JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Original Article

An immunoinformatic investigation on Rift Valley fever virus protein reveals possible epitopes for vaccines

Tanjir Hosen[1], Saaimatul Huq[2], Mohammad Abdullah-Al-Shoeb[1], Shahidul Islam[3], Muhammad Abul Kalam Azad[1,3]

[1] Department of Biochemistry and Molecular Biology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh
[2] Department of Radiation Oncology, Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR 72205, United States
[3] Department of Agriculture/Agricultural Regulations, 1200 N. University Drive, University of Arkansas at Pine Bluff, AR 71601, United States

Abstract

Introduction: This immunoinformatic study identified potential epitopes from the envelopment polyprotein (Gn/Gc) of Rift Valley fever virus (RVFV), a pathogenic virus causing severe fever in humans and livestock. Effective vaccination is crucial for controlling RVFV outbreaks. The identification of suitable epitopes is crucial for the development of safe and effective vaccines.
Methodology: Protein sequences were obtained from the UniProt database, and evaluated through VaxiJen v2.0 to predict the B and T-cell epitopes within the RVFV glycoprotein. Gn/Gc protein sequences were analyzed with bioinformatics tools and algorithms. The predicted T-cell and B-cell epitopes were evaluated for antigenicity, allergenicity, and toxicity by the VaxiJen v2.0 system, AllerTop v2.0, and ToxinPred server, respectively.
Results: We employed computational methods to screen the RVFV envelopment polyprotein encompassing N-terminal and C-terminal glycoprotein segments, to discover antigenic T- and B-cell epitopes. Our analysis unveiled multiple potential epitopes within the RVFV glycoprotein, specifically within the Gn/Gc protein sequences. Subsequently, we selected eleven cytotoxic T-lymphocytes (CTL) and four helper T-lymphocytes (HTL) for population coverage analysis, which collectively extended to cover 97.04% of the world's population, representing diverse ethnicities and regions. Notably, the CTL epitope VQADLTLMF exhibited binding affinity to numerous human leukocyte antigen (HLA) alleles. The identification of glycoprotein (Gn/Gc) epitopes through this immunoinformatic study bears significant implications for advancing the development of an effective RVFV vaccine.
Conclusions: These findings provide valuable insights into the immunological aspects of the disease and may contribute towards the development of broad-spectrum antiviral therapies targeting other RNA viruses with similar polymerase enzymes.

Key words: immunoinformatic; RVFV; glycoprotein: Gn/Gc; B-cell epitopes; T-cell epitopes; vaccine.

J Infect Dev Ctries 2024; 18(7):1090-1099. doi:10.3855/jidc.19005

(Received 11 August 2023 – Accepted 27 November 2023)

Copyright © 2024 Hosen *et al*. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

The Rift Valley fever virus (RVFV) is responsible for causing Rift Valley fever (RVF) which is an emerging disease transmitted by mosquitoes, and poses a significant health risk to both humans, and animals (wild and domestic) in sub-Saharan Africa and the Arabian Peninsula [1]. This virus belongs to the Bunyaviridae family and the genus *Phlebovirus*. Its genome is divided into three parts, which have been given the descriptive labels "L," for long, "M," for medium, and "S" for short [2]. The M segment encodes enveloped glycoproteins (Gn, Gc) [2,3]. This virus affects humans, cattle, sheep, goats, and camels, with 10% of humans developing severe illnesses like hemorrhage, encephalitis, or eye disease; 2% experiencing vision loss; and another 1% dying [4,5]. The virus can lead to economic losses and hamper food security [6,7]. Therefore, it is crucial to take measures to prevent and control the spread of this virus, both for public health and economic reasons. It is also important to develop research-based strategies to mitigate its impact on livestock and food security. RVFV-IgG antibodies have been detected in Bangladeshi nationals employed in Saudi Arabian slaughterhouses, highlighting their biothreat potential in densely populated regions like East Asia, including Bangladesh [8]. The government has classified RVFV as a bioterrorism agent because there is such a high risk of its spreading to other countries [4].

The first human deaths from RVFV occurred in South Africa between 1974 and 1976 [9]. Subsequent outbreaks were documented in Saudi Arabia and Yemen (2000), Sudan (2007), South Africa (2010), Uganda (2016), Kenya (2018), and Mayotte (2018–2019), as reported by Nanyingi *et al.* and Gerken *et al*. [10,11]. Therefore, an effective vaccine against RVFV is essential for preventing and controlling the spread of the virus [12,13]. Efforts are underway to develop and test human RVFV vaccines. Still, more research is needed to ensure their safety and effectiveness [13]. A vaccine can help control outbreaks in endemic areas and reduce the risk of the virus's global spread. However, a licensed vaccine for human protection against the RVF virus is not yet available.

A safe and effective RVFV vaccine will protect human and animal populations from the disease, reducing the burden of illness and death [11,14]. It is essential for health organizations to implement a comprehensive strategy to mitigate the increasing risk of RVFV infection by enhancing surveillance and monitoring in animal populations, conducting public health education and awareness campaigns, and establishing early detection and response systems to promptly identify outbreaks [15,16].

Numerous vaccination technologies, such as DNA vaccines, inactivated whole virus vaccines, modified live vaccines, and classical vaccines, are employed to prevent and treat microbial infections, offering enduring protection [17]. Despite their effectiveness in controlling RVF in endemic African countries, the slow and limited production capabilities of these vaccines, coupled with associated risks, pose challenges for their application in non-endemic regions due to the lack of essential characteristics [18].

Some newer vaccine platforms, such as mRNA vaccines, have shown promise in producing highly effective vaccines against various infectious diseases [19,20]. The mRNA vaccines instruct the body's cells to produce viral proteins, stimulating the immune system to mount a protective response. These vaccines can be produced quickly and at scale, making them a promising tool for preventing and controlling infectious diseases [20]. Depending on in silico approaches and database information, a vaccine based on T helper cell-mediated immune responses can be developed to replace conventional vaccines which could offer several advantages over conventional vaccines, including greater efficacy, longer-lasting protection, and the ability to tailor to specific populations [21,22]. Multiepitope vaccines contain small peptides of immunogenic stimuli, triggering a robust immune response and reducing the host system's allergic response significantly [14]. Due to their validity, safety, and economic rationale, epitope-based vaccines leverage immunogenic proteins or epitopes to effectively induce an immune response and enhance potency [23]. We employed a computational approach to screen the envelopment polyprotein (EPP) sequences of RVFV, as it has the highly conserved glycoprotein Gn/Gc sequence. This glycoprotein segment was selected as it has a high annotation score based on the protein sequence data and experimental results available in the UniProt database [24].

Our aim was to identify the most antigenic T- and B- cell epitopes. Immunogenic and antigenic cytotoxic T-lymphocytes (CTL), helper T-lymphocytes (HTL), and B-cell epitopes were initially identified and subsequently shortlisted for toxicity and allergenicity analysis, along with their corresponding major histocompatibility (MHC) alleles. We identified eleven CTL epitopes and four HTL epitopes, through a conservancy analysis, that displayed promising immunogenic and antigenic characteristics. These epitopes have the potential to elicit a robust immune response, making them promising candidates for the development of an effective RVFV vaccine. The cross-reactivity of these selected epitopes with human proteomes and human MHC class alleles revealed that the CTL epitope VQADLTLMF demonstrated binding capability to numerous human leukocyte antigen (HLA) alleles. As a result, we gained valuable insights into population coverage analysis, which can facilitate the development of a more effective and broadly protective RVFV vaccine. In a nutshell, our study successfully identified the candidate epitopes as not only effective, but also safe and stable, thus making it a promising choice for further vaccine development.

Methodology

Acquisition of protein sequences

The protein sequences were sourced from the UniProt database [24], a reliable resource offering meticulously curated and non-redundant protein sequences, experimental findings, computed features, and scientific insights. Epp (Uniprot entry: P03518) are RVFV proteins retrieved in FASTA format [25]. Figure 1 presents the overall study design for the identification of RVFV epitopes.

