Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

European Journal of Pharmaceutical Sciences 151 (2020) 105387

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Design of a Novel Multi Epitope-Based Vaccine for Pandemic Coronavirus Disease (COVID-19) by Vaccinomics and Probable Prevention Strategy against Avenging Zoonotics

Sajjad Ahmad, Afifa Navid, Rabia Farid, Ghulam Abbas, Faisal Ahmad, Naila Zaman, Nousheen Parvaiz, Syed Sikander Azam*

Computational Biology Lab, National Center for Bioinformatics (NCB), Quaid-i-Azam University, Islamabad, 45320, Pakistan.

ARTICLE INFO

Keywords:
Coronavirus disease (COVID-19)
Vaccinomics
Non-structural protein 8
3C-like proteinase
Spike glycoprotein

ABSTRACT

The emergence and rapid expansion of the coronavirus disease (COVID-19) require the development of effective countermeasures especially a vaccine to provide active acquired immunity against the virus. This study presented a comprehensive vaccinomics approach applied to the complete protein data published so far in the National Center for Biotechnological Information (NCBI) coronavirus data hub. We identified non-structural protein 8 (Nsp8), 3C-like proteinase, and spike glycoprotein as potential targets for immune responses to COVID-19. Epitopes prediction illustrated both B-cell and T-cell epitopes associated with the mentioned proteins. The shared B and T-cell epitopes: DRDAAMQRK and QARSEDKRA of Nsp8, EDMLNPNYEDL and EFTPFDVVR of 3C-like proteinase, and VNNSYECDIPI of the spike glycoprotein are regions of high potential interest and have a high likelihood of being recognized by the human immune system. The vaccine construct of the epitopes shows stimulation of robust primary immune responses and high level of interferon gamma. Also, the construct has the best conformation with respect to the tested innate immune receptors involving vigorous molecular mechanics and solvation energy. Designing of vaccination strategies that target immune response focusing on these conserved epitopes could generate immunity that not only provide cross protection across Betacoronaviruses but additionally resistant to virus evolution.

1. Introduction

A recent outbreak of pneumonia in Wuhan, China, is associated with Betacoronavirus of group 2B from family Coronaviridae and the order Nidovirales [1] [2]. The viruses are positive-sense RNA, enveloped and non-segmented [2]. This coronavirus disease (COVID-19) is known as a third human zoonosis of the 21st century and is caused by a new strain not previously identified in humans [1]. The coronaviruses causing minor infections of the respiratory tract in humans are NL63, OC43, hCoV-229E, and HKU1 while, the lethal coronavirus infections that emerged in this century are the Middle East respiratory syndrome coronavirus (MERS-CoV), severe acute respiratory syndrome coronavirus (SARS-CoV) and the recent SARS-CoV 2 or COVID-19 [2]. The source of the COVID-19 is still not confirmed but some evidence suggests that the source may be in the seafood market of Huanan in Wuhan, China [3] [4]. The Center for Disease Control and Prevention (CDC) reported that the recent COVID-19 is caused by Betacoronavirus just like the previous two outbreaks of coronaviruses; MERS and SARS, the source of which is camels and bats, respectively [5]. The first transmission of CoV from animals to humans was notified in 2002 causing SARS-CoV with a 10% mortality rate [6]. It was suggested that the virus needs some intermediate reservoir to infect humans efficiently as confirmed later by a thorough investigation revealing palm civets and raccoon dogs of the wet market carried SARS-CoV viral RNA and might act like intermediate reservoirs [3].

The COVID-19 RNA virus carries a high mutation rate and ability to transfer from person to person as compared to other coronaviruses. According to the World Health Organization (WHO), till 23rd February 2020, the COVID-19 affected a total of 78,811 individuals across globally of which, 77,042 were reported in China while 1769 were reported in other countries. The death toll in China is 2445 and 17 deaths have been reported in the rest of the world. According to the reports till 2nd January 2020, 41 patients have been admitted to the hospital of which most of the patients were men and 66% of them had exposure to the Huanan seafood market and the median age of patients was 49 years [2]. The health-care workers are also diagnosed with the

* Corresponding author: National Center for Bioinformatics, Quaid-i-Azam University, Islamabad 45320, Pakistan
E-mail address: syedazam2008@gmail.com (S.S. Azam).

https://doi.org/10.1016/j.ejps.2020.105387
Received 14 March 2020; Received in revised form 22 April 2020; Accepted 19 May 2020
Available online 23 May 2020
0928-0987/

infection including those working in similar wards [6]. The time taken by COVID-19 to infect other individuals is similar to that of SARS. It is estimated that on average each infected person infects 2-3 persons and this occurrence increases two-folds, every 6.4 days [7]. It was observed that people with mild infection are more actively spread the infection [8]. Due to critical nature of the outbreak, the virus was sequenced on urgent basis and the first sequence was available on 10th January, 2020 online at virological.org. It was noted that COVID-19 has much resemblance to SARS-CoV at the genomic level [4]. Symptoms of COVID-19 include fever, dry cough, shortness of breath and dyspnea, sore throat and leukopenia. To date no vaccine COVID-19 is available and is need of an hour to develop a vaccine to prevent further spread of the disease. To this end, immuno-informatics can be applied to a complete protein data set of the virus for deep antigen analysis and thus can save time and cost for designing a vaccine against COVID-19. This will ease the early development of a vaccine and proposed design can be subjected immediately to experimental trials.

2. Material and Methods

The stepwise flow of the methodology followed to design a vaccine against COVID-19 is illustrated in Figure 1Fig. 1.

2.1. Prioritization of Potential COVID-19 Vaccine Candidates

The complete dataset of proteins available in (NCBI) [9] coronavirus data hub (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType_s=Nucleotide&VirusLineage_ss=Wuhan%20seafood%20market%20pneumonia%20virus,%20taxid:2697049&utm_campaign=wuhan_nCoV&utm_source=insights&utm_medium=referral) was retrieved and subjected to screening phase to identify potential vaccine candidates. First, host non-similar proteins of the pathogen were filtered that show no homology to the human host (taxonomic id: 9606). Proteins having sequence E score $< 1.0\ E^{-5}$, bit score > 100, and sequence identity ≤ 30 % were selected [10]. Again, a BLASTp search against the Mouse (Mus musculus, taxonomy id: 10090) was performed using host non-similar proteins keeping the input parameters E score cut-off 0.005, bit score > 100, and identity $< 30\%$ [11]. The screened mouse non-similar proteins were then subjected to TMHMM 2.0 [12] and HMMTOP 2.0 [13] for observing the number of transmembrane helices. Proteins having less than two transmembrane helices were subjected to SPAAN [14] for predicting the adhesive proteins as they have the potential to facilitate attachment to the host tissues [15]. Selected vaccine candidates were then used in BLASTp tool to align the selected adhesive protein candidates with the probiotic bacteria proteome including three Lactobacillus species: L. rhamnosus (tax id: 47715), L. johnsonii (tax id: 33959), and Lactobacillus casei (tax id: 1582) to avoid accidental inhibition of the useful gut bacteria [11].

2.2. B and T-cell Epitopes Mapping

Selected vaccine candidates were then subjected to the immune epitope database (IEDB) Bepipred Linear Epitope Prediction 2.0 [16–18]. The threshold of 0.5 was used for the prediction of linear B-cells epitopes which were then utilized in T-cell epitopes mapping to

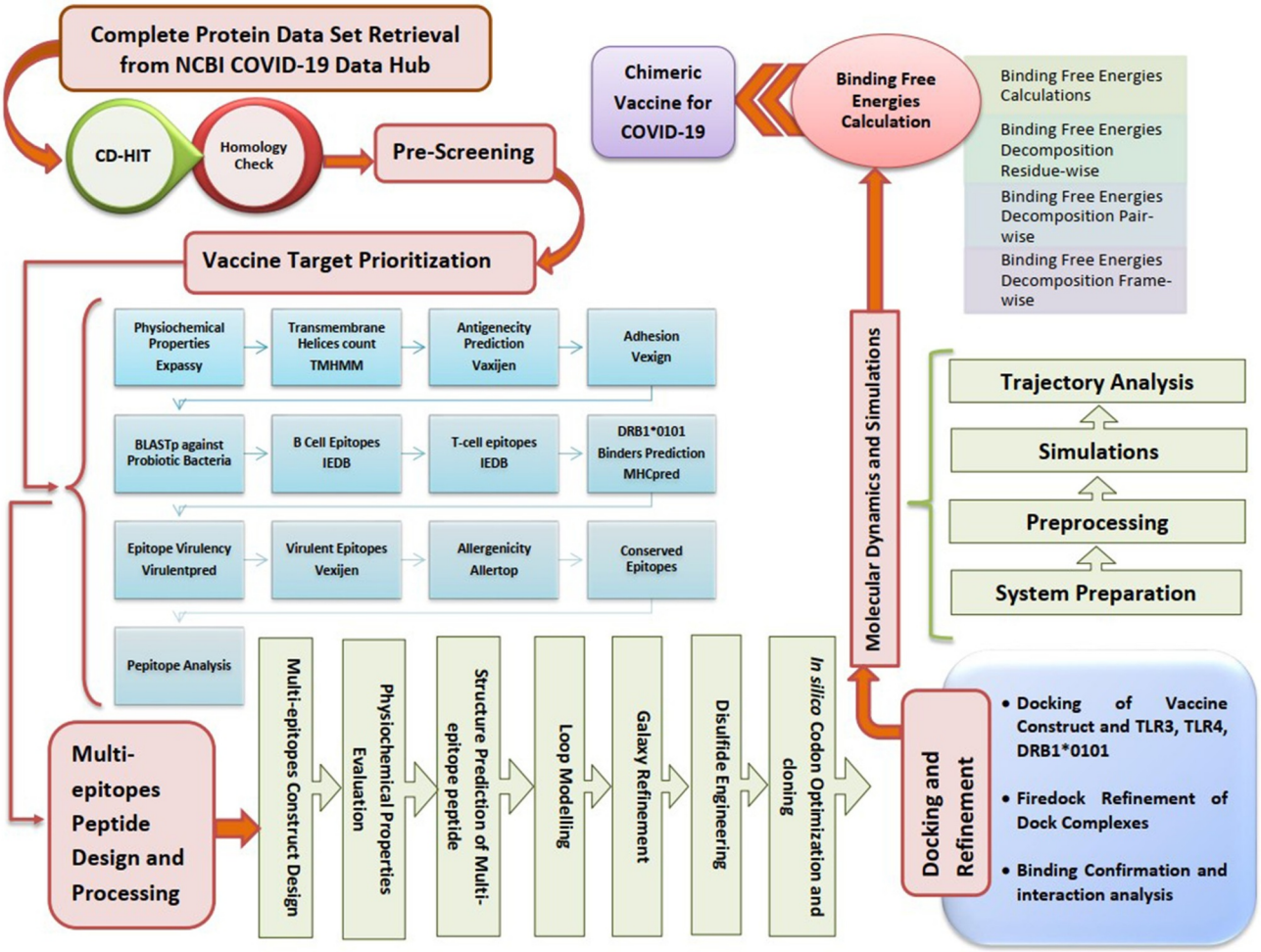

Fig. 1. Designed workflow for in silico vaccine engineering against COVID-19.

identify subsequences with the potential to bind reference set of major histocompatibility complex (MHC) class I and II alleles [16]. Epitopes were ranked according to their percentile score, the ones with the low percentile were considered as high affinity binders. The selected B-cell derived T-cell epitopes were then subjected to MHCPred 2.0 [19] analysis for interpreting their binding affinity potential. The cut-off criterion was set to IC_{50} values < 100 nM for DRB*0101[20]. Following this, the VirulentPred [21] and VaxiJen 2.0 [22] were utilized to validate the virulence and antigenicity of the selected epitopes, respectively. AllerTOP 2.0 [23] was applied to remove the allergic epitopes. CLC main workbench was employed to check the conservation of non-allergic epitopes required for designing an effective broad-spectrum vaccine.

