

Reverse vaccinology approach towards the in-silico multiepitope vaccine development against SARS-CoV-2

Vipul Kumar

Indian Institute of Technology Delhi, India

Manoj Jena (✉ manoj.20283@lpu.co.in)

Lovely Professional University, Punjab, India

Research Article

Keywords: SARS-CoV-2, COVID-19, Immunoinformatics, multi-epitope, docking, MD simulations

Posted Date: May 27th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-31779/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at F1000Research on January 23rd, 2021. See the published version at <https://doi.org/10.12688/f1000research.36371.1>.

Abstract

The novel severe acute respiratory syndrome related corona virus-2 (SARS-CoV-2) belongs to the "Coronaviridae" family and order "Nidovirales" cause coronavirus disease (COVID19). The SARS-CoV-2 has been spread in more than a hundred countries, and more than a million have lost their lives. Recently COVID-19 has been declared as pandemic worldwide. Vaccination and immunization could only be an effective strategy to combat this fatal COVID-19. For identification of the effective vaccine candidate against COVID-19, various immunoinformatics online tools and software were used to predict the epitopes. The cytotoxic T cell epitopes, Helper T cell epitopes, and B cell epitopes from three structural polyproteins (Spike, Membrane, and Nucleocapsid (SMN)) based on the binding affinity towards MHC , antigenicity, non-allergenicity, and non-toxicity were identified for vaccine development. The multiepitope based vaccine was constructed linking two additional adjuvants human betadefensin-3 and human beta-defensin-2 at N and C terminal, respectively. Constructed vaccine sequence was found to be a good antigen and non-allergen for the human body. The constructed vaccine was docked with the TLR-3 receptor. And the docked complex then further taken for Molecular dynamics simulations and RMSD was calculated, which showed stable binding of the complex. The codon adaptation index (CAI) of 0.92 and GC content of 55.5% for E.coli (k12) strain suggested the efficient expression of the predicted vaccine. The current study can be helpful in the reduction of time and cost for further experimental validations and could give a valuable contribution against this pandemic.

1. Introduction

Coronaviruses (CoVs) belong to the family of coronaviridae in the order Nidovirales, have single-strand positive-sense RNA [1, 2]. The size of the RNA of coronavirus is the largest among the viruses (~30 kb) [2]. They have glycoprotein projections on the envelope, which gives the corona appearance. CoVs are the pathogens that mainly involved in respiratory and gastrointestinal diseases in the wide range of animals and humans [1, 2]. CoVs are divided into four sub-categories, namely alpha, beta, gamma, and delta, out of which alpha and beta coronavirus are known to infect humans [1, 3]. From alpha and Beta, four strains are responsible for the common cold, and two strains were found to be responsible for severe acute respiratory syndrome (SARS-CoV) and Middle East respiratory syndrome CoV (MERS-CoV) [4, 5]. Recently in December 2019, a novel coronavirus SARS-CoV-2 was detected from patients of novel coronavirus disease (COVID-19) in the Wuhan, China [6-8]. The symptoms of COVID-19 infection include Headache, fever, pneumonia, asthenia, and more [9, 10]. A severe and fatal outbreak of this virus has taken many lives and has made an enormous economic loss worldwide. The treatment and prevention from this infection is the need of the hour.

To date, no specific drug or vaccine is known for the SARS-CoV-2. So here, an attempt has been given to construct an in-silico vaccine, which can be further validated through experimental assays and could play a major role in the management of this pandemic. In this study, three structural proteins from SARS-CoV-2, based on the antigenicity has been selected for the construction of the vaccine. The first structural protein is Spike (S) glycoprotein, which has been reported to be a crucial surface protein of SARS-CoV-2,

which facilitates the entry of the virus inside the host cell. It has been reported that for the entry of the SARS-CoV-2, S protein first binds with Angiotensin-Converting Enzyme-2 (ACE-2) receptor, and then it gets primed by the host serine protease (TMPRSS2) [11, 12]. This priming of the S protein lets it to fuse into the host cell membrane and entry inside the cell. The second crucial structural protein, which induces a strong immune response, is Membrane (M) glycoprotein. It plays a crucial role in virus morphogenesis and assembly by interacting with several other viral proteins [13]. The third structural protein chosen for vaccine construct is Nucleocapsid (N) phosphoprotein; it links the viral genome to the envelope. It consists of two domains, i.e., N terminal and C terminal, and both can bind to RNA. It has been reported that the C terminal domain facilitates the physical interaction of the RNA genome and envelope [14, 15]. All these three structural proteins are predicted to be good antigens and could induce the immune response.

In this pandemic situation, the immunoinformatics approach could be a fast, scientifically sound, and reliable option for quicker vaccine development. These all proteins were chosen were predicted to be a good antigen, which gives the opportunity to predict the B and T cell epitopes. When naive B cells interact with the antigenic B cell epitopes via its transmembrane bound antibody, they differentiate into two types of cells plasma and memory cells [16, 17]. Plasma cells lack receptors, but they produce a large number of antibodies against the antigen. Memory cells express membrane-bound antibody molecules, but they are functionally inactive unless they encounter the same antigen again [18-20]. Furthermore, T cell epitopes are recognized by Major Histocompatibility Class (MHC), a glycoprotein present on the variety of the cells, which display the antigen to T cells [21, 22]. Antigen-presenting MHCs are divided into major two classes, MHC class-I are expressed on nucleated cells while MHC class-II are only expressed by antigen-presenting cells. Class II MHC interacts with T helper cells and activates B cells via cytokines, while Class-I MHC interacts with cytotoxic T cells, which kills virus-infected host cells [23, 24]. Further to know the ability of constructed vaccine for inducing innate as well as antigen-specific acquired immunity, constructed vaccine docked with Toll-Like Receptor-3 (TLR-3). TLRs mainly expressed on various leukocytes such as dendritic cells, natural killer cells, and cells of adaptive immunity such as T cells and B cells [25, 26]. Hence, in this study, an attempt has been made to construct the multiepitope vaccine consists of Helper T cells (HTLs), Cytotoxic T Cells (HTLs), and B cell epitopes, which could interact with TLR-3 and generate the immune response. This constructed multiepitope vaccine could induce both humoral as well as cell-mediated immune responses.