Physicochemical properties analysis and antigenicity prediction

The antigenic tendency of the retrieved protein sequences was assessed using VaxiJen v2.0

(http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html/), with a threshold set at 0.4 [25]. Various physical and chemical parameters, including target protein molecular weight, extinction coefficient, instability index, theoretical isoelectric point (pI), anticipated half-life, aliphatic index, and grand average of hydropathicity (GRAVY), were analyzed using the widely-used web-based tool ExPASy ProtParam (https://web.expasy.org/protparam/) [26]. Additionally, the DeepLoc-1.0 (https://services.healthtech.dtu.dk/service.php?DeepLoc-1.0) online prediction tool was utilized to predict the subcellular localization of the targeted proteins [27].

T cell epitopes prediction

The chosen protein sequences underwent analysis using the NetCTL 1.2 server (http://www.cbs.dtu.dk/services/NetCTL/) [28] with default settings: 0.75 threshold for epitope identification, 0.15 weight on C-terminal cleavage, and 0.05 weight on transporter associated with antigen processing (TAP) transport efficiency. This allowed us to predict CTL epitopes for 12 distinct MHC class-I supertypes (A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, and B62) through the integration of MHC class-I binding, proteasomal C-terminal cleavage, and TAP protein transport efficiency predictions [28]. The Epitope Database 2.22 (IEDB, http://tools.iedb.org/) method was employed for identifying the 15 amino acids (aa) long helper T-lymphocytes (HTL) epitopes within the selected proteins [29]. A total of 27 human alleles were utilized, ensuring coverage of 97.04% of the world's population, encompassing various ethnicities and regions. Epitopes with half-maximal inhibitory concentration (IC50) values below 50 nM were categorized as high affinity, below 500 nM as intermediate affinity, and below 5000 nM as low affinity. This study focused on epitopes with IC50 values less than 50 nM for further investigation.

B-cell epitopes prediction

The ABCpred server (https://webs.iiitd.edu.in/raghava/abcpred/ABC_submission.html), with a default threshold of 0.51, was utilized to identify the 16-mer linear B-cell epitopes within the targeted proteins [30]. Various factors, such as hydrophilicity, flexibility, antigenic propensity, surface accessibility, and the Parker hydrophilicity prediction technique, were considered to determine the linear B-cell epitopes [31]. The Kolaskar and Tongaonkar antigenicity scale, Karplus and Schulz flexibility prediction tool [32], Emini surface accessibility prediction method [33], and Chou and Fasman beta-turn prediction algorithm were employed in this assessment [34].

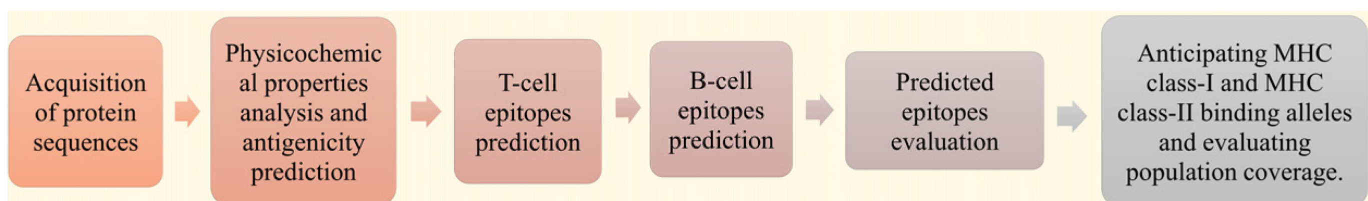
Predicted epitopes evaluation

The most favored T-cell and B-cell epitopes were selected based on their allergenicity, antigenicity, and toxicity profiles. The VaxiJen v2.0 system was employed to assess the epitopes' antigenicity. AllerTop v2.0 (https://www.ddg-pharmfac.net/AllerTOP/) and ToxinPred server (https://webs.iiitd.edu.in/raghava/toxinpred/design.php/) were utilized to evaluate allergenicity and toxicity, respectively [35]. To analyze the induction of cytokines such as interferon-γ (IFN-γ), interleukin-4 (IL-4), and interleukin-10 (IL-10) by the predicted epitopes, IFNepitope (https://webs.iiitd.edu.in/raghava/ifnepitope/developer.php) a server for predicting and designing IFN-gamma inducing epitopes, IL4pred (https://webs.iiitd.edu.in/raghava/il4pred/design.php/), and IL10pred (https://webs.iiitd.edu.in/raghava/il10pred/predict3.php/) web servers were utilized [36]. The conservancy analysis of the chosen CTL and HTL epitopes was performed using the IEDB Epitope Conservancy Analysis server (http://tools.iedb.org/conservancy/) to determine the epitope conservancy level [29].

Anticipating MHC class-I and MHC class-II binding alleles and evaluating population coverage

For each CTL and HTL epitope, the MHC class-I and MHC class-II binding alleles were predicted using the IEDB recommended method and the consensus prediction method (percentile rank ≤ 5) [36]. The

Figure 1. Illustration of the overall study design used to identify River Valley fever virus (RVFV) epitopes.



population coverage tool from IEDB (population coverage (www.iedb.org)) was employed to determine the percentage of population coverage for the selected CTL and HTL epitopes, along with their respective MHC binding alleles, both individually and in combination [29].

Results

Selection of antigenic proteins and their physicochemical characterization

A conserved protein sequence of RVFV envelopment polyprotein (glycoprotein: Gn/Gc) was acquired from the Uniprot database through the accession number P03518.1 in FASTA format. The retrieved envelopment polyprotein was subjected to analysis of various physical and chemical parameters using the online tool ExPASy-ProtParam. The protein was estimated to have an in vitro half-life of 30 hours in mammalian reticulocytes. The GRAVY value indicated a negative hydrophobicity, signifying its hydrophilic nature (Table 1).

Table 1. The physicochemical properties of envelopment polyprotein (glycoprotein: Gn/Gc).

Protein name	Envelopment polyprotein RVFV (Glycoprotein: Gn/Gc)
UniProt entry	P03518
Subcellular localization and score	Plasma membrane and 2.496
Number of amino acids	1206
Antigenicity score	0.5145
Molecular weight (Daltons)	132053.50
Theoretical Pi	7.74
Chemical formula	$C_{5802}H_{9174}N_{1600}O_{1747}S_{86}$
Estimated half-life (mammalian reticulocytes, in vitro)	30 h
Instability index	46.02
Aliphatic index	81.88
Grand average of hydropathicity (GRAVY)	-0.106

RVFV: River Valley fever virus.

Forecasting and selecting T lymphocyte epitopes from target proteins

CTL play a crucial role in the cellular immune response, as they have the ability to recognize immunogenic antigens presented on virus-infected cell surfaces. This recognition involves the binding of the antigen-specific T-cell receptor to CTL epitopes and MHC class-I molecules, thereby forming a complex on the surfaces of virus-infected cells. Consequently, using the NetCTL 1.2 server with the default threshold score, it was possible to predict 9-mers CTL epitopes for each of the 12 different MHC class-I supertypes. A total of 269 Epp potential CTL epitopes were selected. Based on the antigenicity, allergenicity, and toxicity, 90 epitopes were selected and used for further analysis (Supplementary Table 1).

The HTL plays a crucial role in the adaptive immune response, facilitating B-cell stimulation for antibody production and engaging in other T-cell activation processes. In our study, 44 HTL epitopes from the envelopment polyprotein were selected, each comprising 15 amino acids, and predicted using the NN-align method against the complete set of 27 human alleles through the IEDB server. Following epitope selection, we conducted an analysis of antigenicity, allergenicity, and toxicity. Among the indicated epitopes, only ten were found to be antigenic, non-allergenic (except for two epitopes), and non-toxic (Table 2). These selected epitopes of envelopment polyprotein (Epp) were predicted by NetCTL 1.2 server. Further investigation revealed that out of these ten epitopes, only four demonstrated cytokine-inducing ability (Figure 2), inducing IL-4 and IL-10.

