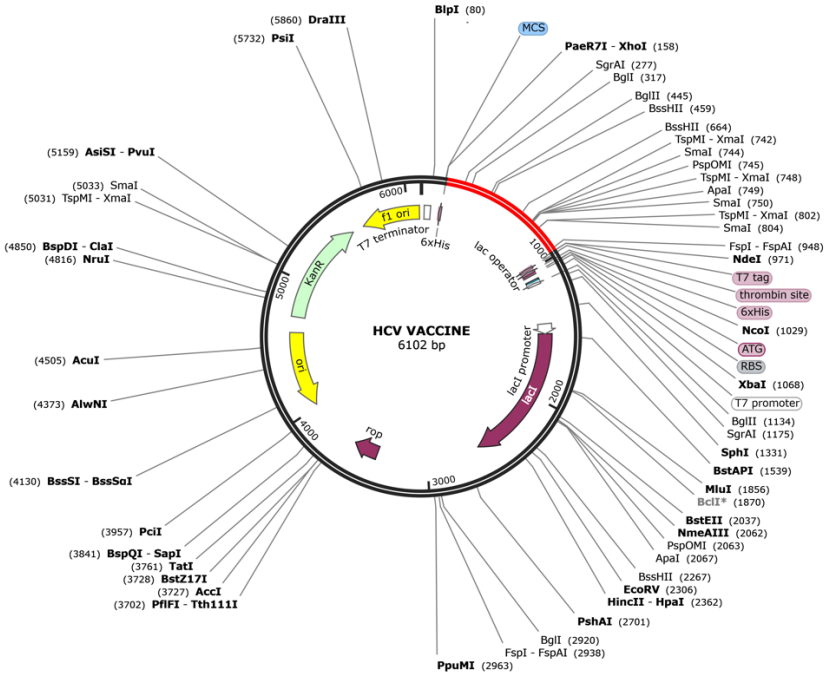

Figure 5: *In silico* cloning of HCV vaccine constructs. The optimized HCV VC cDNA sequence (red) was inserted between *XhoI* (158) and *NdeI* (971) cleavage sites in the pET-28a (+) vector.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of Department of Science and Technology (DOST) – Virology and Vaccine Institute of the Philippines Program, DOST – S&T Fellow Program, Career Incentive Program (CIP) under DOST – Science Education Institute (SEI), Taguig City, Philippines, and DOST- Advanced Science and Technology Institute (ASTI), Taguig City, Philippines.

CONFLICT OF INTEREST

There is no conflict of interest among the authors.

CONTRIBUTIONS OF INDIVIDUAL AUTHORS

Nyzar Mabeth Odchimar: conceptualization and design of the study, investigation, data curation, methodology, software, resources, validation, writing-original of draft and writing - review & editing. Fredmoore Orosco: Conceptualization and design of the study, writing-original of draft, and writing - review & editing, and supervision. Ella Mae Joy Sira: methodology, software, resources, validation, and writing - review & editing. Edward Banico: methodology, software, resources, validation, and writing - review & editing. Alea Maurice Simbulan: methodology, software, resources, validation, and writing - review & editing.

REFERENCES

Adugna, A. (2023). Therapeutic strategies and promising vaccine for hepatitis C virus infection. Immunity, Inflammation and Disease, 11(8), e977. https://doi.org/10.1002/iid3.977

Ahmad, S., Shahid, F., Tahir Ul Qamar, M., Rehman, H. U., Abbasi, S. W., Sajjad, W., Ismail, S., Alrumaihi, F., Allemailem, K. S., Almatroudi, A., & Ullah Saeed, H. F. (2021). Immuno-Informatics Analysis of Pakistan-Based HCV Subtype-3a for Chimeric Polypeptide Vaccine Design. Vaccines, 9(3). https://doi.org/10.3390/vaccines9030293

Aiman, S., Alhamhoom, Y., Ali, F., Rahman, N., Rastrelli, L., Khan, A., Farooq, Q. U. A., Ahmed, A., Khan, A., & Li, C. (2022). Multi-epitope chimeric vaccine design against emerging Monkeypox virus via reverse vaccinology techniques- a bioinformatics and immunoinformatics approach. Frontiers in Immunology, 13, 985450. https://doi.org/10.3389/fimmu.2022.985450

Alican Gulsevin, Bing Han, Jason C. Porta, Hassane S. Mchaourab, Jens Meiler, & Anne K. Kenworthy. (2022). Template-free prediction of a new monotopic membrane protein fold and oligomeric assembly by Alphafold2. bioRxiv, 2022.07.12.499809. https://doi.org/10.1101/2022.07.12.499809

Ayyagari, V. S., T C, V., K, A. P., & Srirama, K. (2022). Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: An immunoinformatics approach. Journal of Biomolecular Structure & Dynamics, 40(7), 2963–2977. https://doi.org/10.1080/07391102.2020.1850357

Aziz, S., Waqas, M., Halim, S. A., Ali, A., Iqbal, A., Iqbal, M., Khan, A., & Al-Harrasi, A. (2022). Exploring whole proteome to contrive multi-epitope-based vaccine for NeoCoV: An immunoinformtics and in-silico approach. Frontiers in Immunology, 13, 956776. https://doi.org/10.3389/fimmu.2022.956776

Bailey, J. R., Barnes, E., & Cox, A. L. (2019). Approaches, Progress, and Challenges to Hepatitis C Vaccine Development. Gastroenterology, 156(2), 418–430. https://doi.org/10.1053/j.gastro.2018.08.060

Behmard, E., Abdulabbas, H. T., Abdalkareem Jasim, S., Najafipour, S., Ghasemian, A., Farjadfar, A., Barzegari, E., Kouhpayeh, A., & Abdolmaleki, P. (2022). Design of a novel multi-epitope vaccine candidate against hepatitis C virus using structural and nonstructural proteins: An immunoinformatics approach. PloS One, 17(8), e0272582. https://doi.org/10.1371/journal.pone.0272582

Bhattacharjee, C., Singh, M., Das, D., Chaudhuri, S., & Mukhopadhyay, A. (2021). Current therapeutics against HCV. Virusdisease, 32(2), 228–243. https://doi.org/10.1007/s13337-021-00697-0

Borgia, S. M., Hedskog, C., Parhy, B., Hyland, R. H., Stamm, L. M., Brainard, D. M., Subramanian, M. G., McHutchison, J. G., Mo, H., Svarovskaia, E., & Shafran, S. D. (2018). Identification of a Novel Hepatitis C Virus Genotype From Punjab, India: Expanding Classification of Hepatitis C Virus Into 8 Genotypes. The Journal of Infectious Diseases, 218(11), 1722–1729. https://doi.org/10.1093/infdis/jiy401