2. Methods

2.1 Retrieval of Structural polyproteins of the SARS-CoV-2

The complete sequence of all three structural polyproteins from SARS-CoV-2 reference sequence (NC045512.2) were retrieved from NCBI on the basis of their antigenicity. The spike (S) glycoprotein (YP_009724390.1), Membrane (M) glycoprotein (YP_009724393.1) and Nucleocapsid (N) phosphoprotein (YP_009724397.2) were retrieved in FASTA format. These three proteins together are referred to as SMN polyprotein in this study.

2.2 Cytotoxic T cell epitopes (CTL) prediction

First of all, the CTL epitopes for SMN polyproteins were predicted using Netctl 1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/>) [27]. Prediction of the epitopes depends on three major attributes firstly binding affinity of MHC-1 class, second, ability of the proteasome cleavage, and third, TAP transport efficiency; the first two are predicted with the artificial neural network algorithm while third one using weight matrix. For the prediction of the epitopes threshold for epitopes identification was chosen to be 0.75, weight on C terminal cleavage was set on 0.15 while the weight on TAP transport efficiency was set on 0.05. The predicted epitopes were ranked according to the combined score.

2.3 Helper T cell (HTL) epitope prediction

For the prediction of HTL epitopes, the IEDB MHC II server (<http://tools.iedb.org/mhcii/>) was used [28]. The species/locus was selected as Human/HLA-DR, and a 7-allele HLA reference set was selected for the prediction. Further, 15 mer length of the epitopes were retrieved and ranked according to the percentile. The percentile rank is given after comparing the peptides score with five million 15 mers from the SWISSPROT database. The higher percentile value means a lower binding affinity of MHC-II . For further refinements of the HTL epitopes, these selected HTL epitopes were subjected to investigate whether they can induce IFN gamma immune response using the IFN epitope server (<http://crdd.osdd.net/raghava/ifnepitope/predict.php>) [29]. For IFN gamma inducing epitopes selection, the Motif/SVM hybrid approach was chosen, and the model was set to be IFN gamma versus non-IFN gamma. Finally, the epitopes whose results were positive for the IFN gamma response were chosen for the in silico vaccine development.

2.4 B cell epitope prediction

B cell epitopes were predicted using the ABCpred server (<http://crdd.osdd.net/raghava/abcpred/>) [30]. This server predicts B cell epitopes using recurrent neural network algorithm. For the identification of the epitopes, the threshold was set on 0.51, while the window length for the prediction was chosen to be 16, keeping overlapping filter on. Top predicted epitopes having scored more than 0.9 was only chosen for the development of the candidate vaccine. Further, after the construction of the vaccine, linear as well as discontinuous conformational B cell epitopes were identified in the vaccine construct using ElliPro, an online server (<http://tools.iedb.org/ellipro/>) [31]. ElliPro predicts the antibody epitopes taking protein antigen tertiary structure as input.

2.5 Antigenicity, Allergenicity, and Toxicity prediction

The important attributes such as the antigenicity, allergenicity and toxicity were predicted for all the predicted epitopes individually as well as after construction of the vaccine. First of all, the antigenicity was investigated using the vaxijen server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) [32], and only probable antigen epitopes were chosen for the construction of the vaccine. Further, the allergenicity was predicted using the algpred server (<http://www.imtech.res.in/raghava/algpred/>) [33] and

only non-allergenic epitopes were selected. And Finally, all the epitopes were investigated for toxicity using the ToxinPred server (https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php) [34] and non-toxic epitopes were selected. All the predicted epitopes had to cross all these barriers, and finally, the overall construct of the vaccine has also been tested for these attributes.

2.6 Construction of multiepitope vaccine sequence

The vaccine sequence was constructed using the best identified CTL, HTL, and B cell epitopes. For the construction of the sequence at the N terminal and C terminal, an adjuvant was added using EAAAK linkers. While HTL epitopes were linked using GPGPG, linkers and CTL epitopes were linked using AAY linkers. And in the C terminal, HHHHHH was added for the easy purification of the vaccine.

2.7 Physicochemical properties of the vaccine sequence

The physicochemical properties such as molecular weight, PI, half-life, aliphatic index, and hydrophobicity were predicted using online tool protparam (<http://web.expasy.org/protparam/>) [35].

2.8 Secondary structure of the vaccine sequence

Protein secondary structure prediction gives further opportunity to predict the tertiary structure as well as gives information about the activity and function of the protein. The secondary structure of the final multiepitope vaccine sequence was predicted by the free online web tool CFSSP (<http://www.biogem.org/tool/chou-fasman/index.php>) [36].

