Immunoinformatic approach to design a multiepitope vaccine targeting to non-mutational hotspot region of structural and non-structural proteins of the SARS CoV2 (#53343)

First submission

Guidance from your Editor

Please submit by **23 Jan 2021** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

6 Figure file(s)
2 Table file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Immunoinformatic approach to design a multiepitope vaccine targeting to non-mutational hotspot region of structural and non-structural proteins of the SARS CoV2

Vandana Solanki [1], **Vishvanath Tiwari** [Corresp. 2]

[1] Department of Biochemistry, Central University of Rajasthan, Ajmer, Rajasthan, India

[2] Department of Biochemistry, Central University Rajasthan, India, Ajmer, Rajasthan, India

Corresponding Author: Vishvanath Tiwari
Email address: vishvanath@curaj.ac.in

Background: The rapid SARS-CoV-2 outbreak caused severe pandemic infection worldwide. The present article aims to design a potential vaccine construct VTC3 targeting the non-mutational region of structural and non-structural proteins of SARS CoV2.

Methods: The presence of different epitopes like T cell, B cell, and IFN-gamma were estimated along with their antigenicity, allergenicity, and toxicity. The location of all the epitopes was determined in virus proteins. Vaccine constructs were evaluated for antigenicity, allergenicity, physicochemical properties, and structural details. The design vaccine construct was validated using docking and molecular dynamics simulation (MDS). Mutational sensitivity profiling of the designed vaccine was performed, and mutations were reconfirmed from the experimental database.

Results: Results identified ten (structural and non-structural) proteins of this virus that have a role in cell adhesion and infection were taken as a target. The different epitopes were predicted, and only extracellular epitopes were selected that do not have cross-reactivity. Finalized epitopes of all proteins were linked using linkers to designed different vaccine constructs. Docking of different constructs with different TLRs and HLA demonstrated a stable and reliable binding affinity of VTC3 with the TLRs and HLAs. MDS analysis further confirms the interaction of VTC3 with TLR1/2 complex and HLA. The VTC3 does not have similarities with the human microbiome, and interacting residues of VTC3 do not have mutations. The present study designs a multiepitope vaccine targeting the non-mutational region of structural and non-structural proteins of the SARS CoV2 using an immunoinformatic approach, which needs to be experimentally validated.

Title: Immunoinformatic approach to design a multiepitope vaccine targeting to non-mutational hotspot region of structural and non-structural proteins of the SARS CoV2.

Vandana Solanki and Vishvanath Tiwari*

Department of Biochemistry, Central University of Rajasthan, Bandarsindri, Ajmer-305817, India

*Corresponding Author
Dr. Vishvanath Tiwari
Department of Biochemistry, Central University of Rajasthan, Bandarsindri, Ajmer-305817, India E-mail: vishvanath@curaj.ac.in; Mobile: + 91-850-300-2573;

Short Title: Multiepitope vaccine against SARS CoV2

Keywords: Multiepitope vaccine; SARS CoV2; Membrane proteins; ORF8 protein; ORF3a protein; Envelope protein; Surface glycoprotein; ORF1ab polyprotein

Abstract:

Background: The rapid SARS-CoV-2 outbreak caused severe pandemic infection worldwide. The present article aims to design a potential vaccine construct VTC3 targeting the non-mutational region of structural and non-structural proteins of SARS CoV2.

Methods: The presence of different epitopes like T cell, B cell, and IFN-gamma were estimated along with their antigenicity, allergenicity, and toxicity. The location of all the epitopes was determined in virus proteins. Vaccine constructs were evaluated for antigenicity, allergenicity, physicochemical properties, and structural details. The design vaccine construct was validated using docking and molecular dynamics simulation (MDS). Mutational sensitivity profiling of the designed vaccine was performed, and mutations were reconfirmed from the experimental database.

Results: Results identified ten (structural and non-structural) proteins of this virus that have a role in cell adhesion and infection were taken as a target. The different epitopes were predicted, and only extracellular epitopes were selected that do not have cross-reactivity. Finalized epitopes of all proteins were linked using linkers to designed different vaccine constructs. Docking of different constructs with different TLRs and HLA demonstrated a stable and reliable binding affinity of VTC3 with the TLRs and HLAs. MDS analysis further confirms the interaction of VTC3 with TLR1/2 complex and HLA. The VTC3 does not have similarities with the human microbiome, and interacting residues of VTC3 do not have mutations. The present study designs a multiepitope vaccine targeting the non-mutational region of structural and non-structural proteins of the SARS CoV2 using an immunoinformatic approach, which needs to be experimentally validated.