Bui, H.-H., Sidney, J., Li, W., Fusseder, N., & Sette, A. (2007). Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics, 8(1), 361. https://doi.org/10.1186/1471-2105-8-361

Chauhan, V., Rungta, T., Goyal, K., & Singh, M. P. (2019). Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. Scientific Reports, 9(1), 2517. https://doi.org/10.1038/s41598-019-39299-8

Chen, F., Nagy, K., Chavez, D., Willis, S., McBride, R., Giang, E., Honda, A., Bukh, J., Ordoukhanian, P., Zhu, J., Frey, S., Lanford, R., & Law, M. (2020). Antibody Responses to Immunization With HCV Envelope Glycoproteins as a Baseline for B-Cell-Based Vaccine Development. Gastroenterology, 158(4), 1058-1071.e6. https://doi.org/10.1053/j.gastro.2019.11.282

Chen, Y., Lin, J., Zhao, Y., Ma, X., & Yi, H. (2021). Toll-like receptor 3 (TLR3) regulation mechanisms and roles in antiviral innate immune responses. Journal of Zhejiang University. Science. B, 22(8), 609–632. https://doi.org/10.1631/jzus.B2000808

Dehghan, Z., Lari, A., Yarian, F., Ahangarzadeh, S., Sharifnia, Z., Shahzamani, K., & Shahidi, S. (2020). Development of Polyepitopic Immunogenic Contrast Against Hepatitis C Virus 1a-6a Genotype by in Silico Approach. Biomedical and Biotechnology Research Journal (BBRJ), 4(4). https://journals.lww.com/bbrj/fulltext/2020/04040/development_of_polyepitopic_immunogenic_contrast.17.aspx

Donnison, T., McGregor, J., Chinnakannan, S., Hutchings, C., Center, R. J., Poumbourios, P., Klenerman, P., Drummer, H. E., & Barnes, E. (2022). A pan-genotype hepatitis C virus viral vector vaccine generates T cells and neutralizing antibodies in mice. Hepatology (Baltimore, Md.), 76(4), 1190–1202. https://doi.org/10.1002/hep.32470

Donnison, T., von Delft, A., Brown, A., Swadling, L., Hutchings, C., Hanke, T., Chinnakannan, S., & Barnes, E. (2020). Viral vectored hepatitis C virus vaccines generate pan-genotypic T cell responses to conserved subdominant epitopes. Vaccine, 38(32), 5036–5048. https://doi.org/10.1016/j.vaccine.2020.05.042

Dubuisson, J. (2007). Hepatitis C virus proteins. World Journal of Gastroenterology, 13(17), 2406–2415. https://doi.org/10.3748/wjg.v13.i17.2406

Echeverría, N., Comas, V., Aldunate, F., Perbolianachis, P., Moreno, P., & Cristina, J. (2021). In the era of rapid mRNA-based vaccines: Why is there no effective hepatitis C virus vaccine yet? World Journal of Hepatology, 13(10), 1234–1268. https://doi.org/10.4254/wjh.v13.i10.1234

Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Research, 33(Web Server issue), W526-531. https://doi.org/10.1093/nar/gki376

Haq, A. U., Khan, A., Khan, J., Irum, S., Waheed, Y., Ahmad, S., Nizam-Uddin, N., Albutti, A., Zaman, N., Hussain, Z., Ali, S. S., Waseem, M., Kanwal, F., Wei, D.-Q., & Wang, Q. (2021). Annotation of Potential Vaccine Targets and Design of a Multi-Epitope Subunit Vaccine against Yersinia pestis through Reverse Vaccinology and Validation through an Agent-Based Modeling Approach. Vaccines, 9(11). https://doi.org/10.3390/vaccines9111327

Ikram, A., Anjum, S., & Tahir, M. (2014). In Silico Identification and Conservation Analysis of B-cell and T-Cell Epitopes of Hepatitis C Virus 3a Genotype Enveloped Glycoprotein 2 From Pakistan: A Step Towards Heterologous Vaccine Design. Hepatitis Monthly, 14(6), e9832. https://doi.org/10.5812/hepatmon.9832

Irekeola, A. A., Malek, N. A., Wada, Y., Mustaffa, N., Muhamad, N. I., & Shueb, R. H. (2021). Prevalence of HCV genotypes and subtypes in Southeast Asia: A systematic review and meta-analysis. PloS One, 16(5), e0251673. https://doi.org/10.1371/journal.pone.0251673

Izumida, K., & Morita, E. (2024). The roles of HCV core protein and its binding host factor in virus assembly and release. Frontiers in Virology, 4. https://doi.org/10.3389/fviro.2024.1383261

Jyotisha, Singh, S., & Qureshi, I. A. (2022). Multi-epitope vaccine against SARS-CoV-2 applying immunoinformatics and molecular dynamics simulation approaches. Journal of Biomolecular Structure & Dynamics, 40(7), 2917–2933. https://doi.org/10.1080/07391102.2020.1844060

Kar, T., Narsaria, U., Basak, S., Deb, D., Castiglione, F., Mueller, D. M., & Srivastava, A. P. (2020). A candidate multi-epitope vaccine against SARS-CoV-2. Scientific Reports, 10(1), 10895. https://doi.org/10.1038/s41598-020-67749-1

Kayesh, M. E. H., Kohara, M., & Tsukiyama-Kohara, K. (2022). Toll-like Receptor Response to Hepatitis C Virus Infection: A Recent Overview. International Journal of Molecular Sciences, 23(10). https://doi.org/10.3390/ijms23105475

Keck, M.-L., Wrensch, F., Pierce, B. G., Baumert, T. F., & Foung, S. K. H. (2018). Mapping Determinants of Virus Neutralization and Viral Escape for Rational Design of a Hepatitis C Virus Vaccine. Frontiers in Immunology, 9, 1194. https://doi.org/10.3389/fimmu.2018.01194

Keikha, M., Eslami, M., Yousefi, B., Ali-Hassanzadeh, M., Kamali, A., Yousefi, M., & Karbalaei, M. (2020). HCV genotypes and their determinative role in hepatitis C treatment. Virusdisease, 31(3), 235–240. https://doi.org/10.1007/s13337-020-00592-0

Khalid, H., & Ashfaq, U. A. (2020). Exploring HCV genome to construct multi-epitope based subunit vaccine to battle HCV infection: Immunoinformatics based approach. Journal of Biomedical Informatics, 108, 103498. https://doi.org/10.1016/j.jbi.2020.103498