2.9 Tertiary structure prediction

The tertiary structure of the constructed vaccine was predicted using the Rosetta web tool (<https://yanglab.nankai.edu.cn/trRosetta>) [37]. Rosetta tool applies a deep neural network algorithm to predict the inter-residue distances as well as orientations. Then these orientations are converted to smooth inter-residue constraints followed by gradient descent energy minimization. Further, the coarse-grained models are generated, and full atom refinement is done. The validation of the model has been done through Ramachandran plot analysis using vadar web tool (<http://vadar.wishartlab.com/>) [38]. Further modeled structure was validated through the PROSA web tool (<https://prosa.services.came.sbg.ac.at/prosa.php>) [39], which gives the quality Z score of the modeled protein based on the already known similar size of the proteins crystal structures.

2.10 Minimization, equilibration and Molecular dynamics of the predicted structure

For getting a better stable structure, the predicted structure was further taken for molecular dynamics (MD) simulations using Gromacs software [40]. The structure was minimized using the steepest descent algorithm with 50000 steps, followed by NVT and NPT equilibration for 100 picoseconds, followed by MD simulations of 500 microseconds. And the last frame of the MD trajectory was taken for further analysis.

2.11 Molecular docking of constructed vaccine with TLR-3 receptor

For molecular docking, the last frame from the MD simulations of the constructed vaccine was taken, and the TLR-3 structure was retrieved from Protein Data Bank (PDB) having ID 1ZIW. The downloaded structure was prepared and processed for docking using dock prep tool UCSF chimera software. For the docking, the vaccine construct, and TLR-3 was uploaded to patchDock (<http://bioinfo3d.cs.tau.ac.il/PatchDock/>) [41] server. Further, for refinement of the rigid body molecular docking solutions, FireDock (<http://bioinfo3d.cs.tau.ac.il/FireDock/>) was used [42]. It gave the best 10 docked confirmation based on global energy and Vander wall interactions.

2.12 Molecular dynamics simulation of TLR-3-vaccine construct

The docked TLR-3-vaccine complex was further taken for the MD simulation study. For assessment of the dynamic behavior, interactions, and stability of the complex, 300 picoseconds MD simulations were performed using gromacs software [40]. Root Mean Square deviation (RMSD) was calculated from the obtained MD trajectories of the complex to assess the stability.

2.13 Reverse translation and codon optimization

Finally, for expressing the constructed multiepitope vaccine needs to be expressed in the suitable vector inside the prokaryotic system. Hence, reverse translation and codon optimization were analyzed using the Java codon adaptation web tool (Jcat) (<http://www.jcat.de/>) [43]. The codon optimization was performed for *E.coli* strain K12 as a host. Jcat gives the codon adaptation index (CAI) and percentage GC content as output. The CAI gives the information of codon usage, generally score between 1 and 0.8, while GC contents should be between 40 % to 70%, values lying outside the given margin is suggested to be inefficient [44].

3. Results

3.1 Retrieval of the polyproteins and antigenicity

The amino acid sequence of all three (SMN) structural proteins were retrieved from the NCBI database in fasta format. The proteins were investigated for the antigenicity by vaxijen web tool, and it was found that all the three chosen proteins could be good antigens. The default threshold of 0.4 was chosen as the criteria for the antigenicity in the vaxijen tool. The spike protein showed a score of 0.46; Membrane glycoprotein showed a score of 0.51, while nucleocapsid protein showed a score of 0.50. Hence, all three proteins were chosen for further predictions of B cell and T cell epitopes and the construction of the vaccine.

3.2 Prediction of CTL epitopes

CTL epitopes were predicted using Netctl 1.2 server for all the three selected proteins. A total of 38 epitopes was predicted from spike glycoprotein; 10 epitopes were predicted from membrane glycoprotein, while 9 were predicted from nucleocapsid protein. Out of all these predicted CTL epitopes, only 8 were

selected for the construction of the vaccine, based on a high binding affinity towards MHC-I, antigenicity, non-allergenicity, and non-toxicity predictions as shown in **Table1**.

3.3 Prediction of HTL epitopes

HTL epitopes were predicted using the IEDB MHC II server for all the three SMN structural proteins. Finally, 4 HTL epitopes were selected on the basis of binding affinity, antigenicity, non-allergenicity, and non-toxicity, as shown in **Table2**. Four human alleles and position of predicted epitopes are HLA-DRB1*07:01 (166-180), HLA-DRB4*01:01(298-312), HLA-DRB5*01:01 (232-246), HLA-DRB5*01:01 (345-359).

3.4 B cell epitope prediction

For the prediction of B cell epitopes ABCpred server was used. Based on the binding score (>0.9), non-allergenicity and non-toxicity, a total of four B cell epitopes were finally selected, as shown in **Table3**.

3.5 Construction of multiepitope based vaccine

The four B cell epitopes, four HTL epitopes and 8 CTL epitopes were selected for vaccine construction, which fulfilled all the criteria of binding affinity, antigenicity, non-toxicity and non-allergenicity. Besides these epitopes, two adjuvants were also added at the N terminal (human beta defensin-3) and at C terminal (human beta defensin-2) of the vaccine for increasing the antigenicity. Adjuvant were linked via EAAAK linkers to the epitopes, HTL epitopes were linked via GPGPG linkers, while CTL epitopes were linked with AAY linkers as shown in **Figure 1**. The constructed vaccine sequence was again checked for antigenicity, non-allergenic, non-toxicity and it was fulfilling all the criteria.

3.6 Prediction of physicochemical parameters of the constructed vaccine sequence

The physicochemical parameters of the vaccine sequence were predicted by the ProtParam server. The molecular weight of the construct was predicted to be 38.8 KDa, and the theoretical PI value was 9.92. The predicted half-life in *E.coli* was more than 10 hours, and the instability index in the test tube was found to be stable. The aliphatic value of the vaccine sequence was 58.7 and the Grand average of hydropathicity (GRAVY) was -0.348.