50

51

52

53

54 **1. Introduction**

55 Immuno-pathogenesis of the infectious epidemic COVID-19 emerged as serious threat worldwide. In
56 the Coronaviridae family, among four different coronavirus classes (alpha, beta, delta, and gamma), alpha
57 and beta positive-sense RNA virus strains have been confirmed for broadly epidemic infection (de Wilde
58 et al. 2018; Qamar et al. 2020). SARS-CoV2 (Severe Acute Respiratory Syndrome Coronavirus 2) mediated
59 outbreak of the disease was firstly reported in Wuhan city, Hubei province, China, in December 2019
60 (Gorbalenya et al. 2020; Huang et al. 2020; Perlman 2020; Tahir Ul Qamar et al. 2019; World Health 2020;
61 Wu et al. 2020b; Zhu et al. 2020). The outbreak has so far infected more than 10,00,000 patients worldwide
62 on dated 10 May 2020. SARS-CoV2 genome sequence comparison showed that it has almost 96%, 79.5%,
63 and 40% similarities with bat coronavirus, SARS-CoV and MERS-CoV strain respectively (Alamri et al.
64 2020; Benvenuto et al. 2020; Zhou et al. 2020). The clinical symptoms of SARS-CoV2 exhibit up to 14 days
65 in infected people with fever (≥38ºC), diarrhea, dry cough, low peripheral white blood cell count,
66 respiratory disorder, and low lymphocyte count (Huang et al. 2020). SARS-CoV2 mutation prone (Jiayuan
67 et al. 2020) genomic organization is a difficult task to develop a vaccine that is composed of 5′-leader-UTR-
68 replicase-S (Spike)-E (Envelope)-M (Membrane)-ORF6-ORF7a-ORF8-N (Nucleocapsid)-3′UTR-polyA tail.
69 Proteins such as ORF3a, ORF7a, ORF8 function as accessory proteins playing an essential role in viral
70 pathogenesis (Seema 2020; Zhu et al. 2020). Spike protein receptor-binding domain of mutation prone
71 SARS-CoV2 showed the more efficient binding with host's angiotensin-converting enzyme 2 (ACE2) than
72 SARS-CoV (Dong et al. 2020; Gralinski & Menachery 2020; Tian et al. 2020; Wrapp & Wang 2020; WU et al.
73 2020a).

74 To overcome the issues such as cost & time mediated traditional method (monoclonal
75 oligonucleotides, small drug molecules) with rapid detection, isolation, disease prevention, and control
76 measures (Li & De Clercq 2020), virus genome was analyzed by computational methods (Chen et al. 2020).
77 Recently published immune-informatics based vaccine design against MERS virus, Ebola virus
78 chikungunya, and Zika showed the promising potential to fought against disease (Ahmad et al. 2019;

Shahid et al. 2020; Tahir Ul Qamar et al. 2018; Tahir Ul Qamar et al. 2019). Reverse vaccinology includes different software algorithms to evaluate the immunological data that analyze the epitopes binding efficiency with HLA alleles, antigenicity, allergenicity, and toxicity to design the potential chimeric subunit vaccine (De Gregorio & Rappuoli 2012; Khan et al. 2018; Mirza et al. 2016; Patronov & Doytchinova 2013). The final chimeric epitope vaccine is a group of different epitopes joined with the help of linkers that may enhance specific adaptive-immune responses in the host cell (Brennick et al. 2017; Chauhan et al. 2019; Jensen & Andreatta 2018; Lu & Meng 2017; Nain et al. 2020; Purcell et al. 2007; Saadi et al. 2017).

In the recent study, SARS-CoV2 proteome was explored to determine the antigenic proteins, and various T-cell and B-cell epitopes were predicted with their HLA alleles with each epitope antigenicity, allergenicity, and physiochemical properties evaluation. The final epitope constructs molecular docking was performed with different TLRs (Toll Like Receptors), and HLA (human leukocyte antigen) alleles to confirm the stable binding interaction of the multi-epitope vaccine-receptor complex. To activate the host immune system, the interaction of TLRs with the designed vaccine could be a potential approach against viral infection. Intracellular TLRs that present on cell endosomes interact with ssRNA (TLR-7 & TLR-8), dsRNA (TLR-3), and CpG DNA (TLR-9) to activate the NFkβ mediated immune/cell response (Carty & Bowie 2010). With protein-based-epitope vaccine designing, we have focused on TLR-2, TLR-4, TLR2/6, and TLR1/2 heterodimer mediated IRF3/7 and NFKβ mediated immune cell activation. A previous study on TLR-2 and TLR-4 mediated activation showed protective cell immunity and subvert effect on the host cell. Targeting these two TLRs would mold the signaling cascade, beneficial for viral cell replication and survival. TLR4 knockout studies with respect to wild type showed that for host protective immune activation, cells need a certain degree of TLR-4 activation (Olejnik et al. 2018). To overcome the TLRs subvert effect of the host cell, it is better to target TLR1/2 heterodimer mediated activation signalling with mild TLR-4 interaction.

