Acta Biologica Szegediensis

Volume 67(1):11-23, 2023
Acta Biologica Szegediensis
http://abs.bibl.u-szeged.hu/index.php/abs
DOI:10.14232/abs.2023.1.11-23

ARTICLE

In silico immunoinformatics based prediction and designing of multi-epitope construct against human rhinovirus C

Saubashya Sur[1]*, Mritunjoy Ghosh[1], Ritu Rai[2]

[1]Postgraduate Department of Botany, Life Sciences Block, Ramananda College, Bishnupur-722122, West Bengal, India
[2]Department of Botany, Parimal Mitra Smriti Mahavidyalaya, Malbazar-735221, West Bengal, India

ABSTRACT Human rhinovirus C (HRV-C) is an RNA virus infecting human respiratory tract. It is associated with complexities like asthma, chronic obstructive pulmonary disease, and respiratory damage. HRV-C has many serotypes. Till date there is no vaccine. Despite some limitations, corticosteroids, bronchodilators, and common cold medicines are used to treat HRV-C infections. Here, we have used immunoinformatics approach to predict suitable cytotoxic T-cell, helper T-cell and linear B-cell epitopes from the most antigenic protein. VP2 protein of Rhinovirus C53 strain USA/CO/2014-20993 was found to be most antigenic. The multi-epitope construct was designed using the best CTL, HTL and linear B-cell epitopes and attaching them with adjuvant and linkers. Interferon-gamma inducing epitopes and conformational B-cell epitopes were also predicted from the construct. Physicochemical and structural properties of the construct were satisfactory. Binding pockets were identified that could be the targets for designing effective inhibitors. Molecular docking revealed strong binding affinity of the construct with human Toll-like receptors 2 and 4. Normal mode analysis divulged stability of the docked complex. Codon optimization, *in silico* cloning and immune simulation analysis demonstrated suitability of the construct. These findings are likely to aid *in vitro* studies for developing vaccine against HRV-C.

KEY WORDS

Human rhinovirus C
immunoinformatics
linker
molecular docking
multi-epitope
toll-like receptors

ARTICLE INFORMATION

Submitted
30 April 2023
Accepted
30 June 2023
***Corresponding author**
E-mail: saubashya@gmail.com

Introduction

Human rhinoviruses (HRVs), first discovered in 1950's, are non-enveloped RNA viruses belonging to the Picornaviridae family (Glanville and Johnston 2015). It infects the upper and lower respiratory tracts in humans (Arruda et al. 1995; Jakiela et al. 2008) and is accountable for acute respiratory complexities in various ethnicities worldwide (Rotbart and Hayden 2000). HRVs are associated with common cold, wheezing, asthma, pneumonia, chronic obstructive pulmonary disease, and flu-like symptoms (Arden and Mackay 2009; Cordey et al. 2010). They have a high rate of mutation assisting adaptability and transmissibility (Cordey et al. 2010). HRVs are categorized into HRV-A, HRV-B, and HRV-C respectively (Hao et al. 2012). Multiple lines of evidence have revealed that HRV-C is more predominant and virulent compared to HRV-A and HRV-B (Hao et al. 2012). The high virulence of rhinovirus C stems from its ability to bind to host cells using cadherin-related family member 3 receptor (Scully et al. 2018). HRV-C is linked to severe symptoms. It is responsible for greater respiratory damage (Palmenberg et al. 2010). HRV-C has been linked to asthma exacerbations worldwide in children (Bizzintino et al. 2011; Mak et al. 2011). Some workers have found a distinct correlation between maternal atopy and asthma in offspring (Miller et al. 2011). HRVs comprise many serotypes whose categorization is based on factors like receptor specificity, predisposition to antiviral responses, similarity in nucleotide sequences, etc. (Lau et al. 2010). The genome of HRVs is made up of a single gene. Nevertheless, its translated product yields structural and non-structural proteins (Jacobs et al. 2013). The capsid contains structural proteins viz. VP1, VP2, VP3, VP4, and VPg whereas non-structural proteins function in replication and assembly (Palmenberg et al. 2010).

Acute airway infections are the major cause of morbidity and mortality worldwide. Although HRV-C is more virulent and linked to the high incidence of asthma, chronic obstructive pulmonary disease in adults, and severe respiratory complexities in children (Bochkov and Gern 2012), little has been achieved in developing a vaccine in the last 70 years. However, the development of effective vaccines is time-consuming and costly. For clinicians, the antigenic diversity of HRVs, the number of serotypes along with lack of good animal models became stumbling blocks for developing vaccines (Papi and Con-

toli 2011). The capsid proteins of HRV-C viz. VP1, VP2, VP3, VP4, and VPg arranged in an icosahedral structure have been the target of many antiviral studies (Stone and Miller 2016). However, the development of effective therapeutics remained partially successful resulting in seasonal occurrences of HRV-C-mediated illness. Research on immunized rats has illustrated the possibility of developing cross-serotype protection (McLean 2014). A model tested in rhesus macaques indicated the effectiveness of a polyvalent inactivated vaccine in generating virus-neutralizing antibodies against HRV types (Lee et al. 2016). Some workers hypothesized the use of a highly conserved and immunogenic epitope to design a universal rhinovirus vaccine that can protect against all serotypes (Stepanova et al. 2019). However, a lot needs to be done on human models. In the absence of a suitable vaccine, the treatment has been supportive in the form of non-prescription cold remedies, and the use of corticosteroids and bronchodilators targeted at providing symptomatic relief (Glanville and Johnston 2015; Jacobs et al. 2013). Again, inhalation of glucocorticoids failed to provide a positive outcome in acute cases (Papi and Contoli 2011).

The objective of this work has been in silico design of multi-epitope construct housing appropriate properties, using an immunoinformatics approach that may be efficient against HRV-C. With newer technologies in bioinformatics, and the advent of high-quality immunoinformatics tools, designing multi-epitope-based constructs has been successful (Korber et al. 2006; Purcell et al. 2006; Seib et al. 2012). Epitope-based vaccines are known to have improved potency with better safety issues (Majid and Andleeb 2019). In the present study, robust bioinformatics analyses were carried out to predict the most antigenic protein and determine the most suitable cytotoxic T lymphocyte (CTL), helper T cell (HTL), and linear B cell epitopes. The multi-epitope construct was designed with the addition of adjuvant, linkers and tag. Antigenicity, immunogenicity, MHC allele binding capability, non-allergenicity, non-toxicity, and physicochemical and structural properties of the construct were assessed. This was followed by molecular docking, codon optimization, and immune simulation studies. The outcome is aimed at supporting clinicians and experimental biologists to validate the findings from this work and facilitate HRV -C vaccine development.

Materials and methods

Retrieval and screening of sequences

Amino acid sequences of the polyprotein of human rhinovirus C (HRV-C) strains/isolates were visualized in the Bacterial and Viral Protein Resource Center Database (https://www.bv-brc.org/). We strictly restricted the dataset to complete sequences and discarded partial and duplicated sequences. This was vital for quality control. Thus, as of 30/09/022, we retrieved 231 sequences of HRV-C strains/isolates (Supplementary Table 1) for performing the analysis. It was observed that the polyproteins of each of these strains/isolates had structural and non-structural proteins. All the non-structural proteins were omitted and structural proteins like VP1, VP2, VP3, VP4, and VPg were considered for the analysis. Hence, amino acid sequences of the proteins in FASTA format from the selected HRV-C strains/isolates were considered for antigenicity prediction.

Identification of the most antigenic protein

Vaxijen v2.0 (Doytchinova and Flower 2007) was used to predict the most antigenic protein amongst the selected structural proteins. A threshold of 0.4 (Doytchinova and Flower 2007) was set.