3.7 secondary structure prediction of the vaccine sequence

Secondary structure prediction was made using the CFSSP web tool. The result showed the presence of helix: 44.5%, sheet: 35.6%, and turns: 14 %.

3.8 Tertiary structure prediction of the vaccine sequence

The 3D structure of the multiepitope predicted vaccine was predicted using the Rosetta web tool .It uses de-novo structure prediction using deep neural network algorithm to predict the inter-residue distances as well as orientations. Then these orientations are converted to smooth inter-residue constraints followed

by gradient descent energy minimization. Further, coarse-grained models are generated, and full atom refinement is done. It gave 5 best-predicted models, and based on the TM score, one model was selected for further investigation, as shown in **Figure 2A**. Further to validate the predicted model, Ramachandran plot analysis was done, and results showed that 96.3% residues were in the favourable region, 2.5 % were in the allowed region while ~ 1% were in the outlier region as shown in **Figure 2B**. Additionally, the PROSA web tool was used to predict the quality of the modeled vaccine , which predicted the Z score of -6.34 . Ramachandran plot and Z score have suggested that the predicted model of protein was valid and could be taken for further analysis.

3.9 Conformational B cell epitope analysis from modelled vaccine

EliPro predicts the antibody epitopes taking protein 3D structure as input. Linear B, as well as discontinuous conformational epitopes, were identified in the vaccine construct using ElliPro, an online server .A total of 8 linear epitopes were predicted, and the sequence of the top 3 epitopes have been reported in **Table 4** and has been shown structurally in **Figure 2D**. And various discontinuous epitope residues were predicted from vaccine sequence length 232-253 (21 epitope residues), between 299-357 (55 epitope residues), between 1-54 (52 epitope residues), between 69-128 (33 epitope residues) and between 168-176 (9 epitope residues) were predicted. The individual score of each of the discontinuous epitopes has been shown in **Figure 3B**.

3.10 Docking of vaccine with TLR-3 receptor

The modeled structure of the vaccine was taken through the energy minimization, equilibration, and MD simulations before docking. The last frame from the simulated trajectory was taken further for docking . The simulated structure has been compared with the crude modeled structure, as shown in **Figure 2A**. The TLR-3 structure was retrieved from Protein Data Bank (PDB) having ID 1ZIW. The downloaded structure was prepared and processed for docking using the dock prep tool UCSF chimera software [45]. The simulation was done using the PatchDock server and further refined using FireDock. The best-docked complex had global energy of -14.91 Kcal/mol, and attractive Vander wall energy was -18.1 Kcal/mol, which shows a decent binding affinity of the vaccine towards TLR-3. Further, the best binding pose was investigated for polar interactions using discovery studio visualizer [46] between TLR-3 and vaccine, and it was found that GLN352, SER428, ILE370 of TLR-3 was making the hydrogen bond with TYR260, ARG321, and LYS166 of vaccine respectively as shown in **Figure 3A**.

3.11 MD simulation of TLR-3-vaccine complex

The docked complex was further taken for MD simulations in water using gromacs software. The after minimization and NPT and NVT equilibration, the MD production run was performed for 300 picoseconds to assess the dynamic behavior and stability of the docked complex. The RMSD plot of the trajectory showed a stable complex. After 100 picoseconds, the trajectory was converging with fluctuation between 2.25 Angstrom to 2.5 Angstrom, as shown in **Figure 3C**.

3.12 Reverse translation and codon optimization

The Java codon adaptation tool (Jcat) was used for the optimization of the codon for the proper expression of the protein. *E.coli* strain k12 was chosen as a host, with additional options such as avoid rho-independent transcription terminators, avoid prokaryotic ribosome binding sites, and avoid Cleavage Sites of Restriction Enzymes. The CAI gives the information of codon usage, generally score between 1 and 0.8, while GC contents should be between 40 % to 70%, values lie outside the given margin is suggested to be inefficient. The codon adaptation index (CAI) of the optimized nucleotide sequence of the vaccine was found to be 0.92, with a GC content of 55.6%, which indicates the effective expression of the protein in the *E.coli*.

4. Discussion

The structural polyproteins from SARS-CoV-2 were selected for developing the multiepitope based vaccine. Initially, four structural proteins were chosen, namely spike, membrane, envelope, and nucleocapsid protein, based on their antigenicity prediction, role in facilitating the entry of the virus, and packaging inside the host cells. However, when CTL, HTL, and B cell epitopes were predicted, it was found that epitopes from envelope protein were not able to satisfy the criteria of non-allergenicity and antigenicity simultaneously, and hence envelope protein was not considered further in the study. While epitopes selected from SMN structural polyprotein were satisfying, all the criteria such as antigenicity, non-allergenicity, non-toxicity, and high binding affinity towards MHC and also HTL epitopes were able to induce IFN gamma immune response. The constructed multiepitope vaccine from the selected epitopes from all the three SMN polyproteins was again investigated for antigenicity, and it was found that vaccine construct predicted to be a potent antigen with score 0.60 (predicted by vaxijen). Further, the vaccine construct was looked for its allergenicity, and it was found that the vaccine was a non-allergen with a score of -0.59 (threshold was set on -0.4, predicted by AlgPred tool). Further physicochemical parameters were analyzed for vaccine sequence, and it was predicted to have a molecular weight of 38.8 KDa, PI of 9.92, and half-life inside the *E.coli* >10 hours, which shows that protein can easily express and isolated. The 3D model of the constructed vaccine sequence was predicted from the Rosetta web server. And 8 conformational linear B cell epitopes were found in the modeled structure of the vaccine as predicted by the Elipro web tool. It was evident from the prediction that the constructed vaccine model could easily produce adaptive immune response specific to the SARS-CoV-2 antigens. Further, to investigate the ability of the modeled vaccine to interact with TLR receptors on immune cells, the TLR-3 receptor was docked with the modeled vaccine. The results showed that the modeled vaccine had a good binding affinity towards TLR-3, and it was found that GLN352, SER428, ILE370 of TLR-3 was making the hydrogen bond with TYR260, ARG321, and LYS166 of vaccine respectively. This interaction of vaccine with TLR-3 was suggesting that vaccine have the potential to generate both innate as well as adaptive humoral and cell-mediated immune responses. For investigating the stability and dynamic behavior of the TLR-3-vaccine docked complex, MD simulation was performed, and the RMSD plot of the trajectory was indicating the stable binding of the complex. For efficient expression of the protein inside the *E.coli*, codon optimization was done to improve the translation and transcription efficiency. The constructed vaccine sequence was

