



## Original article

# Development of multivalent vaccine targeting M segment of Crimean Congo Hemorrhagic Fever Virus (CCHFV) using immunoinformatic approaches



Maaza Sana<sup>a</sup>, Aneela Javed<sup>a,\*</sup>, Syed Babar Jamal<sup>b</sup>, Muhammad Junaid<sup>c</sup>, Muhammad Faheem<sup>b,\*</sup>

<sup>a</sup> Atta-ur-Rahman School of Applied Biosciences, National University of Science and Technology, Sector H-12, Islamabad, Pakistan

<sup>b</sup> Department of Biological Sciences, National University of Medical Sciences, Abid Majeed Rd, Rawalpindi, Punjab 46000, Pakistan

<sup>c</sup> Precision Medicine Laboratory, Rehman Medical Institute, Hayatabad, Peshawar, KPK, 25000, Pakistan

## ARTICLE INFO

## Article history:

Received 12 August 2021

Revised 25 November 2021

Accepted 4 December 2021

Available online 10 December 2021

## Keywords:

CCHFV

Vaccine

Glycoproteins

Adjuvant

B-cell

T-cell

## ABSTRACT

Crimean–Congo Hemorrhagic Fever (CCHF) is a tick-borne viral infection with no licensed vaccine or therapeutics available for its treatment. In the present study we have developed the first multi-epitope sub-unit vaccine effective against all the seven genotypes of CCHF virus (CCHFV). The vaccine contains five B-cell, two MHC-II (HTL), and three MHC-I (CTL) epitopes screened from two structural glycoproteins (Gc and Gn in M segment) of CCHFV with an N-terminus human  $\beta$ -defensin as an adjuvant, as well as an N-terminus EAAAK sequence. The epitopes were rigorously investigated for their antigenicity, allergenicity, IFN gamma induction, anti-inflammatory responses, stability, and toxicity. The three-dimensional structure of the vaccine was predicted and docked with TLR-3, TLR-8, and TLR-9 receptors to find the strength of the binding complexes via molecular dynamics simulation. After codon adaptation, the sub-unit vaccine construct was developed in a pDual-GC plasmid and has population coverage of 98.47% of the world's population (HLA-I & II combined). The immune simulation studies were carried out on the C-ImmSim in-silico interface showing a marked increase in the production of cellular and humoral response (B-cell and T-cell) as well as TGF $\beta$ , IL-2, IL-10, and IL-12 indicating that the proposed vaccine would be able to sufficiently provoke both humoral and cell-mediated immune responses. Thus, making it a new and promising vaccine candidate against CCHFV.

© 2021 The Authors. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Crimean–Congo Hemorrhagic Fever (CCHF), caused by Crimean–Congo Hemorrhagic Fever Virus (CCHFV), is a deadly zoonotic virus that infects humans worldwide. CCHF ranks at the top of WHO's Blueprint List of Priority Diseases published in 2018, demanding intense research and development (World Health Organization, n.d.). First discovered in Crimea during the 2nd World War, it has

now spread to Africa, Eastern Europe, Central Asia, Middle East, and South Asia (ranked as 2nd most endemic region for CCHF) (World Health Organization, n.d.). The disquieting increase in both geographical distribution and number of cases of CCHF is due to the lack of epidemiological research in animals and humans, which is in-turn linked to absence of affordable and proficient diagnostic as well as therapeutics (both animals and humans) in CCHF-endemic countries (World Health Organization, n.d.).

Molecular as well as serological assays such as nucleic acid tests (NAT) for RNA detection, IgG (for survivors), IgM (acute viral infection) and antigen (Ag) capture ELISA, are typically used to detect CCHF in human and animals (World Health Organization, n.d.; Mazzola and Kelly–Cirino, 2019). However, there are certain limitations, such as viral genetic diversity (World Health Organization, n.d.) and lack of a detectable antibody response in hosts (Drosten et al., 2003; Fernandez-García et al., 2014). Consequently, at present, there is no universally accepted gold standard (reference test) for CCHFV (World Health Organization, n.d.). Furthermore, infected animals remain asymptomatic and can easily migrate from one

\* Corresponding authors at: Department of Biological Sciences, National University of Medical Sciences, Abid Majeed Road, Rawalpindi, Punjab 46000, Pakistan (M. Faheem).

E-mail addresses: [javedaneela@asab.nust.edu.pk](mailto:javedaneela@asab.nust.edu.pk) (A. Javed), [mohammad.fahem@numspak.edu.pk](mailto:mohammad.fahem@numspak.edu.pk) (M. Faheem).

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

region to another, there-by spreading the virus to new locations (Gonzalez et al., 1991; World Health Organization, n.d.). Therefore, absence of standard diagnostic test and uncontrollable migration of infected animals there is an even greater need of accessible prophylaxes. While the World Health Organization recommends ribavirin (Stuart et al., 2009), contradicting evidence regarding ribavirin's efficacy in treating CCHFV are observed, both in the case of human (Johnson et al., 2018) and animal models (Hawman et al., 2018; Oestereich et al., 2014; Bente et al., 2010). Thereby, indicating the pressing need of new effective treatments. Although, a mouse brain-derived vaccine obtained after attenuation of a strain V42/81 of CCHFV, is being used in Bulgaria. But probably due to its safety and practical concerns, it will not have global outreach (Papa et al., 2011; Mousavi-Jazi et al., 2012). Other vaccination platforms such as, vesicular stomatitis virus-based, modified vaccinia Ankara-based, subunit-based, human adenovirus-based, DNA-based, virus-like particle-based (VLP), and transgenic plant-based vaccines, have been investigated in mouse models (Zivcec et al., 2018; Buttigieg et al., 2014; Ghiasi et al., 2011; Garrison et al., 2017; Kortekaas et al., 2015; Canakoglu et al., 2015; Hinkula et al., 2017; Scholte et al., 2019; Rodriguez et al., 2019) as well as *Cynomolgus* macaque based vaccine has been tested which is a non-human primate's model (Hawman et al., 2021). But the human randomized controlled trials are still lacking. Another facet of slower development of CCHF vaccine is the critical requirement of high-cost BSL-4 facility for its production (World Health Organization, n.d.). In light of these statistics where no licensed therapeutics currently available for CCHF and it becomes even more worrisome when CCHF carries a mortality rate of 10–40% (World Health Organization, n.d.).

The virus is maintained in the zoonotic cycle via tick-mediated transmission (*Hyalomma* (predominantly) as well as *Amblyomma* and *Rhipicephalus* species), which cycles between various vertebrate species (wild and domestic mammals). Despite the animals developing transient viremia, they remain asymptomatic (Gonzalez et al., 1991; Gonzalez et al., 1991; "WHO EMRO | Infectious disease outbreaks reported in the Eastern Mediterranean Region in 2018 | News | Epidemic and pandemic diseases," n.d.; World Health Organization, n.d.).

CCHFV is a negative-sense single-stranded RNA virus((-)ssRNA) classified in the *Nairovirus* genus of the *Bunyaviridae* family. It has a three-segmented genome (Gonzalez et al., 1991) that encode small (S), medium (M), and large (L) RNA segments, which translates into a viral nucleocapsid protein (NP), glycoprotein precursor (Pre-Gn, Pre-Gc ultimately forms Gn and Gc), and polymerase proteins (viral RNA dependent RNA Polymerase), respectively. CCHF has seven known genotypes CCHFV-1, CCHFV-2, CCHFV-3, CCHFV-4, CCHFV-5, CCHFV-6 and CCHFV-7 (Flick and Whitehouse, 2005). Human host contracts infection when nucleolin receptor on endothelium allows the virus to enter the cells (Gaafar et al., 2019). After 2–14 days of incubation period, symptoms start to appear on 1–7th day of pre-hemorrhagic phase. These symptoms vary from fever, myalgia, nausea, vomiting, diarrhea. The classical immunological TH1 and Th2 and CTL responses are generated. Some individuals will clear the infection after 7 days of hemorrhagic phase but if virus persists, it progresses to hemorrhagic phase where hemorrhages, bleeding from various sites, hematemeses, and ultimately causes death (World Health Organization, n.d.).

Tick Borne Viruses (TBV) are recognized by immature Dendritic Cells (DCs) via Toll-like receptors (TLRs) present on the cell surface or inside endosomes (Kochs et al., 2010). In the case of CCHFV, DCs bind to TLR-3, TLR8 and TLR-9 (Lindquist et al., 2018). Once, DCs take up viral antigens they become activated and start migration to lymphoid tissues for maturation. Therefore, for a virus to effectively infect the host DCs must be manipulated by the viruses

(Kazimírová et al., 2017). Activated DCs trigger naïve T cells. Subsequently, the activated T cells may give rise to T helper type 1 (Th1), Th2, and cytotoxic T lymphocyte (CTL) effector cells. These interactions initiate signaling cascades that elicit anti-viral Th1 responses. These responses include, increased expression of major histocompatibility complex (MHC) class II molecules (required for antigen presentation), T cell co-stimulatory molecules (i.e., CD80 and CD86), and proinflammatory cytokines, such as type I interferon (IFN); principally IFN gamma, interleukin (IL) 6, IL12 and TNF-alpha (Johnston et al., 2000; Masson et al., 2008).

In the backdrop of recent pandemic, the fact that CCHFV has much higher fatality rate than SARS CoV2, its zoonotic nature, its rapid spread to more regions and lack of effective therapeutics, the virus is a looming public health threat for many countries of the world. Its control demands research and development of effective vaccine and therapeutics. Current study was thus designed to develop a multivalent vaccine targeting all the seven genotypes, found globally, unlike previous genotype specific vaccines, targeting Gc and Gn (M-segment) of CCHFV. We exploited recently developed immunoinformatics tools to evaluate conserved genetic sequences from different CCHFV genotypes computationally in search of conserved epitopes, which were used to design sub-unit vaccine (Centers for Disease Control and Prevention et al., 2003; Dowall et al., 2017; Flick and Whitehouse, 2005; Gaafar et al., 2019; Larkin et al., 2007). Our designed sub-unit vaccine, when simulated for in-vivo molecular binding to different human immune components, was able to demonstrate humoral and innate immunity in-silico that can be further tested *in vitro* and *in vivo* for the development of multivalent vaccine targeting the envelop glycoprotein Gn and Gc encoded by the M segment of CCHFV genome.

## 2. Material and methods

The sub-sequential steps taken in this study (summarized in **Figure S1**) for the designing of our multivalent vaccine are thoroughly described below.

### 2.1. Epitope selection

#### 2.1.1. Sequences retrieval and sequence alignment

There were 350 sequences of the envelop glycoprotein encoded by M segment of CCHFV genome comprising of the seven genotypes retrieved from NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) and UniProt from EMBL (<https://www.uniprot.org/>). The reference sequence GenBank Accession Numbers are; ABB30027.1 (Genotype-I), ABB30037.1 (Genotype-II) ARB51455.1 (Genotype-III), ABD98124.1 (Genotype-IV), ABB72473.1 (Genotype-V), ABB30025.1 (Genotype-VI) and ABB30028.1 (Genotype-VII) (Wampande et al., 2021). The obtained sequences were subjected to multiple sequence alignment, to detect the similarity, using Clustal Omega. (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). The Clustal Omega aligned the input sequences into 26 Clustals. Then, the conserved sequences were searched manually in these obtained Clustals. They were classified as Highly Conserved (HC), Relatively Conserved (RC), and Less Conserved (LC) based on the percent sequence similarity between the sequences of each Clustal. Therefore, HC Clustals had 90–98%, RC Clustals had 80–89%, and LC Clustals had 70–79% similarity, respectively.

#### 2.1.2. B-Cell epitope prediction

The linear B-cell epitopes were predicted using IEDB (<http://tools.iedb.org/main/bcell/>), BepiPred-2.0 (<https://services.healthtech.dtu.dk/service.php?BepiPred-2.0>) and ABCPred ([https://webs.iitd.edu.in/raghava/abcpred/ABC\\_submission.html](https://webs.iitd.edu.in/raghava/abcpred/ABC_submission.html))

at the default threshold of 0.5 (Jespersen et al., 2017). The peptides which are  $\geq 20$  amino acids long were considered for further analyses (Ikram et al., 2018).

### 2.1.3. T-Cell epitope prediction

The MHC-I restricted epitopes were predicted using IEDB's Proteasomal cleavage/TAP transport/MHC class I combined predictor (<http://tools.iedb.org/processing/>) (Shekhar et al., 2012) and Syfpeithi server (<http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm>). The nanomer peptides were chosen for analysis. Subsequently the MHC-II restricted nanomer epitopes were predicted using NetMHCIIpan (<http://www.cbs.dtu.dk/services/NetMHCIIpan/>) interface. The epitopes predicted for this study used the default allele datasets of each prediction server. In that regard, the IEDB server uses special matrices for 108 MHC-I and 27 MHC-II (15 HLA-DR, 6 HLA-DQ, and 6 HLA-DP) alleles for prediction. The same allele sets from the IEDB allele database are employed by NetMHCIIpan as well. While the Syfpeithi prediction server has a dataset of 39 MHC-I alleles. After completion of the prediction of MHC-I restricted epitopes, the epitopes that were common for most of the alleles were selected. The same methodology was replicated for MHC-II epitopes.

### 2.1.4. Identification of common epitopes

The predicted B and T Cell epitopes (bind to secreted immunoglobulins) were then searched for their presence in the HC, RC and LC Clustals. In addition to checking their sequences in the HC, RC and LC Clustals, the selected B-cell and T-cell epitopes were also compared with one another for scanning the presence of any common epitopes between themselves.

### 2.1.5. Antigenicity, allergenicity, toxicity, anti-inflammatory response prediction and IFN gamma induction analysis of epitopes

The predicted B and T Cell epitope peptides were subjected to rigorous selection criteria. For that purpose, several prediction parameters were applied for validating their eligibility to be a part of vaccine construct. The selection parameters were antigenicity, allergenicity, toxicity, anti-inflammatory response prediction, and IFN gamma induction analysis. Furthermore, once the vaccine construct was obtained, it was also subjected to the same parameters for its evaluation. The antigenicity of the peptides was determined using Vaxijen v2.0 at threshold of 0.5 (<http://www.ddg-pharmfac.net/vaxijen/Vaxijen/Vaxijen.html>) and the allergenicity of the peptides was determined using AllerTOP v2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>) (Dimitrov et al., 2014). For determining the toxicity of the peptides, ToxinPred ([http://crdd.osdd.net/raghava/toxinpred/multi\\_submit.php](http://crdd.osdd.net/raghava/toxinpred/multi_submit.php)) was used (Ayyagari et al., 2020). The anti-inflammatory response prediction was carried out using AIPPred (<http://www.thegelelab.org/AIPpred/>) (Manavalan et al., 2018) and PreAIP (<http://kurata14.bio.kyutech.ac.jp/PreAIP/>) (Khatun et al., 2019). Finally, IFN gamma induction analysis was carried out using IFN epitope (<http://crdd.osdd.net/raghava/ifnepitope/predict.php>) (Dhanda et al., 2013). The default settings were applied for calculating the above properties on the individual peptides as well as on the entire vaccine construct.