B-cell epitope prediction from target proteins

Initially, a total of 118 potential linear B-cell epitopes were detected using the ABCpred server, with a default threshold of 0.51. Subsequently, we identified 71 epitopes as antigenic, non-allergenic, and non-toxic (Supplementary Table 2).

Evaluation of epitope conservancy and selection of top-ranked epitope

Epitope conservancy across different strains plays a crucial role in determining vaccine efficacy and

Table 2. HTL epitopes predicted by MHC class-II.

Epitopes	Start	End	Allergenicity	Antigenicity	Toxicity
PLKLYSSFACMLHYQ	1156	1170	NA	< 0.4	Non-Toxin
MLHYQLGSFSSLYIL	1166	1180	NA	< 0.4	Non-Toxin
CMLHYQLGSFSSLYI	1165	1179	NA	< 0.4	Non-Toxin
FVVVFVFSSIAIICL	586	600	NA	< 0.4	Non-Toxin
VVVFVFSSIAIICLA	587	601	NA	< 0.4	Non-Toxin
YRVLKCLKIAPRKVL	604	618	NA	< 0.4	Non-Toxin
RSTGFKISSAVACAS	494	508	NA	< 0.4	Non-Toxin
VVFVFSSIAIICLAV	588	602	NA	< 0.4	Non-Toxin
VQADLTLMFDNFEVD	1007	1021	A	< 0.4	Non-Toxin
LIVSYASACSELIQA	683	697	A	< 0.4	Non-Toxin

HTL: helper T-lymphocytes; MHC: major histocompatibility.

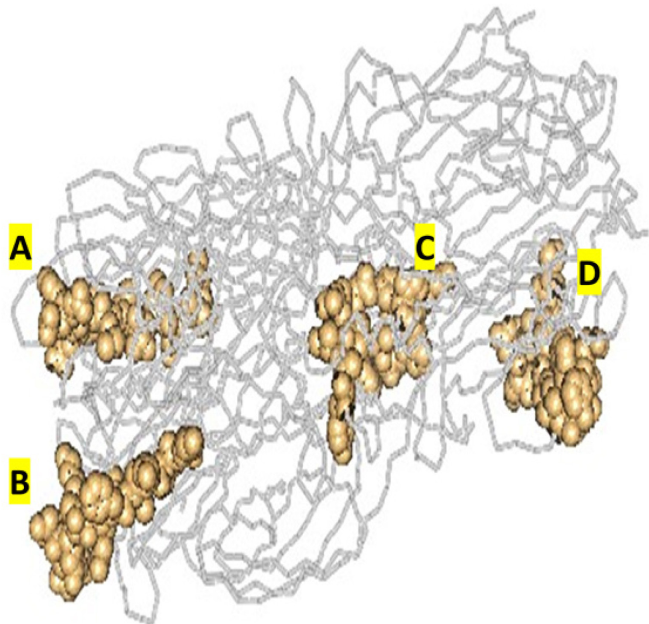
conferring broad-spectrum immunity [35]. In this study, we conducted conservancy analysis on the previously selected ninety CTL epitopes and ten HTL epitopes. From this analysis, eleven CTL epitopes (Table 3) and 4 HTL epitopes (Table 4) were chosen, considering their high conservancy and antigenicity scores. We evaluated the cross-reactivity of the chosen CTL and HTL epitopes with human proteomes and their interaction with human MHC class alleles (Tables 3 and 4).

To predict MHC class-I and MHC class-II binding alleles of the selected CTL and HTL epitopes, we utilized the MHC-I and MHC-II crucial prediction tools from the IEDB web server, respectively. Among all the CTL epitopes, VQADLTLMF exhibited binding capability to numerous alleles, including HLA-B15:01, HLA-A32:01, HLA-A23:01, HLA-A24:02, HLA-A02:06, HLA-A30:02, HLA-B58:01, HLA-B35:01, HLA-B40:01, HLA-B44:03, HLA-B53:01, HLA-A26:01, HLA-B*44:02. Subsequently, the results of MHC-I and MHC-II binding alleles were employed for population coverage analysis.

Population coverage of the epitopes

The immunogenic response of epitope-based vaccines may vary due to the highly polymorphic nature of human HLA across different ethnicities and regions [37].

Figure 2. Selected epitopes (shown in orange) of the envelopment polyprotein (Epp).



The conformation of the envelopment polyprotein (EPP) is a folded structure that can be visualized in 3D using the Pymol molecular visualization tool. The orange regions represent the epitope regions, which are the parts of the protein. The selected epitopes are: A) PLKLYSSFACMLHYQ; B) CMLHYQLGSFSSLYI; C) LIVSYASACSELIQA; D) VQADLTLMFDNFEVD.

Table 3. Final selected CTL epitopes

Epitope	Position	Super type	Combined core	Antigenicity	Immunogenicity	Interaction with MHC class I alleles
RRPHRSVRV	1194-1202	B27	1.1798	0.7047	-0.0113	HLA-A*30:01, HLA-B*08:01
RENKCFEQC	810-818	B44	0.9254	0.7453	-0.14458	HLA-B*40:01, HLA-B*44:02
RQMTGASLK	285-293	B27	1.3561	0.7603	-0.09421	HLA-A*03:01, HLA-A*11:01, HLA-A*30:01, HLA-A*31:01,
REEEMPEEL	48-56	B39	0.9556	0.7447	0.09481	HLA-B*40:01, HLA-B*44:03, HLA-B*44:02
NETSAEFSF	794-802	B44	1.3193	1.5503	-0.01938	HLA-B*44:03, HLA-B*44:02, HLA-B*40:01
VQADLTLMF	1007-1015	B27	0.8195	0.8723	-0.0512	HLA-B*15:01, HLA-A*32:01, HLA-A*23:01, HLA-A*24:02, HLA-A*02:06, HLA-A*30:02, HLA-B*58:01, HLA-B*35:01, HLA-B*40:01, HLA-B*44:03, HLA-B*53:01, HLA-A*26:01, HLA-B*44:02
LMLLLIVSY	679-687	B27	0.7735	0.5708	0.0379	HLA-B*15:01, HLA-A*30:02, HLA-A*32:01
VFALAPVVF	144-152	B62	0.8315	0.8181	0.08816	HLA-A*24:02, HLA-A*23:01, HLA-B*35:01, HLA-B*15:01, HLA-B*08:01, HLA-B*57:01, HLA-A*30:02, HLA-B*58:01
TMAGIAMTV	130-138	B62	0.8172	0.9979	0.08771	HLA-A*02:03, HLA-A*02:01, HLA-A*02:06
KLYSSFACM	1158-1166	B62	0.8171	0.618	-0.21705	HLA-B*15:01
ITSTGTGSL	1052-1060	B62	0.755	1.0818	-0.01316	HLA-A*30:02, HLA-B*57:01, HLA-B*58:01

