

Immunoinformatic design of a COVID-19 subunit vaccine using entire structural immunogenic epitopes of SARS-CoV-2

Esmail Behmard

Medical Biology Research Center, Health Technology Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

Bijan Soleymani (✉ bijansolimani@gmail.com)

Medical Biology Research Center, Health Technology Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

Ali Najafi (✉ najafi74@bmsu.ac.ir)

Molecular Biology Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

Ebrahim Barzegari (✉ e.barzegari@kums.ac.ir)

Medical Biology Research Center, Health Technology Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

Research Article

Keywords: COVID-19, SARS-CoV-2, Multi-epitope Vaccine, Molecular Dynamics Simulations, Toll-like Receptor 3, Immunoinformatics

Posted Date: May 8th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-27260/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 Scientific Reports on November 30th, 2020. See the published version at <https://doi.org/10.1038/s41598-020-77547-4>.

Abstract

Coronavirus disease-19 (COVID-19) is an acute resolved disease, with estimated 3.4% case fatality rate. Due to insufficient data and short onset time of the disease, researchers have to change the strategy against its associated virus, SARS-CoV-2. One of these strategies is the use of computational methods in the field of drug and vaccine design, which can greatly reduce the time and cost of the therapeutic or immunogenic development projects. In this study, we employed various immunoinformatics tools to design a multi-epitope vaccine polypeptide with the highest potential for activating the human immune system against SARS-CoV-2. The initial epitope set was extracted from the whole set of viral structural proteins. Their potential non-toxic and non-allergenic T-cell and B-cell binding and cytokine inducing epitopes were then identified through a priori immunoinformatic prediction. Selected epitopes were bonded to each other with appropriate links. A suitable adjuvant was added to the N-terminus of the vaccine polypeptide sequence to increase its immunogenicity. Molecular modelling of the 3D structure of the vaccine polypeptide, docking, molecular dynamics simulations and free energy calculations confirmed that the designed vaccine had high affinity for Toll-like receptor 3 binding, and that the vaccine-receptor complex was highly stable. Therefore, the designed polypeptide is promising for antigenicity and inducing an effective and safe immune response against SARS-CoV-2 inside the human body.

Introduction

Coronaviruses (CoVs) named for the crown-like spikes on their surface are grouped into four main types, known as alpha, beta, gamma and delta¹. In recent decades, new human beta-CoVs have evolved from animal-infecting types and led to intercontinental pneumonic outbreaks. Severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002, and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 engendered epidemics. SARS-CoV-2 was a return of coronaviruses in December 2019 which causes the ongoing coronavirus disease 2019 (COVID-19). The spill-over was from Chinese horseshoe bat SARS-like coronavirus transmitted to humans from a Wuhan seafood market. By 23th April, there have been 2,544,792 cases with 175,694 deaths globally².

Identifying preventive options, i.e. a COVID-19 vaccine, in a timely manner is critical for the response to the SARS-CoV-2 mass contagion. Live-attenuated vaccines with a long history of success, as well as inactivated whole virus (IWV), DNA, vectored, subunit, and self-assembling virus-like particle vaccines are the diverse candidates for immunogen development, each with particular advantages and disadvantages. Vaccines based on chemically inactivated CoV virions have led to production of neutralizing antibodies with different levels of protection, but they raise potential side effects and biosafety concerns³. DNA vaccines expressing full-length spike protein or its fragments, as well as DNA priming coupled with protein boosting, are also effective against MERS-CoV infection^{4,5}. By intracellular expression of immunogenic antigens, vectored vaccines allow activation of cellular immune response in addition to humoral immunity; they have a proven safety record, however these vectors are limited to presenting one or a reduced number of CoV antigens to the immune system⁶. Subunit vaccines based on recombinant

spike protein has been shown to be very well suited for creating immunogenicity against SARS and MERS^{7,8}. While DNA, vectored and subunit candidates presenting viral epitopes can elicit a focused antibody response, their subviral components may not portray the full antigenic complexity of the virus, resulting in limited protective efficacy or immunopathology due to unbalanced immune responses⁶. One promising option to confront this issue is the subunit vaccines harbouring numerous and diverse antigenic elements, allowing to produce a more comprehensive spectrum of native viral antigens.

The process of vaccine development takes at least ten years from bench research to approved vaccine use⁹. Identification of apt targets for developing vaccine immunogens can be promoted using chemistry and topology, and it is becoming more of a rational design exercise which may considerably reduce both the time and costs required for vaccine development and production¹⁰. In this study, it was hypothesized that advanced computational screening could identify suitable peptides for the construction of a subunit vaccine. We used the full set of the putative structural proteins from SARS-CoV-2 as targets for creating both B-cell and T-cell immunity. By integrating these peptides together, we develop a multi-epitope-based vaccine polypeptide to bind to Toll-like receptor 3 (TLR-3) and elicit an effective immune response without causing toxicity and sensitivity. We also considered incorporating an adjuvant for augmenting the host antigen-specific immune response. As this method has been validated experimentally with other pathogenic epitopes¹¹, we suggest that the proposed structural formulation would be able to generate an effective immune response against the novel virus, as a vaccine *in vivo*. The wet lab researchers are expected to validate our design, hoping to reach protection for the healthy community against the COVID-19 pandemic.

Results

The basic steps of the procedure for designing the multi-epitope vaccine are shown in Fig. 1.

Structural T-cell and B-cell epitopes of SARS-CoV-2. The predicted antigenic CTL epitopes harboured in the structural proteins of SARS-CoV-2 are listed in Supplementary Tables S1-S4 online for S, E, M and N proteins, respectively. As well, the IEDB MHC-II prediction tool was applied to predict HTL (15-mer) epitopes and their MHC-II binding. For linear B-cell epitopes the score value > 0.75 was chosen as criteria for selection. Using appropriate server tools, antigenic, non-toxic and non-allergenic T-cell or B-cell epitopes that were able to induce IL-4, IL-10 and IFN γ cytokines were selected for the design of a multi-epitope vaccine. The final selected CTL, HTL, and linear B-cell epitopes are shown in Supplementary Tables S1-S6 online.

Multi-epitope vaccine polypeptide construction. The total of 34 CTL, and 12 HTL epitopic peptides were fused to each other by KK, and GPGPG linkers, respectively, followed by adjoining a single LBL epitope using KK linker, to create the multi-epitope peptide-based vaccine¹²⁻¹⁴. Furthermore, to boost the immunogenicity of the multi-epitope vaccine, β -defensin (GIINTLQKYYCRVRGGRCVLSCLPKKEEQIGKCSTRGRKCCRRKK) was added as an adjuvant to the amino

terminus of the polypeptide using an EAAAK linker to the first CTL epitope. A total of 694 amino acids were there in the final multi-epitope subunit vaccine construct (Fig. 2 and Supplementary Fig. S1 online).