Krishnadas, D. K., Li, W., Kumar, R., Tyrrell, L. J., & Agrawal, B. (2009). In vitro activation and differentiation of naïve CD4+ and CD8+ T cells into HCV Core- and NS3-specific armed effector cells: A new role for CD4+ T cells. Cellular Immunology, 259(2), 141–149. https://doi.org/10.1016/j.cellimm.2009.06.009

Laidlaw, B. J., Craft, J. E., & Kaech, S. M. (2016). The multifaceted role of CD4(+) T cells in CD8(+) T cell memory. Nature Reviews. Immunology, 16(2), 102–111. https://doi.org/10.1038/nri.2015.10

Laurens, M. B. (2020). RTS,S/AS01 vaccine (MosquirixTM): An overview. Human Vaccines & Immunotherapeutics, 16(3), 480–489. https://doi.org/10.1080/21645515.2019.1669415

Lee, A. J., & Ashkar, A. A. (2018). The Dual Nature of Type I and Type II Interferons. Frontiers in Immunology, 9. https://doi.org/10.3389/fimmu.2018.02061

Lee, S. J., Shin, S. J., Lee, M. H., Lee, M.-G., Kang, T. H., Park, W. S., Soh, B. Y., Park, J. H., Shin, Y. K., Kim, H. W., Yun, C.-H., Jung, I. D., & Park, Y.-M. (2014). A potential protein adjuvant derived from Mycobacterium tuberculosis Rv0652 enhances dendritic cells-based tumor immunotherapy. PloS One, 9(8), e104351. https://doi.org/10.1371/journal.pone.0104351

Li, Y.-D., Chi, W.-Y., Su, J.-H., Ferrall, L., Hung, C.-F., & Wu, T.-C. (2020). Coronavirus vaccine development: From SARS and MERS to COVID-19. Journal of Biomedical Science, 27(1), 104. https://doi.org/10.1186/s12929-020-00695-2

Luna-Cuadros, M. A., Chen, H.-W., Hanif, H., Ali, M. J., Khan, M. M., & Lau, D. T.-Y. (2022). Risk of hepatocellular carcinoma after hepatitis C virus cure. World Journal of Gastroenterology, 28(1), 96–107. https://doi.org/10.3748/wjg.v28.i1.96

Mahmoodi, S., Amirzakaria, J. Z., & Ghasemian, A. (2023). In silico design and validation of a novel multi-epitope vaccine candidate against structural proteins of Chikungunya virus using comprehensive immunoinformatics analyses. PloS One, 18(5), e0285177. https://doi.org/10.1371/journal.pone.0285177

Manns, M. P., Buti, M., Gane, E., Pawlotsky, J.-M., Razavi, H., Terrault, N., & Younossi, Z. (2017). Hepatitis C virus infection. Nature Reviews. Disease Primers, 3, 17006. https://doi.org/10.1038/nrdp.2017.6

Marcu, Ş.-B., Tăbîrcă, S., & Tangney, M. (2022). An Overview of Alphafold's Breakthrough. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.875587

Nakamura, M., Kanda, T., Haga, Y., Sasaki, R., Wu, S., Nakamoto, S., Yasui, S., Arai, M., Imazeki, F., & Yokosuka, O. (2016). Sofosbuvir treatment and hepatitis C virus infection. World Journal of Hepatology, 8(3), 183–190. https://doi.org/10.4254/wjh.v8.i3.183

National Institute of Allergy and Infectious Diseases (NIAID). (2019). A Staged Phase I/II Study, to Assess Safety, Efficacy and Immunogenicity of a New Hepatitis C Prophylactic Vaccine Based on Sequential Use of AdCh3NSmut1 and MVA-NSmut (Clinical Trial Registration NCT01436357). clinicaltrials.gov. https://clinicaltrials.gov/study/NCT01436357

Nielsen, M., Lundegaard, C., Worning, P., Hvid, C. S., Lamberth, K., Buus, S., Brunak, S., & Lund, O. (2004). Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. Bioinformatics, 20(9), 1388–1397. https://doi.org/10.1093/bioinformatics/bth100

Orosco, F. (2024). Recent advances in peptide-based nanovaccines for re-emerging and emer ging infectious diseases. Journal of Advanced Biotechnology and Experimental Therapeutics, 7(1), 106. Crossref. https://doi.org/10.5455/jabet.2024.d10

Page, K., Melia, M. T., Veenhuis, R. T., Winter, M., Rousseau, K. E., Massaccesi, G., Osburn, W. O., Forman, M., Thomas, E., Thornton, K., Wagner, K., Vassilev, V., Lin, L., Lum, P.

J., Giudice, L. C., Stein, E., Asher, A., Chang, S., Gorman, R., … Cox, A. L. (2021). Randomized Trial of a Vaccine Regimen to Prevent Chronic HCV Infection. The New England Journal of Medicine, 384(6), 541–549. https://doi.org/10.1056/NEJMoa2023345

Palatnik-de-Sousa, I., Wallace, Z. S., Cavalcante, S. C., Ribeiro, M. P. F., Silva, J. A. B. M., Cavalcante, R. C., Scheuermann, R. H., & Palatnik-de-Sousa, C. B. (2022). A novel vaccine based on SARS-CoV-2 CD4+ and CD8+ T cell conserved epitopes from variants Alpha to Omicron. Scientific Reports, 12(1), 16731. https://doi.org/10.1038/s41598-022-21207-2

Petruzziello, A., Marigliano, S., Loquercio, G., Cozzolino, A., & Cacciapuoti, C. (2016). Global epidemiology of hepatitis C virus infection: An up-date of the distribution and circulation of hepatitis C virus genotypes. World Journal of Gastroenterology, 22(34), 7824–7840. https://doi.org/10.3748/wjg.v22.i34.7824

Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Research, 48(W1), W449–W454. https://doi.org/10.1093/nar/gkaa379

Rinaldi, L., Perrella, A., Guarino, M., De Luca, M., Piai, G., Coppola, N., Pafundi, P. C., Ciardiello, F., Fasano, M., Martinelli, E., Valente, G., Nevola, R., Monari, C., Miglioresi, L., Guerrera, B., Berretta, M., Sasso, F. C., Morisco, F., Izzi, A., & Adinolfi, L. E. (2019). Incidence and risk factors of early HCC occurrence in HCV patients treated with direct acting antivirals: A prospective multicentre study. Journal of Translational Medicine, 17(1), 292. https://doi.org/10.1186/s12967-019-2033-x