### 2.1.6. Population coverage prediction of epitopes

The frequencies of HLA genotypes differ among different populations of the world. Therefore, the HLA polymorphism affects the binding capability of a specific epitope peptide chosen for the multivalent vaccine to HLA-I and II molecules (Bui et al., 2006). Consequently, it is crucial to determine the degree to which the world population will be covered by the multi-epitope vaccine construct

developed in this study. For this purpose, the IEDB tool (<http://tools.iedb.org/population/>) was employed to predict the population coverage of T cell.

## 2.2. Development of multivalent vaccine design construct

The candidate immunodominant epitopes obtained from the M segment of the CCHFV genome, were selected to form the multi-epitope vaccine construct. The human-defensin was added at the vaccine's N-terminus as an adjuvant, with the (EAAAK) linker serving as an adjuvant. The B Cell epitopes were connected to each other by a (KK) linker. The B Cell epitopes were spaced from HTL (MHC-II epitopes) by (GPGPG) linker. The GPGPG linker was also used to connect HTLs with each other. HTL were linked to CTL (MHC-I epitopes) by (AAV). The AAV Linker was also used to connect CTLs to each other. Finally, the addition of His-tag (six histamines) at the C-terminus to conclude the linking of the final vaccine construct (Shey et al., 2019).

## 2.3. Prediction of tertiary structure and validation of multivalent vaccine design construct

The following facets were considered for tertiary structure prediction and validation of multivalent vaccine polypeptide designed in this study.

### 2.3.1. Prediction and optimization of multivalent vaccine's secondary and tertiary structure

The vaccine sequence was submitted into PDBSum (<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>) for finding details of polypeptide's secondary structure (Laskowski et al., 2018). Subsequently, the prediction of tertiary structure of the protein was carried out using iTASSER (<https://zhanglab.cmb.med.umich.edu/I-TASSER/>) (Roy et al., 2010). The predicted model was refined using 3Drefine (<http://sysbio.rnet.missouri.edu/3Drefine/>) (Bhattacharya et al., 2016). (<https://zhanglab.cmb.med.umich.edu/ModRefiner/>), GalaxyRefine (<http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE>), ReFOLD (<https://www.reading.ac.uk/bioinf/ReFOLD/>) (Adiyaman and McGuffin, 2019) and Yasara Energy Minimization Server (<http://www.yasara.org/minimizationserver.htm>) respectively (Krieger et al., 2009). Further validation of the protein structure was obtained by Ramachandran Plot using the tool Molprobity (<http://molprobity.biochem.duke.edu/index.php?MolProbSID=71h5tp5jihjuri698nmt84d25ki1&eventID=62>) for identifying the number of residues in the favored, allowed and outlier regions (Chen et al., 2010).

The percentage of error comparisons in predicted residues structure of the unrefined compared to refined protein was checked using ERRAT (<https://saves.mbi.ucla.edu/>) (Colovos and Yeates, 1993). The validation of the refined tertiary structure after the refinement process was finalized using ProSA-web (a protein structure analysis tool) available at (<https://prosa.services.came.sbg.ac.at/prosa.php>) (Sippl, 1993).

Finally, the physicochemical properties of the vaccine such as molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, etc. were calculated using ProtParam (<https://web.expasy.org/protparam/>) (Gasteiger et al., 2003).



### 2.3.2. Prediction of Proteasomal, TAP binding protein and cell localization

The MHC-I ligands processing pathway mainly comprises of two steps – initial degradation of proteins by the proteasome, and the transport of the initial degraded products to the endoplasmic reticulum (ER) using the transporter associated with antigen processing (TAP). It is here in the ER, antigen peptides are bounded by MHC class I molecules, and later presented on the cell surface by MHCs (Craiu et al., 1997). Therefore NetChop (<http://www.cbs.dtu.dk/services/NetChop/>) was used for locating proteasomal cleavage sites (Doytchinova et al., 2006). Likewise, the binding affinity of the multivalent vaccine with the TAP was predicted by the tool TAP-REG (<http://imed.med.ucm.es/Tools/tapreg/>). Furthermore, the subcellular localization in eukaryotes of the vaccine polypeptide was predicted by employing the BaCelLo (Sette et al., 1994) (<http://gpcr.biocomp.unibo.it/bacello/>) (Pierleoni et al., 2006) and Phobius (<https://phobius.sbc.su.se/index.html>) (Kotturi et al., 2007).

### 2.3.3. Immune simulation

For immune simulation analysis, C-IMMSIM (<http://150.146.2.1/C-IMMSIM/index.php?page=1>, accessed on January 13, 2021) was used to simulate the immune response of host to the designed vaccine. This tool employs Miyazawa–Jernigan residue–residue potential for scoring the force with which a given peptide–HLA complex interacts with a T-cell receptor (Rapin et al., 2010). The tool was used on default settings, except in the host HLA selection option, the heterozygous combination of HLA-A (HLA-A\*02:01, HLA-A\*30:01), HLA-B (HLAB\*15:01, HLA-B\*13:01) and HLA-II (DRB1\_0104, DRB1\_0107) were randomly selected, from the set of alleles which were also used for the construction of multivalent vaccine peptides.

### 2.3.4. Codon adaptation and in-silico cloning

The vaccine sequence obtained from this study was fed into JCat (<http://www.icat.de/Start.jsp>). For the creation of the optimized DNA sequence, the organism selected for codon-adaptation was *Escherichia coli* (Strain K12). In addition to choosing an organism, other options that were selected from the tool's interface were (i) avoid rho-independent transcription terminators, (ii) avoid prokaryotic ribosome binding sites, and (iii) avoid cleavage sites of restriction enzymes (Grote et al., 2005). These three options were selected for exclusion of these regions from being optimized while undergoing codon adaptation algorithm. As rho-independent transcription terminators consists of RNA structures which are GC-rich that results in formation of stem-loops followed by a series of uridine residues, which correspondingly forms RNA hairpins. The prokaryotic and eukaryotic RNAs are abundant in these loops, which contribute to enhance the stability of RNA hairpins. The RNA hairpins facilitate the nucleation of the nascent RNA structures, which allows transcription termination process to take place. Therefore, by selecting the option, “avoid rho-independent transcription terminators”, uncontrollable transcription is avoided (d Aubenton Carafa et al., 1990). Similarly, options, “avoid prokaryotic ribosome binding sites” and “avoid cleavage sites of restriction enzymes” were selected to keep nucleotide residues unchanged even after codon optimization, to get optimal cloning and expression processes downstream (Grote et al., 2005). The DNA sequence of multivalent vaccine polypeptide obtained after the codon optimization was subjected to in-silico cloning into pDual-GC vector by using SnapGene tool (<https://www.snapgene.com/>). The pDual-GC vector was used because of its capability of heterologous gene expression in both mammalian and prokaryotic systems (Mullinax et al., 2003).

### 2.3.5. Post-translational modifications

In this study, the final vaccine construct is designed using a dual expression system, therefore, for eukaryotic (mammalian) system, the expressed vaccine polypeptide must be examined for the possible post translational modifications. The typically investigated post-translational modifications are di-sulfide engineering, N-glycosylation, O-glycosylation, and ubiquitination. They have also shown to increase the stability of extracellular and secreted protein structures following translation (Craig and Dombkowski, 2013; Lee et al., 2015; Schmidt et al., 2017; Su and Lau, 2009). Therefore, for the di-sulfide engineering the Disulfide By Design 2 (<http://cptweb.cpt.wayne.edu/DbD2/index.php>) (Craig and Dombkowski, 2013), for N-glycosylation (<http://www.cbs.dtu.dk/services/NetNGlyc/>) (Blom et al., 2004), for O-glycosylation (<http://www.cbs.dtu.dk/services/NetOGlyc/>) (Steentoft et al., 2013), for Ubiquitination (<http://www.jci-bioinfo.cn/iUbiq-Lys/>) (Radivojac et al., 2010) and for phosphorylation (<http://www.cbs.dtu.dk/services/NetPhos/>) (Blom et al., 1999) were used.

### 2.3.6. Molecular docking

The immune receptors such as Toll Like Receptors (TLRs) are essential in evoking immune responses against infection by pathogens. The TLR's implicated in eliciting immune responses in humans during CCHFV infection are TLR-3, TLR-8 and TLR-9 (Lindquist et al., 2018). The protein sequences of TLR-3 (PDB ID: 5GS0), TLR-8 (PDB ID: 4R08) and TLR-9 (UniProt ID: Q9NR96) were obtained from RCSB database (<https://www.rcsb.org/structure/4R08>). The docking was performed on TLR-3, TLR-8 and TLR-9 individually, using ClusPro (<https://cluspro.bu.edu/login.php>) (Desta et al., 2020; Kozakov et al., 2017, 2013; Vajda et al., 2017). The docked structures were refined using PatchDock (<https://bioinfo3d.cs.tau.ac.il/PatchDock/php.php>) (Mashiach et al., 2010) and Firedock (<http://bioinfo3d.cs.tau.ac.il/FireDock/php.php>) (Mashiach et al., 2008). The docked output was analyzed in PyMOL molecular graphics system (Zhu et al., 2014).

### 2.3.7. Molecular dynamics analysis of docked structures

The behavior of constructed vaccine with TLR3, TLR8 and TLR9 systems was studied using a detailed molecular dynamics simulation. For MD simulation, the amber14 kit was used (Salomon-Ferrer et al., 2013). The ff14SB force field was used to define both TLRs and vaccines. In the present study, 100 × 3 ns simulations were performed. A rectangular box of TIP3P water molecules was used to solvate the systems. The buffer distance in all rectangular box were kept 12 Å. Counter ions were used to neutralize each system. The steepest descent and conjugate gradient methods were applied to relax each system by removing bad atoms contacts. Subsequently each system was heated up to 300 K. Subsequently, a two-step method was employed (by keeping pressure and temperature constant at 1 atm and 300 K respectively) for balancing each system. Firstly, the equilibration was carried out on the density using weak restraint for 2 ns. Secondly, the system without any restraint for more 2 ns was subjected to equilibration. Finally, each system was subjected to a long run production simulation. The temperature of each system was controlled by the Langevin thermostat (Zwanzig, 1973). The long range electrostatic interactions were treated with Particle Mesh Ewald algorithm (Darden et al., 1993; Ml and Lg, 1995). The SHAKE algorithm was employed for covalent bonds (Ml and Lg, 1995) The construction step of molecular dynamic simulation for each system was performed using the pmemd code (Götz et al., 2012), and the analysis of the trajectories was accomplished via cpptraj package and g\_sham module of gromacs.

**Table 1**  
Details of the selected epitopes and their properties.

Epitope type	Peptide	Start	End	% Homology	Vaxijen Score; ≥ 0.5	Allergenicity	IFN $\gamma$ induction	Toxicity	Anti-inflammatory Response	Total no. of aminoacids	MW	pI
B-Cell	GRSESIMKLEERTGISWDLGVEDASESKLL	1133	1104	40	0.7692	Non-Allergen	+	Non-Toxic	High AIP	30	3336.72	4.49
	QSTYKPTVSTANIALSWSSVEHRGN	1074	1098	48.28	0.8463	Non-Allergen	+	Non-Toxic	High AIP	25	2733.98	8.6
	IFLFLAPFILLILFFMFGWR	1599	1618	39.47	0.6527	Non-Allergen	+	Non-Toxic	High AIP	20	2505.19	9.75
	ISWDLGVEDASESKLITVSV	1118	1137	41.38	0.8403	Non-Allergen	+	Non-Toxic	High AIP	20	2148.39	3.92
	DTEGLLEWCKRNKRLGLDCTD	279	298	46.43	0.50	Non-Allergen	+	Non-Toxic	High AIP	20	2296.51	3.96
T-Cell (MHC-I)	SLMARKLEF	1493	1501	85.71	1.6461	Non-Allergen	+	Non-Toxic	High AIP	9	1067.37	5.52
	LAVCKRMCF	950	958	75	1.2425	Non-Allergen	-	Non-Toxic	High AIP	9	1067.37	5.27
	MAFIWFSE	691	699	70	0.7732	Non-Allergen	+	Non-Toxic	High AIP	9	1195.44	5.28
T-Cell (MHC-II)	LMYAILCLQ	5	13	85.71	1.3232	Non-Allergen	+	Non-Toxic	High AIP	9	1094.34	8.46
	MYAILCLQ	6	14	85.71	1.5426	Non-Allergen	+	Non-Toxic	High AIP	9	1070.39	8.96

**Table 2**

The population coverage of selected epitopes.

MHC Class	Coverage	Average Hit	PC90
Combined	98.47	4.74	2.6

### 3. Results

#### 3.1. Epitope selection for multivalent vaccine design

There were a total of twenty-nine Clustals found from the sequences of M Segment of the CCHFV after its input into the Clustal Omega tool. There were ten HC and RC Clustals and nine LC Clustals discovered. Subsequently, the prediction of B-Cell epitopes was performed by ABCPred, Bepipred & IEDB. While the T-Cell epitopes were obtained using IEDB and Syphpeithi (MHC-I) and NetMHCIIpan (MHC-II). For each IEDB, NetMHCIIpan and Syfpeithi, almost 1700 peptides were predicted for each prediction server. Therefore, the significant peptides were selected on the basis of IC50 (if < 50 nM then highest affinity of binding) and top 1% (have highest affinity of binding) of the predicted peptides (Kotturi et al., 2007; Sette et al., 1994). The peptides common for majority of the alleles were selected for both B and T-Cell epitope prediction. Subsequently, the epitope selection from common peptides was carried out by applying rigorous criteria. This included determination of the %age homology, antigenicity, allergenicity, toxicity, anti-inflammatory response, IFN gamma induction, population coverage and physicochemical properties of epitope. Ultimately, after passing through these parameters, five B-cell, three MHC-I and two MHC-II epitopes were obtained - Table 1, that will bind to secreted immunoglobulins. The population coverage analysis was performed on all the B and T-cell epitopes selected for the construction of the vaccine indicated projected population coverage was 98.47%, in world's population when MHC Class's combined (HLA-I & II) option was selected. While average number of epitope hits or HLA sets of combinations recognized by the population was 4.74 and the minimum number of epitope hits or HLA combinations recognized by 90% of the population (PC90) was 2.6, Table 2. The projected percentage of population coverage was explained graphically in (Fig. 1).