Immunogenicity, allergenicity and physiochemical properties of the vaccine candidate. Assessment of immunogenic, allergenic and solubility indicated a suitable design for the multi-epitope vaccine (Table 1). The calculated molecular weight of the final vaccine (78.35 kDa) indicates its good antigenic nature^{13,15}, and the pI value of 10.38 shows the basic nature of the final vaccine candidate. The instability index of the final polypeptide (36.73) shows its high stability. The GRAVY value was -0.261, showing that the final structure is slightly hydrophilic in nature and can lead to better connection with other proteins. Moreover, the aliphatic index of the peptide is equal to 81.44, which implies the final peptide has a high thermostability (Table 1). Estimated half-life was 30 h in mammalian reticulocytes *in vitro*, and > 20 h in yeast, and > 10 h in *E. coli in vivo*.

Table 1. Antigenic, allergenic and physiochemical assessments of the primary sequence of final vaccine protein

| Features | Assessment |
|--|--|
| Number of amino acids | 694 |
| Molecular weight | 78351.68Dalton |
| Theoretical pI | 10.37 |
| No. of negatively charged residues (Asp + Glu) | 25 |
| No. of positively charged residues (Arg + Lys) | 137 |
| Extinction coefficient (at 280 nm in H ₂ O) | 120305 M ⁻¹ cm ⁻¹ |
| Instability index | 36.73 |
| Aliphatic index | 81.44 |
| Grand average of hydropathicity (GRAVY) | -0.261 |
| Antigenicity | 0.7342 (VaxiJen v.2.0) |
| Allergenicity | Probable non-allergen (AllergenFP v.1.0) Probable non-allergen (AllerTOP v.2.0) |
| Solubility | 0.914870 (SOLpro) |

3D structure of the vaccine polypeptide. The GalaxyWEB server was employed for modelling and refinement of the 3D structure of the multi-epitope vaccine polypeptide. The modelled 3D structures were then validated to gain a best model. ProSA-web tools showed a Z-score of -4.33, indicating the good quality of the final vaccine structure (Supplementary Fig. S1 online). In addition, for the best model (Supplementary Fig. S1 online), RAMPAGE analysis indicated that 92.3% amino acids of the final structure were in the favoured area, 6.1% were in the allowed area, and 1.6% were in the disallowed area of Ramachandran plot, reflecting a high structural quality of the constructed vaccine model (Supplementary Fig. S1 online).

Conformational and linear B-cell epitopes in the vaccine polypeptide structure. The total of 7 conformational epitopes with a score of 0.702 to 0.882, and 9 linear epitopes with a score of 0.7 to 0.878 were selected as the final B-cell epitopes (Fig. 3, Supplementary Table S7 online). PI value (the score given by ElliPro) of 0.882 shows that 88.2% residues are locating in the predicted ellipsoid area of the epitope and this epitope features the highest solvent accessibility.

Binding of the vaccine structure to TLR-3. The best docked complex of the immune receptor TLR-3 with the vaccine model was identified from among various server outputs by comparing the binding free energy. Analysis of the residues contributing at the protein-protein interface showed that the C-terminal domain of the polypeptide is involved in the interaction, where residues from vaccine polypeptide form polar or non-polar contacts with three domains on the TLR-3 structure (Supplementary Fig. S2 online). The docked complex was further applied for running MD simulation investigations.

MD relaxation and analysis of the receptor-vaccine complex. The MD simulations of the docked multi-epitope-based subunit vaccine with TLR-3 as the receptor was done to achieve information about the conformational changes of TLR-3-vaccine polypeptide complex. Such studies were essential for several vital facets: (1) to comprehend whether the designed vaccine is stable at the bound pocket; (2) to validate that the induced conformational mobility of both TLR-3 receptor and the multi-epitope vaccine has not negative impact on the conformation of the docked proteins; and (3) to corroborate that the epitopes of the multi-epitope vaccine are subject to efficient recognition by the human immune system, causing strong immune response.

Three statistical factors were assessed based on 24000 ps of simulation trajectory (Fig. 4). The root mean squared deviation (RMSD) values of TLR-3 and multi-epitope vaccine in the complex reflect the high conformational stability of the docked molecules. An average RMSD of 0.29 nm with maximum of 0.44 nm realized at 14000 ps was noted for the TLR-3 molecule (Fig. 4A). The RMSD value for multi-epitope vaccine (Fig. 4B) showed that it mostly remained stable during simulation time, with a plateau at about 10500 ps. The observed trends can be attributed to the moving multi-epitope vaccine at TLR-3 binding pocket in an effort to obtain a suitable and stable docked conformation.

Region-wise structural fluctuations of the TLR-3-vaccine polypeptide complex were studied by calculating the root mean squared fluctuation (RMSF) parameter (Fig. 4C-D). The mean RMSF calculated for TLR-3 and the multi-epitope vaccine are 0.2 nm and 0.83 nm, respectively, which are overall in favour of protein residues local stability. Along the vaccine sequence, residues located at the loop regions (such as I137, T138, W204, N376 and R654) have high fluctuation (Fig. 4D). The flexibility of the loop regions is essential for proper holding the vaccine at the binding pocket.

The compactness of the complex structure was estimated by calculating the radius of gyration (Rg) of TLR-3 and multi-epitope peptide molecules (Fig. 4E-F). The graph shows that during the simulation time, TLR-3 and multi-epitope vaccine molecules have mean Rg values of about 3.56 nm and 4.87 nm, respectively. Rg fluctuations is an indication of movements and conformational changes in flexible

regions of the multi-epitope vaccine peptide in the TLR-3 binding pocket. This dynamics seems essential to suitably identify the vaccine and incorporating it in the binding pocket.

Free energy of the binding between vaccine polypeptide and TLR-3. To figure out the strength of the contact between multi-epitope vaccine and TLR-3 structures, the binding free energy between the two molecules was calculated using MMPBSA approach. According to Table 2, the nonpolar element (-136.92 kcal/mol) was an important energy term in the binding free energy of the complex. Our findings clarified that the favourable electrostatic energy ($E_{\text{ele}} = -241.5$ kcal/mol) was covered up by the huge polar energy component ($\Delta G_{\text{GB}} = 240.3$ kcal/mol) in the binding process of the multi-epitope vaccine polypeptide. Therefore, the nonpolar energy was known as the main driving force in the vaccine binding to TLR-3, and this hydrophobic contribution leads to a thermodynamically favourable interaction ($\Delta G_{\text{binding}} = -138.11$ kcal/mol). To further clarify the binding mode, the binding free energy of the protein complex was broken down into residue-residue pairs through the binding free energy decomposition analysis. According to the data, there are several residues of vaccine polypeptide with less than -1.5 kcal/mol free energy of contribution in the binding mechanism. The binding pose of vaccine with the key residues are illustrated in Fig. 5.