Prediction of cytotoxic T lymphocytes (CTL) epitopes

CTL epitopes are significant for vaccine design since these peptide fragments have the property of stimulating immune responses (Adhikari et al. 2018). NetCTL 1.2 (http://www.cbs.dtu.dk/services/NetCTL/) server was used to predict interacted cytotoxic T-lymphocyte epitopes. The epitopes were determined based on MHC class I supertypes (A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62). For determining the epitope, a threshold was set at 0.75, having a sensitivity of 0.80 and specificity of 0.97 (Larsen et al. 2007). The initial selection was based on the highest combined score; however, the final selection was made after predicting antigenicity and immunogenicity of the epitopes using Vaxijen v2.0 (Doytchinova and Flower 2007) and IEDB server (www.iedb.org). The epitopes were investigated using stabilized matrix base method (SMM) in the IEDB analysis tool (http://tools.iedb.org/mhci/) (Peters et al. 2005) for determining common and non-frequently occurring MHC-I-binding alleles. To predict MHC-I-binding alleles the cut-off for amino acid length of the peptide was set at 9.0 (Peters et al. 2005) and IC50 value less than 200 since, high binding affinities to MHC-I binding alleles were possible with IC50 values lower than 200 (Khatoon et al. 2017). Peptides having IC50 values < 50nM indicated the highest affinity, followed by <500nM with intermediate affinity, and <5000nM with low affinity respectively (Khatoon et al. 2017). The NetMHCII 2.3 server was utilized to estimate the binding of peptides to HLA-DR, HLA-DQ, and HLA-DP alleles (Nielsen and Lund

Table 1. Physicochemical properties of the multi-epitope construct.

Features	Value
Number of amino acids	250
Molecular weight	26128.81 Da
Theoretical pI	5.28
Total number of negatively charged residues (Asp + Glu)	34
Total number of positively charged residues (Arg + Lys)	27
Molecular formula	C1171 H1877 N307 O361 S3
Total number of atoms	3719
Extinction coefficient	21430
Instability index	20.31
Aliphatic index	91.48
Grand average of hydropathicity (GRAVY)	-0.115

2009) using default parameters. The prediction was based on receptor affinity, inferred from the IC50 values. Conservancy (http://tools.iedb.org/conservancy/) and immunogenicity http://tools.iedb.org/immunogenicity/) tools of IEDB database were used (Nielsen et al. 2007; Calis et al. 2013) to predict epitope conservancy and immunogenicity.

Prediction of helper T lymphocyte epitope

The Net MHCII 2.3 server (https://services.healthtech.dtu.dk/service.php?NetMHCII-2.3) was used to predict HTL epitopes of 15-mer length for human alleles. Furthermore, the binding capacity of the epitopes to HLA-DR, HLA-DQ, and HLA-DP alleles was also determined based on IC50 values.

Identification of linear B-cell epitopes

The identification of suitable B-cell epitopes was crucial for designing vaccines since they are known to interact with B lymphocytes for eliciting immune response (Nair et al. 2002). The BCEPS epitope prediction software (Ras-Carmona et al. 2021) was used to determine the B-cell epitopes from the most antigenic protein using default parameters. The IEDB database was used to predict conservancy.

Evaluation of allergenicity and toxicity

AllergenFP 1.0 (Dimitrov et al. 2014) was employed to evaluate allergenicity or non-allergenicity of the CTL, HTL, and linear B cell epitopes. ToxIBTL was utilized to determine toxicity or non-toxicity of the same (Wei et al. 2022).

Designing multi-epitope construct

The multi-epitope construct was designed using the best CTL, HTL and linear B cell epitopes. Furthermore, adjuvant and linkers were used for effective construction and epitope design. The 50S ribosomal protein L7/L12 (NCBI ID: P9WHE3) was used as adjuvant since it is known to be a human toll-like receptor 4 (TLR4) agonist (Olejink et al. 2018). It was attached to the N-terminal end of the construct with EAAAK linker. The CTL epitopes were linked to the former using AAY (Ala-Ala-Tyr) linkers. Next, the HTL epitopes were linked together with the CTL epitopes using GPGPG (Gly-Pro-Gly-Pro-Gly) linkers. Next, the linear B cell epitope was linked to the HTL epitopes using a KK (Lys-Lys) linker. An EPEA tag was attached to the C-terminal end to facilitate efficient purification (Jin et al. 2017). The AAY, GPGPG and KK linkers are reported to assist in stability, epitope presentation and preservation of immunogenic properties (Abdellrazeq et al. 2020).

Prediction of IFN-γ inducing epitopes

IFN-γ inducing epitopes in the construct were predicted using the hybrid approach in IFNepitope server (Dhanda et al. 2013).

Prediction of antigenicity, allergenicity and toxicity of the multi-epitope construct

Vaxijen v2.0, AllergenFP 1.0, and ToxIBTL were used to predict antigenicity, allergenicity, and toxicity of the multi-epitope construct.

Cross-checking of human homology

BLASTp within the PSI-BLAST algorithm was utilized to detect homology between the multi-epitope construct and human proteome. This was carried out to eliminate the risk of autoimmune response. The BLASTp search was limited to *Homo sapiens* (taxid: 9606).

Prediction of physicochemical properties of the multi-epitope construct

ExPASy ProtParam (Gasteiger et al. 2005) was used to

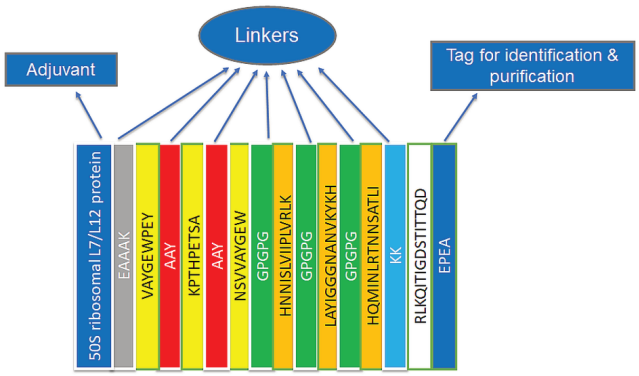

Figure 1. Graphical representation of the multi-epitope construct. CTL, HTL & linear B cell epitopes are colored yellow, red and white.

determine the number of amino acids, molecular weight, extinction coefficient, isoelectric point (pI), instability index, aliphatic index and grand average hydropathicity (GRAVY) values of the multi-epitope construct. The solubility of the construct was predicted using a recombinant protein solubility prediction tool (https://biotech.ou.edu/). It assumed of the protein construct's overexpression in *E. coli* (Diaz et al. 2010).

Extrapolation of the secondary structure and tertiary of the multi-epitope construct

PSIPRED server (McGuffin 2000) was used to predict the secondary structural properties, while DISOPRED3 (Jones and Cozzetto 2015) was used for visualizing disordered regions in the construct. The tertiary structure of the multi-epitope construct was determined using trRosetta (Du et al. 2021). This tertiary structure was refined using GalaxyRefine server (Heo et al. 2013). The refined structure was examined for quality using ProSA (Wiederstein and Sippl 2009), ProQ (Wallner and Elofsson 2003), and SAVES (https://servicesn.mbi.ucla.edu/SAVES/).

Ligand binding pocket prediction

PrankWeb (Jendele et al. 2019) was utilized to predict ligand binding pockets in the multi-epitope construct. It employed machine learning algorithms for determining ligand binding sites with high accuracy.