2.3. Construction of Multi-Epitopes Peptide (MEP) and Adjuvanting

Overlapping immunodominant epitopes were used to construct a multi-epitope peptide (MEP) which is considered a promising strategy to stop viral infections [24]. One of the key issues with the design of peptide vaccine is its weak immunogenicity that can be resolved by designing a MEP with appropriate adjuvants [25]. In the current study, MEP was designed using AAY linkers to combine screened multiple epitopes [26]. To the N-terminal of MEP, an adjuvant in the form of B subunit of cholera toxin was linked to the MEP thus creating a multi-epitope peptide vaccine construct (MEPVC) [27]. The tertiary structure of the MEPVC was created through a software called 3Dpro of SCRATCH protein predictor [28], I-tasser [29], and Swiss-Model [30]. The best model was further loop modelled using GalaxyLoop [31] and refined using GalaxyRefine [32] of GalaxyWeb. To improve the construct's stability, disulfide bonds were introduced in the structure [33] using Design 2.0 [34]. The sequence of the MEPVC was translated in reverse and then optimized for codon usage according to the Escherichia coli, which will end up in the increased expression of the MEPVC sequence cloned in the mentioned expression system [35]. The entire activity was accomplished using Java Codon Adaptation Tool (JCat) server [36]. In order to assess the expression of sequence that have been cloned, the GC content and codon adaptation index (CAI) were measured. The value of 1 CAI is contemplated ideal [37,38] whereas the appropriate GC content should be fluctuated between 30-70% due to favorable transcriptional and translational efficiencies [27]. There were other input factors carefully calculated to prevent rho-independent transcription termination, the binding sites of the prokaryotic ribosome, and the cleavage sites of restriction enzyme. As the final step of this phase, the cloning of the engineered construct was carried out into pET-28a (+) expression vector using SnapGene (https://www.snapgene.com/).

2.4. Physicochemical Characterization of MEPVC

ProtParam tool [39] was applied for analyzing the physical and chemical properties of MEPVC such as amino acid composition, estimated half-life, instability index, extinction coefficient, theoretical pl, atomic composition, molecular weight, and grand average of hydropathicity (GRAVY) to assist experimental studies. Instability index is one of the key parameters that is significantly considered as it helps in discarding the unstable protein candidates (protein instability index > 40).

2.5. MEPVC Immune Simulation

In this step, the vaccine construct underwent immune response profiling and immunogenicity classification, which was done using the C-ImmSim server [40]. In order to predict the immune epitopes, a position-specific scoring matrix (PSSM) employed by C-ImmSim was used. Whereas, the different machine learning procedures were used to forecast the immune connections. This server is concurrently used to execute an immune simulation for 3 compartments such as bone marrow, tertiary lymph nodes, and thymus [41]. Default simulation parameters were used which are as follows: random seed (12345), simulation steps (100), simulation volume (10), host HLA selection (A MHC class I A0101 allele, B MHC class I B0702, DR MHC class II DRB1_0101 allele), and time step of injection was set to 1.

2.6. MEPVC Docking and Refinement

The technique of molecular docking was utilized to predict conformation of the MEPVC with respect to a suitable innate immune receptor. This analysis plays a significant role to determine the high-affinity contacts amid the vaccine construct and the immune receptor. The PDB ID: 4G8A and TLR3 PDB ID: 2A0Z were retrieved from the Protein Data Bank (PDB) for the TLR4 and TLR3, respectively. Blind docking employed to calculate the regular pose of the vaccine construct with the mentioned receptors using an online PATCHDOCK server [42]. The resulting structures were further refined using the Fast Interaction Refinement in Molecular Docking (FireDock) [43]. The top ranked complex with the minimum global energy was selected and used to analyse the binding pose and intermolecular connections using UCSF Chimera 1.13.1 [44], Discovery Studio (DS) Visualizer 17.2.0 [45].16349, and Visual Molecular Dynamics (VMD) 1.9.3 software [46].

2.7. Simulations of MEPVC-Innate Immune Receptors

In order to gain insights into the dynamics of the vaccine construct with the receptors, molecular dynamics (MD) simulations have been carried out. The simulation analysis was also important to validate the exposure of epitopes towards the host structure for identification and handling of a substantial outcome. The MD simulations took place in three different phases: system preparation, pre-processing and production [47] with an Assistant model building with Energy Refinement (AMBER) 16 [48]. The antechamber program [49] was used to build the libraries and parameters for the TLR4 and vaccine construct. The TIP3P solvation box (size 12 Å) was inserted to solvate the construct. To study the intermolecular interactions, force field, ff14SB [50] was used whereas, the system was neutralized by the addition of Na^+ counter ions. In the second phase of MD simulations, the energy minimization of the complexes was carried out. Each complex was minimized using the following steps: the energy minimization of hydrogen atoms (500 cycles), energy minimization of water box (1000 cycles, control of 200 kcal/mol – $Å^2$ on rest of the system), minimization of the whole atoms of the system (1000 cycles with the restraint of 5 kcal/mol –$Å^2$ on Cα atoms), and the rest of the system was subjected to non-heavy atoms minimization with 300 cycles and restraint of 100 kcal/mol –$Å^2$. In the next step, the system was gradually heated from 0 K to 300 K with the time step of 2 femtoseconds and restraints of 5 kcal/mol –$Å^2$ on Cα atoms. In order to sustain the temperature of the system, Langevin dynamics [51] with the gamma value of 1.0 was castoff. The SHAKE algorithm [52] was used to put constraints on the hydrogen bonds of the system for heating. In the next step, systems were equilibrated for 100 ps with a time step of 2 fs followed by pressure equilibrium, which was attained using the NPT ensemble with restraints of 5 kcal/mol – Å 2 on Cα atoms. The same step was extended for 50 ps with the 1 scale down on restraints on carbon atoms. However, the step of system equilibration was carried out for a time scale of 1 nanosecond followed by the production run of 100 ns with a time scale of 2 fs. For the production run, the Berendsen algorithm [53] with the NVT ensemble cast off with a cut-off of 8.0 Å. Simulation trajectories were calculated to investigate the strength of a complex via the CPPTRAJ module [54] of AMBER. However, the visualization of simulation trajectories was done with UCSF Chimera [55], DS Visualizer [45] and VMD [46].

2.8. Estimation of Binding Free Energies

In order to estimate the MMPBSA binding free energies for the receptors and multi-epitope peptide vaccine construct, the MMPBSA.py module [56] of AMBER16 was castoff. The program generated the input files for the complex, receptor and MEPVC molecule using the ante-MMPBSA.py module. To compute the variance between the solvated and un-solvated phases, 100 frames of simulation trajectories were picked and analyzed [57]. For the precise values of binding free energies, the two different conformations were matched to the binding energies of significant residues. To estimate the free binding energy of the anticipated complex, $\Delta G_{bind,\ solv}$ was resolved using the three equations (Eqn 1,Eqn 2,Eqn 3) given below:

$$\mathbf{\Delta G}_{bind,\,solv} = \mathbf{\Delta G}_{bind,\,vaccum} + \mathbf{\Delta G}_{solv,\,complex} - \mathbf{\Delta G}_{solv,\,ligand} - \mathbf{\Delta G}_{solv,\,complex} \quad (1)$$

$$\mathbf{\Delta G}_{solv} = \mathbf{\Delta G}_{electrostatic}(\epsilon_{80-1}) + \mathbf{\Delta G}_{hydrophobic} \quad (2)$$

$$\mathbf{\Delta G}_{vaccum} = \mathbf{\Delta E}_{molecular\ mechanics} - -\mathbf{T}.\ \mathbf{\Delta G}_{normal\ mode\ analysis} \quad (3)$$

The net free binding energy was then decomposition into each residue to highlight the interacting and stable residues.

3. 3. Results and Discussions

3.1. Retrieval of protein sequences

The NCBI most recently dedicated a coronavirus disease data hub containing all nucleotide and protein sequences information published from across the world. In this study, we aimed at in silico prioritization potential vaccine candidates and designing a chimeric peptide vaccine for COVID-19 based on all available protein sequences in the data hub. Several bioinformatic and immunoinformatics techniques are employed with the aim to assist experimentalists in vaccine development against the virus.

3.2. Identification of Potential Vaccine Candidates

Prioritization of potential vaccine candidates could help in minimizing time, labor cost and resources for developing and optimizing the success of getting an effective vaccine against the pathogen. In total, 193 protein entries were retrieved (S-Table 1) and analyzed first for sequence homology with the human host proteome. This was significant to evaluate as homology between virus protein (s) to be used in vaccine designing and the host is likely to cause strong autoimmune reactions in the host [58]. This check identified two proteins: (orf1a polyprotein (Accession id, YP_009725295) and nsp3 (Accession id, YP_009725299) as host homologous thus discarded from further evaluations. Experimental studies of the human vaccines are usually done in mice because of the many practical advantages they provide compared to vaccine research in higher animal models [22]. The appropriateness of mice as an animal model for vaccine research is defined by its ability to reproduce relevant human physiology [59]. Considering this, a homology check was devoted to the pipeline and applied to the filtered human non-similar proteins (191 in number) to ensure the selection of non-similar mice proteins. This assessment resultant into shortlisting of Orf1ab polyprotein (Accession ids, QHQ82463, QHD43415, QHR63279, QHR63259, QHQ71962, QHO62106, QHQ71972, QHR63249, QHO60603, QHO62111, QHN73794, QHR84448, QHR63269, QHN73809, QHR63289, and QHO62876) and nsp3 (Accession id, YP_009724389) as mice similar proteins. The mice non-similar proteins will aid in discouraging false positive results during in vivo experimentations and accurate interpretation of immune protective efficiency of the prioritized vaccine candidates against the virus [22,60]. Next, enumeration of transmembrane helices in the pooled 174 mice non-similar proteins was accomplished. This transmembrane topology characterization is deemed vital in relatedness to the afterward experimental expression studies of the proteins. Protein with transmembrane helices less than 2 in numbers are often considered as best vaccine candidates as multiple helices make recombinant proteins purification and expression difficult in vaccine development [20]. This result 33 proteins spanning across five types, and include: Orf3a polyprotein (Accession id, QHQ71974, QHR84450, QHD43417, QHQ71964, QHQ82465, QHO60595, YP_009724391, QHN73811, QHN73796, QHO62878), nsp4 (Accession id, YP_009725300), membrane glycoprotein (Accession id, QHD43419, QHQ71966, QHQ71976, QHQ82467, YP_009724393, QHO60597, QHN73813, QHN73798, QHO62880, QHR84452), membrane protein (Accession id, QHR63283, QHR63263, QHR63273, QHR63293, QHR63253), matrix protein (Accession id, QHO62109, QHO62114), and nonstructural protein NS3 (Accession id, QHR63251, QHR63281, QHR63261, QHR63271, QHR63291) containing multiple helices therefore not proceeded further. The creation of adhesin-based vaccines is considered an attractive and effective strategy and is being explored as a solution to number of infectious pathogens [61]. The idea behind exploiting adhesin for a vaccine is based on the promising preclinical findings. The aim is to confer protective immunity via two main mechanisms: (i) opsonization driven by opsonising antibodies that is capable of binding the target antigen as an immunological tag leading to activation of other components of the host immune system for enhance recognition of the pathogen and subsequent complement system activity and virus killing by phagocytosis, (ii) neutralization driven by adhesin-specific antibodies that block virus binding ability to host tissues. The adhesion probability computation revealed 26 protein to have adhesion probability value greater than a threshold as tabulated in S-Table 2 and can be ideal putative vaccine candidates against COVID-19. The adhesin probability of protein ranges from 0.593 to 0.796 (mean, 0.645).Antigenicity of proteins was predicted to reflect their ability of binding to products of adaptive immunity: antibodies or T-cell receptors. In total, 7 proteins: nsp8, nsp9, nsp10, 3C-like proteinase, spike glycoprotein, surface glycoprotein, and ORF1ab polyprotein were recognized as antigenic and scored higher than the threshold. Coronavirus nsp8 suggested having diverse activities, including template-dependent RNA polymerase activities, canonical RNA-dependent RNA polymerases, cofactor function of nsp8 for nsp12-mediated RNA-dependent RNA polymerase activity, and metal ion-dependent RNA 3′ polyadenylation activities [62]. Nsp9 is a non-structural protein 9, key to coronavirus replication and is a single-stranded RNA-binding protein [63]. Nsp10 is a critical cofactor that switches on multiple enzymes in replication cycle. It is known to interact with nsp14 and nsp16 subunits activating their respective 3′-5′ exoribonuclease and 2′-O-methyltransferase functions [64]. The 3C-like proteinase is main cysteine protease and is nonstructural protein number 5 (nsp5) and essential in

Table 1
The final set of selected B-cell derived T-cell epitopes for the potential three vaccine candidates against COVID-19.

Protein	Common B and T-cell Epitopes	Antigenicity (cut off score, 0.4)	MHCphred (IC50 score, 100 nM)	Allergenicity	Virulentpred (cut off score, 0.5)
Nsp8	DRDAAMQRK	0.8641	35.81	non-allergen	1.0606
	QARSEDKRA	0.5770	24.15	non-allergen	1.0606
Proteinase	EDMLNPNYEDL	1.0913	21.68	non-allergen	1.0600
	EFTPFDVVR	1.6049	5.07	non-allergen	1.0604
Spike Glycoprotein	VNNSYECDIPI	1.0996	23.93	non-allergen	1.0593

Table 2
Top 10 refined models of the MEPVC along with the initial input structure.