#### 3.2. Physicochemical properties prediction of multivalent vaccine construct

The Physicochemical properties of the full construct was also evaluated, which is summarized in Table 3, indicating the final vaccine construct consists of 243 amino acids, having molecular weight of 27462.29 g/mol, pI of 9.3, IFN  $\gamma$  inducing and is non-toxic. The half-life of the vaccine varies in different invitro systems (Table 3).

#### 3.3. Multivalent vaccine design construct

The final vaccine construct consisted of 243 amino acids with a total of five B-cell epitopes followed by two MHC-II (HTL) epitopes and three MHC-I (CTL) epitopes, separated by respective linkers. Besides these, the N-terminal end of the vaccine has the human  $\beta$ -defensin as adjuvant for increasing the immunogenicity and C-terminal added with six histidine tag (Shey et al., 2019). Furthermore, the use of the EAAK linker in the multivalent vaccine design is due to its capacity to be a rigid linker, as rigid linkers have demonstrated several advantages over flexible linkers. One of the main advantages is the preservation of efficient separation of its functional domains, is to maintain its specific functional properties (a precise distance was maintained with least interference) from

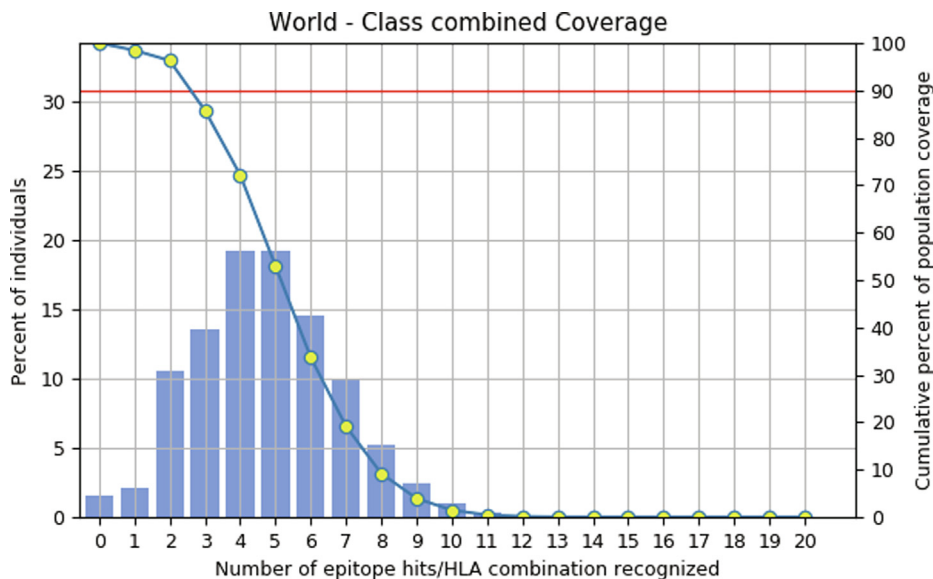


Fig. 1. Percent class combined coverage of the selected epitopes in the world's population.

Table 3 The physicochemical properties of the vaccine.

Toxicity	IFN $\gamma$ induction	Total no. of amino acids	MW (g/mol)	pI	Formula	Total no. of atoms	Approx 1/2 life		
							Mamalian Cell (Invitro)	Yeast (Invivo)	E. coli (Invivo)
Nontoxic	+	243	27462.29	9.3	C <sub>1238</sub> H <sub>1950</sub> N <sub>336</sub> O <sub>332</sub> S <sub>19</sub>	3875	>30 h	>20 h	>10 h

the epitopes (Chen et al., 2013). For the development of a bifunctional fusion protein, it is extremely essential to achieve the effective separation of the domains in their structures (Arai et al., 2001). Subsequently, the KK linker – a Lysine linker, was used to bind the B-Cell epitopes (Gu et al., 2017; Nain et al., 2020; Sarkar et al., 2020). The KK linker was chosen because: (i) It is a target for lysosomal proteases like the Cathepsin B (which is involved in the peptide presentation on the surface of the cell – via a process called antigen processing of the peptides occurring in the MHC-II restricted antigen presentation) (Yano et al., 2005). (ii) It reduces the junctional immunogenicity by circumventing the antibodies induction. Since, the antibodies can be induced for the linear peptides (Yano et al., 2005). (iii) The KK linkers intensify the immunogenic responses (Li et al., 2015). Additionally, a universal spacer which is the GPGPG linker was used to attach the HLA-II epitopes. As, a study performed on mice model demonstrated the ability of GPGPG linker to induce TH lymphocyte (HTL) as well as its capability of splitting the epitopes by maintaining a boundary (junctional immunogenicity), which restores the immunogenicity of the individual epitopes (Livingston et al., 2002). Correspondingly, the HLA-I epitopes were linked via AAY (Ala-Ala-Tyr) linker. Since,

the AAY linker serves as a cleavage site for the proteasomes in mammalian cells. Therefore, effective junctional immunogenicity is maintained between the epitopes using AAY as a linker (Bhatnager et al., 2021). The AAY linker also contributes to increasing the immunogenicity of the multivalent vaccine (Yang et al., 2015). The order and sequence of epitopes of multivalent vaccine construct (Fig. 2).

### 3.4. Prediction, refinement and validation of multivalent vaccine's tertiary structure

The secondary structure prediction of multivalent vaccine construct by PDBsum was first predicted. The output indicated that it contained 9 helices, 17 helix-helix interactions, 18 beta turns and 5 gamma turns (Figure S2).

While the prediction of the tertiary structure of the multivalent vaccine construct was carried out by I-TASSER, which generated top five final models. The Confidence score (C-score) defines the merit of each of the predicted models. Since a higher C-score is the reflection of higher confidence in the generated model. In addition to C-score, I-TASSER calculated TM-score and RMSD value for

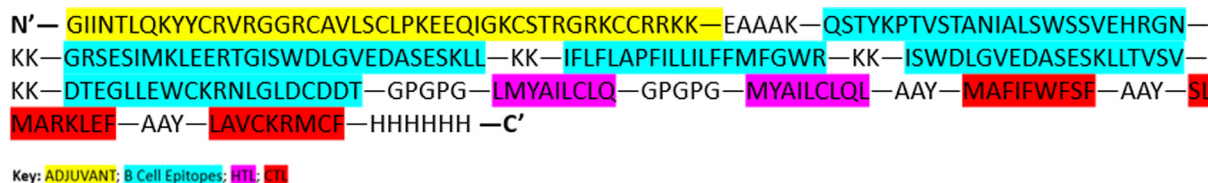


Fig. 2. The sequence of multivalent vaccine construct marked. (Yellow colored residues:  $\beta$ defensin; Cyan colored residues: B cell epitopes; Magenta colored residues: Helper T-cell epitopes; Red colored residues: Cytotoxic cell epitopes).

all the five predicted models, that contributed to overall quality and selection of the model. The tertiary model selected in this study had C-score =  $-3.59$ ; TM-score =  $0.32 \pm 0.11$  and estimated RMSD =  $14.4 \pm 3.7$  Å. Afterward this selected tertiary model obtained from I-TASSER was submitted into 3Drefine. Just like I-TASSER, 3Drefine also generates top 5 models. The 3Drefine Scoring works on the principal of calculating the potential energy of the refined model according to 3Drefine force field. Therefore, the model with better quality carries a lower potential energy score. For additional refinement 3Drefine also used GDT-TS, GDT-HA (both of which are a similarity score in [0,1] co-ordinates of the refined model vs the initial model. Higher the score, higher will be the conservative refinement), RMSD (CA-RMS – a deviation score measured in angstrom (Å) of the refined model vs the initial model. Higher the score, higher will be the aggressive refinement), RWplus (indicates the potential energy of the refined model as per RWplus statistical potential. Lower the score, higher will be the quality of the model) and MolProbity (indicates the physical realism of the refined model in the form of MolProbity score. Lower the score, higher the physically realistic model will be). The tertiary model selected after running 3Drefine for this study had 3Drefine Score = 15970.5; GDT-TS = 0.9979; GDT-HA = 0.9537; RMSD (Å) = 0.420; RWplus = -50525.167242 and MolProbity = 3.675.

The subsequent further refinements were carried out on the generated refined model by using ModRefiner, GalaxyRefine, Yasara Energy Minimization Server and ReFOLD. Ultimately, the refinements were validated by Ramachandran analysis available at Molprobity, ERRAT and ProSA-web plots (Fig. 3).

### 3.5. Prediction of cleavage sites, TAP binding peptides and cell localization

The number of possible cleavage sites predicted by NetChop tool in 243 amino acid long vaccine polypeptide were 87. Similarly, the TAP binding affinity of multivalent vaccine polypeptide was predicted by TAPREG server. This prediction yielded more than 65 TAP binding peptides in the vaccine construct, out of which first three were considered since lower the IC50 value, higher is its binding affinity (Table 4). Furthermore, the subcellular localization of the vaccine polypeptide was determined using BaCelLo (a server for eukaryotes) indicating it to be cytoplasmic. While the details of a combined transmembrane topology and signal peptide were observed via Phobius (eukaryotes). The output indicated the protein to be cytoplasmic with transmembrane domains (Figure S3). The transmembrane domains (TMD's) embeds inside the protein core because they are hydrophobic. Therefore, our vaccine construct demonstrating the hydrophobic TMD's is a very good sign. Since, both signal peptides and TMD's have exceptionally high epitope densities. This fact has been corroborated by previous studies where it was demonstrated that improved MHC binding of TMD domains is attributed to their hydrophobic nature (Kovjazin et al., 2011). This can be further explained in connection with MHC-I molecules and their role in antigen presentation using TAP - proteasomal pathway. For a predicted peptide to work as a good CTL epitope, it must be cleaved at specific regions. Now this cleaved peptide must be transferred to the ER and later bind with sufficient affinity to MHC-I molecules. That is why in eukaryotic systems MHC-I is loaded with peptide repertoire of its cells. Since, the TAP- proteasome pathway is fundamental to the antigen presentation machinery of eukaryotic cells. As it is involved in the breakdown of all intracellular proteins into short peptides. While the TAP itself is a membrane-traversing heterodimer and have preference for transporting the degraded peptides based on their (Kovjazin et al., 2011). Our vaccine demonstrating hydrophobic domains is indicative of its capability to have good CTL epitopes and strong MHC-I presentation.

### 3.6. Codon adaptation and in-silico cloning

The Codon adaptation analysis is essential for the expression of the inserted foreign genes. As the bulk of the synthetic genes (for obtaining the required proteins) are expressed in prokaryotic systems. Therefore, their codon optimization is highly crucial. If the foreign gene possesses codons, which are very rarely employed by the host, it will result in its expression level being extremely minimal. This is one of the drawbacks of heterologous protein expression (Puigbò et al., 2007). In this study, the codon optimization was accomplished by using the JCat tool (Figure S4). Furthermore, the optimized DNA sequence obtained from JCat was cloned into the pDual-GC plasmid using SnapGene (Figure S5). The pDual-GC vector has dual nature and can be expressed both in bacterial and mammalian system. Furthermore, it was used because of the presence of a constitutive promoter (which is a CMV promoter/enhancer slightly altered by mutation) for expression in mammalian cells. While the inducible gene expression in prokaryotes is controlled via hybrid T7/lacO promoters (Mullinax et al., 2003).

### 3.7. Immune-simulation

The C-IMMSIM interface was used for simulation of immune response of designed vaccine construct. The immune responses for vaccine with and without adjuvant (not shown) was predicted. The prediction indicated better response for vaccine with adjuvant. Several immune responses such as IgM and IgG production, population level of B-cell and TH-cell (Helper), TC (Cytotoxic) as well as concentration of Cytokines were observed after the injection of vaccine. The immune simulation output established that the designed vaccine triggered sufficient innate and adaptive immune responses (Fig. 4).

### 3.8. Molecular docking and structural stability of vaccines with TLR proteins

#### 3.8.1. Docking refinements and dynamic stability of all systems

Initial docking prediction was performed using ClusPro on TLR-3, TLR-8 and TLR-9 respectively. Then the docked structures were refined using PatchDock. Finally, the top 10 models of each TLR's obtained from PatchDock were further refined and ranked by FireDock, which allows a high-throughput refinement of docking candidates. The best docked complexes of vaccine – TLR-3, Vaccine – TLR-8 and Vaccine – TLR-9 obtained from FireDock were selected (Fig. 5(a)) The major selection criterion was observing the Global Energy of the docked complex as well as other details. Lowest global energy complex was ranked highest in FireDock (Adiyaman and McGuffin, 2019). The details of top ranked Patchdock models chosen by Firedock after refining and their energies are shown in (Table 5).