reverse translated, and codon adaptation index (CAI) and GC content were assessed, taking *E.coli* (K12) as a host organism. The CAI index of 0.92 and GC content of 55.6 % and half-life was already predicted to be more than 10 hours, suggests the efficient expression of recombinant protein inside the *E.coli*. The overall immunoinformatics study is suggesting that the predicted vaccine could generate specific adaptive immunity against SARS-CoV-2 and could provide a valuable contribution to the management of the COVID-19. However, the major limitation of any in silico study is the predictions, that are based on the tools which work on various machine learning algorithm, and they can't be 100% accurate and same is the case for this study. So these predictions of the modeled vaccine candidate strongly warrant in-vitro and in-vivo study for the practical implications.

5. Conclusions

In this study, a multiepitope (CTL, HTL, and B cell) vaccine construct has been predicted and modeled through immunoinformatic techniques. The predictions were suggesting that the constructed vaccine could generate both humoral and cell-based adaptive immunity towards SARS-CoV-2. Further, it was also predicted that it could easily be expressed inside the *E.coli* strain (K12). This immunoinformatics study can reduce the expenditure and time for vaccine research and could give a significant value in the management of COVID-19. This in silico predictions warrants the in-vitro and in-vivo study to test the practical implications of the predicted vaccine.

6. Declarations

Conflict of interest: None

Ethical approval: We confirm that no experiments have been performed on humans or animals.

7. Results

1. Denison, M.R., et al., *Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity*. RNA Biol, 2011. 8(2): p. 270-9 10.4161/rna.8.2.15013.
2. Fehr, A.R. and S. Perlman, *Coronaviruses: an overview of their replication and pathogenesis*. Methods Mol Biol, 2015. 1282: p. 1-23 10.1007/978-1-4939-2438-7_1.
3. Ye, Z.W., et al., *Zoonotic origins of human coronaviruses*. Int J Biol Sci, 2020. 16(10): p. 1686-1697 10.7150/ijbs.45472.
4. Bhattacharya, S., et al., *Recent apprise on coronavirus and its terrible insinuations*. Virusdisease, 2020: p. 1-7 10.1007/s13337-020-00582-2.
5. Vabret, A., et al., *Human coronavirus NL63, France*. Emerg Infect Dis, 2005. 11(8): p. 1225-9 10.3201/eid1108.050110.
6. Li, Q., et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia*. New England Journal of Medicine, 2020 10.1056/NEJMoa2001316.

7. Hui, D.S., et al., *The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China*. International Journal of Infectious Diseases, 2020. 91: p. 264-266.
8. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China*. The Lancet, 2020.
9. Lechien, J.R., et al., *Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study*. European Archives of Oto-Rhino-Laryngology, 2020: p. 1-11.
10. Nicastrì, E., et al., *National institute for the infectious diseases “L. Spallanzani”, IRCCS. Recommendations for COVID-19 clinical management*. Infectious Disease Reports, 2020. 12(1).
11. Hoffmann, M., et al., *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor*. Cell, 2020 10.1016/j.cell.2020.02.052.
12. Iwata-Yoshikawa, N., et al., *TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection*. J Virol, 2019. 93(6) 10.1128/JVI.01815-18.
13. Prajapat, M., et al., *Drug targets for corona virus: A systematic review*. Indian J Pharmacol, 2020. 52(1): p. 56-65 10.4103/ijp.IJP_115_20.
14. Brunel, J., et al., *Sequence of events in measles virus replication: role of phosphoprotein-nucleocapsid interactions*. J Virol, 2014. 88(18): p. 10851-63 10.1128/JVI.00664-14.
15. Surjit, M. and S.K. Lal, *The SARS-CoV nucleocapsid protein: a protein with multifarious activities*. Infect Genet Evol, 2008. 8(4): p. 397-405 10.1016/j.meegid.2007.07.004.
16. Nemazee, D., *Mechanisms of central tolerance for B cells*. Nature Reviews Immunology, 2017. 17(5): p. 281.
17. Oleinika, K., C. Mauri, and P.A. Blair, *B Cell Activation and B Cell Tolerance*, in *The Autoimmune Diseases*. 2020, Elsevier. p. 171-187.
18. Schmitt, M.E., et al., *The B cell antigen receptor of IgE-switched plasma cells regulates memory IgE responses*. Journal of Allergy and Clinical Immunology, 2020.
19. Gold, M.R. and M.G. Reth, *Antigen receptor function in the context of the nanoscale organization of the B cell membrane*. Annual review of immunology, 2019. 37: p. 97-123.
20. Brink, R. and T.G. Phan, *Self-reactive B cells in the germinal center reaction*. Annual review of immunology, 2018. 36: p. 339-357.
21. Karunakaran, K.P., et al., *Discordance in the Epithelial Cell-Dendritic Cell Major Histocompatibility Complex Class II Immunoproteome: Implications for Chlamydia Vaccine Development*. The Journal of Infectious Diseases, 2020. 221(5): p. 841-850.
22. Peters, B., M. Nielsen, and A. Sette, *T Cell Epitope Predictions*. Annual Review of Immunology, 2020. 38.
23. Gaudino, S.J. and P. Kumar, *Cross-talk between antigen presenting cells and T cells impacts intestinal homeostasis, bacterial infections, and tumorigenesis*. Frontiers in immunology, 2019. 10.