2. Materials and Methods

2.1. Protein sequence collection:

To evaluate the coronavirus suitable antigenic vaccine target, firstly, we have retrieved the different protein sequences from NCBI. Ten different protein FASTA files have been downloaded to cover the complete genome of SARS CoV2. These proteins are ORF6 protein (YP_009724394.1), membrane glycoprotein (YP_009724393.1), ORF3a protein (YP_009724391.1), nucleocapsid phosphoprotein (YP_009724397.2), ORF10 protein (QHI42199.1), ORF7a protein (YP_009724395.1), envelope protein

(YP_009724392.1), ORF1ab (YP_009724389.1), ORF8 protein (QHD43422.1), and surface glycoprotein (YP_009724390.1). All FASTA sequences were used as inputs for further immune-informatics analysis (Seema 2020).

2.2. Analysis of protein Antigenicity and Trans-membrane helicity

In the development of the chimeric multi-epitope vaccine, protein antigenicity, and transmembrane helicity play a key role in vaccine successes. To evaluate the protein antigenicity, we used the Vaxijen server (Doytchinova & Flower 2007) with 0.4 thresholds, while transmembrane helicity was estimated using TMHMM (Krogh et al. 2001) and Protter server (Omasits et al. 2014).

2.3. Cytotoxic T cell epitopes prediction with potential antigenicity, allergenicity, and Toxicity:

Shortlisted seven extracellular protein peptides were evaluated by NetCTLpan version 1.1 to predict promiscuous epitopes that can enhance the immune response in the host cell by interacting with HLA-epitope binding. In this server, we have evaluated the peptide with respect to 12 different HLA supertypes (HLA-A01:01, HLA-A02:01, HLA-A03:01, HLA-A24:02, HLA-A26:01, HLA-B07:02, HLA-B08:01, HLA-B27:05, HLA-B39:01, HLA-B40:01, HLA-B58:01, HLA-B15:01). With the help of the neural network, the algorithm server predicts promiscuous high binding affinity nonameric peptides (Stranzl et al. 2010). Vaxijen further evaluated predicted HLA binder peptides, AllergenFP (Dimitrov et al. 2013) and Toxinpred server used to confirm the epitope's high antigenicity, low allergenicity, and low toxicity level (Gupta et al. 2013).

2.4. Immunogenicity prediction:

Shortlisted NetCTLpan epitopes were used for immunogenicity prediction that confers the property which can elicit the cellular and humoral response in the host cell against viral infection. Promiscuous epitopes across prediction algorithms, ORF3a protein (1 epitope), surface glycoprotein (2 epitopes), and ORF1ab polyprotein (17 epitopes) used as input peptide. High binding capacity epitopes (by IEDB immunogenicity tool server) were selected as positive immunogenicity epitopes. This server predicted the immunogenicity level based on the epitope positions in the expected peptide and physicochemical properties of an amino acid (Calis et al. 2013).

2.5. Helper T cell epitope prediction:

With the help of IEDB MHC II binding prediction tool, we have predicted helper T cell epitopes across HLA alleles reference sets such as HLA-DRB1*01:01, HLA-DRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*12:01, HLA-DRB1*13:02, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB4*01:01, HLADRB5*01:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*05:01/DQB1*03:01, HLADQA1*03:01/ DQB1*03:02, HLADQA1*04:01/DQB1*04:02, HLADQA1*01:01/DQB1*05:01, HLADQA1*01:02/DQB1*06:02, HLADPA1*02:01/DPB1*01:01, HLA-DPA1*01:03/ DPB1*02:01, HLA-DPA1*01:03/DPB1*04:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DPA1*02:01/ DPB1*05:01,HLA-DPA1*02:01/DPB1*14:01. The HLA alleles reference set provides >99% population coverage. IEDB MHC II server is based on a combinational library that generates the percentile rank and IC-50. The lower percentile rank represents a higher affinity of the epitope-HLA complex. The epitope-alleles affinity consensus list was generated for 15 amino acids long epitopes (Wang et al. 2008).