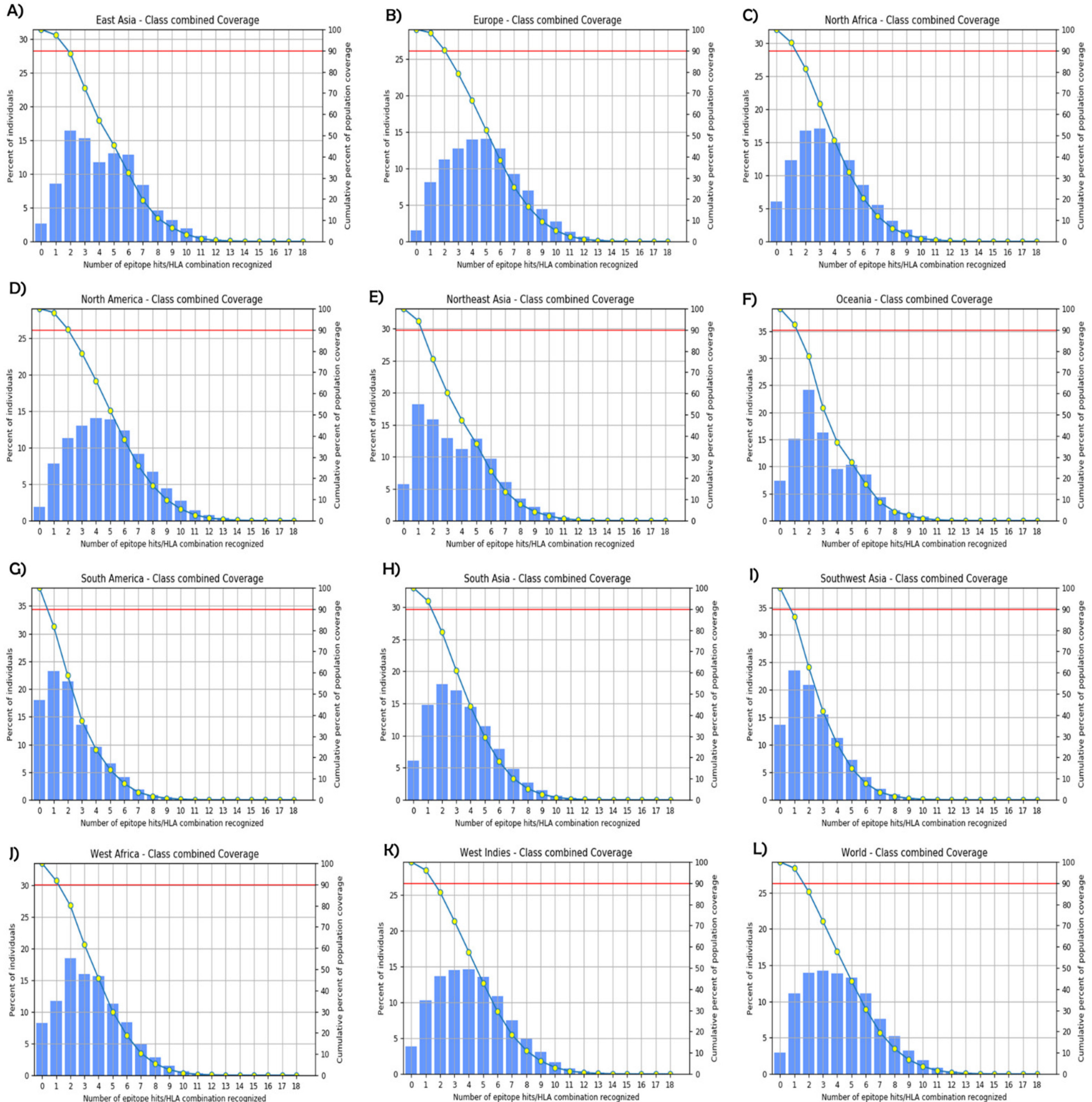
CTL: cytotoxic T-lymphocytes; MHC: major histocompatibility. All selected epitopes were non-toxic and non-allergic. The epitope conservancy for all the selected epitopes was 100%: indicating complete similarity across different variants. Cross-reactivity: on the other hand: was zero: indicating no binding or interaction with other antigens.

Table 4. Final selected HTL epitopes with cytokine-inducing ability IL-4, and IL-10.

Epitopes	Position	smm_align_ic50	Allergenicity	Antigenicity	IL4 Prediction	IL10 Prediction	IFN epitope	Interaction MHC class II alleles
CMLHYQLGSFSSLYI	1165-1179	45	NA	0.9312	Inducer	Inducer	Negative	HLA-DRB1*15:01
PLKLYSSFACMLHYQ	1156-1170	1170	NA	0.6267	Inducer	Inducer	Negative	HLA-DRB1*15:01
VQADLTLMFDNFEVD	1007-1021	74	A	0.495	Inducer	Non-inducer	Positive	HLA-DRB1*03:01
LIVSYASACSELIQA	683-697	231	A	0.1216	Non-inducer	inducer	Positive	HLA-DRB1*15:01

HTL: helper T-lymphocytes; IL: interleukin; IFN: interferon. The epitopes conservancy for these final HTL epitopes was 100% and cross-reactivity: on the other hand: was zero. All final HTL epitopes were non-toxic.

Figure 3. Charts showing population coverage in the different areas in the world.



The charts show the cumulative percentage of population coverage for different ethnicities and areas in the world. A) East Asia, B) Europe (highest percentage of population coverage), C) North Africa, D) North America, E) Northeast Asia, F) Oceania, G) South America, H) South Asia, I) Southwest Asia, J) West Africa, K) West Indies, and L) World population coverage.

To address this, we calculated the combined percentage of population coverage for the selected 11 CTL and 4 HTL epitopes (Figure 3) and their corresponding MHC binding alleles. Additionally, we performed combined calculations using the IEDB-provided population coverage analysis tool.

Remarkably, the highest population coverage was observed in Europe (98.47%) and North America (98.16%), while Central America exhibited the lowest coverage at 25.01% (Table 5). Encouragingly, the combined predictions of epitopes showed a substantial coverage of 97.04% of the global population, indicating their potential effectiveness worldwide.

Discussion

Vaccines are highly effective medical interventions that play a crucial role in preventing communicable diseases and improving human health. While there are conventional vaccines available for RVFV, they are burdened by several limitations, including high cost, restricted accessibility, and safety concerns [38]. Developing a subunit vaccine could overcome these limitations and provide a safer and more cost-effective option for RVFV control. In recent years, researchers have explored various innovative strategies to design effective vaccines, aiming to overcome the limitations of traditional approaches. Immunoinformatics has emerged as a critical tool for peptide-based vaccine development, wherein different antigenic components of pathogens are incorporated to elicit specific immune responses [26,38]. Among these approaches, epitope-based vaccines have gained prominence. They are subunit vaccines that utilize small molecular weight peptides, known as epitopes, which do not cause infection and are less prone to mutation. This method offers advantages such as time efficiency and cost-effectiveness in vaccine development and is capable of generating robust and immunogenic responses [3]. The immunogenicity of multi-epitope vaccines may vary depending on the host’s immune system, age, and health status. Hence, for multi-epitope-driven vaccine design, proper selection of target antigen epitopes is essential [26]. Computational tools have proven successful in predicting epitope-based vaccines against various viruses, including Ebola virus, Chikungunya virus, yellow fever virus, Zika virus, Herpes simplex virus etc. [3].

In our research, we employed computational methodology to screen the protein sequences of RVFV's EPP to identify the most antigenic epitope-based peptides. Our immunoinformatic investigation on the RVFV protein identified potential epitopes for vaccine development. We selected conserved antigenic proteins and analyzed their physicochemical characteristics. This approach allowed us to predict T- and B-cell epitopes, examine protein-ligand interactions, and design protein structures. We predicted T-cell epitopes and selected 90 promising candidates for further analysis. We identified 44 epitopes for helper T-cells. Among the chosen epitopes, ten were found to be antigenic, non-allergenic (except for two), and non-toxic, with four demonstrating cytokine-inducing ability. The EPP protein of RVFV is essential for virus replication. CD8+ T-cells have a significant role in controlling infection by recognizing and eliminating infected cells or releasing specific antiviral cytokines [39]. As a result, T-cell epitope-based vaccination stands out as a unique method for triggering a strong immune response against infectious pathogens like viruses [40]. Identifying the

Table 5. Analysis of population coverage of the selected T-cell epitopes based on their respective HLA alleles.