24. Miller, C.T., *The Expression Of Connexin-43 By CD11c+ Dendritic Cells Is Required to Maintain CD4+ Foxp3+ Regulatory T Cell Population in Peripheral Lymphoid Organs*. 2019.
25. Rich, R.R. and D.D. Chaplin, *The human immune response*, in *Clinical Immunology*. 2019, Elsevier. p. 3-17. e1.
26. Toubi, E. and Z. Vadasz, *Innate immune-responses and their role in driving autoimmunity*. *Autoimmunity reviews*, 2019. 18(3): p. 306-311.
27. Larsen, M.V., et al., *Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction*. *BMC Bioinformatics*, 2007. 8: p. 424 10.1186/1471-2105-8-424.
28. Zhang, Q., et al., *Immune epitope database analysis resource (IEDB-AR)*. *Nucleic Acids Res*, 2008. 36(Web Server issue): p. W513-8 10.1093/nar/gkn254.
29. Dhanda, S.K., P. Vir, and G.P. Raghava, *Designing of interferon-gamma inducing MHC class-II binders*. *Biol Direct*, 2013. 8: p. 30 10.1186/1745-6150-8-30.
30. Saha, S. and G.P. Raghava, *Prediction of continuous B-cell epitopes in an antigen using recurrent neural network*. *Proteins*, 2006. 65(1): p. 40-8 10.1002/prot.21078.
31. Ponomarenko, J., et al., *ElliPro: a new structure-based tool for the prediction of antibody epitopes*. *BMC Bioinformatics*, 2008. 9: p. 514 10.1186/1471-2105-9-514.
32. Doytchinova, I.A. and D.R. Flower, *VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines*. *BMC Bioinformatics*, 2007. 8: p. 4 10.1186/1471-2105-8-4.
33. Saha, S. and G.P. Raghava, *AlgPred: prediction of allergenic proteins and mapping of IgE epitopes*. *Nucleic Acids Res*, 2006. 34(Web Server issue): p. W202-9 10.1093/nar/gkl343.
34. Gupta, S., et al., *In silico approach for predicting toxicity of peptides and proteins*. *PLoS One*, 2013. 8(9): p. e73957 10.1371/journal.pone.0073957.
35. Garg, V.K., et al., *MFPPI - Multi FASTA ProtParam Interface*. *Bioinformatics*, 2016. 12(2): p. 74-77 10.6026/97320630012074.
36. Kumar, T.A., *CFSSP: Chou and Fasman secondary structure prediction server*. *Wide Spectrum*, 2013. 1(9): p. 15-19.
37. Yang, J., et al., *Improved protein structure prediction using predicted interresidue orientations*. *Proc Natl Acad Sci U S A*, 2020. 117(3): p. 1496-1503 10.1073/pnas.1914677117.
38. Willard, L., et al., *VADAR: a web server for quantitative evaluation of protein structure quality*. *Nucleic Acids Res*, 2003. 31(13): p. 3316-9 10.1093/nar/gkg565.
39. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins*. *Nucleic Acids Res*, 2007. 35(Web Server issue): p. W407-10 10.1093/nar/gkm290.
40. Hess, B., et al., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*. *J Chem Theory Comput*, 2008. 4(3): p. 435-47 10.1021/ct700301q.
41. Schneidman-Duhovny, D., et al., *PatchDock and SymmDock: servers for rigid and symmetric docking*. *Nucleic Acids Res*, 2005. 33(Web Server issue): p. W363-7 10.1093/nar/gki481.

42. Mashiach, E., et al., *FireDock: a web server for fast interaction refinement in molecular docking*. Nucleic Acids Res, 2008. 36(Web Server issue): p. W229-32 10.1093/nar/gkn186.
43. Grote, A., et al., *JCat: a novel tool to adapt codon usage of a target gene to its potential expression host*. Nucleic Acids Res, 2005. 33(Web Server issue): p. W526-31 10.1093/nar/gki376.
44. Das, S., B. Chottopadhyay, and S. Sahoo, *Comparative Analysis of Predicted Gene Expression among Crenarchaeal Genomes*. Genomics Inform, 2017. 15(1): p. 38-47 10.5808/GI.2017.15.1.38.
45. Goddard, T.D., C.C. Huang, and T.E. Ferrin, *Software extensions to UCSF chimera for interactive visualization of large molecular assemblies*. Structure, 2005. 13(3): p. 473-82 10.1016/j.str.2005.01.006.
46. BIOVIA, D.S., *Discovery Studio Modeling Environment*. 2020.

Tables

Due to technical limitations the Tables are available as a download in the Supplementary Files.