2.6. B cell epitope prediction

The epitopes of B cells help to detect viral infections by the antibody-based immune response. IEDB BEPIPRED (Jespersen et al. 2017) and ABCpred (Saha & Raghava 2006) were used to analyze the B cell interacting epitopes. Overlapped epitopes from both servers were chosen for further analysis. FASTA files of all seven proteins were used as an input file. These resultant epitopes were further screened by Vaxijen, AllergenFP, and TOXINPRED server to analyze the epitopes antigenicity, allergenicity and toxicity reaction in the host cell.

2.7. Comparative, Cross-reactivity, IFN gamma induction analysis of MHC I, II, and B cell epitopes:

All finalized MHC I, II, and B cell epitope comparative analysis have been done to remove the overlapping epitopes. Non-overlapping epitopes should have a unique presence in the virus. To confirm this, we performed the multiple peptide match against the human proteome (ID 9606) using Protein Information Resource (Chen et al. 2013). Besides, IFN gamma induction has a significant role in the viral elimination and induction of host immune response. Hence, we have predicted the IFN gamma induction efficiency of selected epitopes via the IFNepitope server (http://crdd.osdd.net/raghava/ifnepitope/).

2.8. Chimeric multi-epitope vaccine designing

All screened epitopes that contain the antigenic property with less allergenicity, less toxicity, and less cross-reactivity have been finalized to design the chimeric multi-epitope vaccine against SARS Co-V2. A

potent vaccine should have the capacity to activate the immune response in the host cell but not activates the detrimental immunity. Hence, to maintain the balance immunity in the host cell, we have selected half non-IFN gamma inducible peptide and vice-versa. All selected epitopes were joined with the help of linker EAAK and GGGS (Solanki & Tiwari 2018).

2.9. Evaluation of potential vaccine candiadate or constrcts

Different constructs of the vaccine were further analyzed for antigenicity (via Vaxijen), allergenicity (via AllergenFP), and toxicity (via Toxinpred). A highly antigenic vaccine construct (i.e., VTC3) has been further evaluated by the Protparam tool (Gasteiger et al. 2003). The secondary and tertiary structure of selected chimeric vaccine construct VTC3 was predicted by using PESIPRED (Buchan & Jones 2019) and Phyre2 in intensive modeling mode (Kelley et al. 2015) respectively, and modelled vaccine was validated by PSVS analysis.

2.10. Molecular docking

To eliminate the virus infection, TLR's balanced activation is the primary need of the cell. In recent work, the scientists are targeting TLR1/2 etc. for viral elimination; hence, these TLR heterodimers were also included along with other TLRs for docking study with the chimeric vaccine. For this, we have docked VTC3 with different TLRs (TLR1, TLR2, TLR3, TLR4, TLR1/2, TLR6) by PATCHDOCK server. For efficiency analysis of the HLA-epitope complex, we also docked the VTC3 construct with different HLA alleles.

2.11. Molecular Dynamics Simulation

The MDS analysis was performed by Desmond using the published protocol (Wright et al. 2020) using the OPLS3e force field, TIP3P solvent. MDS was run for 50ns in duplicate. The simulated system was analyzed for the interaction diagram.

2.12. Validation of potential vaccine candidate or constrcuct with the human microbiome

Potential vaccine candidate sequence similarity with the gut microflora would create autoimmunity in the host. To minimize this adverse effect, we compared the vaccine candidate VTC3 against gut flora (226 organisms) sequenced by the Human Microbiome Project (Peterson et al. 2009) using Blast against 226 proteomes, with a significant hit (E-value $\leq 10^{-5}$ and identity $\geq 40\%$) (Ramos et al. 2018).

2.13. Mutational sensitivity profiling and analysis of experimental mutation data

Mutational sensitivity profiling of the designed vaccine was performed using the MAESTROweb as per the published method (Laimer et al. 2016). The amino acid sequence of chimeric multi-epitope vaccine VTC3 was further analysed for any mutation which was originally present in the SARS-CoV2 virus using the CoV-GLUE database.