Rodrigues-da-Silva, R. N., Conte, F. P., da Silva, G., Carneiro-Alencar, A. L., Gomes, P. R., Kuriyama, S. N., Neto, A. A. F., & Lima-Junior, J. C. (2023). Identification of B-Cell Linear Epitopes in the Nucleocapsid (N) Protein B-Cell Linear Epitopes Conserved among the Main SARS-CoV-2 Variants. Viruses, 15(4). https://doi.org/10.3390/v15040923

Rouzbahani, A. K., Kheirandish, F., & Hosseini, S. Z. (2022). Design of a multi-epitope-based peptide vaccine against the S and N proteins of SARS-COV-2 using immunoinformatics approach. The Egyptian Journal of Medical Human Genetics, 23(1), 16. https://doi.org/10.1186/s43042-022-00224-w

Samad, A., Ahammad, F., Nain, Z., Alam, R., Imon, R. R., Hasan, M., & Rahman, M. S. (2022). Designing a multi-epitope vaccine against SARS-CoV-2: An immunoinformatics approach. Journal of Biomolecular Structure & Dynamics, 40(1), 14–30. https://doi.org/10.1080/07391102.2020.1792347

Semmo, N., & Klenerman, P. (2007). CD4+ T cell responses in hepatitis C virus infection. World Journal of Gastroenterology, 13(36), 4831–4838. https://doi.org/10.3748/wjg.v13.i36.4831

Shayeghpour, A., Kianfar, R., Hosseini, P., Ajorloo, M., Aghajanian, S., Hedayat Yaghoobi, M., Hashempour, T., & Mozhgani, S.-H. (2021). Hepatitis C virus DNA vaccines: A systematic review. Virology Journal, 18(1), 248. https://doi.org/10.1186/s12985-021-01716-8

Sher, H., Sharif, H., Zaheer, T., Khan, S. A., Ali, A., Javed, H., & Javed, A. (2023). Employing computational tools to design a multi-epitope vaccine targeting human immunodeficiency virus-1 (HIV-1). BMC Genomics, 24(1), 276. https://doi.org/10.1186/s12864-023-09330-4

Simbulan, A. M., Banico, E. C., Sira, E. M. J. S., Odchimar, N. M. O., & Orosco, F. L. (2024). Immunoinformatics-guided approach for designing a pan-proteome multi-epitope subunit vaccine against African swine fever virus. Scientific Reports, 14(1), 1354. https://doi.org/10.1038/s41598-023-51005-3

Singh, A., Thakur, M., Sharma, L. K., & Chandra, K. (2020). Designing a multi-epitope peptide based vaccine against SARS-CoV-2. Scientific Reports, 10(1), 16219. https://doi.org/10.1038/s41598-020-73371-y

Soltan, M. A., Behairy, M. Y., Abdelkader, M. S., Albogami, S., Fayad, E., Eid, R. A., Darwish, K. M., Elhady, S. S., Lotfy, A. M., & Alaa Eldeen, M. (2022). In silico Designing of an Epitope-Based Vaccine Against Common E. coli Pathotypes. Frontiers in Medicine, 9. https://doi.org/10.3389/fmed.2022.829467

Soto, L. F., Requena, D., & Fuxman Bass, J. I. (2022). Epitope-Evaluator: An interactive web application to study predicted T-cell epitopes. PLOS ONE, 17(8), e0273577. https://doi.org/10.1371/journal.pone.0273577

Sugrue, J. A., Posseme, C., Tan, Z., Pou, C., Charbit, B., Bondet, V., Bourke, N. M., Brodin, P., Duffy, D., & O'Farrelly, C. (2022). Enhanced TLR3 responsiveness in hepatitis C virus resistant women from the Irish anti-D cohort. Cell Reports. Medicine, 3(11), 100804. https://doi.org/10.1016/j.xcrm.2022.100804

Tarrahimofrad, H., Rahimnahal, S., Zamani, J., Jahangirian, E., & Aminzadeh, S. (2021). Designing a multi-epitope vaccine to provoke the robust immune response against influenza A H7N9. Scientific Reports, 11(1), 24485. https://doi.org/10.1038/s41598-021-03932-2

Thimme, R. (2021). T cell immunity to hepatitis C virus: Lessons for a prophylactic vaccine. Journal of Hepatology, 74(1), 220–229. https://doi.org/10.1016/j.jhep.2020.09.022

Tzarum, N., Giang, E., Kadam, R. U., Chen, F., Nagy, K., Augestad, E. H., Velázquez-Moctezuma, R., Keck, Z.-Y., Hua, Y., Stanfield, R. L., Dreux, M., Prentoe, J., Foung, S. K. H., Bukh, J., Wilson, I. A., & Law, M. (2020). An alternate conformation of HCV E2 neutralizing face as an additional vaccine target. Science Advances, 6(30), eabb5642. https://doi.org/10.1126/sciadv.abb5642

Wang, C. Y., Kuo, B.-S., Lee, Y.-H., Ho, Y.-H., Pan, Y.-H., Yang, Y.-T., Chang, H.-C., Fu, L.-F., & Peng, W.-J. (2024). UB-612 pan-SARS-CoV-2 T cell immunity-promoting vaccine protects against COVID-19 moderate-severe disease. iScience, 27(2). https://doi.org/10.1016/j.isci.2024.108887

World Health Organization (WHO). (n.d.). Hepatitis C. Retrieved August 29, 2023, from https://www.who.int/news-room/fact-sheets/detail/hepatitis-c

Ysrafil, Y., Sapiun, Z., Astuti, I., Anasiru, M. A., Slamet, N. S., Hartati, H., Husain, F., & Damiti, S. A. (2022). Designing multi-epitope based peptide vaccine candidates against SARS-CoV-2 using immunoinformatics approach. BioImpacts : BI, 12(4), 359–370. https://doi.org/10.34172/bi.2022.23769

Zhou, Y., Wang, X., Sun, L., Zhou, L., Ma, T.-C., Song, L., Wu, J.-G., Li, J.-L., & Ho, W.-Z. (2016). Toll-like receptor 3-activated macrophages confer anti-HCV activity to

hepatocytes through exosomes. FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology, 30(12), 4132–4140. https://doi.org/10.1096/fj.201600696R