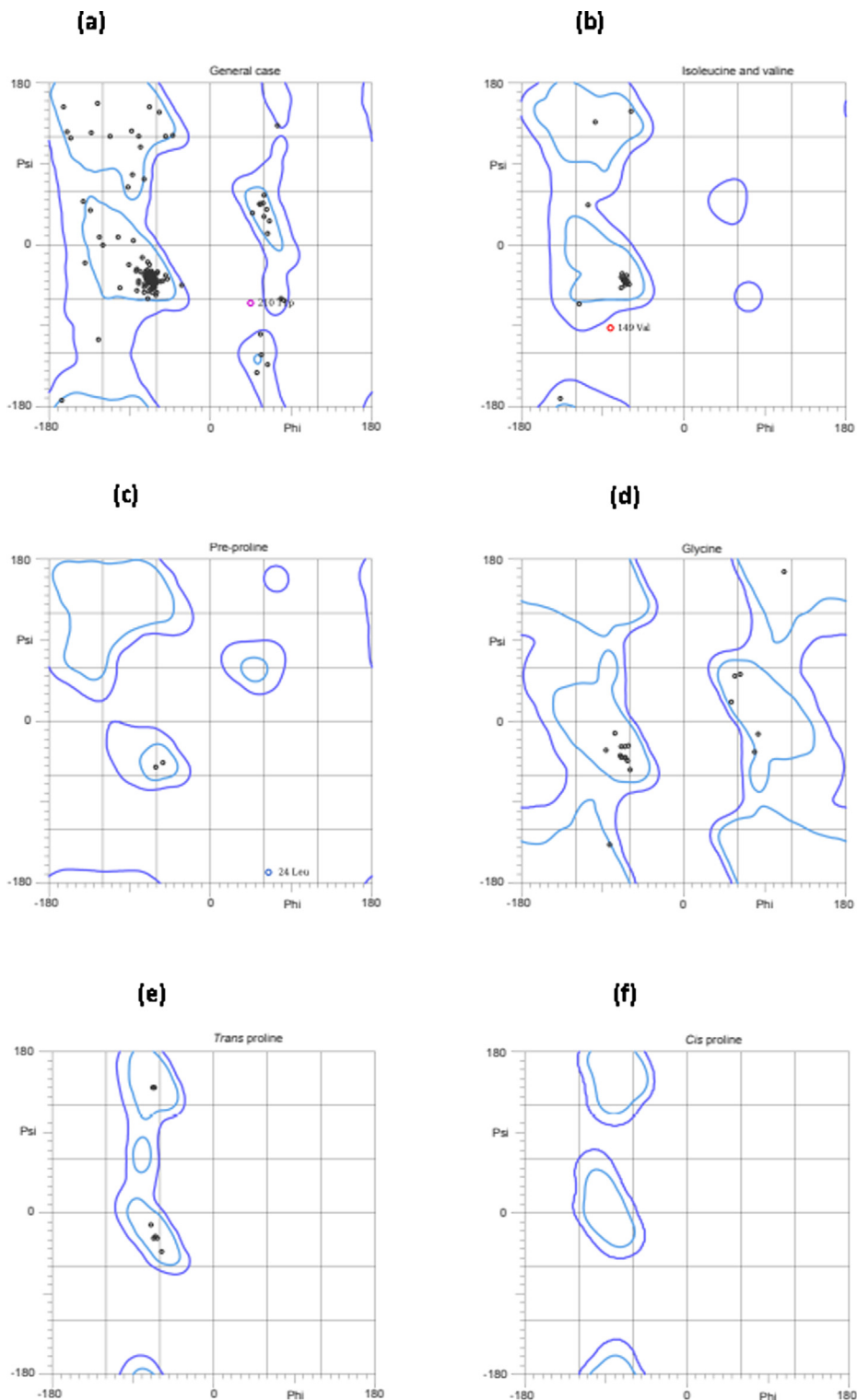
The best complex from each system were subjected to molecular dynamics simulation. Root-mean-square deviations (RMSD) of backbone atoms of all systems were computed with respect to their original configurations in MD simulation to analyze the dynamic stability of all systems (Fig. 5b(i)). The graph shows an increase in RMSD up to 60 ns and then remain constant throughout the simulation with slight fluctuations. During the period of 0 ns to 60 ns, the RMSD raises up to 3.5 Å. It implies that each system converged to their stable form in the first 60 ns time. The root mean square fluctuation (RMSF) was calculated for checking the stability at residue level. The RMSF graph of TLR3-vaccine, TLR8-vaccine and TLR9-vaccine are shown in (Fig. 5b(ii)). The graph shows that the TLR3, TLR8 and TLR9 showed stable behavior along with their vaccine binding partner. In case of TLR3, the C-terminal show higher fluctuations which reflect the adaptation of TLR3 for the binding of its vaccine partner.



3.8.2. Stability at binding region and binding interaction analysis

The stability of each vaccine construct in the binding region of each TLR protein was analyzed by calculating the average distance between each TLR and vaccine (Fig. 5c(i)). In case of TLR3, the initial average distance between TLR3 and vaccine was around 23 Å.

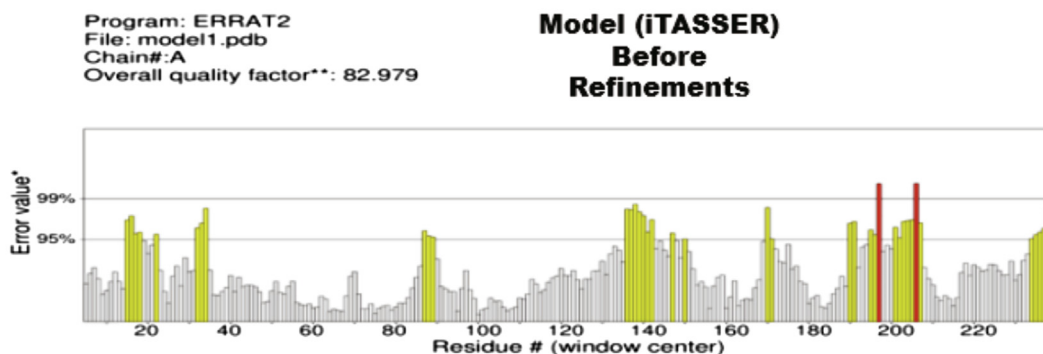
The average distance increases up to 27 Å in the first 20 ns and then reduces to initial value in the rest of simulation. In case of TLR8 and TLR9, the distance showed slight fluctuation around 30 Å. The frequency graph shows that the peak value for TLR3 is at 24 Å (Fig. 5c(ii)). It implies that most of the time the average dis-



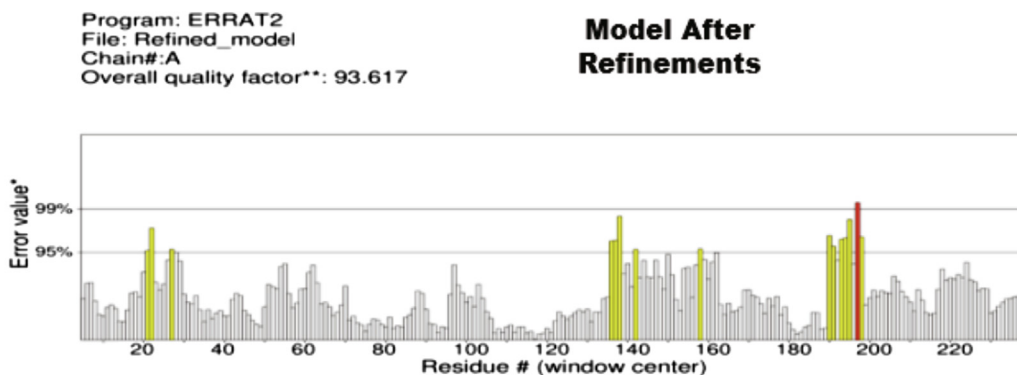
**Fig. 3.** Refinement and validation of predicted tertiary structure of vaccine. (a) to (f) Ramachandran plots for various residues indicating 92.9% of all residues were in favored region and 98.8% of all residues were in allowed regions. The 3 outliers in (phi, psi) region indicated were (24 Leu), (149 Val) and (210 Trp). (g) ERRAT plot of iTASSER generated model without refinement having 82.9% accuracy vs (h) ERRAT plot of vaccine model after series of refinements having 93.67% accuracy. (i) The Z-Score of the vaccine determined using ProSA-web server which is -5.44. (j) The refined 3D structure of the vaccine.



(g)

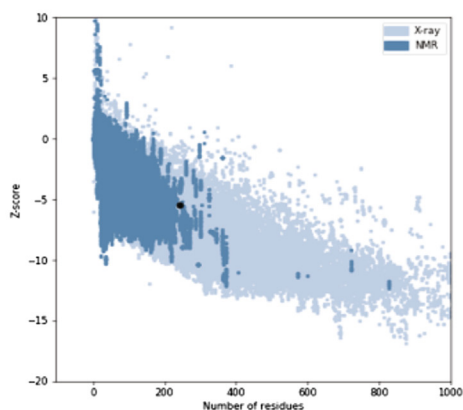


(h)



(i)

Z-Score: -5.44



(j)

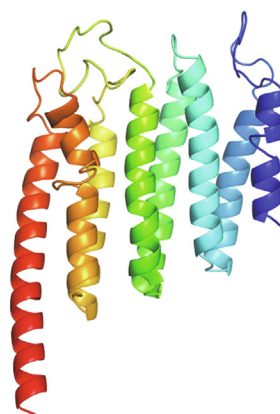


Fig. 3 (continued)

Table 4

Excerpt of MAPREG Output Table Indicating Peptides with Lowest IC50 Value (highest TAP binding affinity).

Peptide	Position	Tap affinityIC50 (nm)
AYMAFIFWF	202	77.61
AAYSLMARK	213	80.39
AAYMAFIFW	201	88.77

tance between TLR3 and vaccine is 24 Å. In case of the TLR8 and TLR9, the peak value was found as 31 Å and 31.3 Å. All these results showed that our constructed vaccine bind well and remain stable in binding region.

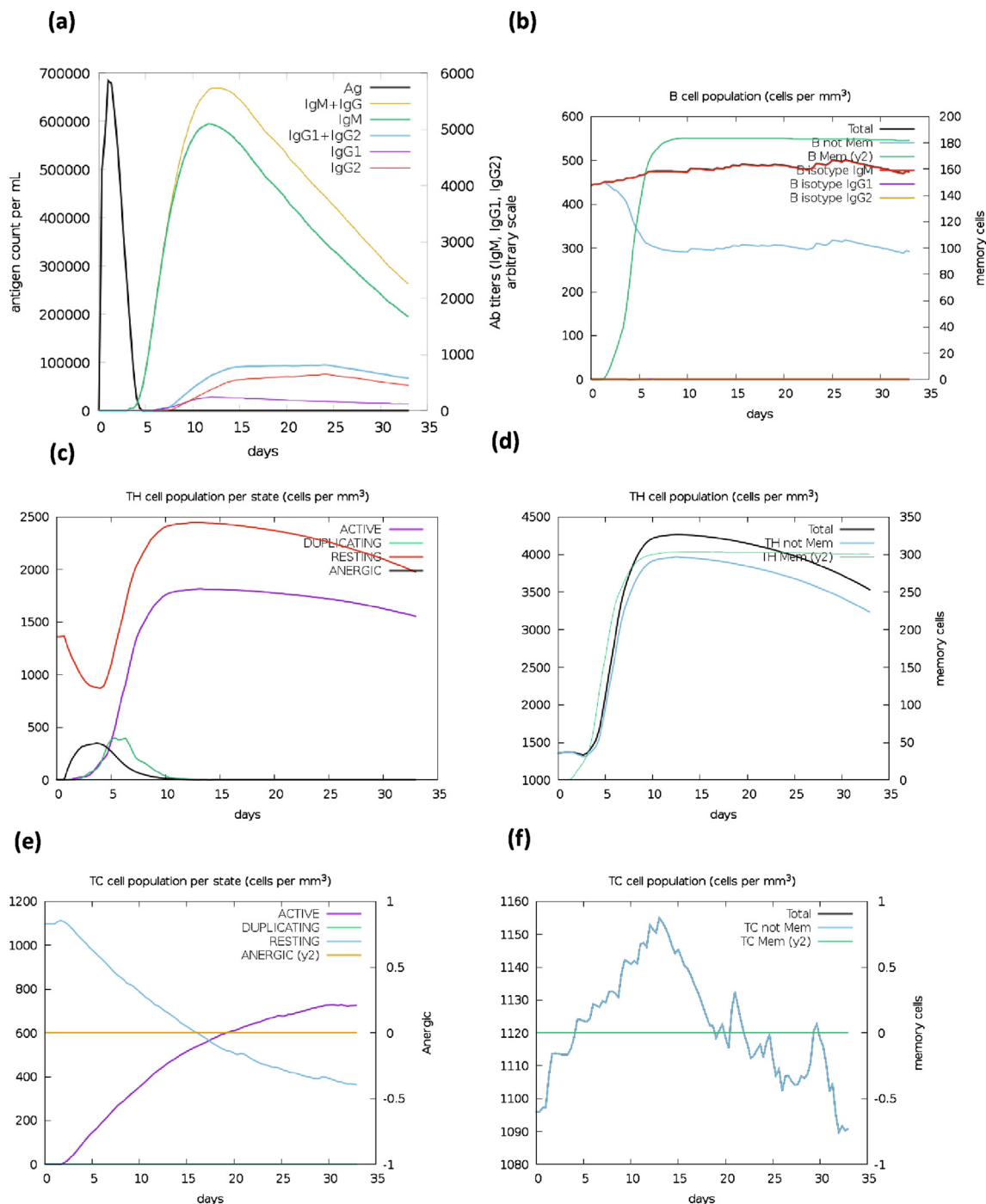
On the other hand, the binding interaction analysis were carried out to gain in-depth insight of binding. First, the representative structure for each system, TLR3, TLR8 and TLR9, were extracted through free energy landscape analysis (FEL). For FEL, the principal components analysis (PCA) was carried out. Based on the principal

components 1 (PC1) and principal components 2 (PC2), the FEL graph were constructed (Fig. 5(d)). The dark region in the graph shows the lowest energy conformer. The structural coordinates were extracted from the dark regions. The respective coordinates for TLR3, TLR8 and TLR9 were analyzed for binding interactions. In case of TLR3, 10 strong hydrogen bonds were observed (Fig. 5d(i)). The strongest hydrogen with -32 kcal were found between the residue of TLR3 Arg462 and vaccine His243. The distance of this hydrogen bond is 2.9 Å. In case of TLR8 the strongest bond has energy -33.5 kcal with bond distance 2.96 Å – Fig. 5d(ii). The strongest bond in case of TLR9 were found between Glu812 of

TLR9 and Lys233 of vaccine (Fig. 5d(iii)). The binding energy of all the hydrogen bonds in TLR3, TLR8 and TLR9 range from -33 to -0.7 kcal. This result implies that all the constructed vaccine will bind in their respective TLR proteins.

### 3.9. Post-translational modifications

In this study vaccine polypeptide was subjected to both prokaryotic and mammalian expression because of employing pDual-GC plasmid for expression. Therefore, post translational modifications which are much more common in nucleated cells,



**Fig. 4.** *In-silico* immune simulation output obtained from C-IMMSIM after the injection of vaccine. (a) Various types of immunoglobulin production triggered by the virus antigen and formation of immunocomplexes. (b) Elevated B-cell population, (c) Elevated active TH cell population per state, (d) Elevated total T-Helper cell population, (e) Elevated population of active TC lymphocytes count per entity-state. (f) Increase in TC cell population. (g) Increase in concentration of cytokines and interleukins. Inset plot shows danger signal together with high production of Leukocyte Growth Factor IL-2.