3. Result

3.1. Analysis of protein antigenicity and trans-membrane helicity

The workflow has been discussed in figure 1, which explains the reverse vaccinology mechanism used in the present study. Amino acid sequences of target proteins were collected in FASTA format to analyse antigenicity and trans-membrane helicity. All the proteins with their antigenicity score such as nucleoprotein phosphoprotein (0.50), ORF10 protein (0.71), ORF8 protein (0.65), ORF7a protein (0.64), ORF6 protein (0.61), membrane glycoprotein (0.51), envelope protein (0.60), ORF3a protein (0.49), surface glycoprotein (0.46), and ORF1ab protein (0.46) antigenic score showed their potential to exceed the immune response in the host cell. To elicit the immune response in the host cell, transmembrane helicity of viral proteins has been evaluated. With the help of TMHMM and Protter servers, we have identified the extracellular, transmembrane, and cytosolic peptides. The protein protter results showed that nucleoprotein phosphoprotein, ORF6 protein, and ORF10 protein are complete cytosolic in nature. Hence, for further study, these three proteins were eliminated, and the rest of the seven proteins was considered for further study in designing the chimeric vaccine against SARS CoV2. The extracellular peptides of these proteins will help to interact with host PAMPs and maximize the solubility of designed vaccines.

3.2. Cytotoxic T cell epitopes prediction with potential antigenicity, allergenicity, and Toxicity:

Cytotoxic T lymphocytes (CTL) epitope predictions of all seven proteins were made using NetCTLpan 1.1 using the same HLA supertypes. Out of all seven, membrane glycoprotein (YP_009724393.1) did not get any potent HLA binder peptide. The rest six proteins showed the peptide binders with different HLA supertypes. From this server results, we have identified different epitopes ORF8 (8 epitopes), ORF7a protein (9 epitopes), an envelope protein (4 epitopes), ORF3a protein (3 epitopes), surface glycoprotein (125 epitopes) and ORF1ab polyprotein (707 epitopes) binders (data not shown). From the results, we manually screened peptides that showed a binding affinity with more than one HLA allele. This approach will minimize the HLA polymorphism so that promiscuous peptide will show binding to HLA of the wide population. After the manual screening of proteins, we have shortlisted ORF8 (1 epitope), ORF7a protein

(2 epitopes), an envelope protein (2 epitopes), ORF3a protein (3 epitopes), surface glycoprotein (43 epitopes) and ORF1ab polyprotein (277 epitopes) epitope binders (data not shown). Based on antigenicity, allergenicity and toxicity analysis of peptide, different epitopes of ORF8 (0 epitopes), ORF7a protein (0 epitopes), an envelope protein (0 epitopes), ORF3a protein (1 epitope), surface glycoprotein (2 epitopes) and ORF1ab polyprotein (17 epitopes) further shortlisted and the potential epitopes were selected that can enhance immune response via HLA activation (Table1).

3.3. Immunogenicity prediction

Immunogenicity prediction of ORF3a protein (1 epitope), surface glycoprotein (2 epitopes), and ORF1ab polyprotein (17 epitopes) epitopes showed that ORF3a protein (1 epitope), surface glycoprotein (1 epitope) and ORF1ab polyprotein (6 epitopes) epitopes had a positive score (Table1). The positive immunogenicity with HLA binding confirmed that these epitopes would elicit a high immune response.

3.4. Helper T cell epitopes prediction with antigenicity, allergenicity, and toxicity:

Helper T-cell mediated 15 amino acid extended epitopes were generated against the HLA allele reference set. In results the percentile rank 0.1 was set as the threshold so that we can screen high binding affinity HLA II interacting epitopes. Based on reference threshold, ORF8 (5 epitopes), ORF7a protein (0 epitopes), membrane protein (0 epitopes), an envelope protein (0 epitopes), ORF3a protein (0 epitopes), surface glycoprotein (19 epitopes) and ORF1ab polyprotein (18 epitopes) showed <0.1 percentile rank (Table2). Low percentile epitopes antigenicity (Table 2), allergenicity, and toxicity level (Table3) further shortlisted the epitopes. The compiled results of all the analysis identified surface glycoprotein (12 epitopes), ORF1ab polyprotein (4 epitopes), and ORF8 (3 epitopes) as HLA-II binders with high antigenicity booster and low allergenicity & low toxicity.