Figures

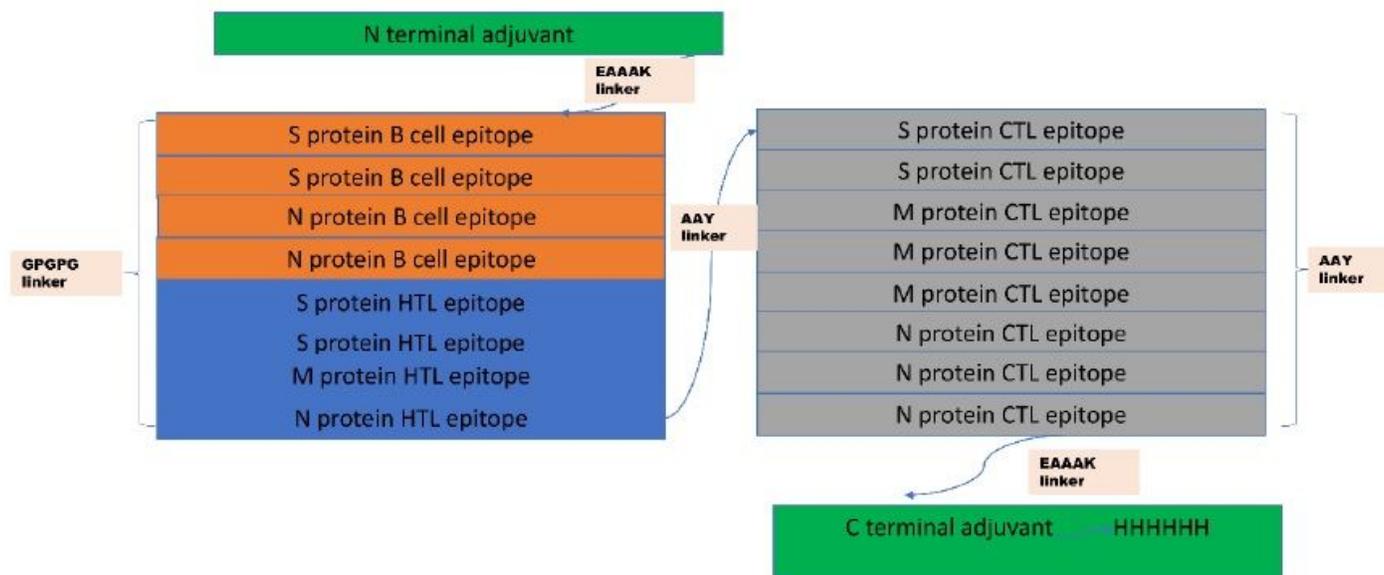


Figure 1

Structure of the final multi-epitope based vaccine. At the C terminal, an adjuvant human B defesin3 has been added, and then it is linked with B cell epitopes using EAAAK linkers. B cell epitopes are linked with HTL with GPGPG linkers, and HTL are linked with CTL with AAY linkers. At N terminal, another adjuvant human B defensin-2 has been linked with six histidine sequences.