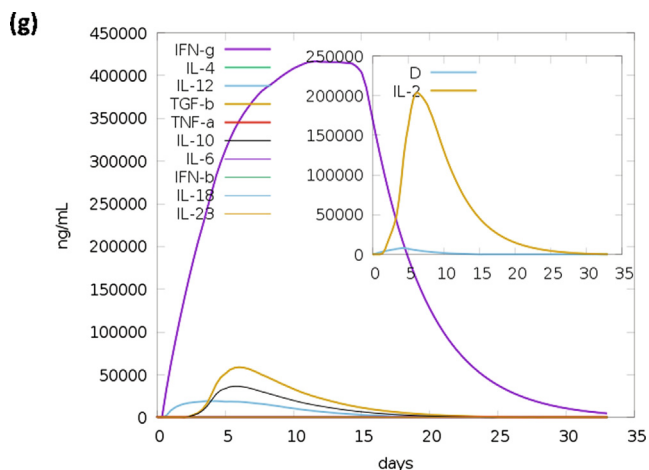


Fig. 4 (continued)

must be observed for the vaccine protein produced. Consequentially, introduction of disulfide bonds, N and O-glycosylation, Ubiquitination and Phosphorylation were some modifications employed to the expressed vaccine. The introduction of novel disulfide bonds into the designed proteins is an important biotechnological advancement. It has been used extensively for, conferring increased protein stability, modification of functional characteristics of the protein, and assisting in the analyses of protein dynamics. The disulfide bonds work on the principle of lowering the entropy (caused due to conformational changes) and increasing the free energy (possessed by denatured state). Therefore, increasing the stability of the folded protein conformation. The disulfide bonds were engineered (using Disulfide by Design 2). This software works by predicting pairs of amino acids that when mutated to cysteines, will most likely form a disulfide bond (Craig and Dombkowski, 2013). The native protein vs mutated vaccine protein obtained by Disulfide by Design 2 (Fig. 6).

The N-glycosylation using NetNGlyc and O-glycosylation using NetOGlyc prediction yielded no potential N-glycosylation and O-glycosylation sites in the vaccine sequence respectively. Similarly, the Ubiquitination prediction by employing iUbiq-Lys yielded no ubiquitination sites as well. While the phosphorylation prediction carried out by NetPhos at 36 sites yielded 17 phosphorylation sites that surpassed the threshold value of 0.5 (Fig. 7).

#### 4. Discussion

The current SARS-CoV-2 pandemic has reminded us of a valuable lesson that vaccines are one of the most potent and economical methods to prevent and control disease outbreaks. A lethal disease such as CCHF, with the unavailability of specific prophylaxis or vaccines, especially when its outbreaks are exponentially surging throughout the globe, demands immediate attention (Formenty, n.d.; World Health Organization, n.d.). In the past Soviet Vaccine was developed in 1970 (Papa et al., 2011). Presently very few vaccines are under clinical trials, but they are only addressing a specific genotype endemic to the respective regions (Centers for Disease Control and Prevention et al., 2003; Dowall et al., 2017). Such as a clinical trial conducted in *Cynomolgus macaque* model of Chinese origin (Hawman et al., 2021) and clinical trials on lambs in Bulgaria (Limon-Vega/The Pirbright Institute and Lyons/The Pirbright Institute, 2019). This study has designed a sub-unit vaccine targeting all the seven genotypes of CCHFV computationally. In this process, a highly scrutinized antigen screening and selection criteria was employed, which led to the identification of 10 epitopes used in the construction of this vaccine (Table 1).

Out of these selected epitopes, a total of five B-Cell epitopes had three epitopes (GRSESIMKLEERTGISWDLGVEDASESKLL; STYKPTV-STANIALSWSSVEHRGN; ISWDLGVEDASESKLLTVSV) in highly conserved and two epitopes (IFLFLAPFILLILFFMFGWR; DTEGLLEWC KRNGLGDCDDT) in relatively conserved region. Whereas, from selected five T-Cell epitopes only one (SLMARKLEF) was found in highly conserved, two (LAVCKRMCF; MAFIFWFS) were in relatively conserved and two (LMYAILCLQ LC; MYAILCLQL) were in least conserved region. While comparison of B-Cell and T-Cell epitopes with each other yielded no common epitopes. Furthermore, the population coverage analysis performed on these epitopes revealed, 98.47% coverage in world's population (HLA-I & II combined). The results propound that our 10 epitopes carrying sub-unit vaccine can cover a very wide range of human population.

The final multivalent vaccine construct consists of 243 amino acids containing the adjuvant and linkers in addition to the B and T-Cell epitopes (Fig. 4). Accordingly, the immunogenicity potential of this multivalent vaccine was increased by linking the  $\beta$ -defensin molecule at the N-terminal end. Furthermore, the addition of  $\beta$ -defensin in our proposed vaccine construct is an unprecedented and distinct aspect of our study. The EAAK linker was not only employed as an adjuvant for boosting its immunogenicity but also as a rigid linker that can perform effective separation (Chen et al., 2013). Subsequently, the KK linker was used to link the B-Cell epitopes (Gu et al., 2017; Nain et al., 2020; Sarkar et al., 2020). The universal spacer which is the GPGPG linker (Li et al., 2015) was used to join the HLA-II epitopes. While for the HLA-I epitopes the AAY linker was used. The physicochemical properties of the multivalent vaccine polypeptide were also predicted. This analysis found the designed vaccine to be non-toxic in nature with a molecular weight 27462.29 g/mol and pI 9.3 respectively. The half-life was estimated to be of 30 h (for mammalian reticulocytes – *in vitro*), >20 h (for yeast – *in vivo*), and >10 h (for *E. coli* – *in vivo*). The computation of the instability index (II) was found to be 47.73 while the calculated aliphatic index was 86.38. The instability index (II) describes *in vitro* stability of peptide, which was improved by di-sulfide engineering. Additionally, the grand average of hydropathicity (GRAVY) was found to be  $-0.064$  with the total number of negatively and positively charged residues were 22 and 36 respectively. The overall results obtained, indicates that our vaccine is good candidate for it to be used for further analysis.

The tertiary structure of the multivalent vaccine construct was predicted by I-TASSER, which generated top five final models. The selected model had the highest Confidence score (C-score) of  $-3.59$ . This model was submitted into 3Drefine for further refinement. The 3Drefine also generated top 5 models. The models were ranked, based on 3Drefine Score which is the indicator of the potential energy of the refined model. Therefore, the model possessing 3Drefine Score of 15970.5 was selected because lower score indicates the better quality of the model. After the additional refinements, the model was validated by Ramachandran analysis available at Molprobit, ERRAT and ProSA-web plots. The Ramachandran analysis revealed that 92.9% of all the residues lied in favored region and 98.8% of all the residues were in allowed regions. These values validated that the overall quality of the refined tertiary structure of the vaccine polypeptide's to be quite good. Similarly, the ERRAT plot on pre and post refinements indicated the improvement in quality of the structure as well. The model before refinement has ramachandran favored residues of 82.9%, while the model after refinements improved to 93.6%. Likewise, the ProSA-web validation was carried out. The quality of the protein was shown by the z-score. The z-score of the refined tertiary structure was found to be  $-5.44$  which also corroborated to the good quality of the tertiary structure of our multivalent vaccine (Fig. 3).

The refined and validated tertiary structure of multivalent vaccine construct was then processed for docking with TLR-3, TLR-8 and TLR-9 respectively. Initial docking predictions were carried out using ClusPro. The obtained output was then refined by PatchDock. Subsequently, all the top 1000 transformations obtained from docking the vaccine with TLR-3, TLR-8 and TLR-9 in PatchDock were input into FireDock. The top-most ranked, refined docked complex by FireDock for all the three TLR receptors was chosen. The transformation numbers obtained from Patchdock, that were ranked number 1 by Firedock after refinement, of docked complexes of vaccine with TLR-3, TLR-8 and TLR-9 are 599, 354 and 162 respectively. These transformation models were refined based on the parameters, such as: Global Energy (principal parameter), Attractive and Repulsive Van der Waal Forces, Atomic Contact Energy and Hydrogen and Disulfide Bonds. The details of these

energy parameters are shown in (Table 5) which consolidates that the selected vaccine:TLR complexes are very stable. The stability at binding region in MD simulations analysis demonstrated that our constructed vaccine binds extremely well and remained stable in the binding region. Furthermore, it was observed that the strongest hydrogen having energy of  $-32$  kcal were found between the residue of TLR3 Arg462 and vaccine His243 and the distance of this hydrogen bond was  $2.9$  Å. While for TLR8 the strongest bond had energy of  $-33.5$  kcal with bond distance  $2.96$  Å and the strongest bond for that of TLR9 were found between Glu812 of TLR9 and Lys233 of vaccine with bond distance of  $2.8$  Å. Since, the binding energy of all the hydrogen bonds in TLR3, TLR8 and TLR9 range from  $-33$  to  $-0.7$  kcal. Thereby, indicating that the constructed vaccine predicted to bind in their respective TLR proteins (Fig. 5).

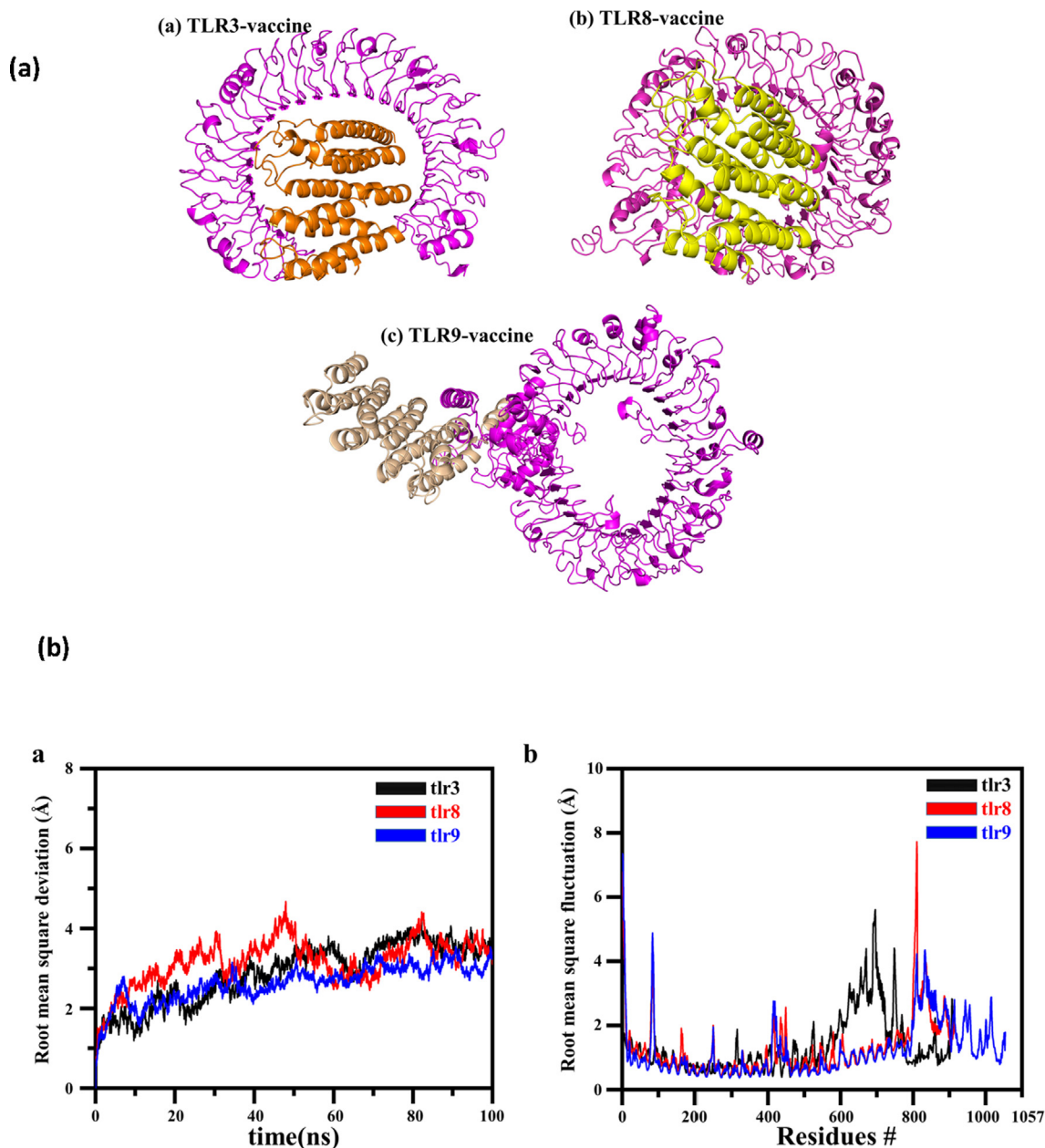
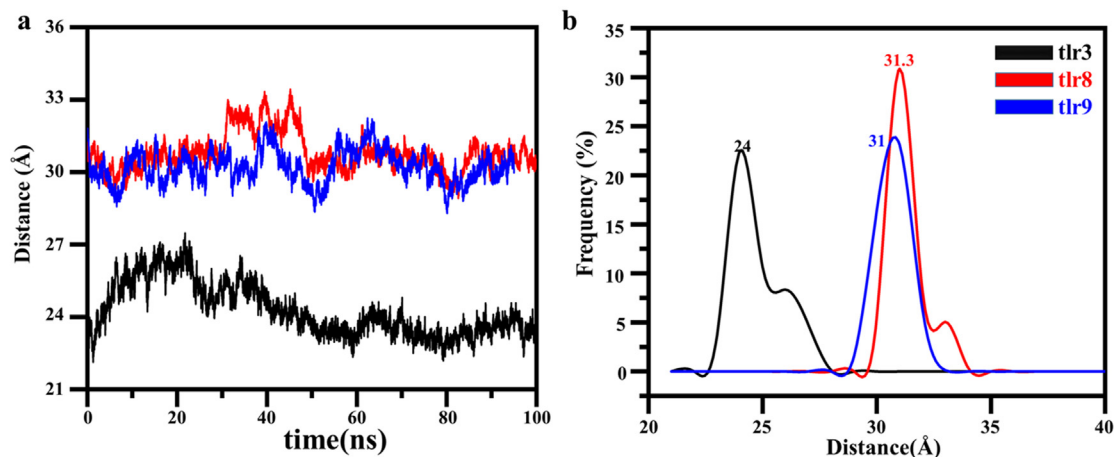


Fig. 5. Docked complexes and MD simulations. (a) Best complexes of TLR protein with constructed vaccine: TLR3-vaccine, TLR8-vaccine and TLR9-vaccine. (b) Dynamic stability given as; (i) RMSD of all systems. (ii) RMSF graph of all complexes. (c) Distance analysis between TLR proteins and constructed vaccine. The color scheme is shown inside the figure. (d) Free energy landscape and interaction analysis of; (i) TLR3-vaccine (ii) TLR8-vaccine (iii) TLR9-vaccine. The deep dark regions showed the lowest energy conformers.