3.5. B cell epitope comparative prediction with antigenicity, allergenicity, and Toxicity analysis

Using the IEDB Bepipred server, we got the different epitopes from selected proteins like ORF8 protein (2 epitopes), ORF7a protein (2 epitopes), membrane glycoprotein (1 epitope), an envelope protein (1 epitope), ORF3a protein(1 epitope), surface glycoprotein (28 epitopes), and ORF1ab protein (98 epitopes) proteins. With ABCpred server based B cell epitope analysis, we have found the different epitope of ORF8 protein (9 epitopes), ORF7a protein (7 epitopes), membrane glycoprotein (1 epitope), and envelope protein (3 epitopes), ORF3a protein (2 epitopes), surface glycoprotein(50 epitopes), and ORF1ab protein (20 epitopes) (data not shown). To filter out the common epitope from both servers, we have manually

screened the result, and shortlist the common overlapped epitopes. This selects different epitopes from Orf8 protein (2 epitopes), ORF7a protein (2 epitopes), membrane glycoprotein (1 epitope), envelope protein (1 epitope), ORF3a protein (1 epitope), surface glycoprotein (16 epitopes), and ORF1ab protein (8 epitopes) (Table 4). The screened epitopes were further analyzed for antigenicity, allergenicity, and toxicity analysis (Table 5). This resulted in shortlisting of ORF8 protein (1 epitope), envelope protein (1 epitope), ORF3a protein (1 epitope), surface glycoprotein(9 epitopes), and ORF1ab protein (6 epitopes) epitopes for further study.

3.6. Comparative Cross-reactivity, IFN gamma induction analysis of MHC I, II, and B cell epitopes:

All finalized MHC I, II, and B cell epitopes were used for comparative analysis (Table 6) to remove the overlapping sequences. In addition to that, we have predicted the cross-reactivity of selected epitopes with humans as the similarity between virus protein epitopes and host cells peptide, which eliminates autoimmune reactivity. BLAST result of all selected epitopes against the human proteome and the result showed that no epitopes were found to have cross-reactivity reactions in the host cell. In addition to this, it was seen that in virus-human cell immune response activity, IFN gamma plays an important role. High expression of IFN gamma leads to viral infection clearance in the human cells. IFNepitope server positive score shows the capacity of the epitope to induce the IFN secretion via T cells, which was listed in Table 6. All the selected epitopes were used to design the different vaccine constructs.

3.7. Vaccine properties analysis selected VTC3

Designed vaccine constructs VTC1, VTC2, and VTC3 (Table 7) showed antigenic, with no allergenic and toxic nature. Highly antigenic vaccines construct VTC3 physiochemical analysis showed the 227 long amino acid peptide constructs with 24 KDa weight have the instability index 23.36 that represent the protein stable nature. VTC3 construct contains the aliphatic and GRAVY index value of 68.72 and -0.455, respectively. PESIPRED secondary structure analysis of vaccine construct VTC3 showed the 62.56 % alpha helicity, 7.05% extended strands, 5.73% beta-turn with 24.67 random coils. VTC3 tertiary structure has been modeled by the Phyre2 intensive modeling tool, which validation by Ramachandran plot showed 91.1% residue in the favored region (Figure 2).

3.8. Molecular docking analysis confirm the interaction of VTC3 with TLR and HLA.

To be effective against COVID 19, a vaccine should have the capacity to activate the immune response of the human host. The virus can subvert the host protein function, which plays a role in host cell invasion

or virus persistence. In previous study, it has been shown that TLRs have an extensive role in pathogen persistence and clearance. Hence, docking analysis was performed which showed that our construct VTC3 interacts with TLR1/2 complex, TLR1, TLR3, TLR4, TLR6 but has the highest binding affinity with TLR1/2 complex, followed by TLR4 and TLR6 (Table 8). The docked poses are shown in figure 3. As mentioned in the introduction, the interaction of VTC3 with TLR1/2 is important for the immune response for viral infection; hence docking results further confirm it. In additional to that, we have also performed molecular docking of VTC3 with different HLAs, which was important to induce MHC-I and MHC-II to activate the immune response in the host cell.

3.9. Molecular dynamics simulation (MDS) analysis confirms the strong interaction of VTC3 with TLR1/2 heterodimer and HLA.