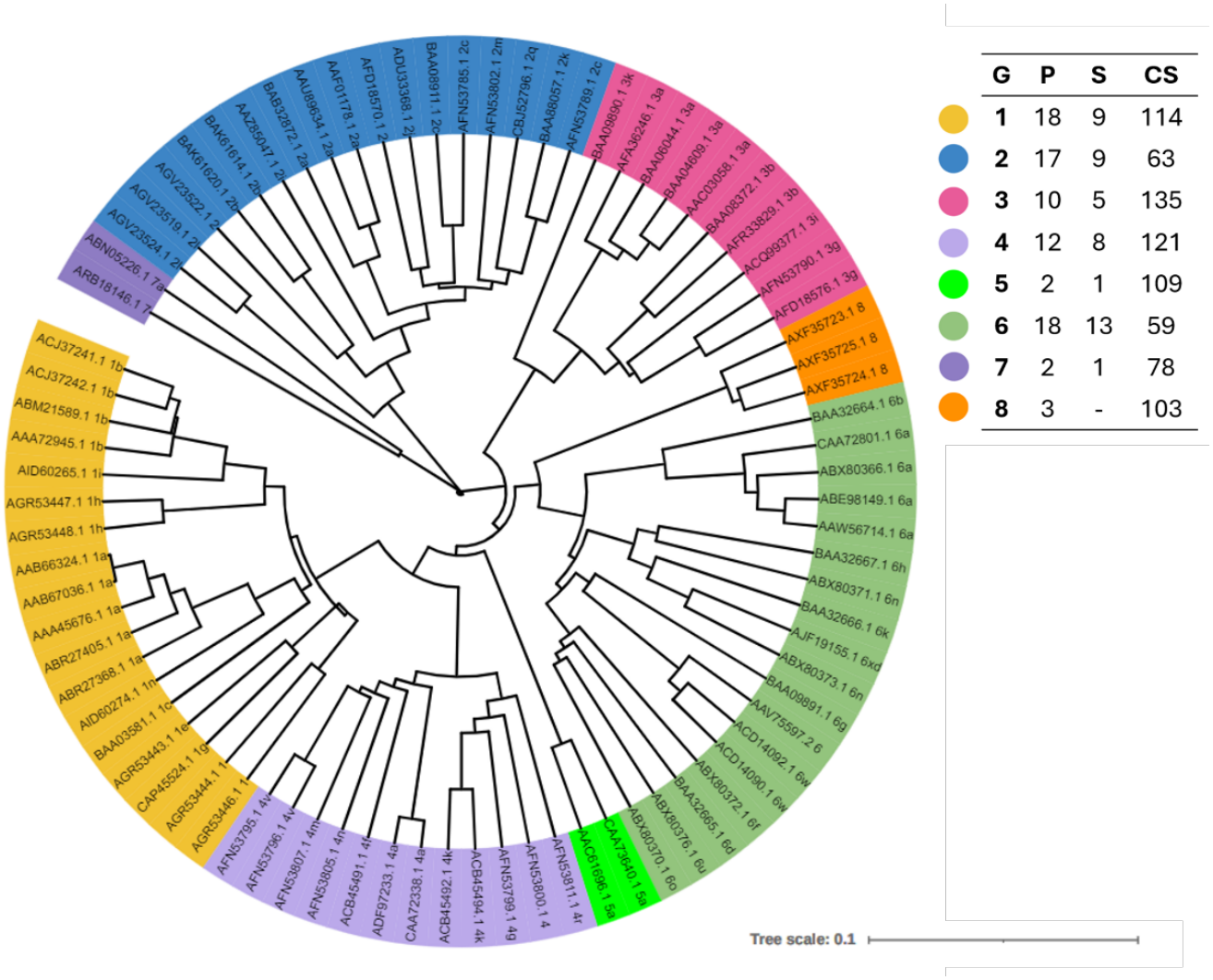

Figure S1: Neighbor-joining phylogenetic tree based on HCV whole proteome sequence showing 8 genotypes and a summary of the number of proteomes, subtypes, and conserved sequences screened for each genotype. Bootstrap method with 1000 replications. G – HCV Genotypes, P – number of Proteomes screened per genotype, S – number of subtypes screened per genotype, CS – number of conserved sequences per genotype.