Model	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored	GALAXY energy
Initial	0.000	3.312	95.9	3.8	93.6	20995.70
MODEL 1	1.734	1.134	1.2	0.0	95.7	-3784.31
MODEL 2	1.920	1.444	2.4	0.0	93.6	-3750.72
MODEL 3	1.766	0.997	0.6	0.0	95.7	-3743.81
MODEL 4	1.667	0.997	0.6	0.0	95.7	-3739.43
MODEL 5	2.611	1.187	0.9	1.3	95.2	-3731.44
MODEL 6	1.782	1.187	0.9	1.3	95.2	-3724.93
MODEL 7	2.907	1.144	1.5	0.0	96.3	-3722.41
MODEL 8	1.999	1.459	3.0	0.6	94.7	-3718.59
MODEL 9	3.194	1.192	1.8	0.0	96.3	-3715.99
MODEL 10	2.235	1.310	2.7	0.6	96.3	-3713.14

mediating cleavage of nsp4 to nsp16 [65]. The trimeric transmembrane coronavirus spike glycoprotein initiates infectious cycle by binding to a specific receptor on the host membrane followed by viral fusion [66]. The surface glycoprotein was analyzed to be different in a sequence patch (Leu3-Phe11) at the start of the spike glycoprotein. The ORF1ab is replicase polyprotein cleaved by papain-like protease and 3C-like protease at specific cleavage sites to yield 15 to 16 non-structural proteins (nsps) [67]. The final numbers of potential vaccine candidates obtained in this step by step subtraction phase are presented in Figure 2.

3.3. B and T-cell Epitopes Mapping

Identification of epitopes in given antigens is vital for a number of practical reasons, including understanding etiology of a disease, monitoring of immune system, development of diagnostic assays, and epitope-based vaccines designing [68]. From vaccine designing point of view, host adaptive immunity is highly specific and is able to recognize and destroy the invading pathogen [69]. Additionally, adaptive immunity is able to remember the pathogens, creating long-lasting pathogen-specific protective memory enabling stronger attacks against the pathogen reencountered on successive times [70]. This arm of host immune system is driven by lymphocytes of two types: B and T-cells responsible for the humoral and cell-mediated immunity, respectively [71]. Both cells recognize pathogen molecular components called antigens. The antigens interact with specific receptors present on the surface of B and T-cells. The activation of both these cells required antigen recognition by these receptors, in addition, to the second activation signals from the innate system. The vaccine candidates prioritized in the first phase were deeply investigated for B-cell epitopes. Different lengths of linear B-cell epitopes were predicted for each protein and only recurrent epitopes simultaneously predicted by different servers were selected for chimeric vaccine designing. The B-cell epitopes predicted for the vaccine candidates were in the following order: nine for Nsp8 and 3C-like proteinase, five for Nsp9, eight for Nsp10, 34 for spike glycoprotein and surface glycoprotein, and four for ORF1ab polyprotein| partial. These B-cells epitopes are recognized as solvent-exposed antigens through B-cell receptors (BCR) and upon activation, B-cells secrete antibodies. Antibodies have different functions including neutralizing pathogens, toxins and labeling pathogens for destruction [72]. The proteins were also analyzed for T-cell epitopes through very stringent criteria of p-value less than 0.005. The epitopes were of different lengths and interact with several different alleles of MHC-I and MHC-II. For Nsp8, 94 epitopes predicted whereas Nsp9, Nsp10, 3C-like proteinase, spike glycoprotein, surface glycoprotein and ORF1ab polyprotein| partial protein were mapped for 61, 76, 153, 657, 651, and 44, respectively T-cell epitopes. These epitopes are presented on the surface through specific receptors known as T-cell receptor (TCR) allowing recognition of these antigens when displayed by antigen-presenting cells bound to MHC molecules [73]. Epitopes presented by MHC I are recognized by CD8 (cytotoxic T lymphocytes) [74] whereas those presented by MHC II are recognized by CD4 T-cells [75]. The CD4 T-cells later become helper T-cells that amplify the immune responses against the pathogen. Comparative analysis of the predicted epitopes was further carried out to select epitopes that are common to B-cell, CD4 T-cell and CD8 T-cell alleles in order to design a specific, effective, and strong vaccine. On this basis, three epitopes from Nsp8 (DRDAAMQRK, QARSEDKRA, EQAVANGDSEV), none for Nsp9 and ORF1ab polyprotein| partial protein, four for Nsp10 (GCSCDQLREP, YLASGGQPIT, YLASGGQPI, TVTPEANMDQESFG), three for 3C-like proteinase (EDMLNPNYEDL, KYNYEPLTQDHV, EFTPFDVVR), five for spike glycoprotein (RVYSTGSNVFQ, VNNSYECDIPI, LADAGFIKQYGDCLG, GQSKRVDFC, RNFYEPQIITTD) and surface glycoprotein (RVYSTGSNVFQ, VNNSYECDIPI, LADAGFIKQYGDCLG, GQSKRVDFC, RNFYEPQIITTD). The epitopes were then reevaluated in antigenicity check to make

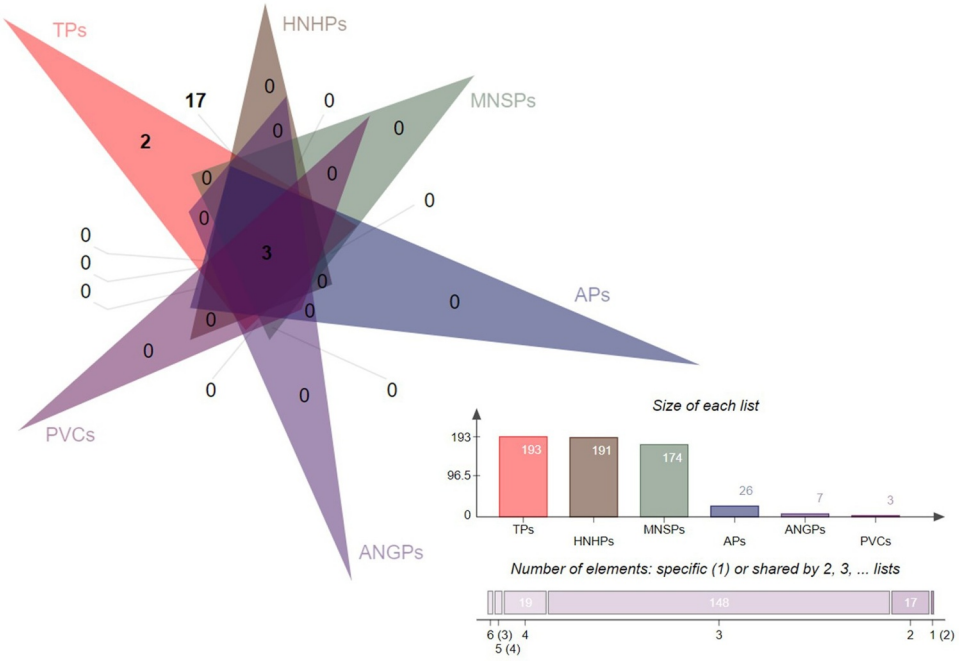

Fig. 2. The final set of potential vaccine candidates filtered through several in silico checks presented as interactive Venn. The total number of proteins (TPs), filtered to 191 host non-homologous proteins (HNHPs), 174 mouse non-similar proteins (MNSPs), 26 adhesive proteins (APs), 7 antigenic proteins and finally to 3 potential vaccine candidates (PVCs).

sure their binding potential of binding to immune cells. For Nsp8, DRDAAMQRK and QARSEDKRA were found antigenic with score higher than the default threshold of 0.4 whereas the third epitope EQAVANGDSEV was found non-antigenic hence removed. All the four epitopes of Nsp9 revealed non-antigenic therefore not processed further. In case of 3C-like proteinase, KYNYEPLTQDHV was found non-antigenic whereas EDMLNPNYEDL and EFTPFDVVR were antigenic therefore considered in afterward analysis. The spike glycoprotein contains epitopes VNNSYECDIPI, and RNFYEPQIITTD as antigenic whereas none of the surface glycoprotein provided epitopes were antigenic. Following, the affinity of the filtered antigenic epitopes for the most prevalent DRB*0101 allele in humans was evaluated through IC_{50} value and those with value < 100 nM were classified as high affinity binders. All the pooled antigenic epitopes were found to have great ability of binding to the mentioned allele. Similarly, these epitopes were evaluated in allergenicity and virulent potential check and only virulent and non-allergen were selected. Virulent check was significant in ensuring selection of epitopes mediating infectious pathways in the host. The final selected epitopes that cleared all these checks are tabulated in Table 1.

3.4. Construction of MEPVC

A MEP was constructed first comprising epitopes finalized in the previous phase. MEP based vaccines are considered an ideal approach to prevent and treat viral infections [24]. The epitopes shown in Table 1 were linked to each other through flexible AAY linkers as such it allows efficient separation required for the effective working of each epitope. Once the MEP was designed, to its N-terminus an adjuvant of Cholera toxin subunit B (CTB) was added [76]. The schematic representation of the MEPVC is shown in Figure 3. CTB is nontoxic part of cholera toxin and is considered an accelerator in protective immunity and a break in auto-immunity. It shows high affinity for monosialotetrahexosylganglioside displayed on variety of cell types, including gut epithelial cells, antigen-presenting cells (dendritic and macrophages) and B-cells [77]. CTB is a preferred choice as an adjuvant because of its ability of self-expression in variety of organisms and can be coupled to antigens through several approaches involving chemical manipulation and genetic fusion resulting in strong immunological responses against the antigens to which it is attached. The vaccine construct was then used in a comparative 3D structure prediction to ensure confidentiality in selection of the most suitable model for the construct with minimum structural errors.

3.5. Evaluation of Physicochemical Properties

Several different physicochemical properties of the MEPVC were deduced from its sequence. The vaccine construct is 189 amino acids long with total number of 2984 atoms. The molecular weight of the construct is ideal i.e. 21.36 kD as small size construct is easy to handle and purify during experimental evaluation. The construct has instability index value of 35.43, signifying its high stability. The aliphatic index computed for the construct is 79.68, reflecting high thermostability. The estimated half-life in mammals, yeast, and Escherichia coli is 30 hours, > 20 hours, and > 10 hours, respectively. The Grand average of hydropathicity (GRAVY) score is -0.315 which highlights hydrophilic nature of the construct. The theoretical pI is 6.10 pointing to construct slightly acidic nature.

3.6. MEPVC Secondary and Tertiary Structure

The secondary structure elements of the vaccine construct can be divided into the following order: alpha helix (56.08 %), 3_{10} helix (0 %), Pi helix (0 %), beta bridge (0 %), extended strand (16.40 %), beta-turn (7.94 %), bend region (0 %), and random coil (19.58 %). Compared to the I-tasser, phyre2, and Swiss-model, the 3Dpro predicted structure was determined as the most suitable structure based on the complete modeling of the given length of the amino acid sequence. Loop modeling was done at Leu29-Gln37, Ser51-Gln70, Glu72-Gln77, Glu57-Ser76, Val103-Thr113, and Glu167-Asp186. The model was refined to minimum RMSD of 1.734 Å and molprobity score of 1.134 that is quite low compared to the original structure score of 3.312, reflecting good quality of the modelled structure. Similarly, the clash score in contrast to the original structure is 94.7 times lower demonstrating steric clash free structure. The galaxy energy of the structure is very stable (-3784.31) and Ramachandran favored distribution increased from 93.6 to 95.7 %. The top 10 refined models of the MEPVC are tabulated in Table 2. The 3D models of the vaccine construct after loop modelling and refinement is presented in Figure 4.A. The overall Z-score of the modelled structure is -4 and the score is within the range of same size proteins in the pdb illustrating good quality as depicted in Figure 4.B. Refinement of the structure Ramachandran plot demonstrated the construct to contain 93.2 % of its residues in the most favored regions, while 5.7%, 0.0%, 1.1% residues are in additional allowed region, generously allowed region, and disallowed regions, respectively (Figure 4.C). The overall average G-factor of the construct is 0.05.