Table 2. Binding free energy calculation for the multi-epitope vaccine candidate-TLR-3 complex

| | Energy values (kcal/mol) |
|--|--------------------------|
| $\Delta E_{\text{ele}}^{\text{a}}$ | -241.5 +/- 14.81 |
| $\Delta E_{\text{vdW}}^{\text{b}}$ | -118.81 +/- 9.0 |
| $\Delta G_{\text{GB}}^{\text{c}}$ | 240.3 +/- 16.71 |
| $\Delta G_{\text{SA}}^{\text{d}}$ | -18.11 +/- 1.1 |
| $\Delta E_{\text{polar}}^{\text{e}}$ | -1.2 +/- 0.4 |
| $\Delta E_{\text{non-polar}}^{\text{f}}$ | -136.92 +/- 4.07 |
| $\Delta G_{\text{binding}}$ | -138.11 +/- 9.46 |

^a Electrostatic contribution, ^b van der Waals contribution, ^c Polar contribution of the solvation effect, ^d Non-polar contribution of solvation effect, ^e $\Delta E_{\text{polar}} = \Delta E_{\text{ele}} + \Delta G_{\text{GB}}$, ^f $\Delta E_{\text{non-polar}} = \Delta E_{\text{vdW}} + \Delta G_{\text{SA}}$.

Discussion

Since the beginning of SARS and MERS epidemics, respectively in 2002 and 2012, until very recently, several research reports and patents have concerned the design of appropriate vaccines¹⁶⁻²¹. Nonetheless, no specific vaccine is presently available in the markets for prevention against these respiratory infectious diseases^{17,22}. This implies an urgent demand for changing the paradigms in the

case of COVID-19. Research on COVID-19 prevention was begun immediately after release of relevant basic data such as its full genomic sequence. Specific epitope regions in SARS-CoV-2 with high homology to SARS-CoV, the best-characterized coronavirus in terms of epitope responses, were identified²³⁻²⁵. Antigenic properties of spike glycoprotein were more focused by theory and experimental researchers²⁶⁻²⁸. A multi-epitope design based on S-protein was proposed²⁹. Despite the current highly active research community as well as the efforts in the industry section in the road to find an optimal immunogenic formulation, the extreme diversity in available design choices with no *a priori* image of their effectiveness is the fundamental obstacle in this route. To overcome this issue, mass design of various vaccine species and development of appropriate screening methods can be proposed as a solution. To date, limited number of designs have been reported. This research offered a first step in designing an alternative candidate for developing a COVID-19 vaccine, and aims to instigate further exploration of design and screening strategies.

Deployment of classic live or attenuated vaccines is associated with biosafety issues such as autoimmune or strong allergic responses, plus difficulties with their synthesis and manufacture. These drawbacks might be addressed by developing fully synthetic peptide-based vaccines³⁰. Integrating multiple immunodominant sites of viral pathogens in the vaccine structure helps augment the antigenic effect. In this study, the amino acid sequence of the whole set of structural proteins of SARS-CoV-2 was included for epitope identification. There is much evidence on immunity-related advantages of each of the S, E, M and N proteins of SARS-COV or SARS-CoV-2^{23,31-34}. This choice, besides, aimed to cope with potential vaccine side effects. Antibody-dependent enhancement of infectivity (ADEI) is one great challenge of subunit vaccines, which refers to the reduced specificity in response elicitation because of the numerous conformations adopted by peptide vaccines¹⁰. Strategies to mitigate this concern include: 1- immunofocusing by considering several most antigenic epitopes³⁵, 2- inclusion of structural proteins: structurally flexible coronaviral epitopes may be of limited value for *in vivo* immunotargeting and require to be replaced by conserved epitopes with low structural plasticity⁶. Structural proteins are characterized by highly conserved sequences, thus the linear epitopes remain highly stable in these proteins and the conformational epitopes preserve their structural patterns during various steps of the virus cycle. Our choice to include the full set of SARS-CoV-2 structural proteins has thus combined the advantage of augmented immunogenicity with both strategies proposed for reducing ADEI.

Reports revealed that multiple antigenic peptides induce stronger B and T cells immune responses than un-conjugated peptide epitopes³⁶⁻³⁸. Thus, consecutive sequences of HTL and CTL epitopic peptides were fused to each other using accepted linkers. The T-cell and B-cell binding, and IFN- γ inducing potential of epitopes were then identified.

Epitope-based peptide vaccines induce relatively weak immune response, when used alone¹⁹. The immunoreactivity could be improved by adding proper adjuvants to this type of vaccines. The vaccine construct in this study was prepared by fusing the epitopic peptides to an immunogenic adjuvant. Use of appropriate adjuvants has also been shown to help induce durable IFN- γ responses; furthermore, lung

eosinophilic immunopathology, though less indicated for subunit vaccines, can be avoided by administration of TLR agonist adjuvants⁶. In this study, the binding of the built vaccine model with its specific immune receptor TLR-3 was validated by performing molecular docking, MD simulation and free energy calculations. While we note that other Toll-like immune receptors such as TLR-2, -4 and -5 may be activated by COVID-19 virus, they were not considered here as their exact roles are unclear with some functioning even to the advantage of the coronavirus³⁹.

Conclusion

In the ongoing urgent situation brought about by SARS-CoV-2, it is hard to fast counteract the circulating disease through preventative or therapeutic measures. The multi-epitope-based subunit vaccine design presented in this study can be promising, as it incorporates *a priori* bioinformatics predictions and up-to-date immunological knowledge. By proposing this vaccine polypeptide, this study also hopes to encourage a rapid design and screening strategy to find one formulation with the highest immunogenicity and biosafety.

Methodology

SARS-CoV-2 structural protein sequences. The amino acid sequences of SARS-CoV-2 structural proteins, including the spike glycoprotein (S), envelope protein (E), membrane protein (M), and nucleocapsid phosphoprotein (N), were retrieved using the NCBI reference genome of the virus (accession number NC_045512.2).

Identifying cytotoxic T lymphocyte (CTL) epitopes. Predicting peptides that are capable of inducing CTL responses is a crucial step in the design of epitope-based vaccine. The MHC-I binding tool of Immune Epitope Database and Analysis Resource (IEDB; <http://tools.iedb.org/mhci>) was used to predict the CD8⁺ T-cell epitopes borne in S, E, M and N proteins^{12,40,41}. In this step, the ANN 4.0 method was set as the prediction method. The human was selected as the source species. The maximum IC₅₀ value was set to 50 nM, and percentile rank < 1 was considered since lower score indicates high affinity^{42,43}.

Identifying helper T lymphocyte (HTL) epitopes. IEDB (<http://www.iedb.org>) was used to predict MHC-II binding of 15-mer epitopes for viral structural proteins against human HLAs such as HLA-DRB1*15:01, HLA-DRB4*01:01, HLA-DRB3*01:01, HLA-DRB5*01:01, HLA-DRB1*03:01, HLA-DRB3*02:02 and HLA-DRB1*07:01 (Supplementary Table S8 online), using NN-align 2.3 method^{13,14,44,45}. Epitopes with the IC₅₀ values < 50 nM have high affinity. The maximum value of 500 nM and 5000 nM indicate intermediate and low affinity of epitopes, respectively⁴⁶. The 15-mer epitopes with IC₅₀ values < 50 nM were considered for next analysis⁴⁶⁻⁴⁸.