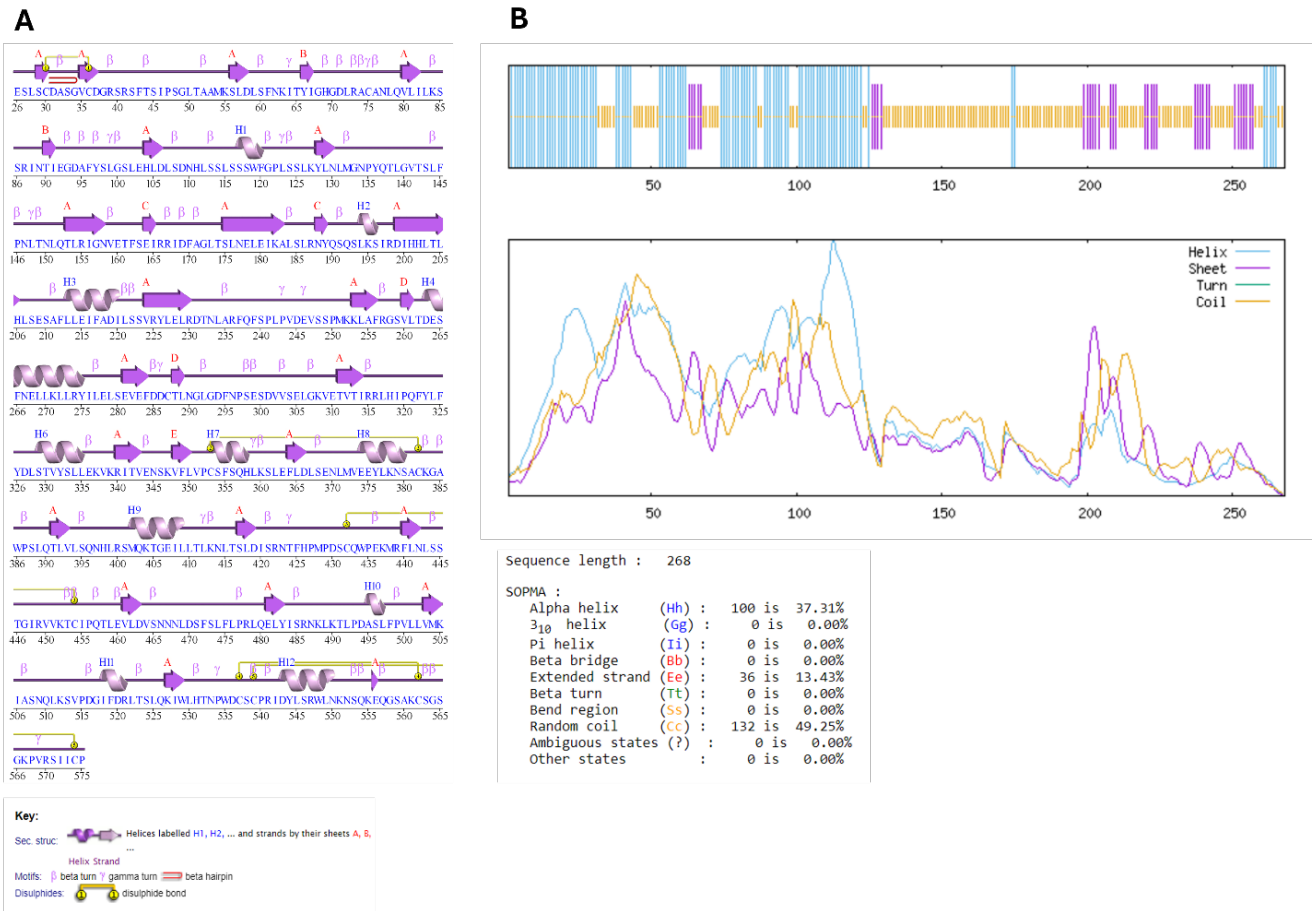

Figure S2: Predicted secondary structure of the final HCV multi-epitope subunit vaccine and its composition. (A) PDBSUM generate 2D structure prediction (B) SOPMA 2D structure prediction.

Table S1: Panel of the 82 HCV proteomes with their corresponding description, genotype, subtype, accession numbers, and the source profile by country.

No	DESCRIPTION	GENOTYPE/ SUBTYPE	DATABASE	PROTEIN ID	GENOME ID	COUNTRY (sourced/published or published only)
1	Hepacivirus C genotype 1_isolate 160526	1	NCBI	AGR53444.1	KC248195.1	UK/USA
2	Hepacivirus C subtype 1n_isolate QC74	1n	NCBI	AID60274.1	KJ439781.1	CANADA/USA
3	Hepacivirus C subtype 1a_isolate 5003	1a	NCBI	ABR27368.1	EF407419.1	USA
4	Hepacivirus C subtype 1a_isolate 1003	1a	NCBI	ABR27405.1	EF407457.1	USA
5	hepatitis C virus genotype 1a	1a	NCBI	AAA45676.1	M62321.1	USA
6	Hepatitis C virus subtype 1a H77	1a	NCBI	AAB66324.1	AF009606.1	USA
7	Hepatitis C virus strain H77 pCV-H77C	1a	NCBI	AAB67036.1	AF011751.1	USA
8	Hepatitis C virus subtype 1b	1b	NCBI	AAA72945.1	M58335.1	JAPAN
9	Hepatitis C virus subtype 1b isolate BR1427_P1_10.7.03	1b	NCBI	ABM21589.1	EF032892.1	USA
10	Hepatitis C virus subtype 1b isolate L801,	1b	NCBI	ACJ37241.1	EU781827.1	USA
11	Hepatitis C virus subtype 1b isolate TN28	1b	NCBI	ACJ37242.1	EU781828.1	USA
12	Hepatitis C virus (isolate HC-G9)	1c	NCBI	BAA03581.1	D14853.1	JAPAN
13	Hepatitis C virus isolate 148636	1e	NCBI	AGR53443.1	KC248194.1	UK/USA
14	Hepatitis C virus subtype 1g	1g	NCBI	CAP45524.1	AM910652.2	SPAIN
15	Hepatitis C virus isolate EBW9	1h	NCBI	AGR53448.1	KC248199.1	CAMEROON/USA
16	Hepatitis C virus isolate EBW443	1h	NCBI	AGR53447.1	KC248198.1	CAMEROON/USA
17	Hepatitis C virus isolate QC181	1i	NCBI	AID60265.1	KJ439772.1	CANADA/USA
18	Hepatitis C virus isolate EBW424	1l	NCBI	AGR53446.1	KC248197.1	CAMEROON/USA
19	Hepatitis C virus isolate QC297	2	NCBI	AFD18570.1	JF735117.1	CANADA/USA
20	Hepatitis C virus isolate MRS40	2	NCBI	AGV23522.1	KC197238.1	FRANCE
21	Hepatitis C virus (isolate JFH-1)	2a	NCBI	BAB32872.1	AB047639.1	JAPAN
22	Hepatitis C virus subtype 2a strain HC-J6CH clone pJ6CF	2a	NCBI	AAF01178.1	AF177036.1	USA
23	Hepatitis C virus genotype 2a	2a	NCBI	AAU89634.1	AY746460.1	JAPAN
24	Hepatitis C virus subtype 2b, patient No.14	2b	NCBI	BAK61614.1	AB661382.1	JAPAN
25	Hepatitis C virus subtype 2b, patient No.26	2b	NCBI	BAK61620.1	AB661388.1	JAPAN
26	Hepatitis C virus subtype 2c isolate HCV-2c/GB/BID-G1238	2c	NCBI	AFN53785.1	JX227949.1	UK
27	Hepatitis C virus (isolate BEBE1)	2c	NCBI	BAA08911.1	D50409.1	ITALY/JAPAN
28	Hepatitis C virus (isolate D54)	2i	NCBI	AAZ85047.1	DQ155561.1	VIETNAM/NETHERLANDS
29	Hepatitis C virus isolate C1292	2j	NCBI	ADU33368.1	HM777359.1	VENEZUELA
30	Hepatitis C virus subtype 2k isolate HCV-2k/GB/BID-G1242	2k	NCBI	AFN53789.1	JX227953.1	UK/USA
31	Hepatitis C virus (isolate VAT96)	2k	NCBI	BAA88057.1	AB031663.1	MOLDOVA/JAPAN
32	Hepatitis C virus isolate MRS89	2l	NCBI	AGV23519.1	KC197235.1	FRANCE
33	Hepatitis C virus isolate PTR7904	2l	NCBI	AGV23524.1	KC197240.1	FRANCE
34	Hepatitis C virus subtype 2m isolate HCV-2m/CA/BID-G1314	2m	NCBI	AFN53802.1	JX227967.1	CANADA/USA
35	Hepatitis C virus subtype 2q, isolate 852	2q	NCBI	CBJ52796.1	FN666429.1	SPAIN
36	Hepatitis C virus subtype 3a isolate RASILBS2-SR-PO	3a	NCBI	AFA36246.1	JN714194.1	INDIA
37	Hepatitis C virus (isolate HCV-K3a/650)	3a	NCBI	BAA06044.1	D28917.1	JAPAN
38	Hepatitis C virus subtype 3a strain CB	3a	NCBI	AAC03058.1	AF046866.1	AUSTRALIA
39	Hepatitis C virus (isolate NZL1)	3a	NCBI	BAA04609.1	D17763.1	NEW ZEALND/JAPAN
40	Hepatitis C virus isolate SH37	3b	NCBI	AFR33829.1	JQ065709.1	CHINA
41	Hepatitis C virus (isolate Tr Kj)	3b	NCBI	BAA08372.1	D49374.1	JAPAN
42	Hepatitis C virus isolate QC260	3g	NCBI	AFD18576.1	JF735123.1	CANADA/USA
43	Hepatitis C virus subtype 3g isolate HCV-3g/GB/BID-G1243	3g	NCBI	AFN53790.1	JX227954.1	UK/USA
44	Hepatitis C virus (isolate JK049)	3k	NCBI	BAA09890.1	D63821.1	INDONESIA/JAPAN