3.7. Disulfide Engineering of MEPVC

Enhancing protein stability is important in many biomedical applications and is an appealing approach to emulate nature stabilizing molecular interactions [78]. The covalent disulfide bonds provide substantial stability to target proteins and disulfide engineering had achieved considerable success in broad range of applications [33]. In

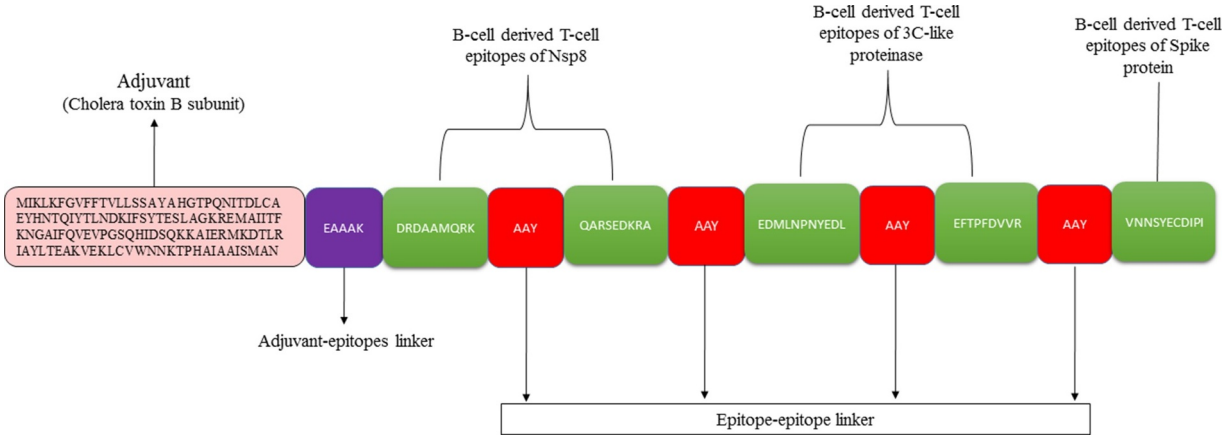

Fig. 3. Schematic representation of the designed MEPVC.

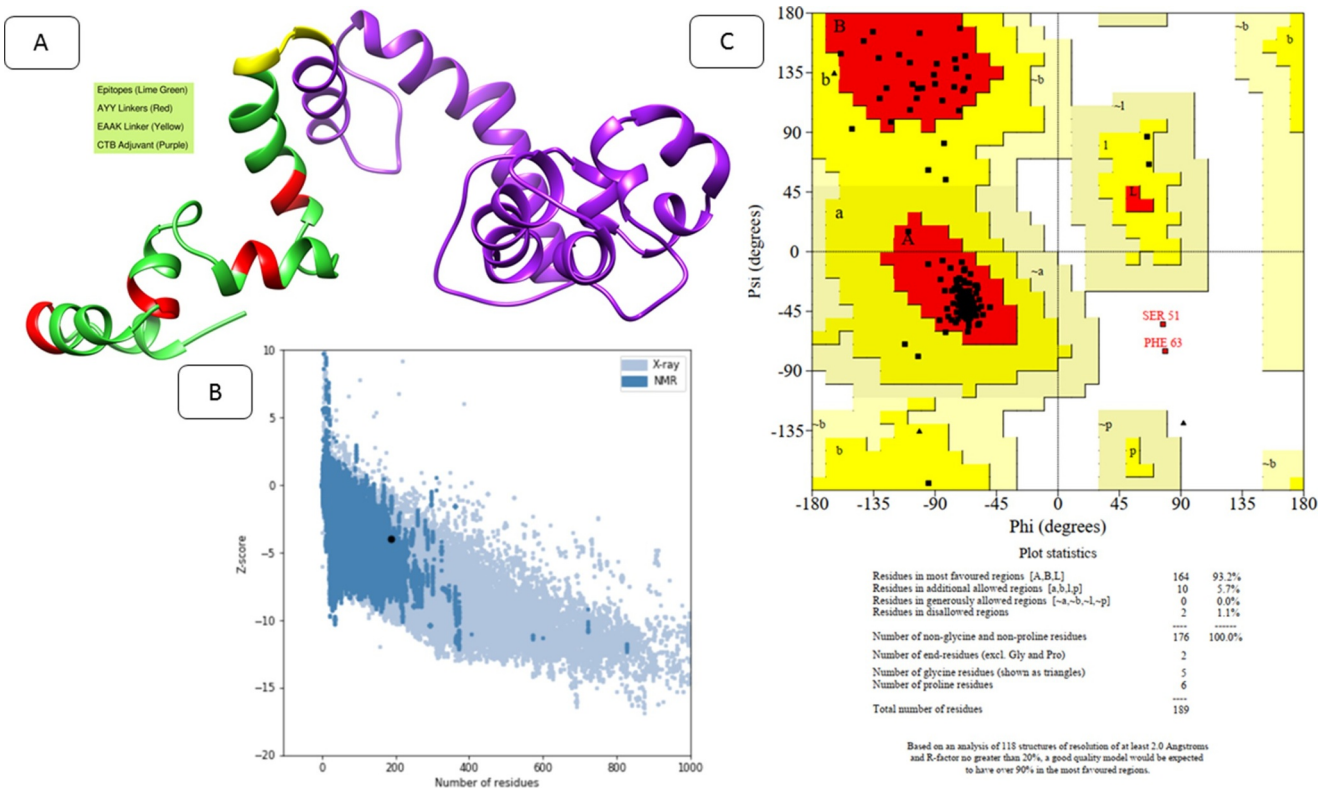

Fig. 4. A. The tertiary structure of the designed MEPVC. The red, green, purple, and yellow colors represent AAY linkers, epitopes, adjuvant, and EAAK linkers, respectively. **B.** The z-score plot indicating the overall good quality of the MEPVC. The z-score of the input MEPVC is shown by a black dot, validating the score within the range revealed for similar size proteins. **C.** Ramachandran plot for the MEPVC illustrating distribution of torsion angles (blue squares) comparative to the core (shown in red) and allowed (shown in brown) regions. Residues in the generously allowed region are shown in dark yellow whereas disallowed regions are depicted in pale yellow.

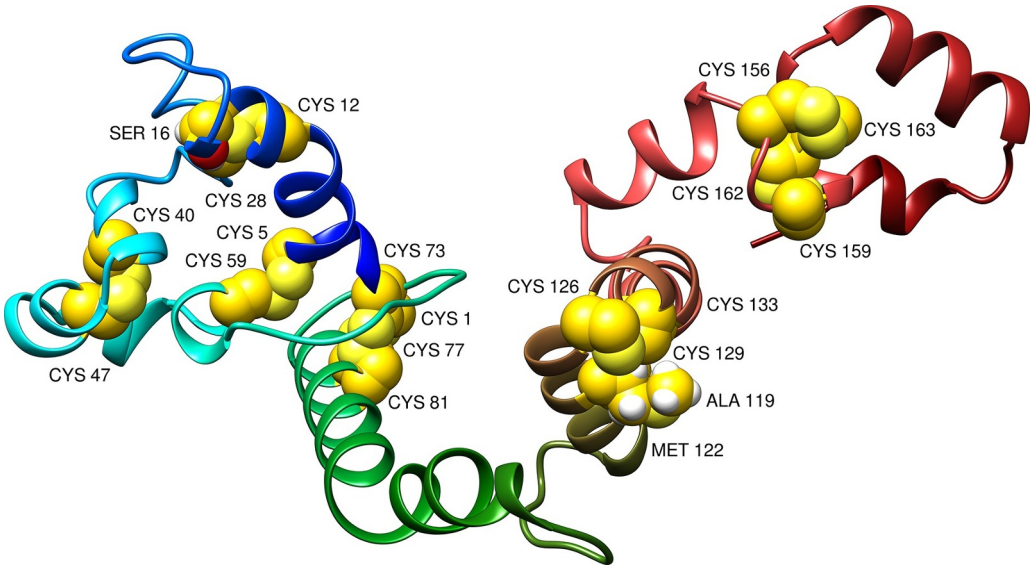

Fig. 5. Disulfide engineering of the MEPVC. The yellow spheres represent mutated residues.

total, 11 pairs of residues were selected for the purpose of disulfide engineering. These include Met1-Ser81, Lys5-Ala59, Val12-Asp28, Ser16-Asp28, Thr40-Ser47, Val73-Gln77, Ala119-Ala133, Met122-Ala133, Ala126-Asp129, Leu156-Leu163, and Asn159-Asp162. The average Chi3 and energy value for the pairs is 12.54 (max, 108.65 and min, -110.64) and 3.12 (max, 4.39 and min, 1.14). The disulfide engineered MEPVC structure is presented in Figure 5.

3.8. Codon Optimization and In Silico Cloning

In the follow up experimental studies, the maximum expression of MEPVC is highly desirable [35]. One requirement for that is the codon usage of MEPVC that must be adapted according to the expression system, for instance, here we used E. coli K12 as a MEPVC expression system. The codon adaptation index (CAI) and GC content revealed for the improved sequence are highly satisfactory with value of 0.96 and

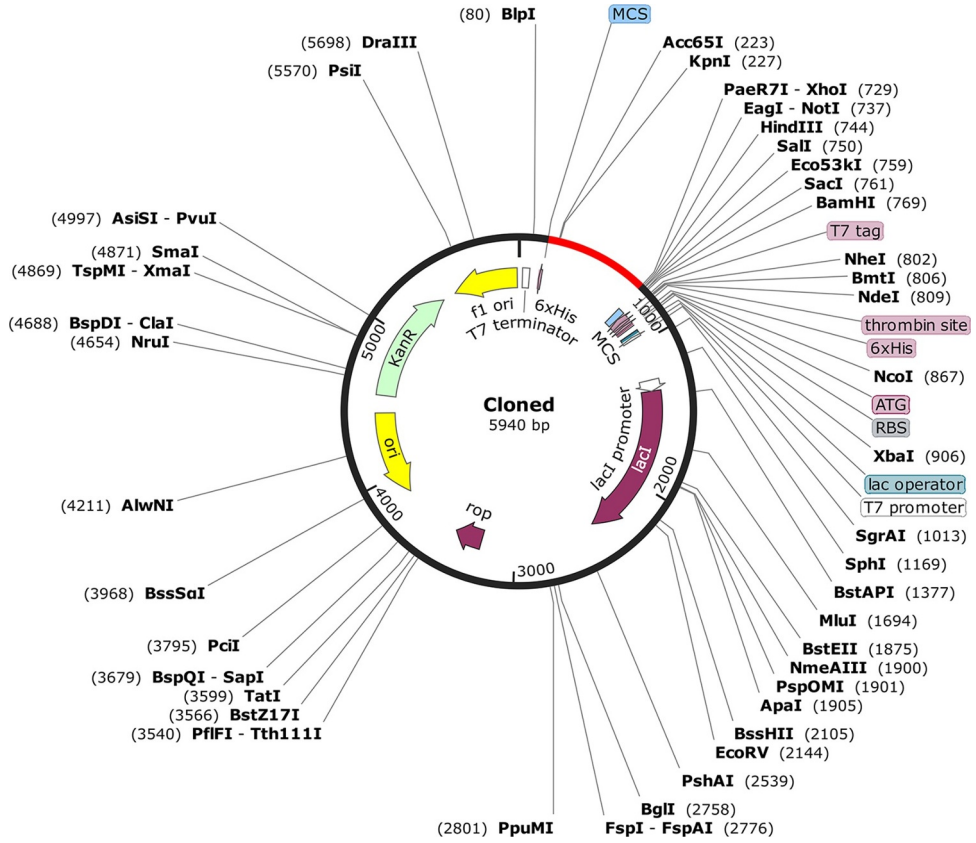

Fig. 6. In silico restriction cloning of the MEPVC into pET28a(+) vector. The insert is shown in red.

48.85, respectively strongly indicating high MEPVC expression. The MEPVC then enclosed on both sites by 6x histidine tag to ease its purification process and inserted at appropriate sites of pET28a(+) vector as shown in Figure 6.