B-cell epitopes prediction. The ABCpreds server was used to predict 16-mer linear B-lymphocyte (LBL) epitopes with a threshold of 0.51⁴⁹. In addition, the ElliPro tool of IEDB was utilized to predict

conformational and linear B-cell epitopes of the vaccine polypeptide.

Assessment of identified epitopes for antigenicity, allergenicity, and toxicity. The antigenic potential of each of the T and B cells epitopes was predicted by VaxiJen v2.0 with a threshold of 0.4⁵⁰. The predicted T and B cells epitopes were then further evaluated to check their toxicity and allergenicity, with ToxinPred and AllergenFP v1.0 server, respectively^{51,52}. Then, the ability of each of the HTL and B cell epitopes (CD4⁺) to induce the secretion of cytokines, such as interferon-gamma (IFN- γ), interleukin-4 (IL-4) and interleukin-10 (IL-10), to overcome the inflammatory response and prevent tissue damage was predicted by IFNepitope, IL4pred and IL10pred server tools, respectively⁵³⁻⁵⁵. The IL4pred and IL10pred operations were carried out based on SVM method and hybrid method respectively, with other parameters kept as default^{12,13}.

Designing the multi-epitope vaccine polypeptide construct. To develop a multi-epitope vaccine, we selected those predicted epitopes with high antigenic potential, which were not identified as allergic or toxic, and had good solubility when highly expressed. To construct the vaccine polypeptide, the selected CTL, HTL and LBL epitopes were fused together using appropriate linkers^{56,57}. An adjuvant was also considered to increase the immunogenic capacity of the multi-epitope vaccine^{12,46}. Accordingly, a 45-amino acid sequence prepared from β -defensin-2 protein was added to the N terminus of the vaccine sequence using the EAAAK linker⁵⁷.

Immunogenic, allergenic and physicochemical evaluation of vaccine construct. The antigenicity of the multi-epitope vaccine polypeptide was predicted utilizing the VaxiJen v2.0 tool, with the threshold value 0.4⁵⁰. The allergenicity of the vaccine was analysed using AllerTOP v.2.0 and AllergenFP v.1.0 servers^{51,58}. The ProtParam server was employed to evaluate the physical chemistry properties of the vaccine construct, such as amino acid composition, molecular weight, theoretical isoelectric point (pI), grand average of hydropathicity (GRAVY), aliphatic and instability index, and *in vitro* and *in vivo* half-life⁵⁹.

Vaccine polypeptide structure modelling, refinement and validation. The SOPMA server was applied for analysing the secondary structural properties of the multi-epitope vaccine polypeptide⁶⁰. The GalaxyWEB server was employed for modelling and refinement of the 3D structure model⁶¹. The server relaxes the model structure using repacking and molecular dynamics (MD) simulation. Next, RAMPAGE server and ProSA-web were used to validate the refined 3D model^{62,63}. All of these tools give us the overall quality of the 3D structure of the peptide vaccine.

Molecular docking of the vaccine polypeptide to TLR-3. The binding of antigenic agents with the target immune cell protein is crucial for the creation of a suitable immune system response. For analysing the binding pattern of the multi-epitope vaccine polypeptide with TLR-3 (PDB ID: 2A0Z)⁶⁴, molecular docking analysis was performed by Hdock, Zdock, Cluspro and Hawkdock⁶⁵⁻⁶⁸. Among the molecular species docked by each of these applications, the best outputs were extracted, followed by the four docked

molecules uploaded into the Hawkdock program, and the free energy of binding of each complex calculated. Based on this screening, it was determined that the model selected from the Cluspro output had the best binding free energy. Therefore, this molecular species was chosen as the primary structure for initiating the molecular dynamics (MD) simulation.

MD simulation of the vaccine-receptor complex. To determine the structural stability and to study the molecular details of the interactions between TLR-3 and the multi-epitope vaccine polypeptide in the docked conformation, MD simulation was performed. Briefly, the system including vaccine polypeptide-TLR-3 was simulated by the Gromacs-2020 package applying OPLS-AA force field⁶⁹. The complex system was solvated using TIP3P water model. Then, the genion module was utilized to neutralize the whole system. Next, the conjugate gradient algorithm was applied to minimize the energy of the system. In the NVT ensemble, the temperature of the system gradually increased from 0 to 310 K during 400 ps. Subsequently, in the NPT ensemble, 500 ps simulation was carried out at a pressure of 1 atm and a temperature of 310 K. Production simulation for 24000 ps was then implemented. The particle-mesh Ewald (PME) and the LINCS algorithms were applied to assess all electrostatic connections and to restrain all bond lengths in the protein, respectively. Moreover, periodic boundary condition was utilized during the simulation. The final coordinates obtained for the complex system were analysed with classic MD analyses, plus the MMPBSA method for calculating the free energy of intermolecular interactions⁷⁰. The results were visualized by Pymol (Schrodinger L.L.C.).

Declarations

Author Contributions

EB (First author) performed data curation, analysis, validation and visualization; and wrote the original draft. BS contributed in resource and funding acquisition, and writing the original draft. AN contributed in project administration, resources and funding acquisition, and validation. EB (Last author) conceptualized the work, defined methodology, and contributed in project administration, supervision, visualization, completing the draft. All authors reviewed the final manuscript.

Competing Interests Statement: The authors declare no competing interests.

References

1. Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nature reviews. Microbiology* **17**, 181-192, doi:10.1038/s41579-018-0118-9 (2019).
2. World Health Organization. *Coronavirus disease 2019 (COVID-19), Situation report-94, 23 April 2020*, <www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf?sfvrsn=b8304bf0_4> (2020).
3. Tseng, C. T. *et al.* Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus. *PloS one* **7**, e35421,