No	DESCRIPTION	GENOTYPE/ SUBTYPE	DATABASE	PROTEIN ID	GENOME ID	COUNTRY (sourced/published or published only)
45	Hepatitis C virus isolate IND-HCV-3i	3i	NCBI	ACQ99377.1	FJ407092.1	INDIA
46	Hepatitis C virus genotype 4 isolate BID-G1253	4	NCBI	AFN53800.1	JX227964.1	UK/USA
47	Hepatitis C virus ED43	4a	NCBI	CAA72338.1	Y11604.1	EGYPT/UK
48	Hepatitis C virus ED43	4a	NCBI	ADF97233.1	GU814265.1	EGYPT/UK
49	Hepatitis C virus isolate CM_SP1578 genotype	4f	NCBI	ACB45491.1	EU392170.1	USA
50	Hepatitis C virus subtype 4g isolate HCV-4g/GB/BID-G1252	4g	NCBI	AFN53799.1	JX227963.1	UK/USA
51	Hepatitis C virus isolate PB65185	4k	NCBI	ACB45492.1	EU392171.1	USA
52	Hepatitis C virus isolate PS3	4k	NCBI	ACB45494.1	EU392173.1	USA
53	Hepatitis C virus subtype 4m isolate HCV-4m/GB/BID-G1657	4m	NCBI	AFN53807.1	JX227972.1	UK/USA
54	Hepatitis C virus subtype 4n isolate HCV-4n/GB/BID-G1655	4n	NCBI	AFN53805.1	JX227970.1	UK/USA
55	Hepatitis C virus subtype 4r isolate HCV-4r/GB/BID-G2036	4r	NCBI	AFN53811.1	JX227976.1	UK/USA
56	Hepatitis C virus subtype 4v isolate HCV-4v/GB/BID-G1248	4v	NCBI	AFN53795.1	JX227959.1	UK/USA
57	Hepatitis C virus subtype 4v isolate HCV-4v/GB/BID-G1249	4v	NCBI	AFN53796.1	JX227960.1	UK/USA
58	Hepatitis C virus (isolate SA13)	5a	NCBI	AAC61696.1	AF064490.1	USA
59	Hepatitis C virus (isolate EUH1480)	5a	NCBI	CAA73640.1	Y13184.1	UK
60	Hepatitis C virus isolate GZ52557	6	NCBI	AAV75597.2	DQ278892.1	CHINA
61	Hepatitis C virus (isolate EUHK2)	6a	NCBI	CAA72801.1	Y12083.1	HONGKONG/UK
62	Hepatitis C virus strain D9	6a	NCBI	ABX80366.1	EU246930.1	VIETNAME/NETHERLANDS
63	Hepatitis C virus (isolate 6a33)	6a	NCBI	AAW56714.1	AY859526.1	HONGKONG
64	Hepatitis C virus subtype 6a strain 6a35	6a	NCBI	ABE98149.1	DQ480513.1	HONGKONG
65	Hepatitis C virus (isolate Th580)	6b	NCBI	BAA32664.1	D84262.2	JAPAN
66	Hepatitis C virus (isolate VN235)	6d	NCBI	BAA32665.1	D84263.2	CANADA
67	Hepatitis C virus strain TH52	6f	NCBI	ABX80372.1	EU246936.1	THAILAND/NETHERLANDS
68	Hepatitis C virus strain TH24	6i	NCBI	ABX80371.1	EU246935.1	THAILAND/NETHERLANDS
69	Hepatitis C virus strain TH22	6n	NCBI	ABX80373.1	EU246937.1	THAILAND/NETHERLANDS
70	Hepatitis C virus strain D85	6o	NCBI	ABX80370.1	EU246934.1	VIETNAME/NETHERLANDS
71	Hepatitis C virus strain D83	6u	NCBI	ABX80376.1	EU246940.1	VIETNAME/NETHERLANDS
72	Hepatitis C virus isolate HCV-6-D140	6w	NCBI	ACD14090.1	EU643834.1	TAIWAN
73	Hepatitis C virus isolate HCV-6-D370	6w	NCBI	ACD14092.1	EU643836.1	TAIWAN
74	Hepatitis C virus isolate km98	6xd	NCBI	AJF19155.1	KM252792.1	CHINA
75	Hepatitis C virus (isolate VN405)	6k	NCBI	BAA32666.1	D84264.2	VIETNAM/JAPAN
76	Hepatitis C virus (isolate VN004)	6h	NCBI	BAA32667.1	D84265.2	VIETNAM/JAPAN
77	Hepatitis C virus (isolate JK046)	6g	NCBI	BAA09891.1	D63822.1	INDONESIA/JAPAN
78	Hepatitis C virus QC69	7a	NCBI	ABN05226.1	EF108306.2	CANADA
79	Hepatitis C virus genotype 7	7	NCBI	ARB18146.1	KU861171.1	AFRICA/UK
80	Hepatitis C virus genotype 8 (GT8-2)	8	NCBI	AXF35723.1	MH590699.1	CANADA/USA
81	Hepatitis C virus genotype 8 (GT8-3)	8	NCBI	AXF35724.1	MH590700.1	CANADA/USA
82	Hepatitis C virus genotype 8 (GT8-4)	8	NCBI	AXF35725.1	MH590701.1	CANADA/USA

Table S2: Conserved proteins and the selected MHC I, MHC II, and B cell epitopes based on epitope binding threshold scores.