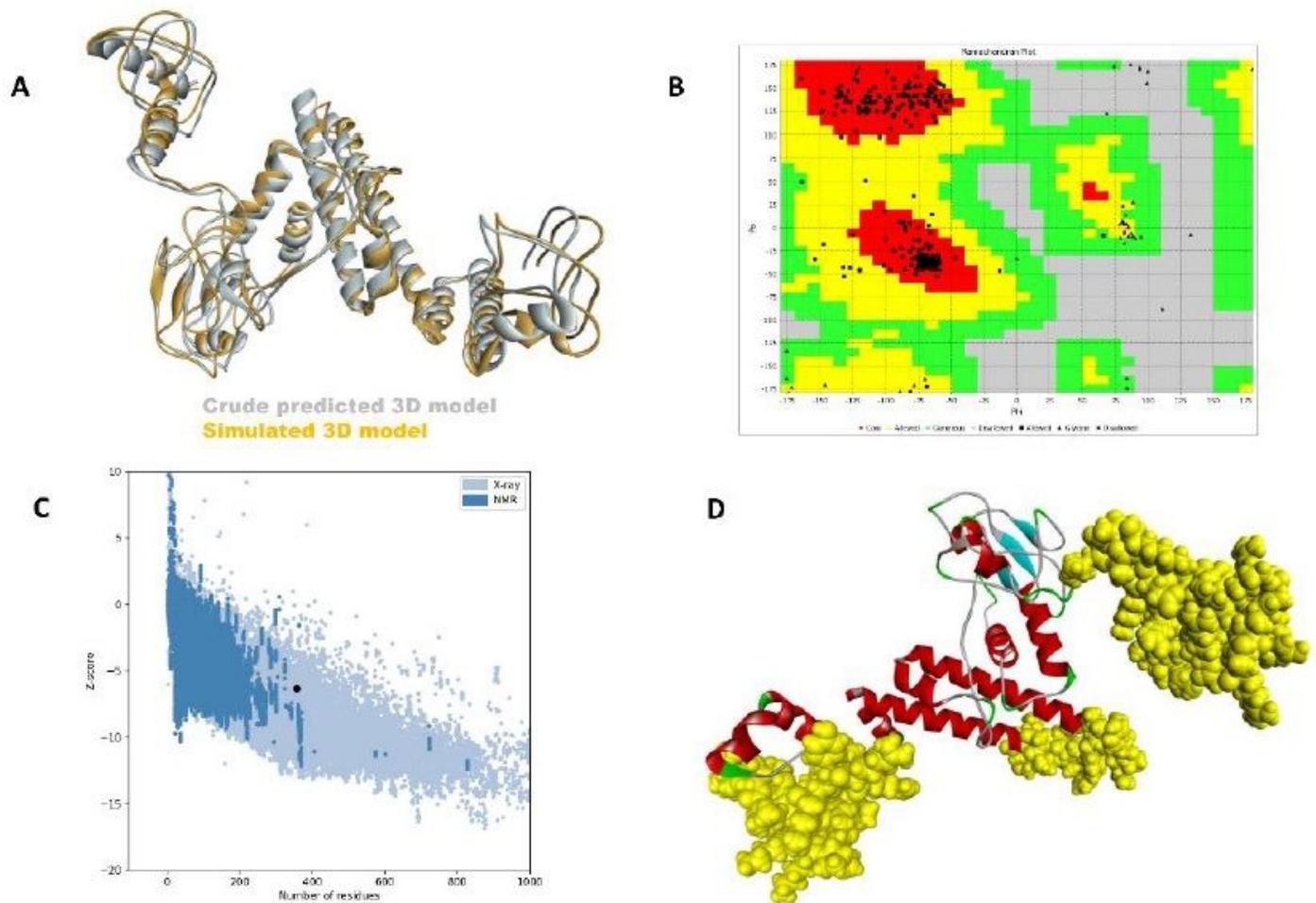


Figure 2

(A) The crude 3D modeled structure of the vaccine (grey) has been superimposed with the simulated model (yellow). (B) Ramachandran plot of the 3D modeled vaccine construct. (C) 3D structure validation by PROSA web tool, Z score (-6.34). (D) The top 3 conformational B cell epitopes predicted in the vaccine has been shown with yellow spheres.

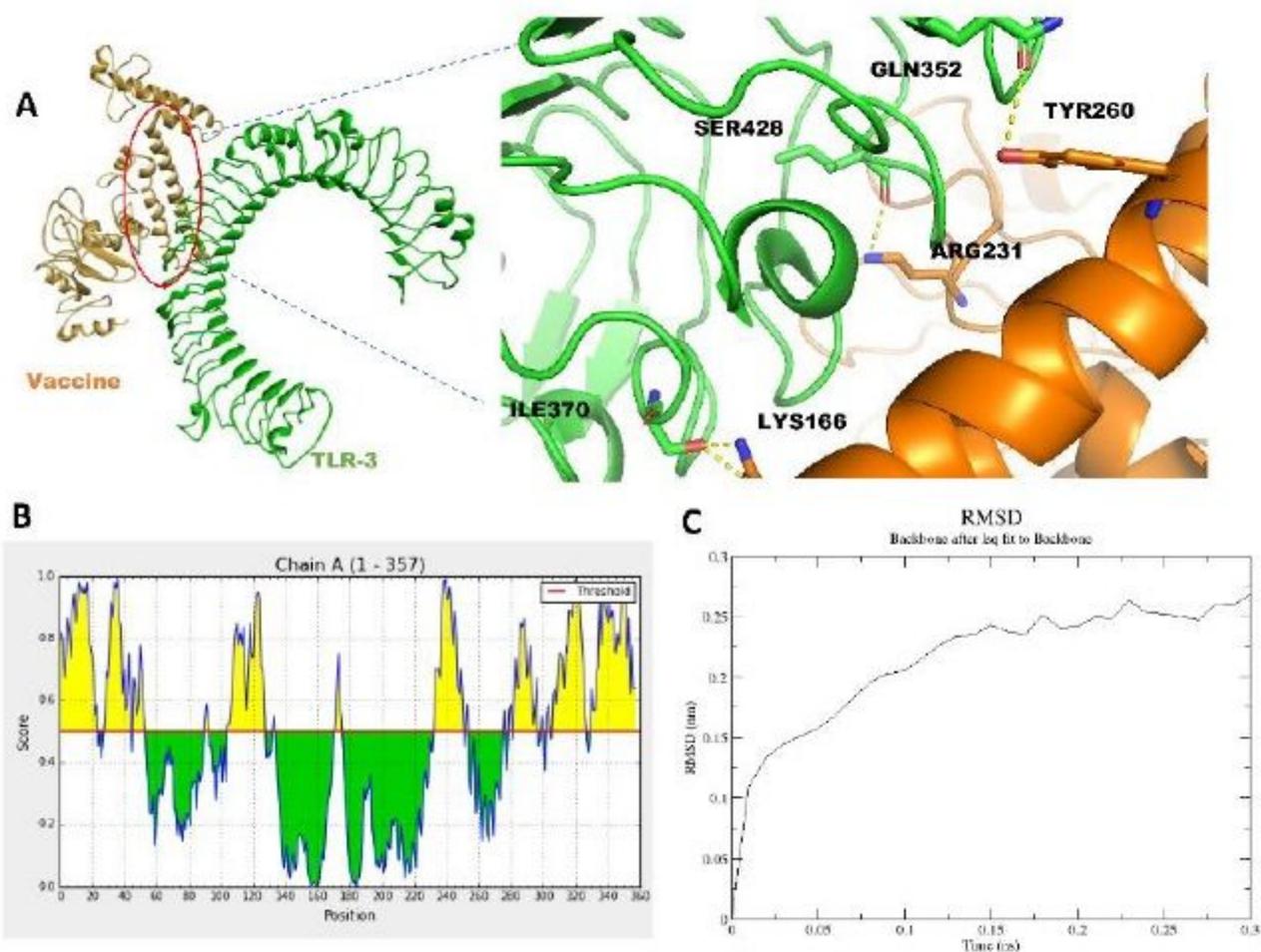


Figure 3

(A) The docked structure of TLR-3 (green color)-modelled vaccine (orange color) complex illustrating critical residues involved in the interactions. (B) The individual score of discontinuous B cell epitopes predicted in the modeled vaccine. (C) RMSD of the simulated TLR3-vaccine shows a stable binding of the complex.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.pdf](#)