3.9. MEPVC Interactions with Immune Receptors

Molecular interactions and binding conformation of the designed MEPVC with TLR3 and TLR4 innate immune receptors were deciphered via a protein-peptide docking approach. Both TLR3 and TLR4 belong to toll-like receptor family of pattern recognition receptor and function to activate intracellular signaling NF-κB pathway and production of inflammatory cytokines responsible for the development of effective innate immunity [79,80]. These receptors recognize viral associated molecular patterns and induce the production of interferon leading to activation of strong host defense responses. Also, the specific adaptive immunity takes time to establish against antigens therefore it's important to evaluate MEPVC affinity for the innate immune receptors. In case of MRPVC-TLR3 complex, the patch dock predicted 10 best solutions sorted based on the docking geometric shape complementarity score (S-Table 3). A high score implies enhanced affinity of the interacting molecules and best docked conformations of the molecules with respect to each other. Solution 3 was visualized for docked conformations and intermolecular forces responsible for such high affinity of the molecules. The selection was based on the FireDock analysis (S-Table 4) which is an efficient package for refinement and reassigning procedure of docking scores to rigid body docking solutions. The global binding energy of solution 3 is better i.e. -6.39 kJ/mol compared to the rest of predicted solutions. The contribution to the total score form attractive van der Waals (VdW) energy is -12.12 kJ/mol, repulsive (VdW) energy (4.79 kJ/mol), hydrogen bond (HB) energy (-1.63 kJ/mol), and atomic contact energy (ACE) (5.95 kJ/mol). The MEPVC, within 3 Å, was noticed to posed right in the center of the TLR3 receptor (Figure 7)

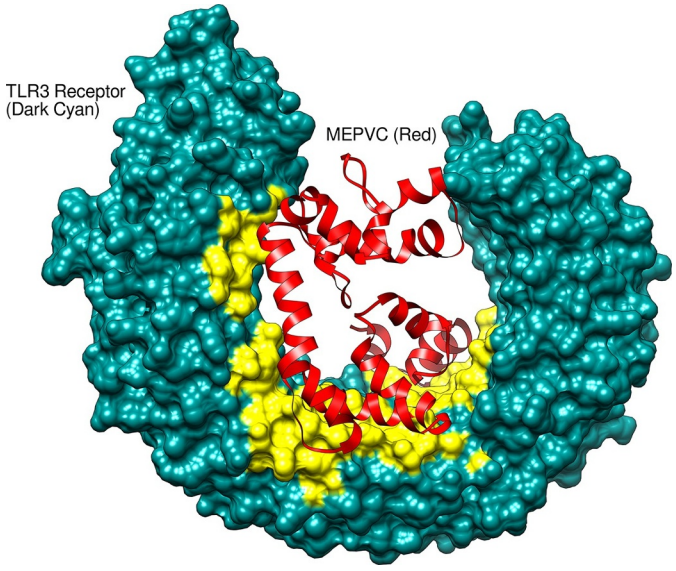

Fig. 7. Binding conformation of MEPVC with respect to TLR3 innate immune receptor. The yellow labeled region pointing to the residues involved in both hydrophobic and hydrophilic interactions with MEPVC.

interacting via hydrogen and hydrophobic bonds with His156, Asp180, Lys201, Glu203, Ser206, Phe227, Asp229, Asp230, Ser254, Ser256, Asp257, Ser282, Tyr283, Asp284, Asp285, Glu301, Tyr302, Phe304, Glu306, Tyr307, Arg325, Glu358, His359, Lys382, Tyr383, P408, His410, Iso411, Gly431, His432, Glu434, Pro408, Asp457, Phe459, Gln483, and Glu533. Similarly, among 10 predicted MEPVC-TLR4 complexes (S-Table 5), solution 7 (S-Table 6) was affirmed as best with total global energy of -10.39 kJ/mol whereas the rest of complexes

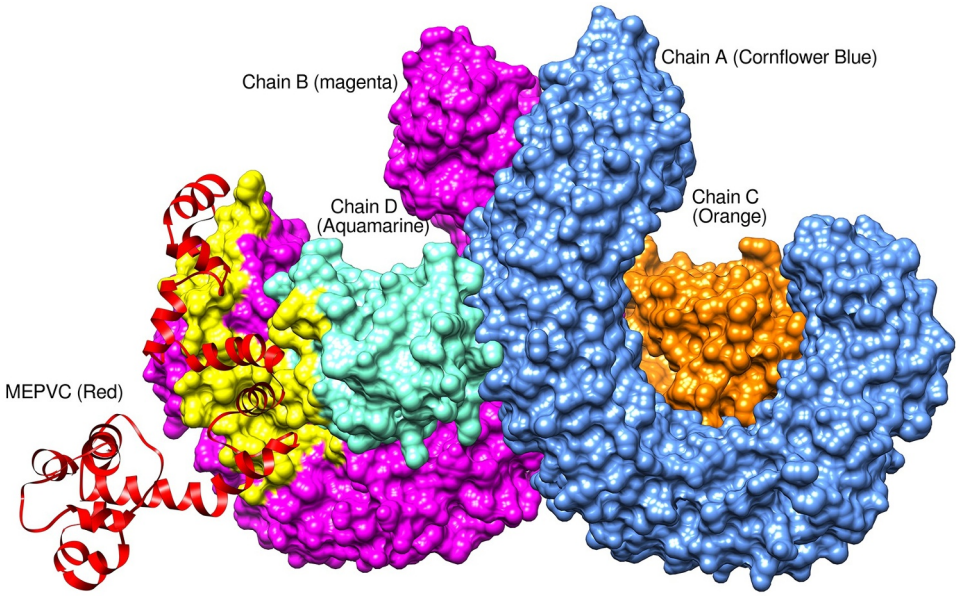

Fig. 8. Binding conformation of MEPVC with respect to TLR4 innate immune receptor. The yellow labeled region pointing to the residues involved in both hydrophobic and hydrophilic interactions with MEPVC.

were noticed as highly unstable with score in positive. The attractive VdW, repulsive VdW, HB and ACE contribution to the global energy is -35.75 kJ/mol, 22.04 kJ/mol, -5.51 kJ/mol, and 12.61 kJ/mol, respectively. Visual analysis of the complex revealed binding of the MEPVC at the interface of chains B and D (Figure 8). The MEPVC is surrounded by chain B residues: Pro23, Glu24, Ser25, Asp44, Lys47, Asp50, Asp51, Arg67, Arg87, Glu89, Pro113, Gln115, Asp137, His159, Asp160, Ser184, and Lys186 and chain D residues: Lys20, Phe64, and Asp114.

3.10. Computational Immune Simulation

The dynamic simulations of the human immune system in response to the designed vaccine construct were deciphered through C-immsim server [40]. The vaccine construct upon administration revealed to generate robust primary immune responses. As can be seen in Figure 9A that combine IgM and IgG antibodies has a titer scale close to 10,000/ml followed by IgM antibody (> 6000 antibody titer per ml). The combined IgG1 and IgG2 and IgG1 were seen to generate high titer scale of around 6300/ml, 3100/ml, and respectively. The IgG2 antibody response revealed to be low throughout post vaccine administration period. The dimerized soluble cytokine IFN-g produced against the antigen is > 400000 ng/ml (Figure 9B).

3.11. Molecular Dynamics Simulation

The stability and dynamics of the designed vaccine construct ensemble docked to innate immune receptors were disclosed through 100-ns of MD production run and interpreted through the root mean square deviation (RMSD) [81], root mean square fluctuation (RMSF) 82, the radius of gyration (Rg) [83] and beta factor (β-factor) [84] assays as depicted in Figure 10.

The average Cα atomic distance over 10,000 frames of TLR3-MEPVC and TLR4-MEPVC was decoded through RMSD assay. An

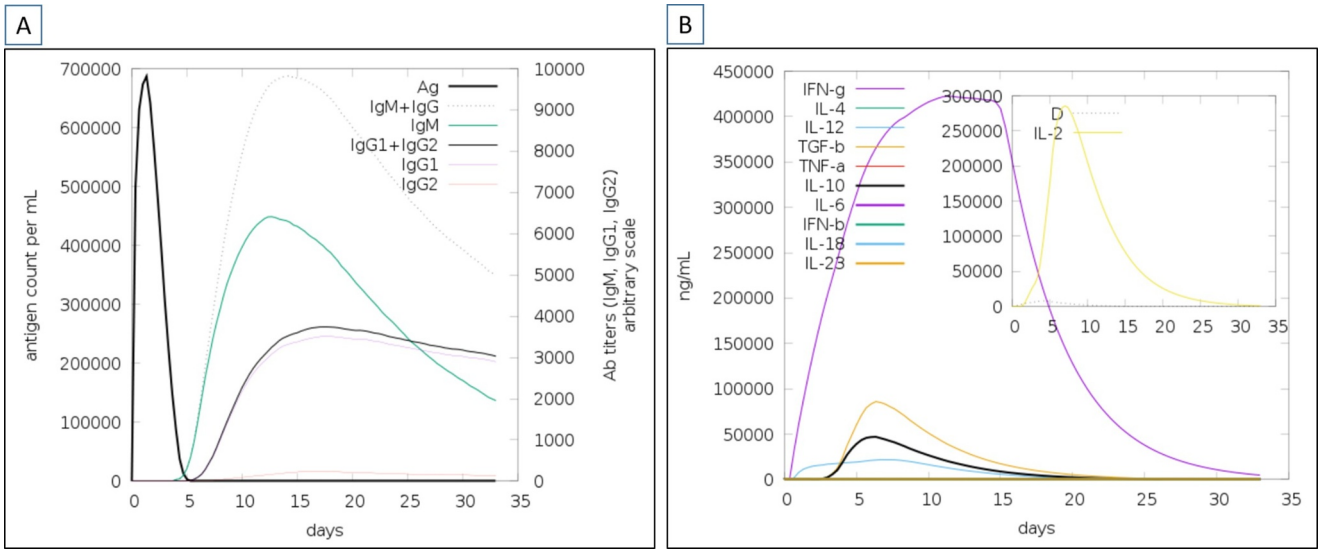

Fig. 9. In silico simulation of the host immune system using MEPVS as an antigen. A. Antibodies titer (A) and cytokines and interleukins (B) in response to the antigen.

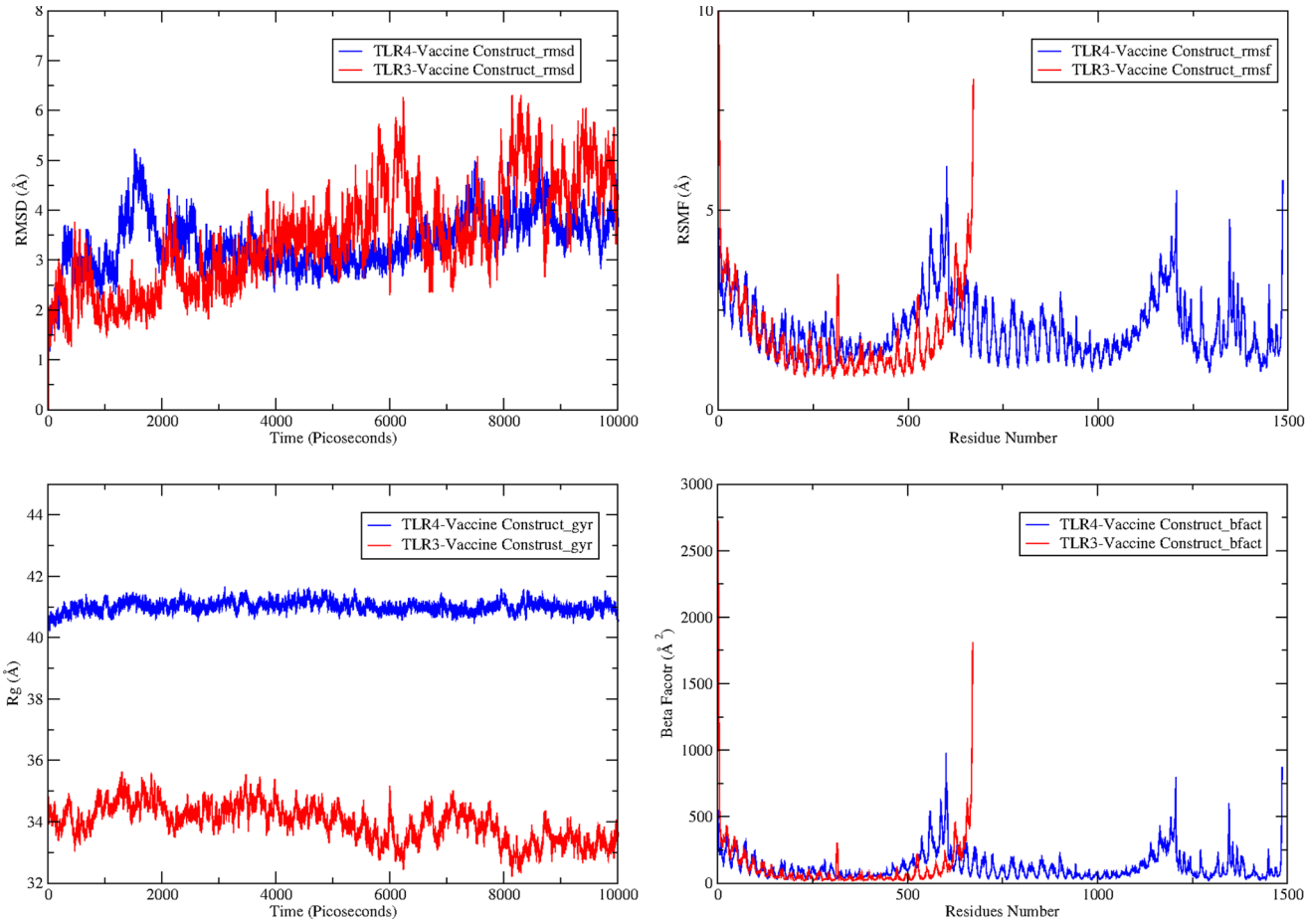

Fig. 10. MD trajectories based calculation of Cα RMSD (top left), RMSD (top right), Rg (bottom left), and β-factor (bottom right).