Population/Area	Combined (class I and class II)		
	Coverage[a]	Average hit[b]	PC90[c]
Central Africa	82.87%	2.70	0.58
Central America	25.01%	0.45	0.13
East Africa	84.48%	2.80	0.64
East Asia	97.36%	4.36	1.86
Europe	98.47%	4.86	2.04
North Africa	93.91%	3.65	1.32
North America	98.16%	4.86	2.03
Northeast Asia	94.32%	3.67	1.24
Oceania	92.61%	3.22	1.17
South Africa	85.85%	3.00	0.71
South America	81.96%	2.30	0.55
South Asia	93.85%	3.45	1.26
Southeast Asia	95.07%	3.92	1.37
Southwest Asia	86.33%	2.46	0.73
West Africa	91.75%	3.48	1.15
West Indies	96.12%	4.25	1.59
World	97.04%	4.32	1.63

HLA: human leukocyte antigen.

immunodominant T-cell response to RVFV EPP is crucial for understanding the protective immune response against RVFV infection and developing effective vaccines and immunotherapies.

The evaluation of epitope conservancy in this study holds significant implications for vaccine design against RVFV. By conducting a conservancy analysis on the selected CTL and HTL epitopes, we identified eleven CTL epitopes and four HTL epitopes that have been predicted to be immunogenic and antigenic. We further assessed the cross-reactivity of the chosen CTL and HTL epitopes with human proteomes and their interactions with human MHC class alleles. Notably, VQADLTLMF, among the CTL epitopes, exhibited binding capability to numerous HLA alleles. These findings offer valuable insights for population coverage analysis, aiding in the development of a more effective and broadly protective RVFV vaccines. In individuals infected with RVFV, both humoral (antibody-mediated) and cellular (T-cell-mediated) immune responses are triggered. Among the significant immunogenic proteins of RVFV, glycoprotein (Gn/Gc) stands out. A previous study with Gn/Gc glycoproteins in human models has demonstrated its capability to elicit robust T-cell responses [41]. Moreover, a DNA vaccine construct containing the glycoprotein Gn gene of RVFV linked to C3d-trimer sequences was shown to increase neutralizing antibody titers in mice [42]. Although this vaccine significantly lessened the disease symptoms in mice, it did not have a substantial impact on survival [42]. On the other hand, an epitope-based vaccine could be very effective due to the potent immune response.

Severe outbreaks of RVFV in humans with high fatality rates have been documented between 2000 and 2016 in various countries, including Saudi Arabia, Kenya, Somalia, Sudan, South Africa, Namibia, Mauritania, Egypt, Senegal, Niger, and Uganda [4]. The emergence and geographical range of RVFV in northern Egypt and the Middle East indicate the possibility of an incursion into Europe [15]. The diversity of human HLA across various ethnicities and regions can significantly impact the immunogenic response of epitope-based vaccines [43]. Therefore, to create an effective vaccine against RVFV, it is crucial to achieve higher population coverage in these regions and countries. By using the IEDB-provided population coverage analysis tool, the combined population coverage analysis showed that Central Africa, East Africa, East Asia, Europe, North Africa, and North America covered approximately 82.87%, 84.48%, 97.36%, 98.47%, 93.91%, and 98.16% of the population, respectively (Table 5). Notably, the combined predictions of epitopes resulted in an impressive population coverage of 97.04%, suggesting their potential efficacy in diverse global populations. These findings underscore the significance of population coverage analysis in optimizing the design and impact of epitope-based vaccines against RVFV. Overall, the immunogenic properties, including antigenicity and allergenicity, along with stable HLA allele binding capability, broaden and validate the potential of these epitopes as promising, safe, and effective vaccine candidates.

Conclusions

This immunoinformatics study on the RVFV protein identified potential epitopes that can serve as vaccine candidates. The utilization of computational methods and bioinformatics tools offered a cost-effective and efficient approach for screening large datasets and predicting immunogenicity. The identified epitopes demonstrated promising results concerning antigenicity, conservancy, and binding affinity to MHC molecules, suggesting their potential use in vaccine development against RVFV. In order to advance vaccine development, researchers must carry out additional experimental validation to verify the immunogenicity and safety of the predicted epitopes. Furthermore, due to the virus's evolution and mutation, ongoing surveillance and updating of the epitope repertoire will be necessary to ensure the vaccine's sustained effectiveness. Overall, this study constitutes a valuable contribution to the fields of immunoinformatics and vaccine development for RVFV, laying the groundwork for further research.

Acknowledgements

We would like to express our sincere gratitude to the Department of Biochemistry and Molecular Biology, Shahjalal University of Science and Technology (SUST), Sylhet, Bangladesh, for their valuable support and resources during the course of this research.

References

1. Ikegami T, Makino S (2011) The pathogenesis of Rift Valley fever. Viruses 3: 493–519. doi: 10.3390/v3050493.
2. Ly HJ, Ikegami T (2016) Rift Valley fever virus NSs protein functions and the similarity to other bunyavirus NSs proteins. Virol J 13: 118. doi: 10.1186/s12985-016-0573-8.
3. Ashik AI, Hasan M, Tasnim AT, Chowdhury MB, Hossain T, Ahmed S (2020) An immunoinformatics study on the spike protein of SARS-CoV-2 revealing potential epitopes as vaccine