VTC3 has the best interaction with the TLR1/2 complex, followed by TLR4; hence both complexes were used for MDS analysis. MDS was performed till 10ns for both the complex and results were analyzed for RMSD, RMSF, etc. RMSD calculation showed that the VTC3-TLR1/2 complex is stable throughout simulation with RMSD of around 3Å (Figure 4A), and the VTC3-TLR4 was found to be unstable with RMSD 7Å (Figure 4C). Similarly, the RMSF analysis showed that the VTC3-TLR1/2 complex has less fluctuation with RMSF of around 3Å (Figure 5B) while the VTC3-TLR4 complex showed that the vaccine has RMSD of 7Å (Figure 5D). MDS result suggests chimeric vaccine construct VTC3 showed a stronger and stable interaction with TLR1/2 complex as compared to TLR4 (Figure 5), this showed that this vaccine construct induces immune response suitable for clearance of SARS-coV2. For a vaccine construct, it is also important to interact with HLA. MDS analysis of the VTC3-HLA complex was investigated till 50ns, and the result was analysed for RMSD and RMSF. RMSD calculation showed that the VTC3-HLA complex is stable throughout simulation with RMSD of around 5Å at 50ns (Figure 5), and RMSF analysis showed that the VTC3-HLA complex has some fluctuation at the terminal. Both the data suggest the interaction between VTC3 and HLA.

3.10. Chimeric subunit vaccine VTC3 does not significant similarity to the human gut microbiome.

NCBI blast of VTC3 construct against 226 gut flora showed that there is no significant similarity (Supplementary Table 1) between amino acid sequence of the chimeric vaccine construct VTC3 and amino acid sequence of the human gut microbiome that further minimize the cross reactivity in the host cell.

3.11. Mutational profiling of chimeric vaccine VTC3 showed less mutation sensitivity

Mutational sensitivity profiling of VTC3 vaccine construct (figure 6) showed that only a few chimeric epitope residues have $\Delta\Delta G_{pred} > 0$, suggesting that our vaccine construct VTC3 has less mutation sensitivity, which enhances the possibility of our vaccine constructs to be effective. Similarly, analysis of our vaccine construct VTC3 in the CoV-GLUE database showed less mutation hot spots in our vaccine construct (Supplementary Table 2). This database contains all replacements, insertions, and deletions, which have been observed in the GISAID hCoV-19 sequence sampled from the pandemic. We have also identified the interacting residues of vaccine construct VTC3 with the TLR1/2 complex, TLR4, and HLA. The result showed that the residues of VTC3 with mutation does not involve in the physical interaction with these proteins. These results further support the efficacy of our vaccine construct.

4. Discussion

In the present study, we are proposing a chimeric vaccine constructs VTC3 for SARS-CoV2 with the help of reverse vaccinology, compiles to the outer surface-exposed epitopes of the structural and non-structural proteins of this virus. Different therapeutic strategies have been tried to control SARS-CoV2 (Tiwari 2020b) (Tiwari 2020a). SARS-CoV2 structural proteins like surface glycoprotein, membrane, and envelope protein have a significant contribution in surface adherence and internalization, while non-structural protein is virulence-associated factors that cause immune-pathogenesis. SARS-CoV2 host cell internalization is facilitated by surface glycoprotein (spike protein) interaction with the ACE2 receptor. Targeting the outer exposed peptides of both structural and non-structural protein as a vaccine target would be a promising approach to induce both humoral and cellular immune response that recognizes the virus and killed it. These proteins were shortlisted by antigenicity score and trans-membrane helicity. The shortlisted proteins were used to identify surface-exposed peptides. Surface exposed peptides of these proteins were further analyzed to determine MHC I, II, and B cell-mediated epitopes. The selected epitopes were shortlisted by their antigenicity, allergenicity, toxicity, cross-reactivity, immunogenicity, and IFN gamma secretion scores. These epitopes were used to design different vaccine constructs (VTC1, VTC2, and VTC3) with the help of linkers. Antigenicity, allergenicity, and physiochemical analysis of vaccine constructs were further analyzed to enhance the peptide potency. The finalized construct VTC3 should have the potential to interact with pathogen-associated molecular patterns activates the host innate immune system and consonance of the adaptive immunity via TLRs (TLR1, 2, 3, 4, 6, 7, 8 and 9). It has been observed that the TLRs mediated virus interaction not-only combat virus virulence and infection but also sometimes initiate the host system to overturn downstream signals for the benefit (replication and survival) of the virus-cell. During the viral infection, membrane-associated proteins interaction with TLR2 and TLR4