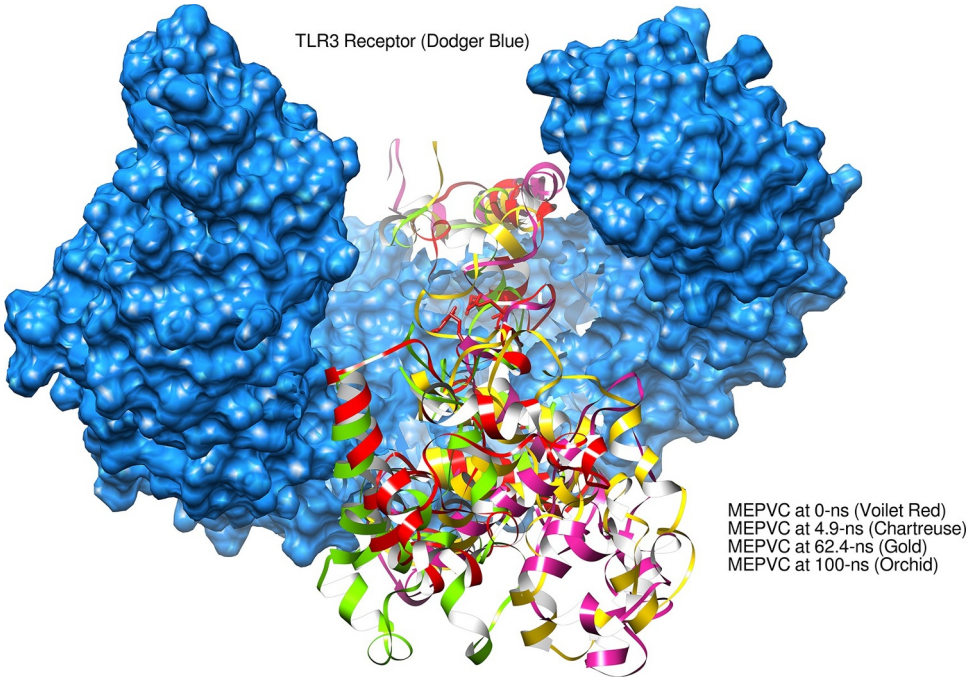

Fig. 11. Binding mode of MEPVC at different snapshots at the TLR3 docked side.

average RMSD of 3.44 Å (maximum, 6.30 Å) and 3.36 Å (maximum, 5.22 Å) were estimated for TLR3-MEPVC and TLR4-MEPVC, respectively. The TLR3 and TLR4 receptors are observed more compact than the MEPVC and as a result, continuous movements of the vaccine construct through its length are noticed at the exposed regions though rooted stable at the docked position. This seems to be responsible for

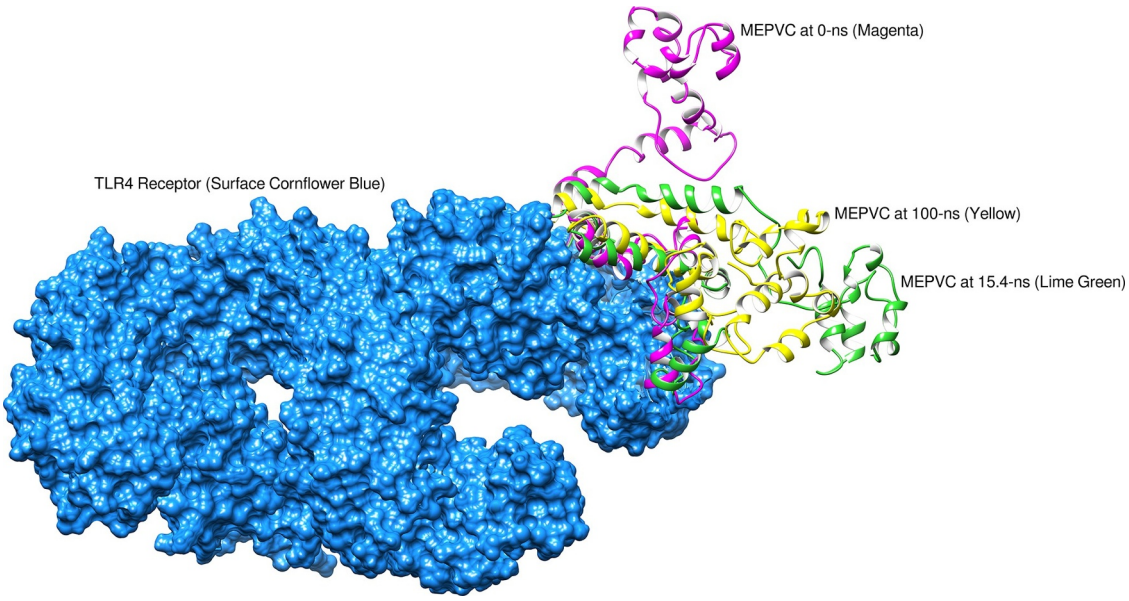

Fig. 12. Binding mode of MEPVC at different snapshots at the TLR4 docked side.

bringing small structural deviation as probed by RMSD. Visual inspection of the trajectories illustrated no major global and local secondary structure conformation changes in the receptor TLR3 and TLR4 structure. The MEPVC movements with respect to the TLR3 are shown at different snapshots as illustrated in Figure 11. The MEPVC seems responsible for the deviations in the receptor TLR4 at positions Glu1-Ser79, Hie540-Asn614, Gln1144-Gln1209, Gln1347-Pro1371, and Lys1489-Lys1491. Continuous flexibility of the loops of MEPVC exposed region is observed and is highly flexible responsible for the movements of the MEPVC at the docked site though rooted stably (Figure 12).

The second statistical parameter computed for both complexes was the root mean square fluctuations (RMSF) that demonstrate average dynamical residue fluctuations over a specific length of time. The mean RMSF concluded for TLR3-MEPVC system is 2.1 Å whereas, an average RMSF of 2.03 Å for TLR4-MEPVC was demonstrated. Majority of the receptors residues are showing less variability and are satisfactory stable with mean RMSF < 3 Å. The residues range mentioned above were reported to have higher RMSF values and revealed fluctuating throughout the simulation time. These fluctuations are in all likelihood as an outcome of the moving adjuvant of the MEPVC. This is supported by the fact that the mean RMSF of the MEPVC is comparatively very high than that of receptors. The RMSF of MEPVC in TLR3-MEPVC and TLR4-MEPVC is 1.86 Å (maximum, 10.16 Å) and 7.45 Å̊ (maximum, 13.04 Å̊). The receptors stability in both complexes is reaffirmed by radius of gyration (Rg) analysis that depicted stable plot with mean value of 33.97 Å (maximum, 35.62 Å) and 41.01 Å̊ (maximum, 41.65 Å̊) for TLR4-MEPVC. The Rg findings are coherent with that of RMSD in interpreting systems stable behavior and compact nature of the receptors. Lastly, thermal beta factor (β-factor) indicated the same pattern of residues dynamics as shown by RMSF. An average β-factor derived from TLR3-MEPVC and TLR4-MEPVC is 121.34 $Å^2$ and 123.47 $Å^2$, respectively.

3.12. Hydrogen bond analysis

Hydrogen bonding results when a hydrogen atom attached to a highly electronegative atom is attracted by another electronegative atom [85]. In a biological system, these hydrogen bonds are vital in determining specificity and directionality fundamental in molecular recognition [86]. The patterns of hydrogen bonds for both complexes were illustrated in each frame within 3 Å in order to probe the strength of intermolecular association across the simulation period. The maximum number of hydrogen bonds of MEPVC with TLR3 and TLR4 are 11 and 12, respectively demonstrating the high strength of interactions. The number of hydrogen bonds for both complexes are shown in Figure 13.

3.13. Analysis of Salt bridges

In a protein molecule, salt bridges are formed between charged side chains of amino acids at neutral pH. The residues mainly involved in these interactions include negative full electron charge glutamine and aspartate and opposite positive full electron charge arginine and lysine [87]. The presence of salt bridges between the interacting molecules is a clear sign of strengthening interaction stability. For TLR3-MEPVC complex high numbers of salt bridges were estimated within 3.2 Å between receptor Glu8, Glu276, Arg306, Glu333 with MEPVC Lys3, Lys8, Glu83, Arg16, respectively. In case of TLR4-MEPVC complex, receptor residues Arg646, Asp268, Arg1396 are involved in salt bridging with Asp146, Arg175 and Glu184 of the MEPVC, respectively.

3.14. Estimation of Binding Free Energies

The binding free energies of both TLR3-MEPVC and TLR4-MEPVC complexes were computed using continuum solvation MMPBSA and its complementary MMGBSA. The binding free energies of the vaccine construct for the receptors were estimated considering molecular mechanics energies as well as solvation energies. The net binding free energy for both complexes revealed to be very high illustrating the high interacting affinity of the molecules. For TLR3-MEPVC, the total free energy of binding is -41.4273 kcal/mol in MMGBSA while in MMPBSA it is -84.4908 kcal/mol (Table 3). In case of TLR4-MEPVC, the net binding energy is -13.9690 kcal/mol and -21.2289 kcal/mol in MMGBSA and MMPBSA, respectively. The gas phase energy in TLR3-MEPVC complex is highly dominating and contributes significantly to the overall binding energy. The total gas phase energy of TLR3-MEPVC is -530.5792 kcal/mol in MMGBSA and -530.5792 kcal/mol in MMPBSA. The electrostatic contribution to the net gas phase energy is high i.e. -393.2724 kcal/mol in both MMGBSA and MMPBSA compared to van der Waals energy (-137.3068 kcal/mol). The electrostatic energy contribution to the solvation energy is 507.5596 kcal/mol in MMGBSA and 462.9381 kcal/mol in MMPBSA and is highly non-favorable to the

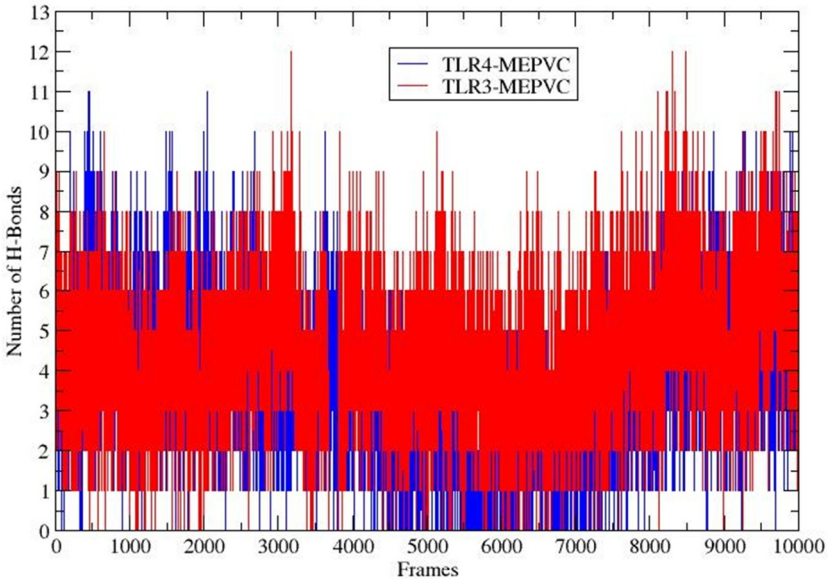

Fig. 13. The number of hydrogen bonds formed during simulation between MEPVC and TLR3 and TLR4 receptors.

Table 3
Differences of binding free energy (Complex - Receptor - Ligand) in MMGB/PBSA for the TLR3-MEPVC complex.

Method	Net Energy Component	Average	Std. Dev.	Std. Err. of Mean
MMGBSA	DWAALS	-137.3068	7.0875	0.7088
	EEL	-393.2724	43.8913	4.3891
	EGB	507.5596	41.1568	4.1157
	ESURF	-18.4077	0.8700	0.0870
	DELTA G gas	-530.5792	41.8517	4.1852
	DELTA G solv	489.1519	41.4134	4.1413
	DELTA TOTAL	-41.4273	7.7453	0.7745
MMPBSA	VDWAALS	-137.3068	7.0875	0.7088
	EEL	-393.2724	43.8913	4.3891
	EPB	462.9381	39.2924	3.9292
	ENPOLAR	-16.8497	0.6609	0.0661
	EDISPER	0.0000	0.0000	0.0000
	DELTA G gas	-530.5792	41.8517	4.1852
	DELTA G solv	446.0884	39.5288	3.9529
	DELTA TOTAL	-84.4908	10.3941	1.0394

Table 4
Differences of binding free energy (Complex - Receptor - Ligand) in MMGB/PBSA for the TLR4-MEPVC complex.