Determination of conformational B-Cell epitopes

IEDB conformational B-cell prediction tool ElliPro (Ponomarenko et al. 2008) was employed to determine the discontinuous B-cell epitopes of the multi-epitope construct. This tool incorporated various algorithms for stabilizing the shape of proteins and determining PI (protrusion index) (Ponomarenko et al. 2008). The minimum score and maximum distance were set to default scores of 0.5 and 6, respectively (Ponomarenko et al. 2008).

Molecular docking of the construct with human TLR-2 and TLR-4

ClusPro 2.0 server (Kozakov et al. 2017) was used to execute molecular docking of the multi-epitope construct with human TLR-2 (PDB ID: 6NIG) and TLR-4 (PDB ID: 4G8A). Based on energy scores, the cluster having the lowest energy was selected.

Normal mode analysis of the docked complexes

iMODS (Lopez-Blanco et al. 2014) was employed to investigate the collective motion of the complexes formed by the construct and human TLR-2 and 4, respectively, using normal mode analysis.

Codon optimization and in silico cloning

Java Codon Adaptation Tool (JCAT) (http://www.prodoric.de/JCat) was utilized for performing reverse translation of the multi-epitope construct protein sequence as well as codon optimization. Optimization is important for the protein to be expressed as a foreign gene in a host e.g., *E. coli* (strain K12) (Mittal et al. 2020). The output from JCAT reveals codon adaptation index (CAI) value and GC percentage. The former implied protein expression. A CAI value higher than 0.8 suggested suitability (Majid and Andleeb 2019). Likewise, GC content within 30-70% emphasized a positive effect on the transcription and translation (Majid and Andleeb 2019). The optimized sequence was subjected to in silico cloning (Adam et al. 2021) into *E. coli* ELF pET-28a vector and EcoRI and HincII restriction sites were exploited.

Immune profile of the multi-epitope construct

C-ImmSim server (Rapin et al. 2010) was used to analyse immune response of the construct. The default immune simulation parameters viz., random seed: 12,345, simulation steps: 100 and simulation volume: 10 μ L were used. The server used, antigen sequence as input and four matrices for MHC class I, two HLA-A and two HLA-B, and two matrices for MHC class II in addition to other variables and parameters to perform *in silico* simulation. The multi-epitope construct was injected at the time steps of 1, 42, and 84, respectively.

Results

Retrieval of amino acid sequences

The amino acid sequences of the structural proteins from HRV-C strains/isolates, (Supplementary Table 1) were retrieved from the Bacterial and Viral Protein Resource Center Database.

Table 2. Predicted binding pockets in the multi-epitope construct.

Pocket	Pocket score	Probability score	Amino acid count	Residues
1	10.42	0.560	19	156, 157, 158, 159, 160, 162, 168, 169, 170, 171, 174, 175, 176, 178, 179, 180, 181, 182, 248
2	2.94	0.095	10	53, 57, 135, 136, 138, 210, 211, 212, 213, 214,
3	2.66	0.078	11	49, 187, 189, 190, 192, 205, 207, 211, 212, 247, 249
4	1.95	0.040	9	159, 162, 164, 166, 168, 182, 183, 184, 250

Selection of the highest antigenic protein

The structural proteins were analysed with Vaxijen 2.0, and it resulted in the prediction of an overall antigenic score for each protein. It also indicated which protein sequences were antigenic or non-antigenic. It was found that VP2 protein of Rhinovirus C53 strain USA/CO/2014-20993, (GenBank: MF775366.1/protein_id="ASW23000.1") was the most antigenic protein with an overall score of 0.5835. This protein was selected, and it contained 261 residues of amino acid.

Identification of Cytotoxic T lymphocytes (CTL) epitopes and their characterization

NetCTL 1.2 server predicted 9-mer CTL epitopes from the most antigenic protein. It was observed that VAYGEWPEY, KPTHPETSA, and NSVVAYGEW revealed antigenicity and immunogenicity scores of 0.848, 0.6956, 0.5408; and 0.388, 0.064, 0.1766 respectively. These indicated that they possessed an increased probability of eliciting a greater immune response. Interestingly, they were non-allergenic and non-toxic. These CTL epitopes were recognised by different MHC-I alleles from human, i.e., HLA-A, HLA-B, and HLA-C. VAYGEWPEY showed strong binding affinities to the MHC-I alleles, HLA-C*03:03, HLA-C*12:03, HLA-B*35:01, HLA-C*14:02, HLA-A*30:02, HLA-A*29:02. While, KPTHPETSA had strong binding affinities to HLA-C*12:03, HLA-C*03:03, NSVVAYGEW elicited high binding affinities to HLA-C*12:03, HLA-C*03:03, HLA-B*58:01 and HLA-B*57:01. The highest binding to the corresponding alleles were based on IC50 values less than 200 (Khatoon et al. 2017). Lower IC50 values designated greater binding capacity of the epitopes to the MHC-I molecules. However, the outcome of NetMHCII 2.3 server did not indicate suitable binding to the HLA-DR, HLA-DQ, and HLA-DP alleles because of poor IC50 scores. These CTL epitopes were considered for the multi-epitope construct.

Helper T lymphocyte (HTL) epitope estimation

The Net MHCII 2.3 server predicted 15-mer HTL epitopes. The HTL epitopes HNNISLVIIPLVRLK, LAYIGGGNANVKYKH, HQMINLRTNNSATLI were ultimately considered for the multi-epitope construct. They were found to be antigenic, non-allergenic and non-toxic.

Linear B-cell epitope identification

BCEPS epitope prediction tool predicted linear B-cell epitopes. The linear B-cell epitope RLKQITIGDSTITTQD showed the highest antigenicity of 1.2918 and exhibited the properties of non-allergenicity, non-toxicity, and conservancy. It had better properties compared to others. Hence, it was shortlisted for designing the multi-epitope construct.

Preparation of the multi-epitope construct

The multi-epitope construct (Fig. 1) was prepared using the best CTL, HTL, and linear B-cell epitopes, derived from the most antigenic protein. The proper functioning of the epitopes was guaranteed by the insertion of AAY, GPGPG, and KK linkers. These linkers connected the CTL, HTL, and linear B-cell epitopes. The EAAAK linker was used to link the 50S ribosomal protein L7/L12 at the N terminal end. Additionally, the C-terminal end of the multi-epitope construct consisted of an EPEA tag. The formulated multi-epitope construct, after the addition of the adjuvant, linkers, epitopes, and tag comprised of 250 amino acid residues. The final construct was MAKLSTDELLDAFKEMTLLELSDFVKKFEETFEVTAAAPVAVAAAGAAPAGAAVEAAEEQSEFDVILEAAGDKKIGVIKVVREIVSGLGLKEAKDLVDGAPKPLLEKVAKEAADEAKAKLEAAGATVTVKEAAAKVAYGEWPEYAAYKPTHPETSAAAYNSVVAYGEWGPGPGHNNISLVIIPLVRLKGPGPGLAYIGGGNANVKYKHGPGPGHQMINLRTNNSATLIKKRLKQITIGDSTITTQDEPEA.

IFN-γ inducing epitopes

IFN-γ is known to catalyze macrophages and natural killer cells during immune response (Dhanda et al. 2013). Around 242 IFN-γ inducing epitopes (Supplementary Table 2), were predicted from the multi-epitope construct using the SVM methodology in IFN epitope server. All of them were found to be positive.