candidates. Heliyon 6: e04865. doi: 10.1016/j.heliyon.2020.e04865.
4. Linthicum KJ, Britch SC, Anyamba A (2016) Rift Valley fever: an emerging mosquito-borne disease. Annu Rev Entomol 61: 395–415. doi: 10.1146/annurev-ento-010715-023819.
5. Uwishema O, Chalhoub E, Torbati T, David SC, Khoury C, Ribeiro LLPA, Nasrallah Y, Bekele BK, Onyeaka H (2022) Rift Valley fever during the COVID-19 pandemic in Africa: a double burden for Africa's healthcare system. Health Sci Rep 5: e468. doi: 10.1002/hsr2.468.
6. Ahmed A, Mahmoud I, Eldigail M, Elhassan RM, Weaver SC (2021) The emergence of Rift Valley fever in Gedaref State urges the need for a cross-border one health strategy and enforcement of the international health regulations. Pathogens 10: 885. doi: 10.3390/pathogens10070885.
7. Kim Y, Métras R, Dommergues L, Youssouffi C, Combo S, Le Godais G, Pfeiffer DU, Cêtre-Sossah C, Cardinale E, Filleul L, Youssouf H, Subiros M, Fournié G (2021) The role of livestock movements in the spread of Rift Valley fever virus in animals and humans in Mayotte, 2018-19. PLoS Negl Trop Dis 15: e0009202. doi: 10.1371/journal.pntd.0009202.
8. Azhar EI, Madani TA, Kao M, Jamjoom GA (2010) A seroepidemiological study of Rift Valley fever virus among slaughter houses' workers in Saudi Arabia. Int J Med Med Sci 2: 111–115.
9. Wright D, Kortekaas J, Bowden TA, Warimwe GM (2019) Rift Valley fever: biology and epidemiology. J Gen Virol 100: 1187–1199. doi: 10.1099/jgv.0.001296.
10. Nanyingi MO, Munyua P, Kiama SG, Muchemi GM, Thumbi SM, Bitek AO, Bett B, Muriithi RM, Njenga MK (2015) A systematic review of Rift Valley Fever epidemiology 1931–2014. Infect Ecol Epidemiol 5. doi: 10.3402/iee.v5.28024.
11. Gerken KN, LaBeaud AD, Mandi H, L'Azou Jackson M, Breugelmans JG, King CH (2022) Paving the way for human vaccination against Rift Valley fever virus: a systematic literature review of RVFV epidemiology from 1999 to 2021. PLoS Negl Trop Dis 16: e0009852. doi: 10.1371/journal.pntd.0009852.
12. Socha W, Kwasnik M, Larska M, Rola J, Rozek W (2022) Vector-borne viral diseases as a current threat for human and animal health-one health perspective. J Clin Med 11: 3026. doi: 10.3390/jcm11113026.
13. Petrova V, Kristiansen P, Norheim G, Yimer SA (2020) Rift valley fever: diagnostic challenges and investment needs for vaccine development. BMJ Glob Health 5: e002694. doi: 10.1136/bmjgh-2020-002694.
14. Bhuiyan MA, Quayum ST, Ahammad F, Alam R, Samad A, Nain Z (2021) Discovery of potential immune epitopes and peptide vaccine design - a prophylactic strategy against Rift Valley fever virus. F1000Res 9: 1–28. doi: 10.12688/f1000research.24975.1.
15. Rolin AI, Berrang-Ford L, Kulkarni MA (2013) The risk of Rift Valley fever virus introduction and establishment in the United States and European Union. Emerg Microbes Infect 2: e81. doi: 10.1038/emi.2013.81.
16. Wichgers Schreur PJ, Bird BH, Ikegami T, Bermúdez-Méndez E, Kortekaas J (2023) Perspectives of next-generation live-attenuated Rift Valley fever vaccines for animal and human use. Vaccines (Basel) 11: 707. doi: 10.3390/vaccines11030707.
17. Pollard AJ., Bijker EM (2021). A guide to vaccinology: from basic principles to new developments. Nat Rev Immunol 21: 83–100. doi: 10.1038/s41577-020-00479-7.
18. Faburay B, LaBeaud AD, McVey DS, Wilson WC, Richt JA (2017) Current status of Rift Valley fever vaccine development. Vaccines (Basel). 5: 29. doi: 10.3390/vaccines5030029.
19. Albutti A (2022) Proteome-wide and protein-specific multi-epitope vaccine constructs against the Rift Valley fever virus outbreak using integrated omics approaches. Front Microbiol 13: 921683. doi: 10.3389/fmicb.2022.921683.
20. Zhang C, Maruggi G, Shan H, Li J (2019) Advances in mRNA vaccines for infectious diseases. Front Immunol 10: 594. doi: 10.3389/fimmu.2019.00594.
21. Nielsen SS, Alvarez J, Bicout DJ, Calistri P, Depner K, Drewe JA, Garin-Bastuji B, Gonzales Rojas JL, Gortázar Schmidt C, Herskin M, Michel V, Miranda Chueca MÁ, Pasquali P, Roberts HC, Sihvonen LH, Stahl K, Calvo AV, Viltrop A, Winckler C, Gubbins S, Antoniou SE, Broglia A, Abrahantes JC, Dhollander S, Van der Stede Y (2020) Rift Valley fever — assessment of effectiveness of surveillance and control measures in the EU. EFSA J 18: e06292. doi: 10.2903/j.efsa.2020.6292.
22. Milne G, Hames T, Scotton C, Gent N, Johnsen A, Anderson RM, Ward T (2021) Does infection with or vaccination against SARS-CoV-2 lead to lasting immunity? Lancet Respir Med 9: 1450–1466. doi: 10.1016/S2213-2600(21)00407-0.
23. Hajissa K, Zakaria R, Suppian R, Mohamed Z (2019) Epitope-based vaccine as a universal vaccination strategy against *Toxoplasma gondii* infection: a mini-review. J Adv Vet Anim Res 6: 174–182. doi: 10.5455/javar.2019.f329.
24. UniProt Consortium (2023) UniProt: the universal protein knowledgebase in 2023. Nucleic Acids Res 51: D523–D531. doi: 10.1093/nar/gkac1052.
25. Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8: 4. doi: 10.1186/1471-2105-8-4.
26. Imran MA, Islam MR, Saha A, Ferdousee S, Mishu MA, Ghosh A (2022) Development of multi-epitope-based subunit vaccine against Crimean-Congo hemorrhagic fever virus using reverse vaccinology approach. Int J Pept Res Ther 28: 124. doi: 10.1007/s10989-022-10430-0.
27. Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, Lu CH (2014) CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. PLoS One 9: e99368. doi: 10.1371/journal.pone.0099368.
28. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 8: 424. doi: 10.1186/1471-2105-8-424.
29. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B (2019) The immune epitope database (IEDB): 2018 update. Nucleic Acids Res. 47: 339–343. doi: 10.1093/nar/gky1006.
30. Yu R, Dong S, Chen B, Liu Y, Li F, Si F, Xie C, Li Z (2022) Antigenicity alternations of variant PEDV S protein disclosed by linear B cell epitope mapping. Viruses 14: 1371. doi: 10.3390/v14071371.
31. El-Manzalawy Y, Dobbs D, Honavar V (2008) Predicting flexible length linear B-cell epitopes. Comput Syst

Bioinformatics Conf 7: 121–132. doi: 10.1142/9781848162648_0011.

32. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins. Naturwissenschaften 72: 212–213. doi: 10.1007/BF01195768.
33. Emini EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol 55: 836–839. doi: 10.1128/jvi.55.3.836-839.1985.
34. Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol 47: 45–148. doi: 10.1002/9780470122921.ch2.
35. Dhanda SK, Vir P, Raghava GP (2013) Designing of interferon-gamma inducing MHC class-II binders. Biol Direct 8: 30. doi: 10.1186/1745-6150-8-30.
36. Dhanda SK, Gupta S, Vir P, Raghava GP (2013) Prediction of IL4 inducing peptides. Clin Dev Immunol 2013: 263952. doi: 10.1155/2013/263952.
37. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics 7: 153. doi: 10.1186/1471-2105-7-153.
38. Kitandwe PK, McKay PF, Kaleebu P, Shattock RJ (2022) An overview of Rift Valley fever vaccine development strategies. Vaccines (Basel) 10: 1794. doi: 10.3390/vaccines10111794.
39. Appay V, Nixon DF, Donahoe SM, Gillespie GM, Dong T, King A, Ogg GS, Spiegel HM, Conlon C, Spina CA, Havlir DV, Richman DD, Waters A, Easterbrook P, McMichael AJ, Rowland-Jones SL (2000) HIV-specific CD8(+) T cells produce antiviral cytokines but are impaired in cytolytic function. J Exp Med 192: 63–75. doi: 10.1084/jem.192.1.63.
40. Hajissa K, Zakaria R, Suppian R, Mohamed Z (2019) Epitope-based vaccine as a universal vaccination strategy against *Toxoplasma gondii* infection: a mini-review. J Adv Vet Anim Res 6: 174–182. doi: 10.5455/javar.2019.f329.
41. López-Gil E, Lorenzo G, Hevia E, Borrego B, Eiden M, Groschup M, Gilbert SC, Brun A (2013) A single immunization with MVA expressing GnGc glycoproteins promotes epitope-specific CD8+-t cell activation and protects immune-competent mice against a lethal RVFV infection. PLoS Negl Trop Dis 7: e2309. doi: 10.1371/journal.pntd.0002309.
42. Boshra H, Lorenzo G, Rodriguez F, Brun A (2011) A DNA vaccine encoding ubiquitinated Rift Valley fever virus nucleoprotein provides consistent immunity and protects IFNAR−/− mice upon lethal virus challenge. Vaccine 29: 27. doi: 10.1016/j.vaccine.2011.04.043.
43. Oyarzun P, Kashyap M, Fica V, Salas-Burgos A, Gonzalez-Galarza FF, McCabe A, Jones AR, Middleton D, Kobe B (2021) A proteome-wide immunoinformatics tool to accelerate t-cell epitope discovery and vaccine design in the context of emerging infectious diseases: an ethnicity-oriented approach. Front Immunol 12: 598778. doi: 10.3389/fimmu.2021.598778.

Corresponding author

Muhammad Abul Kalam Azad, PhD.
Department of Agriculture, University of Arkansas at Pine Bluff, 1200 N. University Drive, Pine Bluff, AR 71601, USA.
Tel: +1(870) 575-8153
Fax: 8705754629
Email: makazad-bmb@sust.edu, azadm@uapb.edu

Conflict of interests: No conflict of interests is declared.

Annex – Supplementary Items

Supplementary Table 1. CTL epitopes were predicted by MHC class-I.