Method	Net Energy Component	Average	Std. Dev.	Std. Err. of Mean
MMGBSA	DWAALS	-35.8526	3.0453	0.3045
	EEL	267.1101	29.5270	2.9527
	EGB	-240.2983	27.5172	2.7517
	ESURF	-4.9282	0.3849	0.0385
	DELTA G gas	231.2576	29.5487	2.9549
	DELTA G solv	-245.2265	27.4463	2.7446
	DELTA TOTAL	-13.9690	4.2201	0.4220
MMPBSA	VDWAALS	-35.8526	3.0453	0.3045
	EEL	267.1101	29.5270	2.9527
	EPB	-248.2779	27.2761	2.7276
	ENPOLAR	-4.2085	0.2122	0.0212
	EDISPER	0.0000	0.0000	0.0000
	DELTA G gas	231.2576	29.5487	2.9549
	DELTA G solv	-252.4865	27.2706	2.7271
	DELTA TOTAL	-21.2289	5.3805	0.5380

net solvation energy. For TLR4-MEPVC, the net MMPBSA is the comparatively high (-21.2289 kcal/mol) than the MMGBSA (-13.9690 kcal/mol) with solvation energy noticed to dominate the overall interaction energies (Table 4). The solvation energy in MMGBSA and MMPBSA is -245.2265 kcal/mol (electrostatic contribution, -240.2983 kcal/mol) and -252.4865 kcal/mol (electrostatic contribution, -248.2779 kcal/mol). The net gas phase energy is 231.2576 kcal/ mol in both MMGBSA and MMPBSA and major favorable contribution is from van der Waals is -35.8526 kcal/mol opposed to the non-favorable contributions is 267.1101 kcal/mol. Residue wise, most of the interacting residues of TLR3 reported in docking assay have very low binding energy reflecting the highly stable nature of these residues in interaction and the key role they are playing in holding the MEPVC at the docked region. The binding energy of the hotspot residues are as follow: His156 (-2.14 kcal/mol), Asp180 (-6.21 kcal/mol), Lys201 (-1.23 kcal/mol), Glu20 (-2.13 kcal/mol), Ser206 (-0.43 kcal/mol), Phe227 (-3.45 kcal/mol), Asp229 (-1.51 kcal/mol), Asp230 (-1.88 kcal/mol), Ser254 (-2.98 kcal/mol), Ser256 (-1.52 kcal/mol), Ser282 (-1.21 kcal/mol), Asp284 (-1.99 kcal/mol), Asp285 (-0.23 kcal/mol), Glu301 (-3.13 kcal/mol), Tyr302 (-1.48 kcal/mol), Phe304 (-3.52 kcal/mol), Glu306 (-3.43 kcal/mol), Tyr307 (-2.89 kcal/mol), Arg325 (-8.12 kcal/mol), His359 (-0.17 kcal/mol), Lys382 (-0.43 kcal/mol), Tyr383 (-3.27 kcal/mol), Pro408 (-2.22 kcal/mol), His410 (-1.16 kcal/mol), Iso411 (-1.81 kcal/mol), His432 (-1.76 kcal/mol), Glu434 (-2.77 kcal/mol), Pro408 (-3.11 kcal/mol), Asp457 (-1.15 kcal/mol), Phe459 (-.5.12 kcal/mol), Gln483 (-4.11 kcal/mol), and Glu533 (-1.23 kcal/mol). In case of TLR4-MEPVC, docking residues Pro23 (-1.54 kcal/mol), Glu24 (-2.31 kcal/mol), Ser25 (-3.33 kcal/mol), Asp44 (-4.43 kcal/mol), Lys47 (-5.23 kcal/mol), Asp50 (-2.21 kcal/mol), Asp51 (-1.99 kcal/mol), Arg67 (-5.65 kcal/mol), Arg87 (-1.21 kcal/mol), Glu89 (-2.11 kcal/mol), Pro113 (-0.12 kcal/mol), Gln115 (-1.21 kcal/mol), Asp137 (-0.23 kcal/mol), His159 (-3.22 kcal/mol), and Asp160 (-2.43 kcal/mol) from chain B and Lys20 (-2,89 kcal/mol), Phe64 (3.21 kcal/mol), and Asp114 (-2.11 kcal/mol) from chain D are hotspot residues.

3.15. Probable Prevention Strategy

After performing the immune system simulation to MEPVC and a thorough peptide dynamics analysis, a novel MEPVC is proposed as a probable solution to the widely spread viral coronaviruses outbreak. Meanwhile, this seems necessary to have a deep down lesson as COVID-19 viral outbreak is propagating to the human species on this part of the universe called Earth. As we are already living under a serious threat of antibiotic and antimicrobial resistance [88] which if ignored can cause havoc and the current spread of COVID-19 is a clear example of even a viral resistance coming in line. Referring to the introduction section where the probable cause of this viral attack is discussed, this is

mandatory to mention that there is retaliation from animals and animal-associated viruses or bacteria observed against Homo sapiens quite frequently in the recent past [89]. This not only alarms the way of consuming animals in our food in the form of livestock but at the same time explains the advancing defense system of various species around. Either it is horizontal gene transfer among bacteria or transient nature of viral transmission throughout the world as one way or the other this is a form of advanced defense of organisms. It is meant to state that even before and after the Cognitive Revolution, Homo sapiens cannot be exempted from biological laws. While considering under the realm of these biological laws this is much expected from all other species to have the right of keeping and exercising a defense mechanism [90]. Instead of spending on wars among humans there is a dire need to be equipped for all wars to come against human species under the disguise of either environmental hazards or antimicrobial resistance. Even viruses are finding new ways of infecting hosts. Under this perspective the role of avian has been observed as much dominating in the recent past [91]. After having insights from comparative advancing mechanism we mention here and thereafter the preliminary existence of a "Theory of Retaliation" on the basis of currently available facts and observations. This theory of retaliation will play its role in coming future. The continuous attacks in the form of outbreaks by them have raised many challenges and threats to human beings. A recent mechanism study by Bazel et al reported H5N1 influenza viruses is posing threat to human and animal health and is currently unclear what restricts these interspecies jumps on the host side. It is further signified that PB1-F2 (a short viral peptide) assists H5N1 bird influenza viruses to overcome a human restriction factor of the viral polymerase complex HAX-1. It is also evidenced that a functional PB1-F2 aids in direct transmission of viruses from birds to humans [92]. Additionally, there is already a mechanism existing for the antimicrobial defense of avian eggs which illustrates their efficacy in defense mechanisms [93]. Furthermore, discovery of avian antimicrobial peptides, classified as B-defensins, present in chicken and turkey found active against bacteria, fungi, and yeast. This is another example of advance mechanism showed by avian and the possibility of a common ancestral gene between avian and other mammalian peptides seems obvious [94]. WHO alarmed quite often that influenza viruses with a vast silent reservoir in aquatic birds are impossible to eradicate while avian influenza is proved to be a threat to human health [95]. Even having the phylogenetic identity between SARS-CoV-2 and SARS-CoV, some clinical characteristics differentiate SARS-CoV-2 from SARS-CoV, MERS-CoV, and seasonal influenza infections [2,96]. This leads to the fact that viruses are attacking humans with different clinical features every time. The complex relationships between the human and animal species never faced a halt in evolving giving rise to numerous infectious pathogens. Whatsoever is the case, the dramatic impact of infectious diseases affecting the modern human population worldwide is evident and unexpected rise of coronavirus infections raised to 100000 while writing this research and getting beyond the normal control (https://www.nature.com/articles/d41586-020-00154-w). After all we belong to the kingdom Animalia with even ten on ten embarrassing similarities to chimpanzees [97]. Probable prevention in this context is to somehow avoid the excruciating utilization and consumption of other mammals as edibles from Homo sapiens and stop becoming a reason for the extinction of certain species. Avenging from these species from time to time by utilizing the microbes associated with them against human has become obvious and vaccine design and discovery is of utmost importance. This is crucial to strategize a preventive control not only for symptomatic relief but in general taking steps to prevent hunting and strategy required to maintain a balanced ecosystem where specifically avian will not dwell under threat by sapiens. The advent of various new mechanisms of viral survival has remained sapiens perplexed and a broader strategy is required to circumvent this problem.

4. Conclusions

In this study, we used available immunoinformatics approaches for the purpose to prioritize potential vaccine candidates against COVID-19 considering their ease of use in experimental investigations. Bearing in mind, the wide immunological applications of peptide vaccines only highly antigenic, virulent, conserved and non-allergic epitopes targeted by both B and T-cells were disclosed. A multi-epitope peptide was constructed and an appropriate adjuvant was added to allow suitable delivery and efficient immune processing of the epitopes. These promising computational findings might deliver preliminary epitopes set for a vaccine against the COVID-19 notwithstanding the experimental testing in appropriate animal models to unravel real effectiveness. Meanwhile, probable prevention discussed in this study for Homo sapiens is to avoid becoming a reason for the extinction of various species either by hunting and/or over utilizing other mammals as this may be a reason for resistance both from microbes and other animal species of this kingdom. There must be strategic studies keeping in view the advancements in defense mechanism of avian and avian related microbes avenging humans. Preventive use of animals or avian in human diet and avoiding hunting can be preventive options for self-defense in this connection and/or maintenance of a balanced ecosystem should be reinvestigated.

5. Disclosure Statement

No conflict of interest was reported by the authors.

Credit Author Statement

Sajjad Ahmad: Conceptualization, Methodology, Writing- Original Draft Preparation, Visualization, Investigation. Afifa Navid: Data Curation, Writing- Original Draft Preparation, Figures Preparation. Rabia Farid: Data Curation, Figures Preparation, Writing- Original Draft Preparation. Ghulam Abbas: Methodology, Data Curation. Faisal Ahmad: Methodology, Figures Preparation. Naila Zaman: Writing-Reviewing and Editing. Nousheen Parvaiz: Writing. Syed Sikander Azam: Supervision, Conceptualization, Methodology, Softwares, Resources, Funding Acquisition, Writing- Reviewing and Editing, Project Management.

Acknowledgment

Authors are highly grateful to the Pakistan-United States Science and Technology Cooperation Program (Grant No. Pak-US/2017/360) for granting financial assistance.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejps.2020.105387.

References

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2019. A Novel Coronavirus from Patients with Pneumonia in China. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2001017. 0 (n.d.) null.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet https://doi.org/10.1016/S0140-6736(20)30183-5.

Gralinski, L.E., Menachery, V.D., 2020. Return of the Coronavirus: 2019-nCoV. Viruses 12, 135.

Hui, D.S., Azhar, E.I, Madani, T.A., Ntoumi, F., Kock, R., Dar, O., Ippolito, G., Mchugh, T.D., Memish, Z.A., Drosten, C., Zumla, A., Petersen, E., 2020. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China. Int. J. Infect. Dis 91, 264–266. https://