(c)



(d)

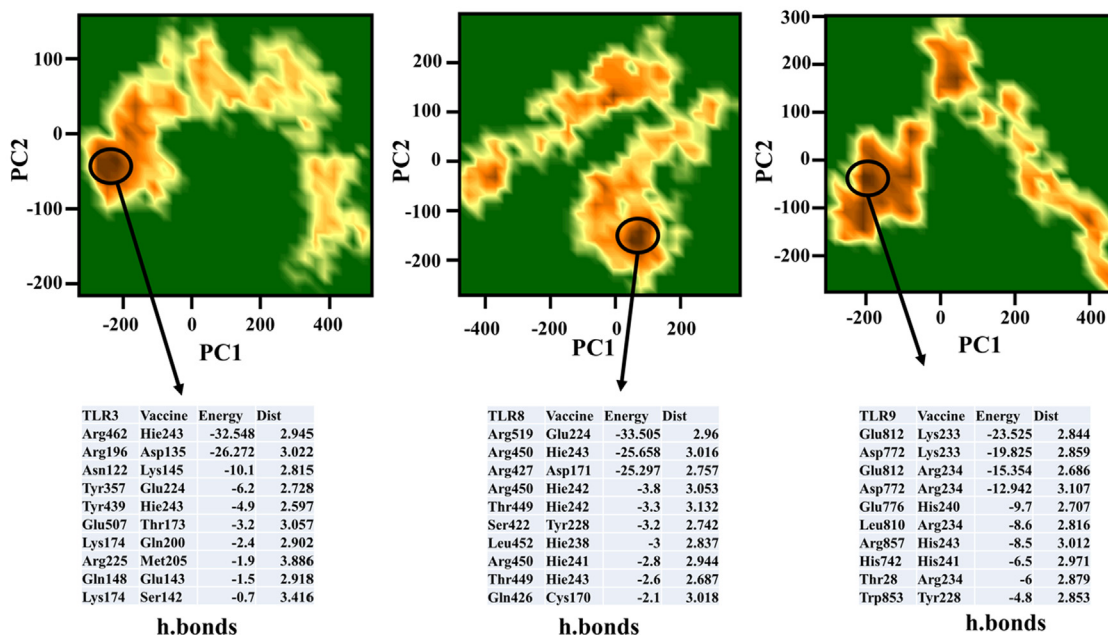


Fig. 5 (continued)

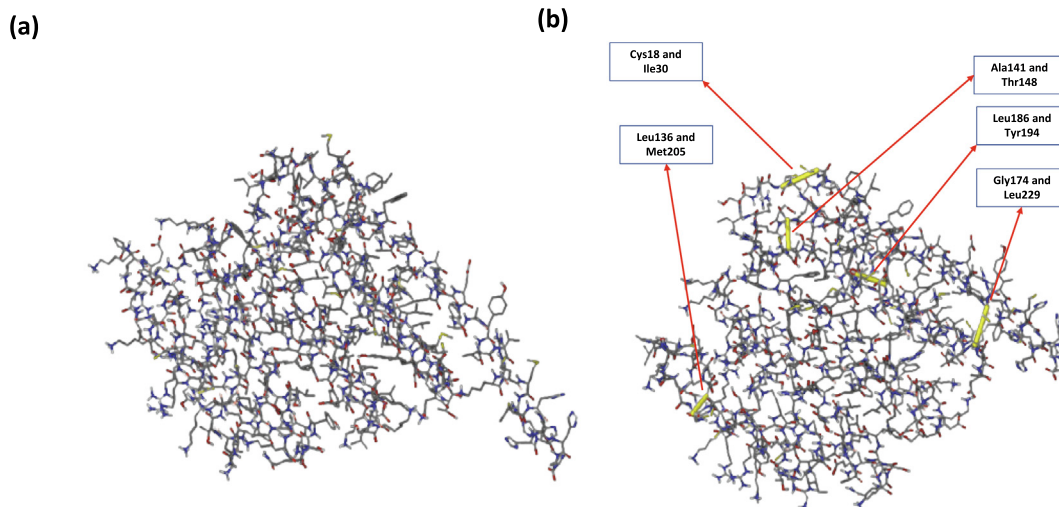
Table 5

The details of selected top ranked models based on energy parameters.

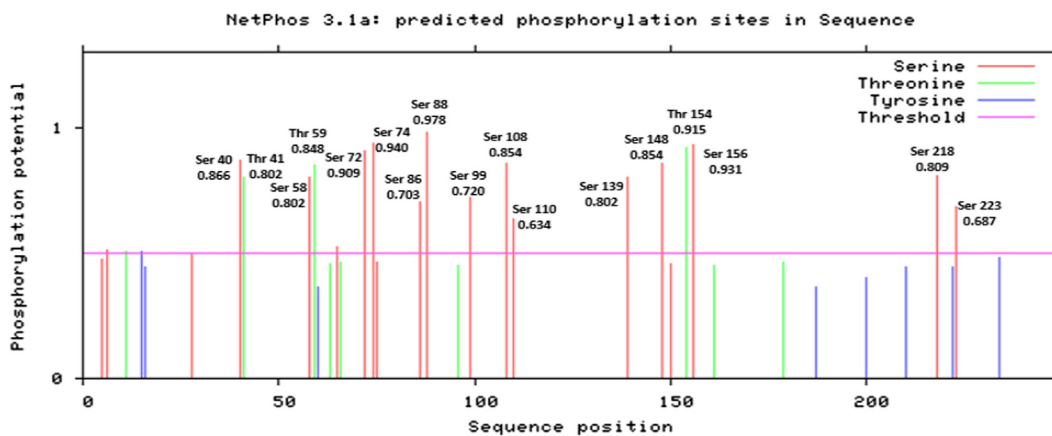
TLR's	Patchdock Transformation No. which is Ranked #1 by Firedock	Global Energy	Attractive Van der Waal Forces (Awdw)	Repulsive Van der Waal Forces (rvdW)	Atomic Contact Energy (ACE)	Hydrogen and Disulfide Bonds (HB)
TLR-3	599	-41.09	-23.10	5.50	-0.82	-4.93
TLR-8	354	-58.61	-38.41	11.17	-3.36	-2.76
TLR-9	162	-31.15	-26.81	9.97	-2.89	-1.02

Moreover, the analysis obtained from NetChop revealed 87 cleavage sites in the vaccine polypeptide. Besides the cleavage site predictions, TAP binding peptides were computed using TAPREG. The TAPREG output calculated 65 TAP binding peptides, but only 3 peptides were considered depending upon lowest IC50 value, as shown in (Table 4). Since, lower the IC 50 value, higher is the TAP binding affinity of that peptide. While, for overcoming the obstacle of reduced heterologous gene expression, codon adapta-

tion is conducted to access the feasibility of expression of the designed vaccine polypeptide (Puigbò et al., 2007). The codon adaptation studies were performed using JCAT interface, which showed that the CAI value and GC-content of the improved sequence were 1.0 and 49.9 % respectively (Figure S5). Therefore, the CAI value and GC-content obtained from JCat analysis for the multivalent vaccine were found to be very suitable for the expression of the multi-epitope vaccine polypeptide in the *E. coli* (strain



**Fig. 6.** Mutating di sulfide bonds to confer stability to the vaccine construct. (a) Original vaccine construct. (b) Mutated vaccine construct with yellow lines indicating the artificial construction of di sulfide bonds between Cys18 and Ile30; Leu136 and Met205; Ala141 and Thr148; Gly174 and Gly229 and Leu186 and Tyr194 having torsion angle ( $\chi_3$ ) values; +110.62, +123.93, +120.39, +114.16 and +115.79.



**Fig. 7.** NetPhos predicted phosphorylation sites. The 17 sites that surpassed the threshold are Ser 40, 58, 72, 74, 86, 88, 99, 108, 110, 139, 148, 156, 218, 223 and Thr 41, 59, 154.

K12) (Narula et al., 2018; Nezafat et al., 2016). The optimized DNA sequence of 729 bp length obtained from JCat was successfully cloned into pDual-GC vector using SnapGene between BclI\*-AccI-SnaBI restriction sites, located in the vector backbone from 4036 bp to 4765 bp (Figure S5). Therefore, this in-silico cloning demonstrated that our vaccine construct can be successfully cloned in *E. coli*. The pDual-GC vector can be used for heterologous expression of recombinant protein in both mammalian and prokaryotic systems.

Meanwhile, the cell localization calculations were carried out using Phobius and BaCELLO. The Phobius output demonstrated that vaccine’s amino acid residue from 1 to 109 and 217 to 243 were cytoplasmic, while the residue 110 to 128 and 194 to 216 were having transmembrane domain (Figure S4). Side by side, BaCelLo indicated the vaccine polypeptide’s localization to be cytoplasmic. The proteins having cytoplasmic localization with transmembrane domains having hydrophobic TMD’s usually are a good candidate for CTL epitopes attachment (Kovjazin et al., 2011; Lucchiari-Hartz et al., 2003). Additionally, the N-glycosylation using NetNGlyc and O-glycosylation using NetOGlyc

tool yielded no potential N-glycosylation and O-glycosylation sites in the vaccine sequence respectively. Apart from N and O glycosylation analysis, other post translational modifications such as Ubiquitination and Phosphorylation was also analyzed using iUbiq-Lys and NetPhos servers, respectively. The iUbiq-Lys yielded no ubiquitination sites. While NetPhos predicted 17 phosphorylation sites. The sites: Ser40, Ser58, Ser72, Ser74, Ser86, Ser88, Ser99, Ser108, Ser110, Ser139, Ser148, Ser156, Ser218, Ser223, Thr41, Thr59 and Thr154 surpassed the threshold and were selected (Fig. 7). The presence of so many serine residues at the phosphorylation site is a normal occurrence as a study conducted in eukaryotic cells, showed that approximately 86.4 % of the protein phosphorylation occurs on Serine residue (Ardito et al., 2017; Deribe et al., 2010). Furthermore, these observations also support the idea that designed protein may participate in signal transduction and serves as an easy target for TAP-Proteasome pathway. Thus, indicating these regions of vaccine possesses effective CTL epitopes capacity. Another post translational modification that was conducted on the designed vaccine was mutating disulfide bonds artificially, to confer structural stability to the construct, using Disulfide by Design 2

tool. The disulfide bonds were constructed between Cys18 and Ile30; Leu136 and Met205; Ala141 and Thr148; Gly174 and Gly229; Leu186 and Tyr194 possessing the torsion angle ( $\chi_3$ ) values: +110.62, +123.93, +120.39, +114.16 and +115.79 respectively (Fig. 7). They added stability to the protein structure. This engineering was performed to enhance the *in vitro* stability of our vaccine polypeptide as shown by the instability index (II) values. If these mutations will be added to our vaccine, then it will confer stability to its structure. Hence, the stability of vaccine peptide in test-tube or *in vitro* can be enhanced.

Finally, the results of immune simulation obtained from the C-IMMSIM interface showed the decline in elevated Antigen (Ag) persistence after five days because of exponential production of IgM+IgG, and IgM which by themselves reduced in about 35 days. Simultaneously, a considerable amount of the antibodies IgG1 +IgG2, IgG1, and IgG2, were also produced (Fig. 4(a)). The total increase in the B-Cell and T-lymphocytes population count caused this decrease in antigen (Fig. 4(b)). After the vaccine injection, the increase in the population of T-lymphocytes, active – (Fig. 5(c)), total T-helper (TH) and memory TH cell (Fig. 4(d)) along with B-Cell population was demonstrated. The increase in the active T-cytotoxic (TC) cell population per state was also observed (Fig. 4 (e)). In addition, TC cell population per state (cells per mm<sup>3</sup>) was very high indicating that they are highly responsive to *in-silico* developed multivalent vaccine candidate (Fig. 4(f)). While, according to the simulation plot of cytokines, remarkably high levels of IFN-g and a considerable amount of TGF $\beta$ , IL-2, and IL-12 production were observed. While very little IL-10 was produced (Fig. 4 (f)), signifying the candidate vaccine's ability to generate an evident and appropriate immune response. Thereby demonstrating that the final vaccine construct aptly provoked a high anti-CCHFV immunity which ultimately leads to later antigen clearance from the system.

The present study's next objective is to conduct *in vitro* and *in vivo* validation and efficacy investigations on our produced vaccine. The primary challenge that an *in-silico* vaccine faces in this regard is dosage optimization, as well as the selection of an optimal animal model for validation tests (Hwang et al., 2021). To address these limitations, the generation of an optimal human PBPK model for better assessment of dosage adjustment and physiological distribution of vaccine in the human body must be established before entering clinical trials (Gómez-Mantilla et al., 2010). Furthermore, ADME studies for PK/PD modelling can be employed to predict the most immunogenic vaccine dose in humans. The PK/PD assessment will also give information in choosing better animal model (Tegenge and Mitkus, 2013). These analyses if performed correctly will greatly support experimentations. Since in all vaccine development, acceleration of preclinical and clinical testing and a reduction in the cost of clinical trials are very valuable.

## 5. Conclusion

The present study indicated that computer-aided approach towards vaccine designing, has a great promise in accelerating and aiding vaccine development for pathogens. This method resulted in the identification of specific epitopes, capable of binding to B-cells and T-cells in the Gc and Gn regions of CCHFV. The epitopes were connected to each other through adjuvant and various linkers to form a multivalent subunit vaccine for CCHFV. Furthermore, the results of this study suggested that *in-silico* both the humoral and cellular immune responses were stimulated by the proposed subunit vaccine. Therefore, it possesses the exquisite potential to serve as a vaccine to counter CCHFV infection. However, for validating the efficacy of this vaccine, rigorous immunological experimentations (*in vitro* and *in vivo*) are needed.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We would like to acknowledge National University of Sciences and Technology and National University of Medical Sciences for providing support to perform this study.