Antigenicity, non-allergenicity and non-toxicity of the multi-epitope construct

The antigenicity of the multi-epitope construct predicted using VaxiJen 2.0 server was found to be 0.5117. VaxiJen score greater than 0.4 specified it to be a probable anti-

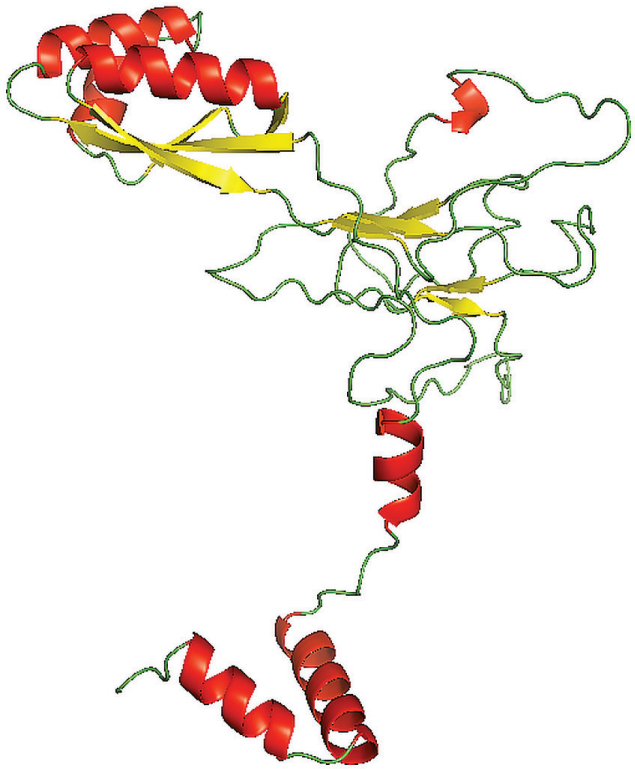

Figure 2. Ribbon diagram of the predicted tertiary structure of the multi-epitope construct.

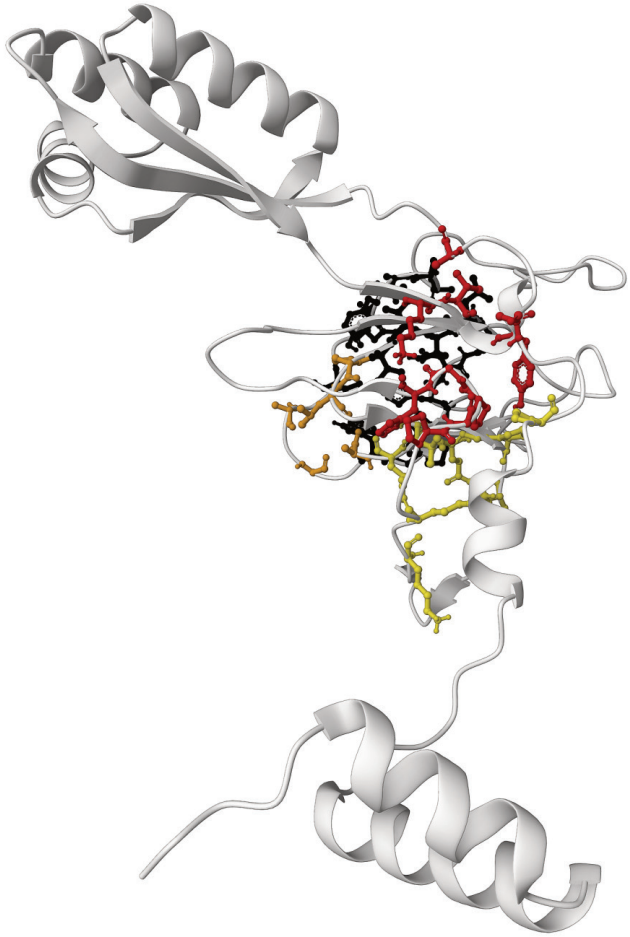

Figure 3. Binding pockets in the multi-epitope construct. Pockets 1 and 2 colored black and red are conserved.

gen. The multi-epitope construct was found to be non-allergenic and non-toxic. These observations indicated that this antigenic construct is a good candidate.

BLAST homology assessment

The sequence homology between the multi-epitope construct and the proteome sequence of *Homo sapiens* revealed no significant similarity. The outcome of BLAST homology assessment pointed out that, the multi-epitope construct is not likely to cause autoimmune responses in humans.

Physicochemical analysis of the multi-epitope construct

Table 1 represented the physicochemical properties of the multi-epitope construct. The molecular weight of the construct was 26128.81 Da. The theoretical pI (isoelectric point) was 5.28. It indicated the acidic nature of the construct. The instability index of 20.31 confirmed it to be a stable protein. The total number of negative and positively charged residues incident in the multi-epitope construct vaccine were 34 and 27, respectively. The aliphatic index was predicted to be 91.48, while the Grand average of hydropathicity (GRAVY) was -0.115. The former guaranteed its thermo-stability, while the latter revealed hydrophilic properties of the construct, and its capability to interact with water molecules. The solubility of the multi-epitope construct was estimated to be 99.9% when overexpressed in *E. coli*.

Secondary and tertiary structure prediction

The secondary structure of the multi-epitope construct revealed 103, 109, and 38 amino acid residues in the formation of alpha-helix, coil, and beta-strand, respectively. Hence, the construct consisted of 41.2%, 43.6%, and 15.2% residues as alpha-helix, coil, and beta-strand. 6.8% of the residues were disordered.

The tertiary structure (Fig. 2) of the multi-epitope construct revealed satisfactory model quality. ProSa-web gave a Z score of -3.61. This score was in the range of proteins of comparable size. The predicted LGscore from ProQ was 10.173 for the model. The ERRAT server revealed 87.70% overall quality factor. Additionally, the Ramachandran plot analysis showed 92.9%, 4.3% and 2.9% residues in the favoured, allowed, and generously

Table 3. Predicted conformational epitopes from the multi-epitope construct

No.	Residues	Score	3D structure
1	A:M1, A:A2, A:K3	0.979	Fig. 4a
2	A:L4, A:S5, A:T6, A:D7, A:E8, A:L9, A:L10, A:D11, A:A12, A:F13, A:E15, A:M16, A:T17, A:L18, A:L19, A:E20, A:L21, A:S22, A:D23, A:F24, A:V25, A:K26, A:K27, A:F28, A:E29, A:V34, A:T35, A:A36, A:A37, A:A38	0.861	Fig. 4b
3	A:E30, A:T31, A:F32	0.851	Fig. 4c
4	A:P142, A:E143, A:Y144, A:A145, A:A146, A:Y147	0.729	Fig. 4d
5	A:L67, A:E68, A:A69, A:A70, A:G71, A:D72, A:K73, A:K74, A:I75, A:G76, A:V77, A:I78, A:K79, A:V80, A:V81, A:R82, A:E83, A:I84, A:V85, A:S86, A:G87, A:L88, A:G89, A:L90, A:K91, A:E92, A:A93, A:K94, A:D95, A:L96, A:V97, A:D98, A:G99, A:A100, A:P101, A:K102, A:P103, A:E115, A:A116, A:K117, A:A118, A:K119, A:L120, A:E121, A:A122, A:A123, A:G124, A:A125, A:T126, A:V127	0.681	Fig. 4e
6	A:L194, A:A195, A:Y196, A:I197, A:G198, A:G199, A:G200, A:N201, A:A202, A:N203, A:V204	0.629	Fig. 4f
7	A:K148, A:P149, A:T150, A:H151, A:P152, A:E153, A:T154, A:S155, A:A156, A:A157, A:A158	0.611	Fig. 4g
8	A:G171, A:P172, A:G173, A:H174, A:N175, A:N176, A:I177, A:S178, A:R231, A:L232, A:K233, A:Q234, A:I235, A:T236, A:I237, A:G238, A:D239, A:S240, A:T241, A:I242, A:T243, A:T244, A:Q245	0.564	Fig. 4h

allowed/disallowed regions, respectively. All these portrayed that the predicted 3D structure was acceptable, and most of the evaluation tests indicated it to be a good quality designed model.