- doi:10.1371/journal.pone.0035421 (2012).
4. Muthumani, K. *et al.* A synthetic consensus anti-spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates. *Science translational medicine* **7**, 301ra132, doi:10.1126/scitranslmed.aac7462 (2015).
 5. Wang, L. *et al.* Evaluation of candidate vaccine approaches for MERS-CoV. *Nature communications* **6**, 7712, doi:10.1038/ncomms8712 (2015).
 6. Enjuanes, L. *et al.* Molecular Basis of Coronavirus Virulence and Vaccine Development. *Advances in virus research* **96**, 245-286, doi:10.1016/bs.aivir.2016.08.003 (2016).
 7. Wang, J. *et al.* The adjuvanticity of an *O. volvulus*-derived rOv-ASP-1 protein in mice using sequential vaccinations and in non-human primates. *PloS one* **7**, e37019, doi:10.1371/journal.pone.0037019 (2012).
 8. Zhang, N. *et al.* Identification of an ideal adjuvant for receptor-binding domain-based subunit vaccines against Middle East respiratory syndrome coronavirus. *Cellular & molecular immunology* **13**, 180-190, doi:10.1038/cmi.2015.03 (2016).
 9. Papaneri, A. B. *et al.* Middle East respiratory syndrome: obstacles and prospects for vaccine development. *Expert review of vaccines* **14**, 949-962, doi:10.1586/14760584.2015.1036033 (2015).
 10. Graham, B. S., Gilman, M. S. A. & McLellan, J. S. Structure-Based Vaccine Antigen Design. *Annual review of medicine* **70**, 91-104, doi:10.1146/annurev-med-121217-094234 (2019).
 11. Khan, M. K. *et al.* In silico predicted mycobacterial epitope elicits in vitro T-cell responses. *Molecular immunology* **61**, 16-22, doi:10.1016/j.molimm.2014.04.009 (2014).
 12. Abdulla, F., Adhikari, U. K. & Uddin, M. K. Exploring T & B-cell epitopes and designing multi-epitope subunit vaccine targeting integration step of HIV-1 lifecycle using immunoinformatics approach. *Microb Pathog* **137**, 103791, doi:10.1016/j.micpath.2019.103791 (2019).
 13. Khatoon, N., Pandey, R. K. & Prajapati, V. K. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci Rep* **7**, 8285, doi:10.1038/s41598-017-08842-w (2017).
 14. Narula, A., Pandey, R. K., Khatoon, N., Mishra, A. & Prajapati, V. K. Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **61**, 4-15, doi:10.1016/j.meegid.2018.03.007 (2018).
 15. Berzofsky, J. A. & Berkower, I. in *Fundamental Immunology* Vol. 631-684 (ed W.E. Paul) (Raven Press, 1993).
 16. Oany, A. R., Emran, A. A. & Jyoti, T. P. Design of an epitope-based peptide vaccine against spike protein of human coronavirus: an in silico approach. *Drug design, development and therapy* **8**, 1139-1149, doi:10.2147/DDDT.S67861 (2014).
 17. Srivastava, S. *et al.* Design of novel multi-epitope vaccines against severe acute respiratory syndrome validated through multistage molecular interaction and dynamics. *Journal of*

- Biomolecular Structure and Dynamics* **37**, 4345-4360 (2019).
18. Srivastava, S., Kamthania, M., Singh, S., Saxena, A. K. & Sharma, N. Structural basis of development of multi-epitope vaccine against Middle East respiratory syndrome using in silico approach. *Infection and drug resistance* **11**, 2377-2391, doi:10.2147/IDR.S175114 (2018).
 19. Tahir Ul Qamar, M. *et al.* Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study. *Journal of translational medicine* **17**, 362, doi:10.1186/s12967-019-2116-8 (2019).
 20. Choi, J., Kim, M. G., Oh, Y. K. & Kim, Y. B. Progress of Middle East respiratory syndrome coronavirus vaccines: a patent review. *Expert opinion on therapeutic patents* **27**, 721-731, doi:10.1080/13543776.2017.1281248 (2017).
 21. Pallesen, J. *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E7348-E7357, doi:10.1073/pnas.1707304114 (2017).
 22. Yong, C. Y., Ong, H. K., Yeap, S. K., Ho, K. L. & Tan, W. S. Recent Advances in the Vaccine Development Against Middle East Respiratory Syndrome-Coronavirus. *Frontiers in microbiology* **10**, 1781, doi:10.3389/fmicb.2019.01781 (2019).
 23. Grifoni, A. *et al.* A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell host & microbe* **27**, 671-680 e672, doi:10.1016/j.chom.2020.03.002 (2020).
 24. Kumar, S., Maurya, V. K., Prasad, A. K., Bhatt, M. L. B. & Saxena, S. K. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *Virusdisease* **31**, 13-21, doi:10.1007/s13337-020-00571-5 (2020).
 25. Ahmed, S. F., Quadeer, A. A. & McKay, M. R. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* **12**, doi:10.3390/v12030254 (2020).
 26. Baruah, V. & Bose, S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *Journal of medical virology* **92**, 495-500, doi:10.1002/jmv.25698 (2020).
 27. Lucchese, G. Epitopes for a 2019-nCoV vaccine. *Cellular & molecular immunology*, doi:10.1038/s41423-020-0377-z (2020).
 28. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, doi:10.1016/j.cell.2020.02.058 (2020).
 29. Bhattacharya, M. *et al.* Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): Immunoinformatics approach. *Journal of medical virology*, doi:10.1002/jmv.25736 (2020).
 30. Skwarczynski, M. & Toth, I. Peptide-based synthetic vaccines. *Chemical science* **7**, 842-854 (2016).
 31. Liu, J. *et al.* The membrane protein of severe acute respiratory syndrome coronavirus acts as a dominant immunogen revealed by a clustering region of novel functionally and structurally defined

- cytotoxic T-lymphocyte epitopes. *The Journal of infectious diseases* **202**, 1171-1180, doi:10.1086/656315 (2010).
32. Ng, O. W. *et al.* Memory T cell responses targeting the SARS coronavirus persist up to 11 years post-infection. *Vaccine* **34**, 2008-2014, doi:10.1016/j.vaccine.2016.02.063 (2016).
33. Peng, H. *et al.* Human memory T cell responses to SARS-CoV E protein. *Microbes and infection* **8**, 2424-2431, doi:10.1016/j.micinf.2006.05.008 (2006).
34. Surjit, M. & Lal, S. K. The SARS-CoV nucleocapsid protein: a protein with multifarious activities. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **8**, 397-405, doi:10.1016/j.meegid.2007.07.004 (2008).
35. Ma, C. *et al.* Searching for an ideal vaccine candidate among different MERS coronavirus receptor-binding fragments—the importance of immunofocusing in subunit vaccine design. *Vaccine* **32**, 6170-6176, doi:10.1016/j.vaccine.2014.08.086 (2014).
36. Tam, J. P. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proceedings of the National Academy of Sciences* **85**, 5409-5413 (1988).
37. Li, W., Joshi, M. D., Singhanian, S., Ramsey, K. H. & Murthy, A. K. Peptide vaccine: progress and challenges. *Vaccines* **2**, 515-536 (2014).
38. Reche, P. A., Fernandez-Caldas, E., Flower, D. R., Fridkis-Hareli, M. & Hoshino, Y. Peptide-based immunotherapeutics and vaccines. *Journal of immunology research* **2014** (2014).
39. Olejnik, J., Hume, A. J. & Muhlberger, E. Toll-like receptor 4 in acute viral infection: Too much of a good thing. *PLoS pathogens* **14**, e1007390, doi:10.1371/journal.ppat.1007390 (2018).
40. Larsen, M. V. *et al.* Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC bioinformatics* **8**, 424, doi:10.1186/1471-2105-8-424 (2007).
41. Yadav, G., Rao, R., Raj, U. & Varadwaj, P. K. Computational modeling and analysis of prominent T-cell epitopes for assisting in designing vaccine of ZIKA virus. *Journal of Applied Pharmaceutical Science* **7**, 116-122, doi:10.7324/japs.2017.70816 (2017).
42. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32**, 511-517, doi:10.1093/bioinformatics/btv639 (2016).
43. Nielsen, M. & Lund, O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC bioinformatics* **10**, 296, doi:10.1186/1471-2105-10-296 (2009).
44. Greenbaum, J. *et al.* Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics* **63**, 325-335, doi:10.1007/s00251-011-0513-0 (2011).
45. Weiskopf, D. *et al.* Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E2046-2053, doi:10.1073/pnas.1305227110 (2013).