Epitope	Super Type	Combined Score	Antigenicity	Toxicity	Allergenicity	Epitope Conservancy (%)	Cross-reactivity
AIICLAVLY	A1	0.875	< 0.4	Non-toxin	NA	100	Zero
YSSFACMLH	A1	0.8494	< 0.4	Non-toxin	NA	100	Zero
STAHEVVPF	A1	0.8456	< 0.4	Non-toxin	NA	100	Zero
DATCKPVTY	A1	0.8412	< 0.4	Non-toxin	NA	100	Zero
LMLLLIVSY	A1	0.813	< 0.4	Non-toxin	NA	100	Zero
KLYSSFACM	A2	1.2391	< 0.4	Non-toxin	NA	100	Zero
TMAGIAMTV	A2	1.2078	< 0.4	Non-toxin	NA	100	Zero
FVFSSIAII	A2	1.1961	< 0.4	Non-toxin	NA	100	Zero
MLLLIVSYA	A2	1.1725	< 0.4	Non-toxin	NA	100	Zero
RVFNCIDWV	A2	1.0863	< 0.4	Non-toxin	NA	100	Zero
AVFALAPVV	A2	1.0125	< 0.4	Non-toxin	NA	100	Zero
SIAIICLAV	A2	0.9928	< 0.4	Non-toxin	NA	100	Zero
IAMTVLPAL	A2	0.9772	< 0.4	Non-toxin	NA	100	Zero
SLKCGLLPL	A2	0.9575	< 0.4	Non-toxin	NA	100	Zero
LLIVSYASA	A2	0.9422	< 0.4	Non-toxin	NA	100	Zero
LAVFALAPV	A2	0.877	< 0.4	Non-toxin	NA	100	Zero
HMAHDDQSV	A2	0.863	< 0.4	Non-toxin	NA	100	Zero
LLVKGTLIA	A2	0.7774	< 0.4	Non-toxin	NA	100	Zero
FALAPVVFA	A2	0.771	< 0.4	Non-toxin	NA	100	Zero
RQMTGASLK	A3	1.5793	< 0.4	Non-toxin	NA	100	Zero
KIGGHGSKK	A3	1.5444	< 0.4	Non-toxin	NA	100	Zero
QMTGASLKK	A3	1.284	< 0.4	Non-toxin	NA	100	Zero
RAPNLISYK	A3	1.2394	< 0.4	Non-toxin	NA	100	Zero
AIICLAVLY	A3	1.0689	< 0.4	Non-toxin	NA	100	Zero
LMLLLIVSY	A3	1.0655	< 0.4	Non-toxin	NA	100	Zero
GAEACLMLK	A3	0.9528	< 0.4	Non-toxin	NA	100	Zero
KLYSSFACM	A3	0.9473	< 0.4	Non-toxin	NA	100	Zero
KCLKIAPRK	A3	0.9089	< 0.4	Non-toxin	NA	100	Zero
VLPSENGTK	A3	0.8589	< 0.4	Non-toxin	NA	100	Zero
TMDSGQTKR	A3	0.8445	< 0.4	Non-toxin	NA	100	Zero
EVVPFAVFK	A3	0.8114	< 0.4	Non-toxin	NA	100	Zero
RYSTYLMLL	A24	1.8042	< 0.4	Non-toxin	NA	100	Zero
VFALAPVVF	A24	1.7143	< 0.4	Non-toxin	NA	100	Zero
ISGSNSFSF	A24	1.2712	< 0.4	Non-toxin	NA	100	Zero
VFVFSSIAI	A24	1.2127	< 0.4	Non-toxin	NA	100	Zero
VQADLTLMF	A24	1.1694	< 0.4	Non-toxin	NA	100	Zero
TYLMLLLIV	A24	1.1312	< 0.4	Non-toxin	NA	100	Zero
VYLDKLDLK	A24	0.9674	< 0.4	Non-toxin	NA	100	Zero
STAHEVVPF	A26	1.5074	< 0.4	Non-toxin	NA	100	Zero
EGISGSNSF	A26	1.3847	< 0.4	Non-toxin	NA	100	Zero
AIICLAVLY	A26	1.1968	< 0.4	Non-toxin	NA	100	Zero
NCIDWVHKL	A26	1.127	< 0.4	Non-toxin	NA	100	Zero
ETMAGIAMT	A26	1.1003	< 0.4	Non-toxin	NA	100	Zero
QTKRELKSF	A26	0.8764	< 0.4	Non-toxin	NA	100	Zero
FVFSSIAII	A26	0.8284	< 0.4	Non-toxin	NA	100	Zero
RQMTGASLK	B27	1.3561	< 0.4	Non-toxin	NA	100	Zero
RRPHRSVRV	B27	1.1798	< 0.4	Non-toxin	NA	100	Zero
SRFTNWGSV	B27	1.1302	< 0.4	Non-toxin	NA	100	Zero
PRYSTYLML	B27	1.0751	< 0.4	Non-toxin	NA	100	Zero
KRELKSFDI	B27	1.0048	< 0.4	Non-toxin	NA	100	Zero
KRELSAKPI	B27	0.9837	< 0.4	Non-toxin	NA	100	Zero
VQADLTLMF	B27	0.8195	< 0.4	Non-toxin	NA	100	Zero
LMLLLIVSY	B27	0.7735	< 0.4	Non-toxin	NA	100	Zero
PRYSTYLML	B27	1.1501	< 0.4	Non-toxin	NA	100	Zero
FSSIAIICL	B27	1.1275	< 0.4	Non-toxin	NA	100	Zero
CNAGARVCL	B27	1.1	< 0.4	Non-toxin	NA	100	Zero
LPALAVFAL	B27	1.0156	< 0.4	Non-toxin	NA	100	Zero
REEEMPEEL	B27	0.9556	< 0.4	Non-toxin	NA	100	Zero
LLEKGKFPL	B27	0.9231	< 0.4	Non-toxin	NA	100	Zero
LQSVRKEAL	B27	0.7675	< 0.4	Non-toxin	NA	100	Zero
REEEMPEEL	B44	1.8504	< 0.4	Non-toxin	NA	100	Zero
NETSAEFSF	B44	1.3193	< 0.4	Non-toxin	NA	100	Zero
LEKGKFPLF	B44	0.9929	< 0.4	Non-toxin	NA	100	Zero
RENKCFEQC	B44	0.9254	< 0.4	Non-toxin	NA	100	Zero
PEVEEEFMY	B44	0.7723	< 0.4	Non-toxin	NA	100	Zero
VQADLTLMF	B62	1.4433	< 0.4	Non-toxin	NA	100	Zero

Epitope	Super Type	Combined Score	Antigenicity	Toxicity	Allergenicity	Epitope Conservancy (%)	Cross-reactivity
LMLLLIVSY	B62	1.4184	< 0.4	Non-toxin	NA	100	Zero
VQIQVSGVW	B62	1.2248	< 0.4	Non-toxin	NA	100	Zero
STAHEVVPF	B62	1.2121	< 0.4	Non-toxin	NA	100	Zero
LSCREGQSY	B62	1.2054	< 0.4	Non-toxin	NA	100	Zero
MLHYQLGSF	B62	1.2006	< 0.4	Non-toxin	NA	100	Zero
KGNRGVQAF	B62	1.1804	< 0.4	Non-toxin	NA	100	Zero
GLVVRSTGF	B62	1.1591	< 0.4	Non-toxin	NA	100	Zero
QTKRELKSF	B62	1.0852	< 0.4	Non-toxin	NA	100	Zero
SLKKGSYPL	B62	1.0074	< 0.4	Non-toxin	NA	100	Zero
ISGSNSFSF	B62	0.9712	< 0.4	Non-toxin	NA	100	Zero
KLTLEITDF	B62	0.9549	< 0.4	Non-toxin	NA	100	Zero
RDNETSAEF	B62	0.9446	< 0.4	Non-toxin	NA	100	Zero
AIICLAVLY	B62	0.9388	< 0.4	Non-toxin	NA	100	Zero
LQSVRKEAL	B62	0.9314	< 0.4	Non-toxin	NA	100	Zero
SLKCGLLPL	B62	0.9096	< 0.4	Non-toxin	NA	100	Zero
PQTRNDKTF	B62	0.9009	< 0.4	Non-toxin	NA	100	Zero
EGISGSNSF	B62	0.8708	< 0.4	Non-toxin	NA	100	Zero
LQSAHGNPC	B62	0.8323	< 0.4	Non-toxin	NA	100	Zero
VFALAPVVF	B62	0.8315	< 0.4	Non-toxin	NA	100	Zero
TMAGIAMTV	B62	0.8172	< 0.4	Non-toxin	NA	100	Zero
KLYSSFACM	B62	0.8171	< 0.4	Non-toxin	NA	100	Zero
VVFAEDPHL	B62	0.7725	< 0.4	Non-toxin	NA	100	Zero
HMAHDDQSV	B62	0.7645	< 0.4	Non-toxin	NA	100	Zero
ITSTGTGSL	B62	0.755	< 0.4	Non-toxin	NA	100	Zero

CTL: cytotoxic T-lymphocytes; MHC: major histocompatibility.