Assessment of ligand binding pockets

The outcome of Prankweb analysis revealed four ligand binding pockets (Fig. 3). Out of these pockets 1 and 2 were found to be conserved. Table 2 exhibited the pockets ranked by score. The surface volumes of the binding pockets were assessed in the centre x-axis, y-axis, and z-axis. Numerous residues associated with the pocket formation may be future targets of suitable ligands for docking simulation studies. The outcome of binding pockets pointed out that the amino acid residues might assist in designing inhibitors.

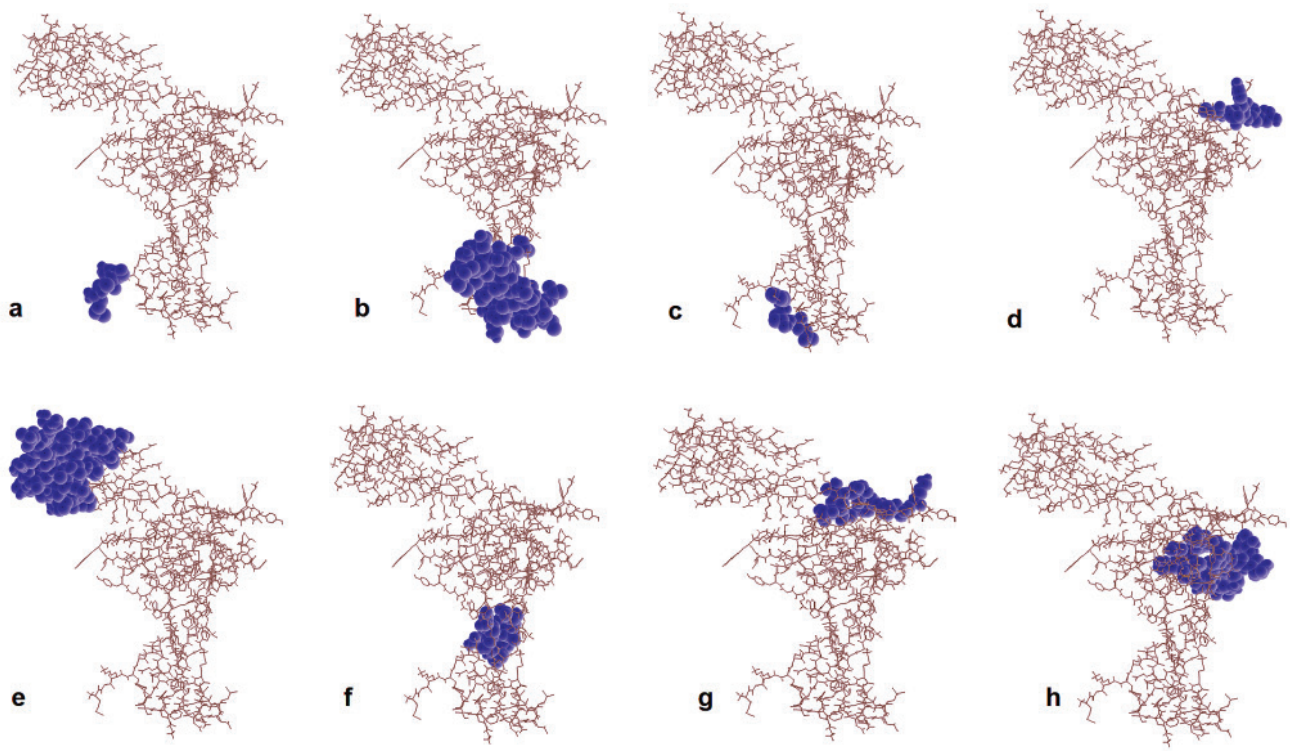

Figure 4. Predicted eight (a-h) conformational B-cell epitopes in the multi-epitope construct. The epitopes are represented by blue spheres.

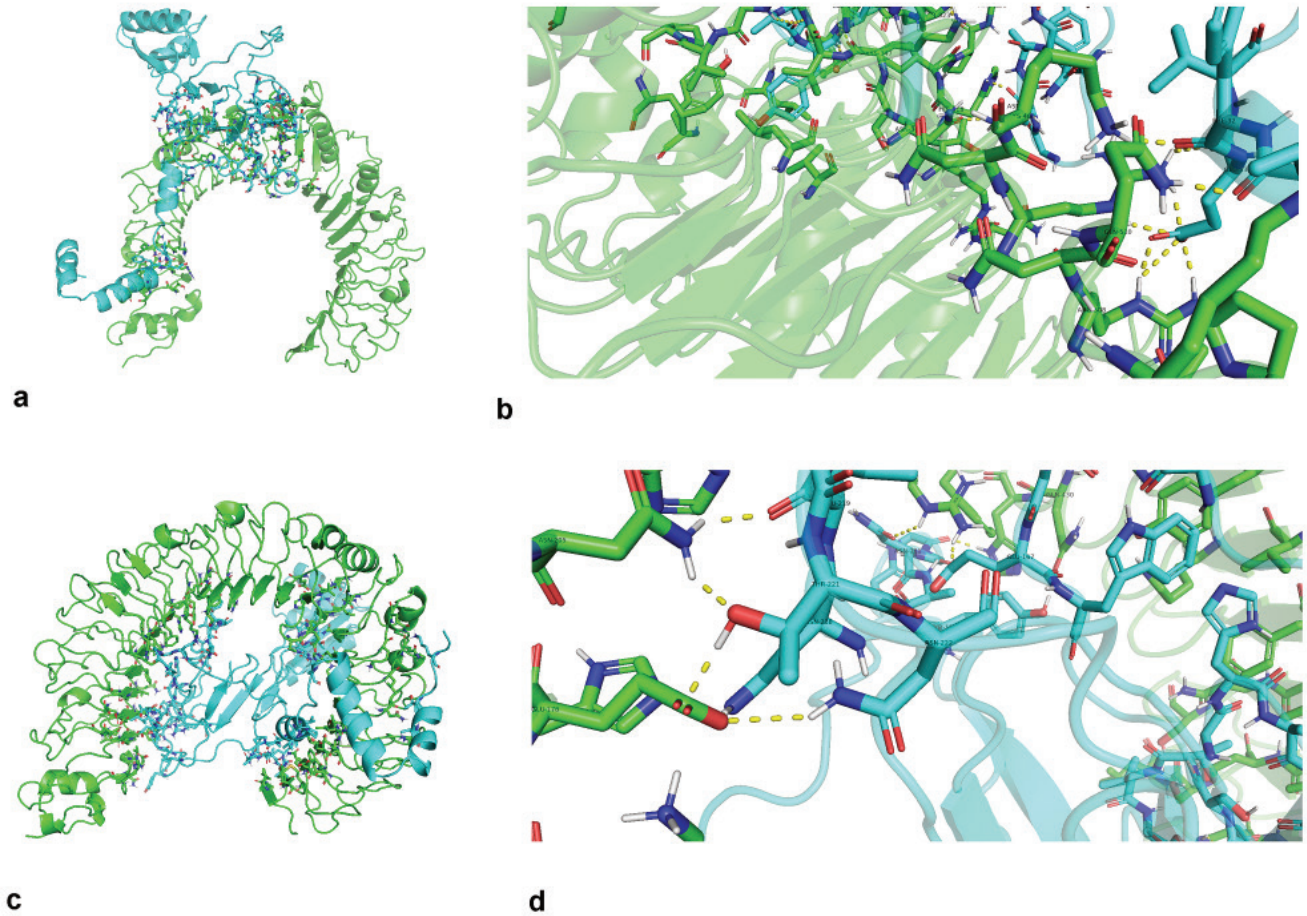

Figure 5. Molecular docking of the multi-epitope construct with (a-b) human TLR2 and (c-d) human TLR4. TLR2 and TLR4 are colored green while the multi-epitope construct is colored sky blue.