46. Duvvuri, V. R. *et al.* Preexisting CD4+ T-cell immunity in human population to avian influenza H7N9 virus: whole proteome-wide immunoinformatics analyses. *PLoS one* **9**, e91273, doi:10.1371/journal.pone.0091273 (2014).
47. Shey, R. A. *et al.* In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. *Sci Rep* **9**, 4409, doi:10.1038/s41598-019-40833-x (2019).
48. Zhou, W. Y. *et al.* Therapeutic efficacy of a multi-epitope vaccine against Helicobacter pylori infection in BALB/c mice model. *Vaccine* **27**, 5013-5019, doi:10.1016/j.vaccine.2009.05.009 (2009).
49. Saha, S. & Raghava, G. P. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* **65**, 40-48, doi:10.1002/prot.21078 (2006).
50. Doytchinova, I. A. & Flower, D. R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics* **8**, 4, doi:10.1186/1471-2105-8-4 (2007).
51. Dimitrov, I., Naneva, L., Doytchinova, I. & Bangov, I. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* **30**, 846-851, doi:10.1093/bioinformatics/btt619 (2014).
52. Gupta, S. *et al.* In silico approach for predicting toxicity of peptides and proteins. *PLoS one* **8**, e73957, doi:10.1371/journal.pone.0073957 (2013).
53. Dhanda, S. K., Gupta, S., Vir, P. & Raghava, G. P. Prediction of IL4 inducing peptides. *Clinical & developmental immunology* **2013**, 263952, doi:10.1155/2013/263952 (2013).
54. Dhanda, S. K., Vir, P. & Raghava, G. P. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct* **8**, 30, doi:10.1186/1745-6150-8-30 (2013).
55. Nagpal, G. *et al.* Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci Rep* **7**, 42851, doi:10.1038/srep42851 (2017).
56. Chen, X., Zaro, J. L. & Shen, W. C. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* **65**, 1357-1369, doi:10.1016/j.addr.2012.09.039 (2013).
57. Mohan, T., Sharma, C., Bhat, A. A. & Rao, D. N. Modulation of HIV peptide antigen specific cellular immune response by synthetic alpha- and beta-defensin peptides. *Vaccine* **31**, 1707-1716, doi:10.1016/j.vaccine.2013.01.041 (2013).
58. Dimitrov, I., Bangov, I., Flower, D. R. & Doytchinova, I. AllerTOP v.2—a server for in silico prediction of allergens. *J Mol Model* **20**, 2278, doi:10.1007/s00894-014-2278-5 (2014).
59. Gasteiger, E. *et al.* in *The Proteomics Protocols Handbook Springer Protocols Handbooks* (ed J.M. Walker) 571-607 (Humana Press, 2005).
60. Geourjon, C. & Deleage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Computer applications in the biosciences : CABIOS* **11**, 681-684, doi:10.1093/bioinformatics/11.6.681 (1995).
61. Ko, J., Park, H., Heo, L. & Seok, C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic acids research* **40**, W294-297, doi:10.1093/nar/gks493 (2012).
62. Wang, W. *et al.* Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in *Gossypium raimondii* and *G. arboreum*. *Data in brief* **9**, 345-348, doi:10.1016/j.dib.2016.05.025

- (2016).
63. Wiederstein, M. & Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research* **35**, W407-410, doi:10.1093/nar/gkm290 (2007).
 64. Bell, J. K. *et al.* The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10976-10980, doi:10.1073/pnas.0505077102 (2005).
 65. Kozakov, D. *et al.* The ClusPro web server for protein-protein docking. *Nature protocols* **12**, 255-278, doi:10.1038/nprot.2016.169 (2017).
 66. Pierce, B. G., Hourai, Y. & Weng, Z. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PloS one* **6**, e24657, doi:10.1371/journal.pone.0024657 (2011).
 67. Weng, G. *et al.* HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA. *Nucleic acids research* **47**, W322-W330, doi:10.1093/nar/gkz397 (2019).
 68. Yan, Y., Zhang, D., Zhou, P., Li, B. & Huang, S. Y. HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic acids research* **45**, W365-W373, doi:10.1093/nar/gkx407 (2017).
 69. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701-1718, doi:10.1002/jcc.20291 (2005).
 70. Kumari, R., Kumar, R. & Lynn, A. g_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *Journal of chemical information and modeling* **54**, 1951-1962, doi:10.1021/ci500020m (2014).

Additional Information

Supplementary information accompanies this paper including Supplementary Tables S1-S8 and Figures S1-S3.

Figures

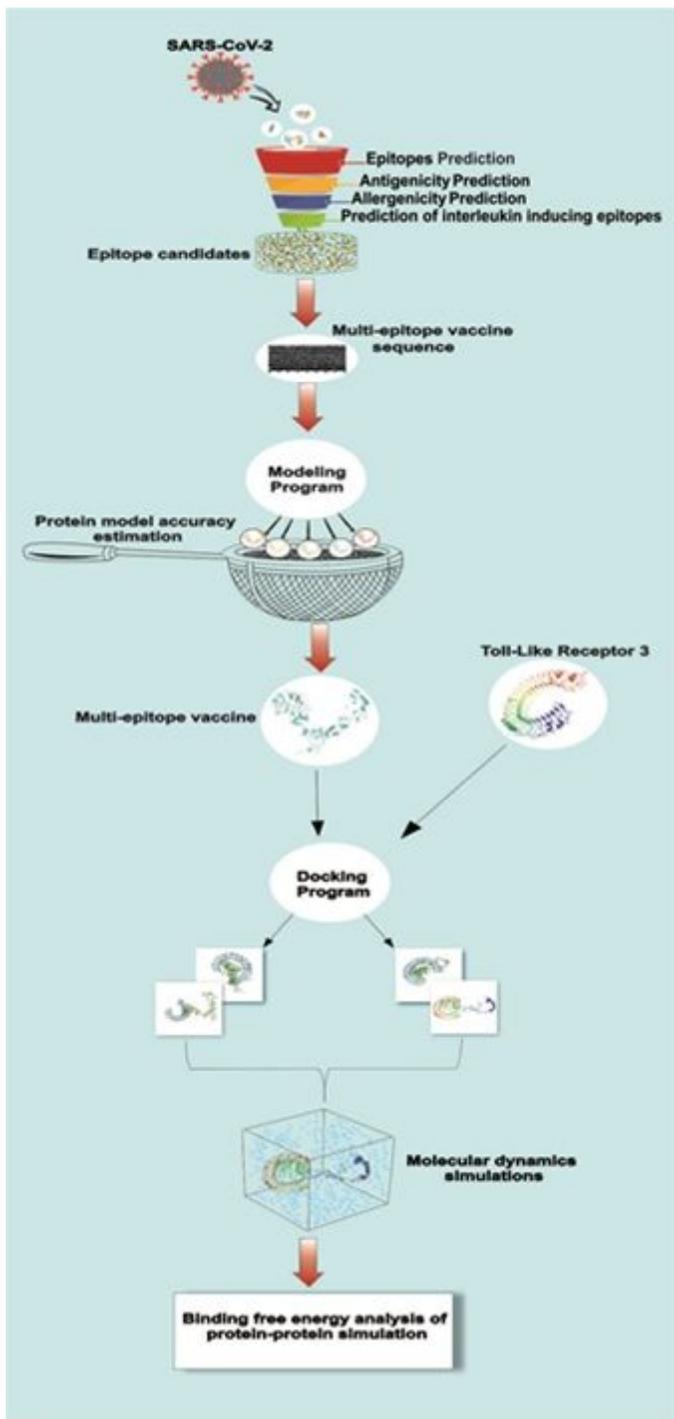


Figure 1

Systemic flowchart of the multi-epitope subunit vaccine building against COVID-19.