## Authors contribution

MS has performed the experiments. MJ and SBJ performed molecular dynamics simulation and data analysis. AJ and MF have designed and supervised the study. MS wrote the manuscript and AJ, MF and SBJ revised the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.sjbs.2021.12.004>.

## References

- Adiyaman, R., McGuffin, L.J., 2019. Methods for the refinement of protein structure 3D models. *Int. J. Mol. Sci.* 20. <https://doi.org/10.3390/ijms20092301>.
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., Nagamune, T., 2001. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* 14 (8), 529–532.
- Ardito, F., Giuliani, M., Perrone, D., Troiano, G., Lo Muzio, L., 2017. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int. J. Mol. Med.* 40, 271–280. <https://doi.org/10.3892/ijmm.2017.3036>.
- Ayyagari, V.S., T C V., K, A.P., Srirama, K., 2020. Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach. *J. Biomol. Struct. Dyn.* 1–15. [10.1080/07391102.2020.1850357](https://doi.org/10.1080/07391102.2020.1850357).
- Bente, D.A., Alimonti, J.B., Shieh, W.-J., Camus, Gaele, Stroher, U., Zaki, S., Jones, S.M., 2010. Pathogenesis and immune response of Crimean-Congo hemorrhagic fever virus in a STAT-1 knockout mouse model. *J. Virol.* 84 (21), 11089–11100. <https://doi.org/10.1128/JVI.01383-10>.
- Bhatnager, R., Bhasin, M., Arora, J., Dang, A.S., 2021. Epitope based peptide vaccine against SARS-COV2: an immune-informatics approach. *J. Biomol. Struct. Dyn.* 39 (15), 5690–5705. <https://doi.org/10.1080/07391102.2020.1787227>.
- Bhattacharya, D., Nowotny, J., Cao, R., Cheng, J., 2016. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Res.* 44 (W1), W406–W409. <https://doi.org/10.1093/nar/gkw336>.
- Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 294 (5), 1351–1362. <https://doi.org/10.1006/jmbi.1999.3310>.
- Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4 (6), 1633–1649. <https://doi.org/10.1002/pmic.200300771>.
- Bui, H.-H., Sidney, J., Dinh, K., Southwood, S., Newman, M.J., Sette, A., 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics* 7, 153. <https://doi.org/10.1186/1471-2105-7-153>.
- Buttigieg, K.R., Dowall, S.D., Findlay-Wilson, S., Miloszewska, A., Rayner, E., Hewson, R., Carroll, M.W., 2014. A novel vaccine against Crimean-Congo Haemorrhagic Fever protects 100% of animals against lethal challenge in a mouse model. *PLoS One* 9, e91516. [10.1371/journal.pone.0091516](https://doi.org/10.1371/journal.pone.0091516).
- Canakoglu, N., Berber, E., Tonbak, S., Ertek, M., Sozdutmaz, I., Aktas, M., Kalkan, A., Ozdarendeli, A., 2015. Immunization of knock-out  $\alpha/\beta$  interferon receptor mice against high lethal dose of Crimean-Congo hemorrhagic fever virus with a cell culture based vaccine. *PLoS Negl. Trop. Dis.* 9, e0003579. [10.1371/journal.pntd.0003579](https://doi.org/10.1371/journal.pntd.0003579).
- Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases, Division of High-Consequence Pathogens and Pathology, Viral Special Pathogens Branch, 2003. Treatment | Crimean-Congo Hemorrhagic Fever (CCHF) | CDC [WWW Document]. Crimean-Congo Hemorrhagic Fever (CCHF). URL <https://www.cdc.gov/vhf/crimean-congo/treatment/index.html> (accessed 8.1.21).
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect.*



- D. Biol. Crystallogr. 66 (1), 12–21. <https://doi.org/10.1107/S0907444909042073>.
- Chen, X., Zaro, J., Shen, W.-C., 2013. In: Fusion Protein Technologies for Biopharmaceuticals: Applications and Challenges. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 57–73. <https://doi.org/10.1002/9781118354599.ch4>.
- Colovos, C., Yeates, T.O., 1993. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2 (9), 1511–1519. <https://doi.org/10.1002/pro.5560020916>.
- Craig, D.B., Dombkowski, A.A., 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* 14, 346. <https://doi.org/10.1186/1471-2105-14-346>.
- Craiu, A., Akopian, T., Goldberg, A., Rock, K.L., 1997. Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. *Proc. Natl. Acad. Sci. USA* 94 (20), 10850–10855. <https://doi.org/10.1073/pnas.94.20.10850>.
- Carafa, Y.D., Brody, E., Thernes, C., 1990. Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J. Mol. Biol.* 216 (4), 835–858.
- Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98 (12), 10089–10092. <https://doi.org/10.1063/1.464397>.
- Deribe, Y.L., Pawson, T., Dikic, I., 2010. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* 17 (6), 666–672. <https://doi.org/10.1038/nsmb.1842>.
- Desta, I.T., Porter, K.A., Xia, B., Kozakov, D., Vajda, S., 2020. Performance and Its limits in rigid body protein-protein docking. *Structure* 28 (9), 1071–1081.e3. <https://doi.org/10.1016/j.str.2020.06.006>.
- Dhanda, S.K., Vir, P., Raghava, G.P.S., 2013. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct* 8, 30. <https://doi.org/10.1186/1745-6150-8-30>.
- Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova, I., 2014. AllerTOP vol 2—a server for in silico prediction of allergens. *J. Mol. Model* 20, 2278. <https://doi.org/10.1007/s00894-014-2278-5>.
- Dowall, S.D., Carroll, M.W., Hewson, R., 2017. Development of vaccines against Crimean-Congo haemorrhagic fever virus. *Vaccine* 35 (44), 6015–6023. <https://doi.org/10.1016/j.vaccine.2017.05.031>.
- Doytchinova, I.A., Guan, P., Flower, D.R., 2006. EpiJen: a server for multistep T cell epitope prediction. *BMC Bioinformatics* 7, 131. <https://doi.org/10.1186/1471-2105-7-131>.
- Drosten, C., Kümmerer, B.M., Schmitz, H., Günther, S., 2003. Molecular diagnostics of viral hemorrhagic fevers. *Antiviral Res.* 57 (1–2), 61–87. [https://doi.org/10.1016/S0166-3542\(02\)00201-2](https://doi.org/10.1016/S0166-3542(02)00201-2).
- Fernandez-García, M., Negro, A., Papa, A., Donoso-Mantke, O., Niedrig, M., Zeller, H., Tenorio, A., Franco, L., the ENIVD members, c. 2014. European survey on laboratory preparedness, response and diagnostic capacity for Crimean-Congo haemorrhagic fever, 2012. *Eurosurveillance* 19, 20844. <https://doi.org/10.2807/1560-7917.ES2014.19.26.20844>.
- Flick, R., Whitehouse, C.A., 2005. Crimean-Congo hemorrhagic fever virus. *Curr. Mol. Med.* 5, 753–760. <https://doi.org/10.2174/156652405774962335>.
- Formenty, D.P., n.d. Introduction to Crimean-Congo Haemorrhagic Fever 23.
- Gaafar, B.B.M., Ali, S.A., Abd-elrahman, K.A., Almofti, Y.A., 2019. Immunoinformatics Approach for Multi-epitope Vaccine Prediction from H, M, F, and N Proteins of Peste des Petits Ruminants Virus. *J. Immunol. Res.* 2019, 1–18. <https://doi.org/10.1155/2019/6124030>.
- Garrison, A.R., Shoemaker, C.J., Golden, J.W., Fitzpatrick, C.J., Suschak, J.J., Richards, M.J., Badger, C.V., Six, C.M., Martin, J.D., Hannaman, D., Zivcec, M., Bergeron, E., Koehler, J.W., Schmaljohn, C.S., 2017. A DNA vaccine for Crimean-Congo hemorrhagic fever protects against disease and death in two lethal mouse models. *PLoS Negl. Trop. Dis.* 11, e0005908. <https://doi.org/10.1371/journal.pntd.0005908>.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. <https://doi.org/10.1093/nar/gkg563>.
- Ghiasi, S.M., Salmanian, A.H., Chinikar, S., Zakeri, S., 2011. Mice orally immunized with a transgenic plant expressing the glycoprotein of Crimean-Congo hemorrhagic fever virus. *Clin. Vaccine Immunol.* 18 (12), 2031–2037. <https://doi.org/10.1128/CVI.05352-11>.
- Gómez-Mantilla, J.D., Trocóniz, I.F., Garrido, M.J., 2010. ADME processes in vaccines and PK/PD approaches for vaccination optimization. In: Gad, S.C. (Ed.), *Pharmaceutical sciences encyclopedia: drug discovery, development, and manufacturing*. John Wiley & Sons Inc, Hoboken, NJ, USA, pp. 1–22. <https://doi.org/10.1002/978047051224.pse558>.
- Gonzalez, J.P., Cornet, J.P., Wilson, M.L., Camicas, J.L., 1991. Crimean-Congo haemorrhagic fever virus replication in adult *Hyalomma truncatum* and *Amblyomma variegatum* ticks. *Res. Virol.* 142 (6), 483–488.
- Götz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., Walker, R.C., 2012. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* 8 (5), 1542–1555. <https://doi.org/10.1021/ct200909j>.
- Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D.C., Jahn, D., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* 33 (Web Server), W526–W531. <https://doi.org/10.1093/nar/gki376>.
- Gu, Y., Sun, X., Li, B., Huang, J., Zhan, B., Zhu, X., 2017. Vaccination with a Paramyxin-Based Multi-Epitope Vaccine Elicits Significant Protective Immunity against *Trichinella spiralis* Infection in Mice. *Front. Microbiol.* 8, 1475. <https://doi.org/10.3389/fmicb.2017.01475>.
- Hawman, D.W., Ahlén, G., Appelberg, K.S., Meade-White, K., Hanley, P.W., Scott, D., Monteil, V., Devignot, S., Okumura, A., Weber, F., Feldmann, H., Sällberg, M., Mirazimi, A., 2021. A DNA-based vaccine protects against Crimean-Congo haemorrhagic fever virus disease in a Cynomolgus macaque model. *Nat. Microbiol.* 6 (2), 187–195. <https://doi.org/10.1038/s41564-020-00815-6>.
- Hawman, D.W., Haddock, E., Meade-White, K., Williamson, B., Hanley, P.W., Rosenke, K., Komeno, T., Furuta, Y., Gowen, B.B., Feldmann, H., 2018. Favipiravir (T-705) but not ribavirin is effective against two distinct strains of Crimean-Congo hemorrhagic fever virus in mice. *Antiviral Res.* 157, 18–26. <https://doi.org/10.1016/j.antiviral.2018.06.013>.
- Hinkula, J., Devignot, S., Åkerström, S., Karlberg, H., Watrang, E., Bereczky, S., Mousavi-Jazi, M., Risinger, C., Lindgren, G., Vernersson, C., Paweska, J., van Vuren, P.J., Blixt, O., Brun, A., Weber, F., Mirazimi, A., García-Sastre, A., 2017. Immunization with DNA plasmids coding for crimean-congo hemorrhagic fever virus capsid and envelope proteins and/or virus-like particles induces protection and survival in challenged mice. *J. Virol.* 91 (10). <https://doi.org/10.1128/JVI.02076-16>.
- Hwang, W., Lei, W., Katritsis, N.M., MacMahon, M., Chapman, K., Han, N., 2021. Current and prospective computational approaches and challenges for developing COVID-19 vaccines. *Adv. Drug Deliv. Rev.* 172, 249–274. <https://doi.org/10.1016/j.addr.2021.02.004>.
- Ikram, A., Zaheer, T., Awan, F.M., Obaid, A., Naz, A., Hanif, R., Paracha, R.Z., Ali, A., Naveed, A.K., Janjua, H.A., 2018. Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. *Sci. Rep.* 8, 16107. <https://doi.org/10.1038/s41598-018-34254-5>.
- Jespersen, M.C., Peters, B., Nielsen, M., Marcantili, P., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45, W24–W29. <https://doi.org/10.1093/nar/gkx346>.
- Johnson, S., Henschke, N., Maayan, N., Mills, I., Buckley, B.S., Kakourou, A., Marshall, R., 2018. Ribavirin for treating Crimean Congo haemorrhagic fever. *Cochrane Database Syst. Rev.* 6, CD012713. <https://doi.org/10.1002/14651858.CD012713.pub2>.
- Johnston, L.J., King, N.J.C., Halliday, G.M., 2000. Langerhans cells migrate to local lymph nodes following cutaneous infection with an arbovirus. *J. Invest. Dermatol.* 114 (3), 560–568. <https://doi.org/10.1046/j.1523-1747.2000.00904.x>.
- Kazimirová, M., Thangamani, S., Bartíková, P., Hermance, M., Holíková, V., Štibráňová, I., Nuttall, P.A., 2017. Tick-Borne Viruses and Biological Processes at the Tick-Host-Virus Interface. *Front. Cell Infect. Microbiol.* 7, 339. <https://doi.org/10.3389/fcimb.2017.00339>.
- Khatun, M.S., Hasan, M.M., Kurata, H., 2019. PreAIP: computational prediction of anti-inflammatory peptides by integrating multiple complementary features. *Front. Genet.* 10, 129. <https://doi.org/10.3389/fgene.2019.00129>.
- Kochs, G., Bauer, S., Vogt, C., Frenz, T., Tschopp, Jürg, Kalinke, U., Waibler, Z., 2010. Thogoto virus infection induces sustained type I interferon responses that depend on RIG-I-like helicase signaling of conventional dendritic cells. *J. Virol.* 84 (23), 12344–12350. <https://doi.org/10.1128/JVI.00931-10>.
- Kortekaas, J., Vloet, R.P.M., McAuley, A.J., Shen, X., Bosch, B.J., de Vries, L., Moormann, R.J.M., Bente, D.A., 2015. Crimean-Congo Hemorrhagic Fever virus subunit vaccines induce high levels of neutralizing antibodies but no protection in STAT1 knockout mice. *Vector Borne Zoonotic Dis.* 15 (12), 759–764. <https://doi.org/10.1089/vbz.2015.1855>.
- Kotturi, M.F., Peters, B., Buendia-Laysa, F., Sidney, J., Oseroff, C., Botten, J., Grey, H., Buchmeier, M.J., Sette, A., 2007. The CD8+ T-cell response to lymphocytic choriomeningitis virus involves the L antigen: uncovering new tricks for an old virus. *J. Virol.* 81 (10), 4928–4940. <https://doi.org/10.1128/JVI.02632-06>.
- Kovjazin, R., Volovitz, I., Daon, Y., Vider-Shalit, T., Azran, R., Tsaban, L., Carmon, L., Louzoun, Y., 2011. Signal peptides and trans-membrane regions are broadly immunogenic and have high CD8+ T cell epitope densities: implications for vaccine development. *Mol. Immunol.* 48 (8), 1009–1018. <https://doi.org/10.1016/j.molimm.2011.01.006>.
- Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., Vajda, S., 2013. How good is automated protein docking? *Proteins* 81, 2159–2166. <https://doi.org/10.1002/prot.24403>.
- Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., Beglov, D., Vajda, S., 2017. The ClusPro web server for protein-protein docking. *Nat. Protoc.* 12 (2), 255–278. <https://doi.org/10.1038/nprot.2016.169>.
- Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., Karplus, K., 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* 77 (Suppl. 9), 114–122. <https://doi.org/10.1002/prot.22570>.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23 (21), 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>.
- Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., Thornton, J.M., 2018. PDBsum: structural summaries of PDB entries. *Protein Sci.* 27 (1), 129–134. <https://doi.org/10.1002/pro.3289>.
- Lee, H.S., Qi, Y., Im, W., 2015. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci. Rep.* 5, 8926. <https://doi.org/10.1038/srep08926>.
- Li, X., Guo, L., Kong, M., Su, X., Yang, D., Zou, M., Liu, Y., Lu, L., 2015. Design and evaluation of a multi-epitope peptide of human metapneumovirus. *Intervirology* 58, 403–412. <https://doi.org/10.1159/000445059>.
- Limon-Vega / The Pirbright Institute, D.G., Lyons / The Pirbright Institute, D.N., 2019. Pirbright Carries Out Vaccine Trials For Crimean-Congo Haemorrhagic