doi.org/10.1016/j.ijid.2020.01.009.
2019 Novel Coronavirus (2019-nCoV) Situation Summary | CDC, (n.d.).
Wang, C., Horby, P.W., Hayden, F.G., Gao, G.F., 2020. A novel coronavirus outbreak of global health concern. Lancet.
C. Seone, ScienceDaily: Your source for the latest research news, 14. 5 (2015) 1–78.
American Thoracic Society, (2020). doi:10.1164/rccm.2014P7.
Coordinators, N.R., 2017. Database resources of the national center for biotechnology information. Nucleic Acids Res 45, D12.
Barh, D., Barve, N., Gupta, K., Chandra, S., Jain, N., Tiwari, S., Leon-Sicairos, N., Canizalez-Roman, A., dos Santos, A.R., Hassan, S.S., 2013. others, Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in Vibrio cholerae targeted by Piper betel derived compounds. PLoS One 8, e52773.
Ahmad, S., Ranaghan, K.E., Azam, S.S., 2019. Combating tigecycline resistant Acinetobacter baumannii: A leap forward towards multi-epitope based vaccine discovery. Eur. J. Pharm. Sci. 132, 1–17.
Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305, 567–580.
Tusnady, G.E., Simon, I., 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849–850.
Sachdeva, G., Kumar, K., Jain, P., Ramachandran, S., 2004. SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. Bioinformatics 21, 483–491.
Wizemann, T.M., Adamou, J.E., Langermann, S., 1999. Adhesins as targets for vaccine development. Emerg. Infect. Dis 5, 395.
Vita, R., Mahajan, S., Overton, J.A., Dhanda, S.K., Martini, S., Cantrell, J.R., Wheeler, D.K., Sette, A., Peters, B., 2018. The immune epitope database (IEDB): 2018 update. Nucleic Acids Res 47, D339–D343.
Larsen, J.E.P., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. Immunome Res 2, 2.
Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 45, W24–W29.
Guan, P., Doytchinova, I.A., Zygouri, C., Flower, D.R., 2003. MHCPred: a server for quantitative prediction of peptide–MHC binding. Nucleic Acids Res 31, 3621–3624.
Naz, A., Awan, F.M., Obaid, A., Muhammad, S.A., Paracha, R.Z., Ahmad, J., Ali, A., 2015. Identification of putative vaccine candidates against Helicobacter pylori exploiting exoproteome and secretome: a reverse vaccinology based approach. Infect. Genet. Evol 32, 280–291.
Garg, A., Gupta, D., 2008. VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. BMC Bioinformatics 9, 62.
He, Y., Xiang, Z., Mobley, H.L.T., 2010. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. Biomed Res. Int (2010).
I. Dimitrov, D.R. Flower, I. Doytchinova, AllerTOP-a server for in silico prediction of allergens, in: BMC Bioinformatics, 2013: p. S4.
Zhang, L., 2018. Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cell. Mol. Immunol 15, 182.
Li, W., Joshi, M., Singhania, S., Ramsey, K., Murthy, A., 2014. Peptide vaccine: progress and challenges. Vaccines 2, 515–536.
Saadi, M., Karkhah, A., Nouri, H.R., 2017. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. Infect. Genet. Evol 51, 227–234.
Ali, M., Pandey, R.K., Khatoon, N., Narula, A., Mishra, A., Prajapati, V.K., 2017. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. Sci. Rep 7, 9232.
Cheng, J., Randall, A.Z., Sweredoski, M.J., Baldi, P., 2005. SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res 33, W72–W76.
Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9, 40.
Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., 2018. others, SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46, W296–W303.
Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., 2005. others, Galaxy: a platform for interactive large-scale genome analysis. Genome Res 15, 1451–1455.
Heo, L., Park, H., Seok, C., 2013. GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res 41, W384–W388.
Creighton, T.E., 1988. Disulphide bonds and protein stability. BioEssays 8, 57–63.
D.B. Craig, A.A. Dombkowski, Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins, BMC Bioinformatics. 14 (2013) 346.
Angov, E., 2011. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol. J 6, 650–659.
Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D.C., Jahn, D., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33, W526–W531.
Pandey, R.K., Bhatt, T.K., Prajapati, V.K., 2018. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. Sci. Rep. 8, 1125.
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.
E. ProtParam, ExPASy-ProtParam tool, (2017).
N. Rapin, O. Lund, F. Castiglione, C-Immsim 10.1 server, (2012).
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.
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PATCHDOCK and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 33, W363–W367.
Andrusier, N., Nussinov, R., Wolfson, H.J., 2007. FireDock: fast interaction refinement in molecular docking. Proteins Struct. Funct. Bioinforma 69, 139–159.
Chen, J.E., Huang, C.C., Ferrin, T.E., 2014. RRDistMaps: a UCSF Chimera tool for viewing and comparing protein distance maps. Bioinformatics 31, 1484–1486.
D.S. Biovia, Discovery studio visualizer, San Diego, CA, USA. (2017).
Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38.
Andleeb, S., Imtiaz-Ud-Din, M.K.Rauf, Azam, S.S., Badshah, A., Sadaf, H., Raheel, A., Tahir, M.N., Raza, S., 2016. A one-pot multicomponent facile synthesis of dihydropyrimidin-2(1: H)-thione derivatives using triphenylgermane as a catalyst and its binding pattern validation. RSC Adv 6, 79651–79661. https://doi.org/10.1039/c6ra19162b.
D.A. Case, D.S. Cerutti, T.E. Cheateham, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, others, AMBER16 Package, Univ. California, San Fr. (2016).
Wang, J., Wang, W., Kollman, P.A., Case, D.A., 2001. Antechamber: an accessory software package for molecular mechanical calculations, J. Am. Chem. Soc. 222, U403.
Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham III, T.E., Darden, T.A., Duke, R.E., Gohlke, H., 2014. others, The FF14SB force field. Amber 14, 29–31.
Izaguirre, J.A., Catarello, D.P., Wozniak, J.M., Skeel, R.D., 2001. Langevin stabilization of molecular dynamics, J. Chem. Phys. 114, 2090–2098.
Kräutler, V., Van Gunsteren, W.F., Hünenberger, P.H., 2001. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, J. Comput. Chem. 22, 501–508.
A.S. Lemak, N.K. Balabaev, On the Berendsen thermostat, Mol. Simul. 13 (1994) 177–187.
Roe, D.R., Cheatham, T.E., 2013. III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9, 3084–3095.
S. Kaliappan, I.I.T. Bombay, UCSF Chimera-Superimposing and Morphing, (2016).
Miller, B.R., McGee, T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E., 2012. MMPBSA.py: An efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8, 3314–3321. https://doi.org/10.1021/ct300418h.
T. Hou, J. Wang, Y. Li, W. Wang, Assessing the Performance of the MM_PBSA and MM_GBSA Methods. 1. The Accuracy.pdf, (2011) 69–82.
Hossain, T., Kamruzzaman, M., Choudhury, T.Z., Mahmood, H.N., Nabi, A.H.M., Hosen, M., 2017. others, Application of the subtractive genomics and molecular docking analysis for the identification of novel putative drug targets against Salmonella enterica subsp. enterica serovar Poona. Biomed Res. Int (2017).
Martins, M.A., Watkins, D.I., 2018. What is the predictive value of animal models for vaccine efficacy in humans? Rigorous simian immunodeficiency virus vaccine trials can be instructive. Cold Spring Harb. Perspect. Biol. 10, a029504.
Abbas, G., Zafar, I., Ahmad, S., Azam, S.S., 2020. Immunoinformatics design of a novel multi-epitope peptide vaccine to combat multi-drug resistant infections caused by Vibrio vulnificus. Eur. J. Pharm. Sci 142, 105160.
Foster, T.J., Geoghegan, J.A., Ganesh, V.K., Höök, M., 2014. Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. Nat. Rev. Microbiol 12, 49–62.
Te Velthuis, A.J.W., van den Worm, S.H.E., Snijder, E.J., 2012. The SARS-coronavirus nsp7+ nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res 40, 1737–1747.
Miknis, Z.J., Donaldson, E.F., Umland, T.C., Rimmer, R.A., Baric, R.S., Schultz, L.W., 2009. Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth, J. Virol 83, 3007–3018.
Bouvet, M., Lugari, A., Posthuma, C.C., Zevenhoven, J.C., Bernard, S., Betzi, S., Imbert, I., Canard, B., Guillemot, J.-C., Lécine, P., 2014. others, Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes, J. Biol. Chem. 289, 25783–25796.
Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J., 2004. others, Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase, J. Biol. Chem. 279, 1637–1642.
Walls, A.C., Tortorici, M.A., Bosch, B.-J., Frenz, B., Rottier, P.J.M., DiMaio, F., Rey, F.A., Veesler, D., 2016. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature 531, 114–117.
Phan, T., 2020. Genetic diversity and evolution of SARS-CoV-2. Infect. Genet. Evol, 104260.
Sanchez-Trincado, J.L., Gomez-Perosanz, M., Reche, P.A., 2017. Fundamentals and methods for T-and B-cell epitope prediction, J. Immunol. Res (2017).
Naz, K., Naz, A., Ashraf, S.T., Rizwan, M., Ahmad, J., Baumbach, J., Ali, A., 2019. PanRV: Pangenome-reverse vaccinology approach for identifications of potential vaccine candidates in microbial pangenome. BMC Bioinformatics 20, 123.
Han, X., LaRosa, K.B., Kawai, T., Taubman, M.A., 2014. DNA-based adaptive immunity protect host from infection-associated periodontal bone resorption via recognition of Porphyromonas gingivalis virulence component. Vaccine 32, 297–303.
F.A. Bonilla, H.C. Oettgen, Adaptive immunity, J. Allergy Clin. Immunol. 125 (2010) S33–S40.
Burton, D.R., 2002. Antibodies, viruses and vaccines. Nat. Rev. Immunol. 2, 706–713.
Janeway Jr, C.A., Travers, P., Walport, M., Shlomchik, M.J., 2001. The major histocompatibility complex and its functions. Immunobiol. Immune Syst. Heal. Dis, 5th. Garland Science.
Hewitt, E.W., 2003. The MHC class I antigen presentation pathway: strategies for viral

immune evasion. Immunology 110, 163–169.
Roche, P.A., Furuta, K., 2015. The ins and outs of MHC class II-mediated antigen processing and presentation. Nat. Rev. Immunol 15, 203–216.
Stratmann, T., 2015. Cholera toxin subunit B as adjuvant—-an accelerator in protective immunity and a break in autoimmunity. Vaccines 3, 579–596.
Baldauf, K.J., Royal, J.M., Hamorsky, K.T., Matoba, N, 2015. Cholera toxin B: one subunit with many pharmaceutical applications. Toxins (Basel) 7, 974–996.
Dombkowski, A.A., Sultana, K.Z., Craig, D.B., 2014. Protein disulfide engineering. FEBS Lett 588, 206–212.
Vaure, C., Liu, Y., 2014. A comparative review of toll-like receptor 4 expression and functionality in different animal species. Front. Immunol 5, 316.
Dang, J., Tiwari, S.K., Lichinchi, G., Qin, Y., Patil, V.S., Eroshkin, A.M., Rana, T.M., 2016. Zika virus depletes neural progenitors in human cerebral organoids through activation of the innate immune receptor TLR3. Cell Stem Cell 19, 258–265.
V.N. Maiorov, G.M. Crippen, Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins, (1994).
Kuzmanic, A., Zagrovic, B., 2010. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophys. J. 98, 861–871.
Lobanov, M.Y., Bogatyreva, N.S., Galzitskaya, O.V, 2008. Radius of gyration as an indicator of protein structure compactness. Mol. Biol 42, 623–628.
Haq, F.U., Abro, A., Raza, S., Liedl, K.R., Azam, S.S., 2017. Molecular dynamics simulation studies of novel $β$-lactamase inhibitor, J. Mol. Graph. Model. 74, 143–152.
Fersht, A.R., 1987. The hydrogen bond in molecular recognition. Trends Biochem. Sci 12, 301–304.
R.E. Hubbard, M. Kamran Haider, Hydrogen bonds in proteins: role and strength, E LS. (2001).
Flores, B.N., Li, X., Malik, A.M., Martinez, J., Beg, A.A., Barmada, S.J., 2019. An intramolecular salt bridge linking TDP43 RNA binding, protein stability, and TDP43-dependent neurodegeneration. Cell Rep 27, 1133–1150.
MacLean, R.C., San Millan, A., 2019. The evolution of antibiotic resistance. Science (80-.) 365, 1082–1083.
Hart, B.L., 2011. Behavioural defences in animals against pathogens and parasites: parallels with the pillars of medicine in humans. Philos. Trans. R. Soc. B Biol. Sci 366, 3406–3417.
Aminov, R.I., 2011. Horizontal gene exchange in environmental microbiota. Front. Microbiol 2, 158.
Morens, D.M., Folkers, G.K., Fauci, A.S., 2004. The challenge of emerging and re-emerging infectious diseases. Nature 430, 242–249.
Mazel-Sanchez, B., Boal-Carvalho, I., Silva, F., Dijkman, R., Schmolke, M., 2018. H5N1 influenza A virus PB1-F2 relieves HAX-1-mediated restriction of avian virus polymerase PA in human lung cells, J. Virol 92 e00425–18.
Wellman-Labadie, O., Picman, J., Hincke, M.T., 2008. Antimicrobial activity of the Anseriform outer eggshell and cuticle. Comp. Biochem. Physiol. Part B Biochem. Mol. Biol 149, 640–649.
Sugiarto, H., Yu, P.-L., 2004. Avian antimicrobial peptides: the defense role of $β$-defensins. Biochem. Biophys. Res. Commun 323, 721–727.
Peiris, J.S.M., De Jong, M.D., Guan, Y., 2007. Avian influenza virus (H5N1): a threat to human health. Clin. Microbiol. Rev 20, 243–267.
Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D.S.C., 2020. others, Clinical characteristics of coronavirus disease 2019 in China. N. Engl. J. Med.
Mora-Bermúdez, F., Badsha, F., Kanton, S., Camp, J.G., Vernot, B., Köhler, K., Voigt, B., Okita, K., Maricic, T., He, Z., 2016. others, Differences and similarities between human and chimpanzee neural progenitors during cerebral cortex development. Elife 5, e18683.