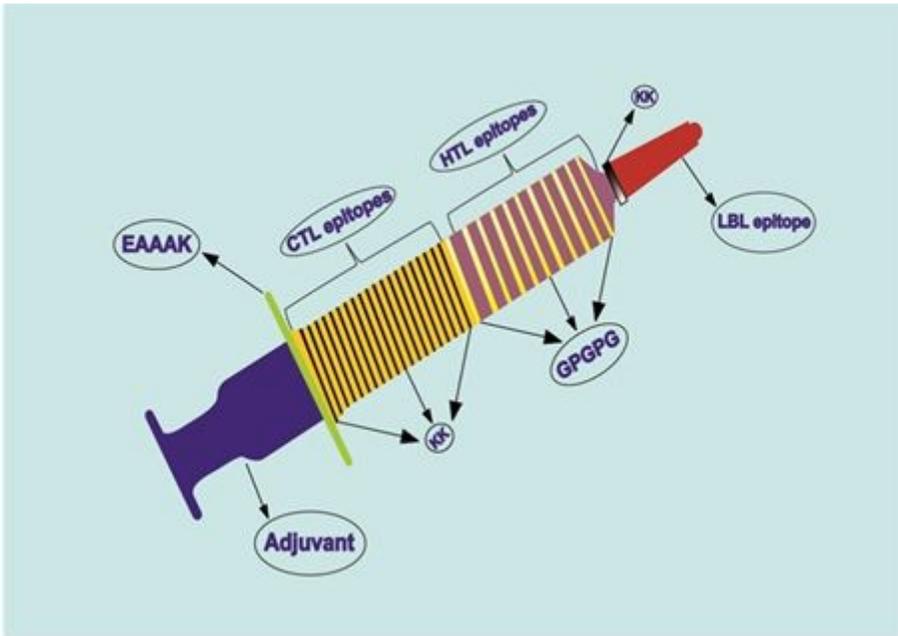


Figure 2

Schematic profile of the multi-epitope subunit vaccine construct of length 694 residues. An adjuvant was added at the N-terminal tail of the vaccine using EAAAK linker, followed by 34 CTL, 12 HTL and 1 LBL epitopes fused by KK and GPGPG linkers.

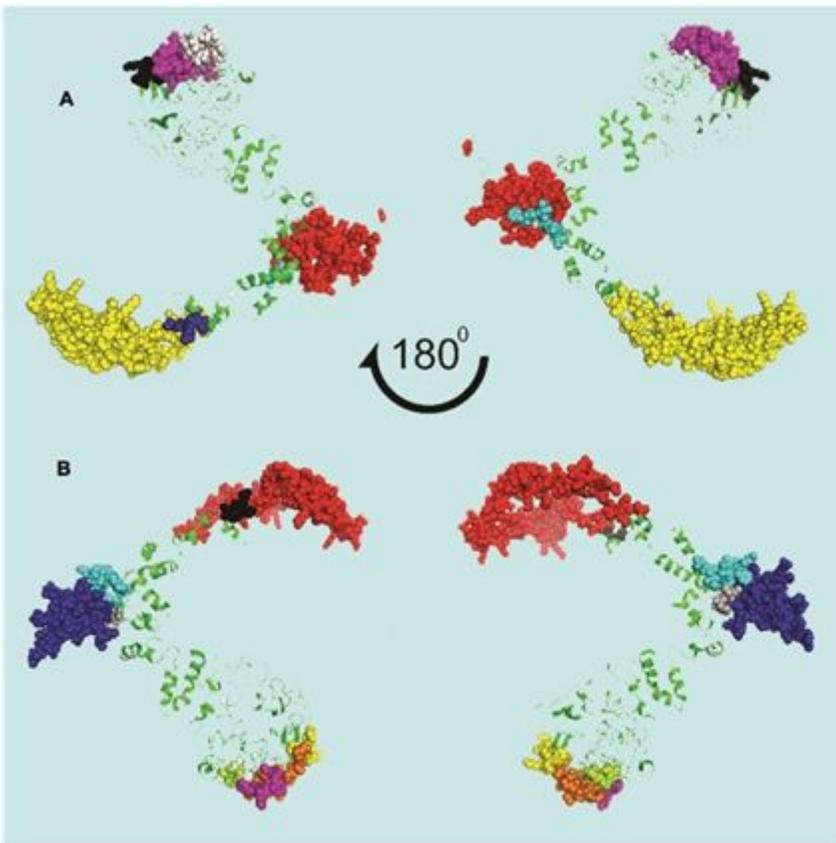


Figure 3

B lymphocyte epitopes present in the designed multi-epitope vaccine. (A) Conformational B cell epitopes shown by spheres, (B) Linear B cell epitopes shown by spheres.