Table S3: Physicochemical Properties, antigenicity, allergenicity, and stability of the HCV vaccine constructs (HCV VC).

adjuvant 1: Human β-defensin						
HCV VC (cut-off scores)	**HCV VC 1**	**HCV VC 2**	**HCV VC 3**	**HCV VC 4**	**HCV VC 5**	**HCV VC 6**
linkers	AAY-GPGPG-KK	KK-GPGPG-KK	AAY-hey-GGGS-hey-KK	KK-hey-GGGS-hey-KK	AAY-hey-GPGPG-hey-KK	KK-hey-GPGPG-hey-KK
length (AA)	184	180	200	206	200	196
Solubility (≥0.95)	0.861051	0.917052	0.950924	0.95169	0.935839	0.951444
Antigenicity (≥0.40)	0.5298	0.5170	0.5374	0.4923	0.5045	0.5258
allergenicity	non-allergen	non-allergen	allergen	non-allergen	non-allergen	non-allergen
molecular weight (kDa)	20084.52	19888.58	21924.33	21835.54	22031.48	21728.39
isoelectric point (pH 9-10)	10.14	10.42	9.86	10.13	9.86	10.13
aliphatic index	60.87	57.77	62.91	59.8	62.6	60.1
Grand Average of Hydropathicity (GRAVY)	-0.627	-0.866	-0.615	-0.842	-0.624	-0.834
Stability (≤40)	44.93	45.39	42.93	42.52	42.15	43.32
adjuvant 2: ribosomal L7/L12 protein						
HCV VC (cut-off scores)	**HCV VC 7**	**HCV VC 8**	**HCV VC 9**	**HCV VC 10**	**HCV VC 11**	**HCV VC 12**
linkers	AAY-GPGPG-KK	KK-GPGPG-KK	AAY-hey-GGGS-hey-KK	KK-hey-GGGS-hey-KK	AAY-hey-GPGPG-hey-KK	KK-hey-GPGPG-hey-KK
length (AA)	266	268	234	284	280	285
Solubility (≥0.95)	0.97	0.97	0.66	0.97	0.96	0.97
Antigenicity (≥0.40)	0.48	0.45	0.47	0.46	0.45	0.44
allergenicity	non-allergen	non-allergen	non-allergen	non-allergen	non-allergen	non-allergen
molecular weight (kDa)	28149.5	28363.81	24915.02	30203.62	30007.67	30310.77
isoelectric point (pH 9-10)	9.27	9.27	9.6	8.72	9.35	8.72
aliphatic index	81.24	80.63	77.69	80.95	79.25	80.67
Grand Average of Hydropathicity (GRAVY)	-0.207	-0.223	-0.413	-0.237	-0.385	-0.245
Stability (≤40)	33.69	31.42	31.11	30.78	30.88	30.27

Table S4: MM/GBSA calculations of the final HCV VC-TLR complexes using HawkDock.

HCV VC-TLR Complexes			
Cluster models	**TLR2**	**TLR3**	**TLR4**
0	-80.09	-117.02	-107.53
1	-79.98	-123.12	-175.68
2	-107.05	**-226.61**	-175.21
3	-207.98	445.74	-180.09
4	**-215.89**	-126.8	-55.76
5	-66.71	-108.14	**-182.94**
6	-219.05	-115.72	-164.63
7	-205.66	-156.89	-142.99
8	-99.1	-113.18	-123.98
9	-161.7	-172.2	343.99

Table S5: Linear B cell epitopes of the final HCV VC predicted by ABCPred server.

	Linear B cell	**ABCPred score**
1	SGKSTAAYQPRGRRQP	0.96
2	VEAAEEQSEFDVILEA	0.86
3	CHSKKKGPGPGQGYKV	0.82
4	AAPAGAAVEAAEEQSE	0.81
5	RRQPIAAYKTSERSQP	0.81
6	AKAPTGSGKSTAAYQP	0.81
7	LEAAGATVTVKEAAAK	0.81
8	SVAATGPGPGGGRHLI	0.79
9	YHLIFCHSKKAAYLIF	0.79
10	KVAKEAADEAKAKLEA	0.79
11	LGLKEAKDLVDGAPKP	0.78
12	GGRHLIFCHSKKKCDK	0.77
13	GGRHLIFCHSKKKCDE	0.76
14	HSKKAAYLIFCHSKKK	0.76
15	DVILEAAGDKKIGVIK	0.75
16	KDLVDGAPKPLLEKVA	0.73
17	AVAAAGAAPAGAAVEA	0.73
18	VFTGLTHIDAHFLSQT	0.71
19	FCHSKKKCDEKKVFTG	0.7
20	EVTAAAPVAVAAAGAA	0.68
21	PGPGQGYKVLVLNPSV	0.68
22	VVREIVSGLGLKEAKD	0.67
23	KTSERSQPRAAYHLIF	0.65
24	HSKKKCDKKGGRHLIF	0.63
25	KVLVLNPSVAATGPGP	0.63
26	AFKEMTLLELSDFVKK	0.63
27	KLSTDELLDAFKEMTL	0.62
28	ADEAKAKLEAAGATVT	0.62

Table S6: Discontinuous B cell epitopes of the final HCV VC predicted by ElliPro server.

No.	**Residues**	**Number of residues**	**Score**
1	A:L265, A:S266, A:Q267, A:T268	4	0.988
2	A:K252, A:V253, A:F254, A:T255, A:G256, A:L257, A:T258, A:H259, A:I260, A:D261, A:A262, A:H263, A:F264	13	0.918
3	A:R150, A:G151, A:R152	3	0.914
4	A:M1, A:A2, A:K3, A:L4, A:S5, A:T6, A:D7, A:E8, A:L9, A:L10, A:D11, A:A12, A:F13, A:K14, A:E15, A:M16, A:T17, A:L18, A:L19, A:E20, A:L21	21	0.779
5	A:S244, A:K246, A:K247, A:C248, A:D249, A:E250, A:K251	7	0.767
6	A:Q154, A:P155, A:I156, A:A157, A:A158, A:Y159, A:K160, A:T161, A:S162, A:E163	10	0.755
7	A:S140, A:G141, A:K142, A:S143, A:T144, A:A145, A:A146, A:Y147, A:Q148, A:P149	10	0.728
8	A:A37, A:A38, A:P39, A:V40, A:A41, A:V42, A:A43, A:A44, A:A45, A:G46, A:A47, A:A48, A:P49, A:A50, A:G51, A:A52, A:A53, A:V54, A:E55	19	0.716
9	A:K230, A:C231, A:D232	3	0.715
10	A:K234, A:G235, A:G236, A:R237, A:H238	5	0.677
11	A:D23, A:F24, A:K27, A:F28, A:E30, A:T31, A:F32	7	0.649