Investigation of conformational or discontinuous B-cell epitopes

Conformational epitopes are known to account for more than 90% of B-cell epitopes (Ponomarenko et al. 2008; Adhikari et al. 2018). ElliPro revealed 8 discontinuous or conformational B-cell epitopes (Fig. 4a-4h) in the structure. Table 3 showed that the protrusion index (PI) scores ranged between 0.564-0.979. Epitopes with PI score above 0.8 tend to have greater acceptability since high scores imply enhanced capacity for solvent accessibility. Taking a cue from this the second epitope with 30 residues and score of 0.861, was regarded as the most appropriate discontinuous B-cell epitope. All the identified discontinuous or conformational epitopes on the surface of the construct denoted accessibility for the virus.

Molecular docking of the multi-epitope construct with immunological receptors

The ClusPro protein-protein docking server, performed the molecular docking between the refined 3D structure of the multi-epitope construct, and immune receptors like TLR2 and TLR4, respectively. The docking program generated 30 clusters for each docked complex. Cluster 8 of TLR2-multi-epitope construct docked complex (Fig. 5a-5b) comprising of 32 members had the lowest energy of -1168.1. On the other hand, Cluster 2 of TLR4 - multi-epitope construct docked complex (Fig. 5c-5d) with 42 members had the lowest energy of -1376.4. Docking simulation revealed satisfactory interactions and several binding residues.

Normal mode analysis

Normal mode analysis of the docked complexes illustrated the stability and movement of the atoms (Supplementary Fig 1- Fig. 2). Supplementary Fig. 1b and 2b, demonstrated regions of deformability in the protein delineated by peaks in the graph. The comparison between NMA and PDB of the TLR2-construct and TLR4-construct complexes is indicated by the B factor graphs (Supplementary Fig. 1c and 2c). Eigenvalues of the TLR2-construct and

TLR4-construct complexes (Supplementary Fig. 1d and 2d) were 1.378714e-05 and 2.793110e-05 respectively. Supplementary Fig. 1e and 2e portrayed individual and cumulative variances. The covariance maps (Supplementary Fig. 1f and 2f) revealed the correlated, uncorrelated and anticorrelated motions between pairs of residues. The elastic maps of the TLR2-construct and TLR4-construct complexes (Supplementary Fig. 1g and 2g) divulged the links between atoms. The dark grey colorations indicated the stiffer residues within the complexes.

Outcome of codon optimization and in silico cloning

Codon optimization is an important step in validating the multi-epitope construct, for screening its immunoreactivity (Majid and Andleeb 2019). The codon adaptation index (CAI) was found to be 1.0, and the GC content was 50.66% for the sequence. Both the CAI and GC values were within the optimal range and signified stable expression of the designed multi-epitope construct and a greater possibility of expression in bacteria. The restriction sites of HincII and EcoR1 were incorporated and ligated into *E. coli* ELF2-pet28a vector for optimal expression by in silico cloning. The cloned construct map is shown in Fig. 6.

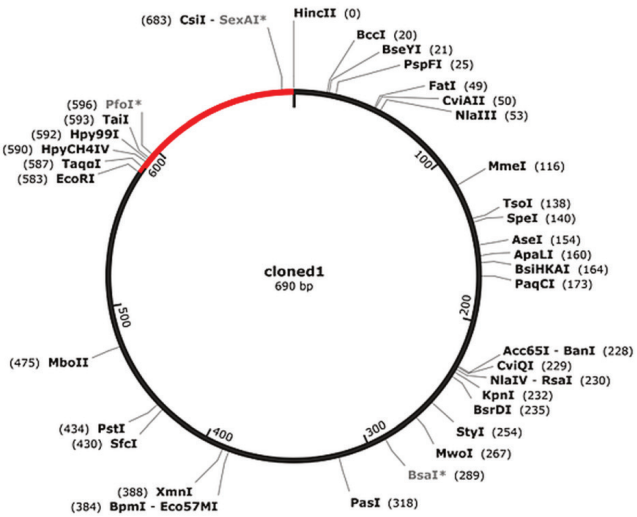

Figure 6. *In silico* restriction cloning of the multi-epitope construct into *E. coli* ELF2-pet28a vector. Red and black color denotes the gene sequence of the construct and backbone of the vector.

Immune simulation

Immune simulation tool C-ImmSim was utilized to determine immune responses fashioned by the multi-epitope construct (Majid and Andleeb 2019). It was observed that there was an increase in the secondary and tertiary

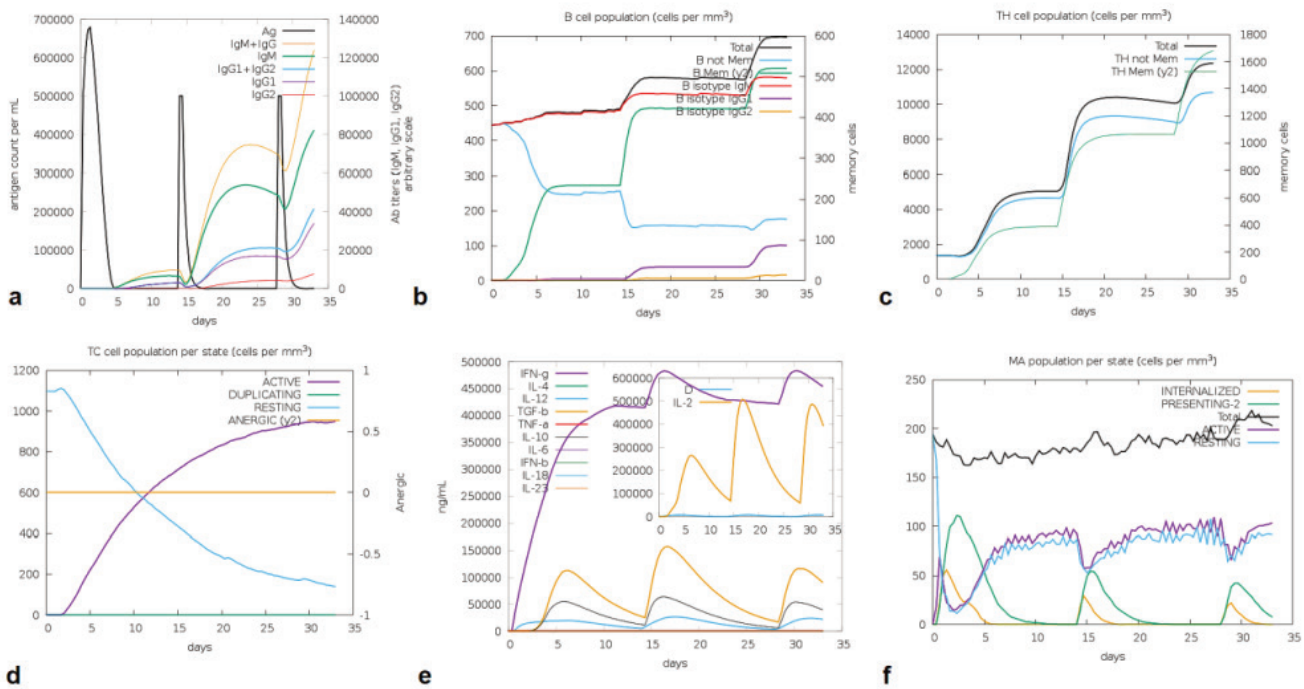

Figure 7. *In silico* immune simulation of the multi-epitope construct. a) Immunoglobulin production triggered by antigen injections (black vertical lines); b) Changes linked to B-cell populations; c) Changes in T-helper cell populations; d) Changes associated with T-cytotoxic cell populations; e) Changes in the concentration of cytokines and interleukins; f) Macrophage population per state.