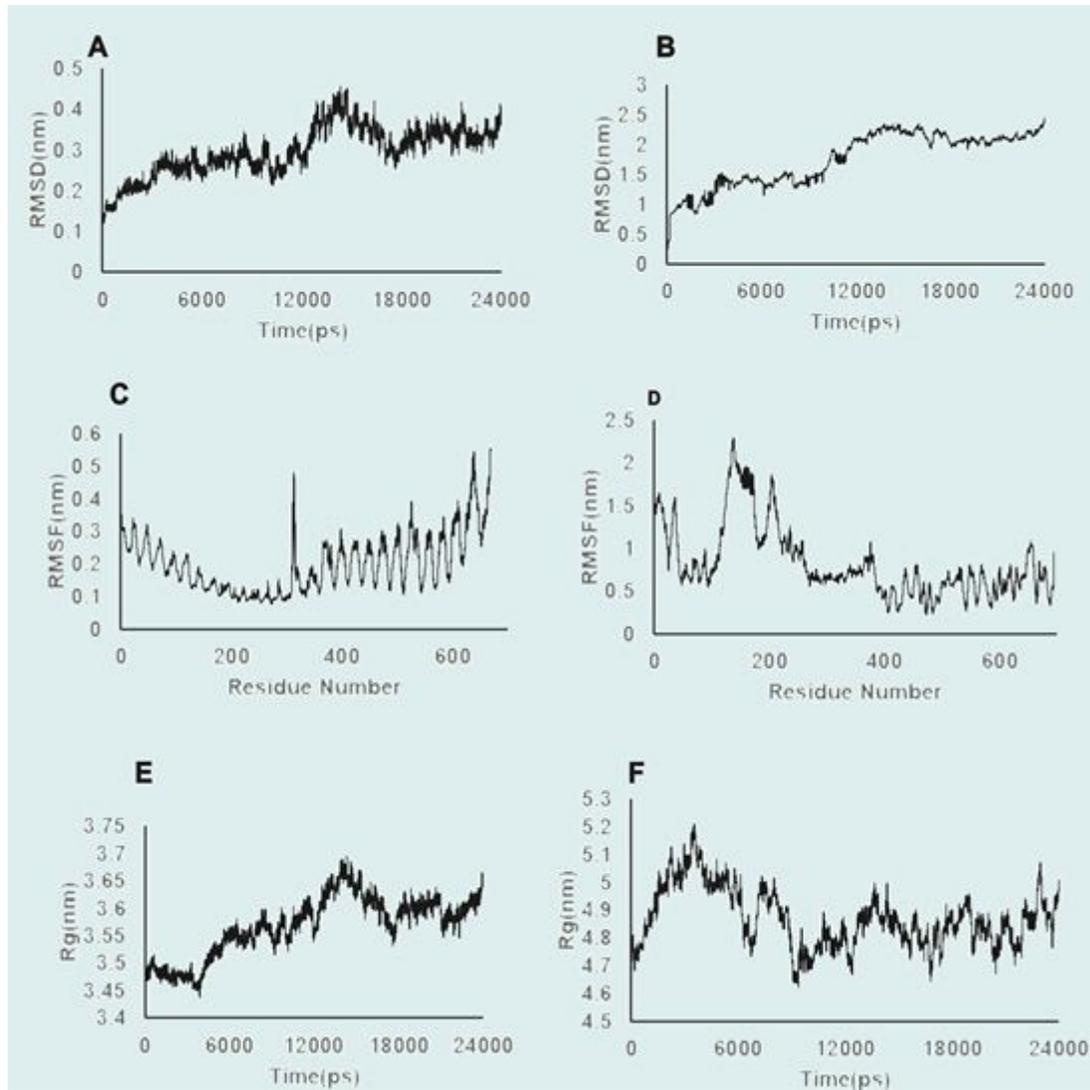


Figure 4

Illustration of the molecular dynamic equilibration for simulation outputs. (A) Root mean squared deviations (RMSDs) of Ca for Toll-like receptor-3 (TLR-3), and (B) for the multi-epitope vaccine polypeptide; (C) Root mean squared fluctuations (RMSFs) of Ca atoms for TLR-3, and (D) for the multi-epitope vaccine polypeptide; (E) Radius of gyration for TLR-3, and (F) for the multi-epitope vaccine polypeptide.

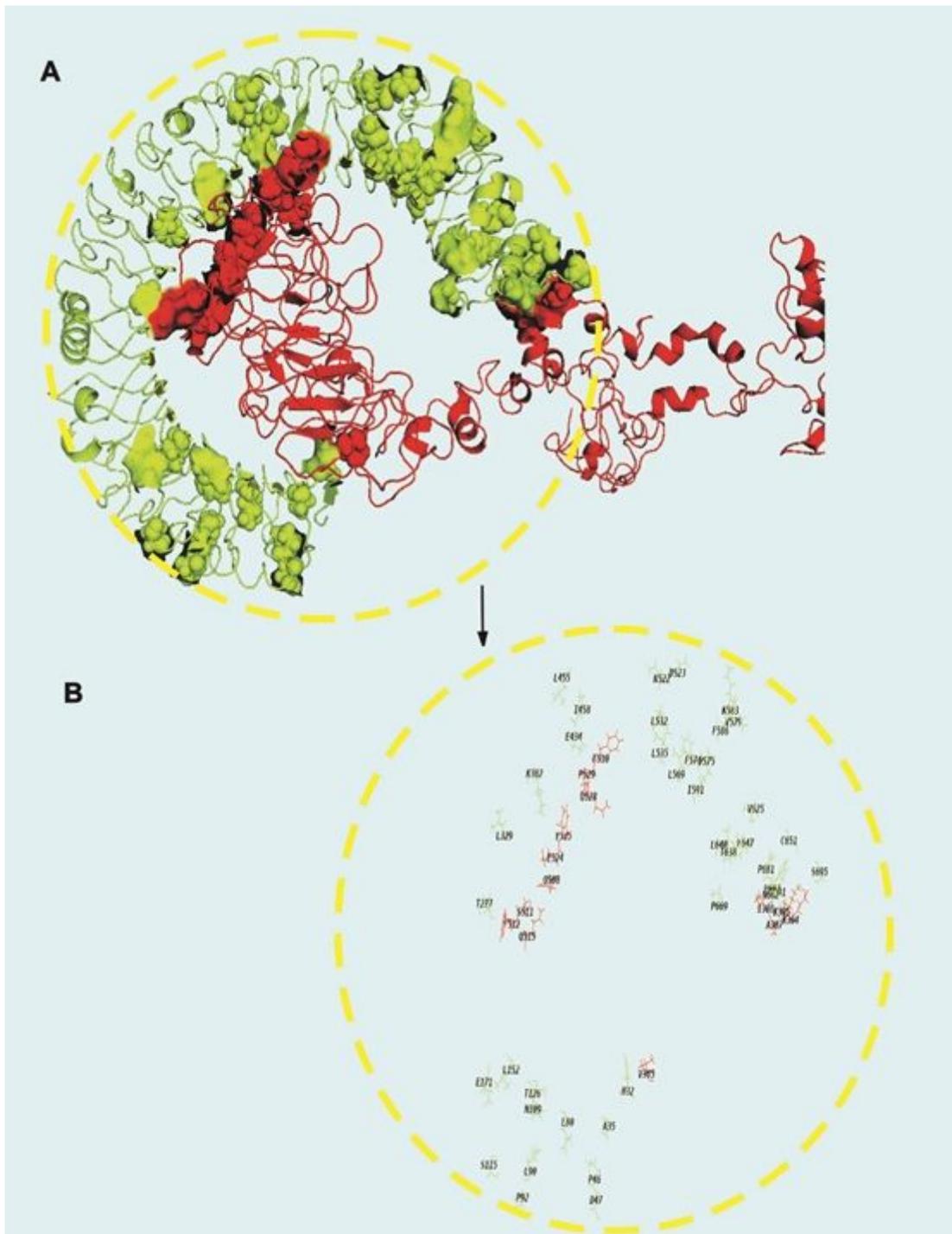


Figure 5

Illustration of the interaction between the vaccine polypeptide and the immune receptor Toll-like receptor-3 (TLR-3). (A) Arrangement of the vaccine protein residues (red spheres) in the binding pocket of TLR-3 (chartreuse spheres); (B) The key residues in the binding free energy. Vaccine polypeptide residues are shown in red, Toll-like receptor-3 residues in chartreuse.

Supplementary Files

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

- [SupplementaryInformation.docx](#)