Supplementary Table 2. B-cell binding epitopes from envelopment polyprotein (glycoprotein: Gn/Gc).

Epitopes	Start position	Score	Antigenicity	Allergenicity	Toxicity
EHKGQYKGTMDSGQTK	376	0.96	< 0.4	NA	Non-toxin
RVLKCLKIAPRKVLNP	605	0.91	< 0.4	NA	Non-toxin
TSTGTGSLSAHNKDGS	1053	0.86	< 0.4	NA	Non-toxin
KGTMDSGQTKRELKSF	382	0.84	< 0.4	NA	Non-toxin
MTQEDATCKPVTYAGA	172	0.84	< 0.4	NA	Non-toxin
GVCVTGSQSPSTEITL	509	0.82	< 0.4	NA	Non-toxin
TALIRAGSVGAEACLM	718	0.81	< 0.4	NA	Non-toxin
SGSWNFFDWFSGLMSW	1137	0.81	< 0.4	NA	Non-toxin
HKLTLEITDFDGSVST	857	0.8	< 0.4	NA	Non-toxin
REAGGESTVVNPKSGS	1124	0.8	< 0.4	NA	Non-toxin
VGIVRCERRRDAKQIG	112	0.8	< 0.4	NA	Non-toxin
EFMYSCDGDERPLLVK	1096	0.8	< 0.4	NA	Non-toxin
FSEIPRQGFLGEIRCN	920	0.79	< 0.4	NA	Non-toxin
ECHVNRCLSWRDNETS	782	0.79	< 0.4	NA	Non-toxin
LKAIIAADGLNNITCH	77	0.79	< 0.4	NA	Non-toxin
TVSSELSCREGQSYWT	749	0.79	< 0.4	NA	Non-toxin
NREIGWMEGGQLVLGN	647	0.79	< 0.4	NA	Non-toxin
GSGIVQIQVSGVWKKP	439	0.79	< 0.4	NA	Non-toxin
LFCQSSEDDGSKLKTK	302	0.79	< 0.4	NA	Non-toxin
HFTVPEVEEEFMYSCD	1087	0.79	< 0.4	NA	Non-toxin
KPMIDQLECTTNLIDP	957	0.78	< 0.4	NA	Non-toxin
SVRKEALRVFNCIDWV	841	0.78	< 0.4	NA	Non-toxin
KGVKEDQTKFLKIKTV	735	0.78	< 0.4	NA	Non-toxin
PSTEITLKYPGISQSS	518	0.78	< 0.4	NA	Non-toxin
CTTCITKCEPHGLVVR	479	0.78	< 0.4	NA	Non-toxin
HKKCDGQLSTAHEVVP	329	0.78	< 0.4	NA	Non-toxin
CDAAFLNLTGCYSCNA	1029	0.78	< 0.4	NA	Non-toxin
QSVSSKIVAHCPPQDP	546	0.76	< 0.4	NA	Non-toxin
KISLIKGPPHKKRVGI	99	0.75	< 0.4	NA	Non-toxin
NITCHGKDPEDKISLI	88	0.75	< 0.4	NA	Non-toxin
IAIICLAVLYRVLKCL	595	0.75	< 0.4	NA	Non-toxin
SCSISGIREVKTSSQE	57	0.75	< 0.4	NA	Non-toxin
VTYAGACSSFDVLLEK	182	0.75	< 0.4	NA	Non-toxin
SGVWKKPLCVGYERVV	448	0.74	< 0.4	NA	Non-toxin
MIEGAWDSLREEEMPE	39	0.74	< 0.4	NA	Non-toxin
DLKTEENLLPDSFVCF	360	0.74	< 0.4	NA	Non-toxin
SVSLSLDAEGISGSNS	887	0.73	< 0.4	NA	Non-toxin
DAKQIGRETMAGIAMT	122	0.73	< 0.4	NA	Non-toxin
CGLLPLRRPHRSVRVK	1188	0.73	< 0.4	NA	Non-toxin

Epitopes	**Start position**	**Score**	**Antigenicity**	**Allergenicity**	**Toxicity**
NDKTFAASKGNRGVQA	986	0.72	< 0.4	NA	Non-toxin
ESSVLSAHESCLRAPN	937	0.72	< 0.4	NA	Non-toxin
LSSRRCHLVGECHVNR	772	0.72	< 0.4	NA	Non-toxin
GVQAFSKGSVQADLTL	998	0.71	< 0.4	NA	Non-toxin
ASSSRFTNWGSVSLSL	877	0.71	< 0.4	NA	Non-toxin
AHGLINYQCHTALSAF	571	0.71	< 0.4	NA	Non-toxin
LSTAHEVVPFAVFKNS	336	0.71	< 0.4	NA	Non-toxin
PSENGTKDQCQILHFT	1074	0.71	< 0.4	NA	Non-toxin
TEGVNTKCRLSGTALI	706	0.7	< 0.4	NA	Non-toxin
VDFVGAAVSCDAAFLN	1020	0.7	< 0.4	NA	Non-toxin
LGEIRCNSESSVLSAH	929	0.69	< 0.4	NA	Non-toxin
AVFKNSKKVYLDKLDL	346	0.69	< 0.4	NA	Non-toxin
EDDGSKLKTKMKGVCE	308	0.69	< 0.4	NA	Non-toxin
TMAGIAMTVLPALAVF	130	0.69	< 0.4	NA	Non-toxin
HGLVVRSTGFKISSAV	489	0.67	< 0.4	NA	Non-toxin
AHESCLRAPNLISYKP	943	0.65	< 0.4	NA	Non-toxin
YKKMVARVAHNINQVN	632	0.65	< 0.4	NA	Non-toxin
SLKKGSYPLQDLFCQS	291	0.65	< 0.4	NA	Non-toxin
KFPLFQSYAHHRTLLE	199	0.65	< 0.4	NA	Non-toxin
ACMLHYQLGSFSSLYI	1164	0.65	< 0.4	NA	Non-toxin
SLSAHNKDGSLHIVLP	1059	0.65	< 0.4	NA	Non-toxin
EVKTSSQELYRALKAI	65	0.63	< 0.4	NA	Non-toxin
VHTYLQSVRKEALRVF	835	0.62	< 0.4	NA	Non-toxin
DNETSAEFSFVGESTT	793	0.62	< 0.4	NA	Non-toxin
YLMLLLIVSYASACSE	678	0.61	< 0.4	NA	Non-toxin
SFSSLYILEEQASLKC	1173	0.61	< 0.4	NA	Non-toxin
NLLPDSFVCFEHKGQY	366	0.6	< 0.4	NA	Non-toxin
STTMRENKCFEQCGGW	806	0.59	< 0.4	NA	Non-toxin
DSLREEEMPEELSCSI	45	0.59	< 0.4	NA	Non-toxin
PLCVGYERVVVKRELS	454	0.57	< 0.4	NA	Non-toxin
YQCHTALSAFVVVFVF	577	0.56	< 0.4	NA	Non-toxin
KKVYLDKLDLKTEENL	352	0.56	< 0.4	NA	Non-toxin