immune response (Fig. 7a). B cell populations, IgG + IgM and IgG1 + IgG2 amplified with the decrease of antigen concentration. Fig. 7b portrayed that constant exposure resulted in memory formation (Majid and Andleeb 2019) in the immune system. Additionally, B cells were found to demonstrate isotype-switching capability. Moreover, cytotoxic and T helper cell populations showed a response (Fig. 7c-7d) on strengthening of memory. High macrophage activity and elevated levels of interferon-gamma were also noticed (Fig. 7e-7f). Immune simulation of the multi-epitope construct revealed that the outcome was in line with immune responses. Immune response amplified with constant exposure to the antigen. An increase in T helper cells (Majid and Andleeb 2019) supported humoral response. In a nutshell, the outcome illustrated that each injection elicited a stronger immune response, a decrease in antigens, and elevated memory.

Discussion

Acute airway infections are a major cause of morbidity and mortality worldwide. The recent pandemic due to novel coronavirus outbreak has shown that vaccine development is very crucial for saving lives. The capsid proteins of HRV-C arranged in an icosahedral structure have been the target of many antiviral studies (Stone and Miller 2016). However, the development of effective therapeutics has been time-consuming and costly. Use of bioinformatics-based approaches may assist in developing different strategies for vaccine design (Seib et al. 2012). With newer technologies in computational immunology, and advent of high quality immunoinformatics tools designing epitope-based constructs in a short time can greatly assist physicians in the long run (Korber et al. 2006; Purcell et al. 2006). Multi-epitope-based constructs are known to have improved potency (Majid and Andleeb 2019).

The focus of this research has been in silico design of multi-epitope construct, that can offer notable defense against HRV-C infection. This construct consisted of CTL, HTL, and linear B-cell epitopes having the ability to initiate humoral and cell-mediated immune response, upon injection within the host. The most antigenic protein amongst the studied HRV-C strains was selected, and CTL, HTL, and linear B-cell epitopes were predicted. The suitable CTL, HTL, and linear B-cell epitopes were screened to design the multi-epitope construct. They were joined by linkers. An adjuvant and an EPEA tag were added at the N terminal and C terminal ends to enhance immune response and purification, respectively. The multi-epitope construct was found to be antigenic, non-allergenic, and non-toxic. Interestingly, the identification of IFN-γ inducing epitopes within the multi-epitope construct, signified that this construct had the capability of IFN-γ induced immune-stimulatory actions.

The average molecular weight of the multi-epitope construct divulged its antigenic nature. That the construct was thermostable at normal body temperature was indicated by its aliphatic index. A lower instability index is associated with the construct protein's stability in biological environments (Sami et al. 2021). Hydrophilicity coupled with effective solubility of the construct upon overexpression in *E. coli,* specified suitability of formulation and purification of the protein (Sami et al. 2021). The greater degree of random coils in the secondary structure of the construct stated its conserved nature and stability. Multiple random coils also designated strong antigenic potential. The 3D structure of the construct exhibited stability and it housed some ligand binding pockets. The 3D structure of this construct was docked with TLR2 and TLR4 receptors. These pattern recognition receptors are well known to elicit innate immune responses (Hamann et al. 2013). Docking studies demonstrated negative binding energy between the construct and TLR2 and TLR4 chains. Thus, docking interactions signified stable and high binding affinity as well as the potential to mount immune response. Furthermore, the outcome of normal-mode analysis portrayed that the docked complexes had strong positive eigenvalues, indicating rigidity and lower chances of deformability (Sami et al. 2021).

In silico cloning expressed the multi-epitope construct in *E. coli*. Codon optimization was executed, and it was found that both the values of codon adaptation index and GC content were optimal indicating maximal expression in bacteria. The outcome of immune simulation studies pointed out that the multi-epitope construct had the potential to generate a strong immune response. The triggering of primary and secondary immune responses coupled with increase in memory B cells, cytotoxic T cells, and helper T cells illustrated a surge defense against the virus.

Conclusions

Human rhinovirus-C related infections are a worldwide problem yet, there has been no permanent cure. Designing appropriate multi-epitope construct augurs well for developing therapy against HRV-C infections. In this study, a series of immunoinformatics tools have been used, to design a multi-epitope construct from the structural proteins of HRV-C. This construct was designed, using the potential CTL, HTL and LBL epitopes after proper screening. An adjuvant was added, and the epitopes were attached by linkers. The construct was antigenic, non-allergenic, and non-toxic, demonstrating a satisfactory

physicochemical and structural profile. Moreover, molecular docking analysis revealed significant binding affinity of the multi-epitope construct to human TLR2 and TLR4. Additionally, the outcome of normal mode analysis, codon optimization, and immune simulation studies demonstrated the capability to elicit effective responses. In a nutshell, this in silico study provides a testable hypothesis that this multi-epitope construct could stimulate an immune response by inducing humoral and cell-mediated immunity and could be a right step forward. Nevertheless, large scale *in vitro* and *in vivo* studies would be required before practical application of these epitopes.

Acknowledgments

SS and MG acknowledges the support of Ramananda College, India for providing infrastructure to carry out the work. The authors thank anonymous reviewers for constructive comments.

References

Abdellrazeq GS, Fry LM, Elnaggar MM, Bannantine JP, Schneider DA, Chamberlin WM, Mahmoud AHA, Park KT, Hulubei V, Davis WC (2020) Simultaneous cognate epitope recognition by bovine CD4 and CD8 T cells is essential for primary expansion of antigen-specific cytotoxic T-cells following ex vivo stimulation with a candidate *Mycobacterium avium* subsp. *paratuberculosis* peptide vaccine. Vaccine 38:2016–2025.

Adam KM (2021) Immunoinformatics approach for multi-epitope vaccine design against structural proteins and ORF1a polyprotein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). Trop Dis Travel Med Vaccines 7(1):22.

Adhikari UK, Tayebi M, Rahman MM (2018) Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. J Immunol Res 2018:6718083.

Arden KE, Mackay IM (2009) Human rhinoviruses: coming in from the cold. Genome Med 1:44.

Arruda E, Boyle TR, Winther B, Pevear DC, Gwaltney JM Jr, Hayden FG (1995) Localization of human rhinovirus replication in the upper respiratory tract by in situ hybridization. J Infect Dis171:1329-1333.

Bizzintino J, Lee WM, Laing IA et al (2011) Association between human rhinovirus C and severity of acute asthma in children. Eur Respir J 37:1037-1042.

Bochkov YA, Gern JE (2012) Clinical and molecular features of human rhinovirus C. Microbes Infect 14:485-194.

Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, Kesmir C, Peters B (2013) Properties of MHC class I presented peptides that enhance immunogenicity. PloS Comput Biol 9:e1003266.

Cordey S, Junier T, Gerlach D, Gobbini F, Farinelli L, Zdobnov EM, Winther B, Tapparel C, Kaiser L (2010) Rhinovirus genome evolution during experimental human infection. PloS One 5:e10588.