- Fever [WWW Document]. URL <https://www.pirbright.ac.uk/news/2019/04/pirbright-carries-out-vaccine-trials-crimean-congo-haemorrhagic-fever> (accessed 11.24.21).
- Lindquist, M.E., Zeng, X., Altamura, L.A., Daye, S.P., Delp, K.L., Blancett, C., Coffin, K.M., Koehler, J.W., Coyne, S., Shoemaker, C.J., Garrison, A.R., Golden, J.W., Dermody, T.S., 2018. Exploring Crimean-Congo Hemorrhagic Fever virus-induced hepatic injury using antibody-mediated type I interferon blockade in mice. *J. Virol.* 92 (21). <https://doi.org/10.1128/JVI.01083-18>.
- Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., Sette, A., 2002. A rational strategy to design multi-epitope immunogens based on multiple Th lymphocyte epitopes. *J. Immunol.* 168 (11), 5499–5506. <https://doi.org/10.4049/jimmunol.168.11.5499>.
- Lucchiari-Hartz, M., Lindo, V., Hitziger, N., Gaedcke, S., Saveanu, L., van Endert, P.M., Greer, F., Eichmann, K., Niedermann, G., 2003. Differential proteasomal processing of hydrophobic and hydrophilic protein regions: contribution to cytotoxic T lymphocyte epitope clustering in HIV-1-Nef. *Proc. Natl. Acad. Sci. USA* 100 (13), 7755–7760. <https://doi.org/10.1073/pnas.1232228100>.
- Manavalan, B., Shin, T.H., Kim, M.O., Lee, G., 2018. AllPpred: sequence-based prediction of anti-inflammatory peptides using random forest. *Front. Pharmacol.* 9, 276. <https://doi.org/10.3389/fphar.2018.00276>.
- Mashiach, E., Schneidman-Duhovny, D., Andrusier, N., Nussinov, R., Wolfson, H.J., 2008. FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Res.* 36, W229–W232. <https://doi.org/10.1093/nar/gkn186>.
- Mashiach, E., Schneidman-Duhovny, D., Peri, A., Shavit, Y., Nussinov, R., Wolfson, H.J., 2010. An integrated suite of fast docking algorithms. *Proteins* 78 (15), 3197–3204. <https://doi.org/10.1002/prot.22790>.
- Masson, F., Mount, A.M., Wilson, N.S., Belz, G.T., 2008. Dendritic cells: driving the differentiation programme of T cells in viral infections. *Immunol. Cell Biol.* 86 (4), 333–342. <https://doi.org/10.1038/icb.2008.15>.
- Mazzola, L.T., Kelly-Cirino, C., 2019. Diagnostic tests for Crimean-Congo haemorrhagic fever: a widespread tickborne disease. *BMJ Glob Health* 4 (Suppl. 2), e001114. <https://doi.org/10.1136/bmjgh-2018-001114>.
- MI, E., 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.*
- Mousavi-Jazi, M., Karlberg, H., Papa, A., Christova, I., Mirazimi, A., 2012. Healthy individuals' immune response to the Bulgarian Crimean-Congo hemorrhagic fever virus vaccine. *Vaccine* 30 (44), 6225–6229. <https://doi.org/10.1016/j.vaccine.2012.08.003>.
- Mullinax, R.L., Wong, D.T., Davis, H.A., Padgett, K.A., 2003. Dual-Expression Vectors for Efficient Protein Expression in Both E. coli and Mammalian Cells. *E. coli Gene Expression* ....
- Nain, Z., Abdulla, F., Rahman, M.M., Karim, M.M., Khan, M.S.A., Sayed, S.B., Mahmud, S., Rahman, S.M.R., Sheam, M.M., Haque, Z., Adhikari, U.K., 2020. Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen Elizabethkingia anophelis using immunoinformatic approaches. *J. Biomol. Struct. Dyn.* 38 (16), 4850–4867. <https://doi.org/10.1080/07391102.2019.1692072>.
- Narula, A., Pandey, R.K., Khatoun, N., Mishra, A., Prajapati, V.K., 2018. Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection. *Infect. Genet. Evol.* 61, 4–15. <https://doi.org/10.1016/j.meegid.2018.03.007>.
- Nezafat, N., Karimi, Z., Eslami, M., Mohkam, M., Zandian, S., Ghasemi, Y., 2016. Designing an efficient multi-epitope peptide vaccine against Vibrio cholerae via combined immunoinformatics and protein interaction based approaches. *Comput. Biol. Chem.* 62, 82–95. <https://doi.org/10.1016/j.compbiolchem.2016.04.006>.
- Oestereich, L., Rieger, T., Neumann, M., Bernreuther, C., Lehmann, M., Krasemann, S., Wurr, S., Emmerich, P., de Lamballerie, X., Ölschläger, S., Günther, S., 2014. Evaluation of antiviral efficacy of ribavirin, arbidol, and T-705 (favipiravir) in a mouse model for Crimean-Congo hemorrhagic fever. *PLoS Negl. Trop. Dis.* 8, e2804. <https://doi.org/10.1371/journal.pntd.0002804>.
- Papa, A., Papadimitriou, E., Christova, I., 2011. The Bulgarian vaccine Crimean-Congo haemorrhagic fever virus strain. *Scand. J. Infect. Dis.* 43 (3), 225–229. <https://doi.org/10.3109/00365548.2010.540036>.
- Pierleoni, A., Martelli, P.L., Fariselli, P., Casadio, R., 2006. BaCellLo: a balanced subcellular localization predictor. *Bioinformatics* 22 (14), e408–e416. <https://doi.org/10.1093/bioinformatics/btl222>.
- Puigbò, P., Guzmán, E., Romeu, A., García-Vallvé, S., 2007. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* 35, W126–W131. <https://doi.org/10.1093/nar/gkm219>.
- Radivojac, P., Vacic, V., Haynes, C., Cocklin, R.R., Mohan, A., Heyen, J.W., Goebel, M.G., Iakoucheva, L.M., 2010. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* 78 (2), 365–380. <https://doi.org/10.1002/prot.22555>.
- 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. <https://doi.org/10.1371/journal.pone.0009862>.
- Rodriguez, S.E., Cross, R.W., Fenton, K.A., Bente, D.A., Mire, C.E., Geisbert, T.W., 2019. Vesicular stomatitis virus-based vaccine protects mice against Crimean-Congo Hemorrhagic Fever. *Sci. Rep.* 9, 7755. <https://doi.org/10.1038/s41598-019-44210-6>.
- Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5 (4), 725–738. <https://doi.org/10.1038/nprot.2010.5>.
- Salomon-Ferrer, R., Case, D.A., Walker, R.C., 2013. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* 3 (2), 198–210. <https://doi.org/10.1002/wcms.1121>.
- Sarkar, B., Ullah, M.A., Johora, F.T., Taniya, M.A., Araf, Y., 2020. Immunoinformatics-guided designing of epitope-based subunit vaccines against the SARS Coronavirus-2 (SARS-CoV-2). *Immunobiology* 225 (3), 151955. <https://doi.org/10.1016/j.imbio.2020.151955>.
- Schmidt, C., Beilstein-Edmands, V., Mohammed, S., Robinson, C.V., 2017. Acetylation and phosphorylation control both local and global stability of the chloroplast F1 ATP synthase. *Sci. Rep.* 7, 44068. <https://doi.org/10.1038/srep44068>.
- Scholte, F.E.M., Spengler, J.R., Welch, S.R., Harmon, J.R., Coleman-McCray, J.D., Freitas, B.T., Kainulainen, M.H., Pegan, S.D., Nichol, S.T., Bergeron, É., Spiropoulos, C.F., 2019. Single-dose replicon particle vaccine provides complete protection against Crimean-Congo hemorrhagic fever virus in mice. *Emerg. Microbes Infect.* 8 (1), 575–578. <https://doi.org/10.1080/22221751.2019.1601030>.
- Sette, A., Vitiello, A., Reheman, B., Fowler, P., Nayarsina, R., Kast, W.M., Melief, C.J., Oseroff, C., Yuan, L., Ruppert, J., Sidney, J., del Guercio, M.F., Southwood, S., Kubo, R.T., Chesnut, R.W., Grey, H.M., Chisari, F.V., 1994. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J. Immunol.* 153, 5586–5592.
- Shekhar, C., Dev, K., Verma, S.K., Kumar, A., 2012. In-silico: screening and modeling of CTL binding epitopes of Crimean congo hemorrhagic fever virus. *Trends Bioinform.*
- Shey, R.A., Ghogomu, S.M., Esoh, K.K., Nebangwa, N.D., Shintouo, C.M., Nongley, N.F., Asa, B.F., Ngale, F.N., Vanhamme, L., Souopgui, J., 2019. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci. Rep.* 9, 4409. <https://doi.org/10.1038/s41598-019-40833-x>.
- Sippl, M.J., 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins* 17 (4), 355–362. <https://doi.org/10.1002/prot.340170404>.
- Steenfot, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.-T.-B.-G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Bennett, E.P., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., Clausen, H., 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 32, 1478–1488. <https://doi.org/10.1038/emboj.2013.79>.
- Stuart, M.C., Kouimtzis, M., Hill, S.R., 2009. WHO Model Formulary 2008, 650.
- Su, V., Lau, A.F., 2009. Ubiquitin-like and ubiquitin-associated domain proteins: significance in proteasomal degradation. *Cell Mol. Life Sci.* 66 (17), 2819–2833. <https://doi.org/10.1007/s00018-009-0048-9>.
- Tegenge, M.A., Mitkus, R.J., 2013. A physiologically-based pharmacokinetic (PBPK) model of squalene-containing adjuvant in human vaccines. *J. Pharmacokinet. Pharmacodyn.* 40 (5), 545–556. <https://doi.org/10.1007/s10928-013-9328-y>.
- Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S.E., Xia, B., Hall, D.R., Kozakov, D., 2017. New additions to the ClusPro server motivated by CAPRI. *Proteins* 85, 435–444. <https://doi.org/10.1002/prot.25219>.
- Wampande, E.M., Waiswa, P., Allen, D.J., Hewson, R., Frost, S.D.W., Stubbs, S.C.B., 2021. Phylogenetic characterization of Crimean-Congo Hemorrhagic Fever Virus detected in african blue ticks feeding on cattle in a ugandan abattoir. *Microorganisms* 9. <https://doi.org/10.3390/microorganisms9020438>.
- WHO EMRO | Infectious disease outbreaks reported in the Eastern Mediterranean Region in 2018 | News | Epidemic and pandemic diseases [WWW Document], n. d. URL <http://www.emro.who.int/fr/pandemic-epidemic-diseases/news/infectious-disease-outbreaks-reported-in-the-eastern-mediterranean-region-in-2018.html> (accessed 8.1.21).
- World Health Organization, n.d. WHO R&D Blueprint: Priority Diagnostics for CCHF Use Scenarios and Target Product Profiles v1.0 14.
- World Health Organization, n.d. WHO EMRO | Infectious disease outbreaks reported in the Eastern Mediterranean Region in 2018 | News | Epidemic and pandemic diseases [WWW Document]. Epidemic and pandemic-prone diseases. URL <http://www.emro.who.int/fr/pandemic-epidemic-diseases/news/infectious-disease-outbreaks-reported-in-the-eastern-mediterranean-region-in-2018.html> (accessed 11.24.21b).
- Yang, Y., Sun, W., Guo, J., Zhao, G., Sun, S., Yu, H., Guo, Y., Li, J., Jin, X., Du, L., Jiang, S., Kou, Z., Zhou, Y., 2015. In silico design of a DNA-based HIV-1 multi-epitope vaccine for Chinese populations. *Hum. Vaccin. Immunother.* 11, 795–805. <https://doi.org/10.1080/21645515.2015.1012017>.
- Yano, A., Onozuka, A., Asahi-Ozaki, Y., Imai, S., Hanada, N., Miwa, Y., Nisizawa, T., 2005. An ingenious design for peptide vaccines. *Vaccine* 23, 2322–2326. <https://doi.org/10.1016/j.vaccine.2005.01.031>.
- Zhu, K., Day, T., Warshaviak, D., Murrett, C., Friesner, R., Pearlman, D., 2014. Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins* 82, 1646–1655. <https://doi.org/10.1002/prot.24551>.
- Zivcec, M., Safronetz, D., Scott, D.P., Robertson, S., Feldmann, H., 2018. Nucleocapsid protein-based vaccine provides protection in mice against lethal Crimean-Congo hemorrhagic fever virus challenge. *PLoS Negl. Trop. Dis.* 12, e0006628. <https://doi.org/10.1371/journal.pntd.0006628>.
- Zwanzig, R., 1973. Nonlinear generalized Langevin equations. *J. Stat Phys* 9, 215–220. <https://doi.org/10.1007/BF01008729>.