Dhanda SK, Vir P, Raghava GP (2013) Designing of interferon-gamma inducing MHC class-II binders. Biol Direct 8:30.

Diaz AA, Tomba E, Lennarson R, Richard R, Bagajewicz MJ, Harrison RG (2010) Prediction of protein solubility in *Escherichia coli* using logistic regression. Biotechnol Bioeng 105:374-383.

Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, Anishchenko I, Baker D, Yang J (2021) The trRosetta server for fast and accurate protein structure prediction. Nat Protoc 16:5634-5651.

Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform 8:4.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In Walker JM, Ed., The Proteomics Protocols Handbook. Humana Press, Totowa, 571–607.

Glanville N, Johnston SL (2015) Challenges in developing a cross-serotype rhinovirus vaccine. Curr Opin Virol 11:83-88.

Hamann L, Koch A, Sur S, Hoefer N, Glaeser C, Schulz S, Gross M, Franke A, Nöthlings U, Zacharowski K, Schumann RR (2013) Association of a common TLR-6 polymorphism with coronary artery disease - implications for healthy ageing? Immun Ageing 10:43.

Hao W, Bernard K, Patel N, Ulbrandt N, Feng H, Svabek C, Wilson S, Stracener C, Wang K, Suzich J, Blair W, Zhu Q (2012) Infection and propagation of human rhinovirus C in human airway epithelial cells. J Virol 86:13524-13532.

Heo L, Park H, Seok C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res 41:W384-W388.

Jacobs SE, Lamson DM, St George K, Walsh TJ (2013) Human rhinoviruses. Clin Microbiol Rev 26:135-162.

Jakiela B, Brockman-Schneider R, Amineva S, Lee WM, Gern JE (2008) Basal cells of differentiated bronchial epithelium are more susceptible to rhinovirus infection. Am J Respir Cell Mol Biol 38:517-523.

Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D (2019) PrankWeb: a web server for ligand binding site prediction and visualization. Nucleic Acids Res 47(W1):W345-W349.

Jin J, Hjerrild KA, Silk SE, Brown RE, Labbé GM, Marshall JM, Wright KE, Bezemer S, Clemmensen SB, Biswas S, Li Y, El-Turabi A, Douglas AD, Hermans P, Detmers FJ,

de Jongh WA, Higgins MK, Ashfield R, Draper SJ (2017) Accelerating the clinical development of protein-based vaccines for malaria by efficient purification using a four amino acid C-terminal 'C-tag'. Int J Parasitol 47:435-446.

Jones DT, Cozzetto D (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics 31:857-863.

López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P (2014) iMODS: internal coordinates normal mode analysis server. Nucleic Acids Res 42 (Web Server issue): W271-276.

Khatoon N, Pandey RK, Prajapati VK (2017) Exploring *Leishmania* secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. Sci Rep 7:8285.

Korber B, LaBute M, Yusim K (2006) Immunoinformatics comes of age. PloS Comput Biol 2(6):e71.

Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S (2017) The ClusPro web server for protein-protein docking. Nat Protoc 12:255-278.

Larsen MV, Lundegaard C, Lamberth K, Buus A, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinform 8:424.

Lau SK, Yip CC, Woo PC, Yuen KY (2010) Human rhinovirus C: a newly discovered human rhinovirus species. Emerg Health Threats J 3:e2.

Lee S, Nguyen MT, Currier MG, Jenkins JB, Strobert EA, Kajon AE, Madan-Lala R, Bochkov YA, Gern JE, Roy K, Lu X, Erdman DD, Spearman P, Moore ML (2016) A polyvalent inactivated rhinovirus vaccine is broadly immunogenic in rhesus macaques. Nat Commun 7:12838.

Majid M, Andleeb S (2019) Designing a multi-epitopic vaccine against the enterotoxigenic *Bacteroides fragilis* based on immunoinformatics approach. Sci Rep 9:19780.

Mak RK, Tse LY, Lam WY, Wong GWK, Chan PKS, Leung TF (2011) Clinical spectrum of human rhinovirus infections in hospitalized Hong Kong children. Pediatr Infect Dis J 30:749-753.

McLean GR (2014) Developing a vaccine for human rhinoviruses. J Vaccines Immun 2:16-20.

McGuffin LJ (2000) The PSIPRED protein structure prediction server. Bioinformatics 16:404-405.

Mittal A, Sasidharan S, Raj S, Balaji SN, Saudagar P (2020) Exploring the Zika genome to design a potential multi-epitope vaccine using an immunoinformatics approach. Int J Pept Res Ther 26:2231-2240.

Nair DT, Singh K, Siddiqui Z, Nayak BP, Rao KV, Salunke DM (2002) Epitope recognition by diverse antibodies suggests conformational convergence in an antibody response. J Immun 168:2371-2382.

Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. BMC Bioinform 8:238.

Nielsen M, Lund O (2009). NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinform 10:296.

Olejnik J, Hume AJ, Mühlberger E (2018) Toll-like receptor 4 in acute viral infection: too much of a good thing. PLoS Pathog 14:e1007390.

Palmenberg AC, Rathe JA, Liggett SB (2010) Analysis of the complete genome sequences of human rhinovirus. J Allergy Clin Immunol 125:1190-1201.

Papi A, Contoli M (2011) Rhinovirus vaccination: the case against. Eur Respir J 37:5-7.

Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger SP, Stewart S, Surko P, Way S, Wilson S, Sette A (2005) The design and implementation of the immune epitope database and analysis resource. Immunogenetics 57:326-336.

Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinform 9:514.

Purcell AW, McCluskey J, Rossjohn J (2006) More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov 6:404-414.

Rapin N, Lund O, Bernaschi M, Castiglione F (2010) Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PLoS One 5:e9862.

Ras-Carmona A, Pelaez-Prestel HF, Lafuente EM, Reche PA (2021) BCEPS: A Web server to predict linear B cell epitopes with enhanced immunogenicity and cross-reactivity. Cells 10:2744.

Sami SA, Marma KKS, Mahmud S, Khan MAN, Albogami S, El-Shehawi AM, Rakib A, Chakraborty A, Mohiuddin M, Dhama K, Uddin MMN, Hossain MK, Tallei TE, Emran TB (2021) Designing of a multi-epitope vaccine against the structural proteins of Marburg virus exploiting the immunoinformatics approach. ACS Omega 6:32043-32071.

Scully EJ, Basnet S, Wrangham RW, Muller MN, Otali E, Hyeroba D, Grindle KA, Pappas TE, Thompson ME, Machanda Z, Watters KE, Palmenberg AC, Gern JE, Goldberg TL (2018) Lethal Respiratory Disease Associated with Human Rhinovirus C in Wild Chimpanzees, Uganda, 2013. Emerg Infect Dis 24:267-274.

Seib KL, Zhao X, Rappuoli R (2012) Developing vaccines in the era of genomics: a decade of reverse vaccinology. Clin Microbiol Infect 18:109-116.

Stepanova E, Isakova-Sivak I, Rudenko L (2019) Overview of human rhinovirus immunogenic epitopes for rational vaccine design. Expert Rev Vaccines 18:877-880.

Stone CA, Miller EK (2016) Understanding the association of human rhinovirus with asthma. Clin Vaccine Immunol 23:6-10.

Wallner B, Elofsson A (2003) Can correct protein models be identified? Protein Sci 12:1073-1086.

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410.

Wei L, Ye X, Sakurai T, Mu Z, Wei L (2022) ToxIBTL: prediction of peptide toxicity based on information bottleneck and transfer learning. Bioinformatics 38:1514-1524.