plays a nuanced role in exceeding inflammatory responses via adhesion and invasion that vitiating the host cell(Olejnik et al. 2018). Simultaneously with TLRs, HLA alleles have a significant contribution in activation of the robust immune response. The strong HLA-epitope interaction would maintain the signaling cascades to activate the immune response for the viral infection clearance. To prove the modeled VTC3 construct, we have docked it with different HLAs and TLRs. Presently, the focus lies on the development of vaccines against viruses that can activate the innate immune system via TLR1/2 (Carty & Bowie 2010; Dowling & Mansell 2016; Jensen et al. 2018) to overcome the subvert effect of the virus on TLR medicated signalling. It is reported that the $CD8^+$ T cell-mediated HLA allele (HLA-B*5801) unique interaction with immune dominant peptide contributes as a potential to control and prevent viral infection (Li et al. 2016). HLA-B*5801 associated patients known as "elite controllers" who become infected but can control viremia. The interaction of VTC3 was investigated with molecular docking, and molecular dynamics simulation and result is an agreement that VTC3 construct has the highest affinity for allele HLA-B*5801. It might also create a ray of hope for the potential creation of vaccines or convalescent serum antibodies against COVID19. The predicted Chimeric vaccine VTC3 should be tested experimentally for therapeutic potency in future studies. Experimental validation is necessary to demonstrate the potency of the designed vaccine in future studies.

5. Conclusion

High mortality and morbidity rate of COVID 19 is unprecedented due to the unavailability of vaccination hence effective treatment strategies (inhibitor, drugs, or vaccine), rapid development, trials, and production needed immediately for this global pandemic disease. To reach out the success, it is necessary to evaluate all possible vaccine candidates to find out the one viable outcome. To increase the chance of success, WHO initiated vaccine solidarity trials to test all vaccine candidates until they fail. In the present study, an attempt was made to design a chimeric multi-epitope vaccine against SARS CoV2 that targets its exposed peptides of structural and non-structural proteins. Immunogenicity, allergenicity, toxicity, and potent IFN-gamma inducer scores have also been analyzed to further narrow down efficient epitopes. Peptide matching with the human proteome showed no indication of possible cross-reactivity. However, current reverse immune-informatics approaches were executed to target surface-exposed proteins for enhancing effective host innate with humoral and cellular immune responses. The present study concludes the design of VTC3 as a chimeric vaccine against SARS CoV2.

Acknowledgments: VT would like to thank the Central University of Rajasthan for providing the Schrodinger suite. VS would like to thank UGC for providing the SRF fellowship.

Authors contributions: Conceived and designed: VT, Performed the experiments: VS & VT. Analyzed the data: VS & VT , Chemical and Resources provided: VT, Wrote the manuscript: VS & VT, Proofread of final version, VT

Ethical approval: The present study does not involve human and animal samples.

Competing financial interests: The authors have declared that no competing interests exist. The experiment was performed in the absence of any financial support. **Data availability:** All the data are available in the manuscript

References

Ahmad B, Ashfaq UA, Rahman MU, Masoud MS, and Yousaf MZ. 2019. Conserved B and T cell epitopes prediction of ebola virus glycoprotein for vaccine development: An immuno-informatics approach. *Microb Pathog* 132:243-253. 10.1016/j.micpath.2019.05.010

Alamri M, Tahir ul Qamar M, and Alqahtani S. 2020. Pharmacoinformatics and Molecular Dynamic Simulation Studies Reveal Potential Inhibitors of SARS-CoV-2 Main Protease 3CL^{pro}. Preprints.org.

Benvenuto D, Giovanetti M, Ciccozzi A, Spoto S, and Angeletti S. 2020. The 2019-new coronavirus epidemic: Evidence for virus evolution. *J Med Virol* 92:455-459. 10.1002/jmv.25688

Brennick CA, George MM, Corwin WL, Srivastava PK, and Ebrahimi-Nik H. 2017. Neoepitopes as cancer immunotherapy targets: key challenges and opportunities. *Immunotherapy* 9:361-371. 10.2217/imt-2016-0146

Buchan DWA, and Jones DT. 2019. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* 47:W402-W407. 10.1093/nar/gkz297

Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, Kesmir C, and Peters B. 2013. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 9:e1003266. 10.1371/journal.pcbi.1003266

Carty M, and Bowie AG. 2010. Recent insights into the role of Toll-like receptors in viral infection. *Clin Exp Immunol* 161:397-406. 10.1111/j.1365-2249.2010.04196.x

Chauhan V, Rungta T, Goyal K, and Singh MP. 2019. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Sci Rep* 9:2517. 10.1038/s41598-019-39299-8

Chen C, Li Z, Huang H, Suzek BE, Wu CH, and Consortium U. 2013. A fast Peptide Match service for UniProt Knowledgebase. *Bioinformatics* 29:2808-2809. 10.1093/bioinformatics/btt484

Chen Y, Liu Q, and Guo D. 2020. Emerging coronaviruses: Genome structure, replication, and pathogenesis. 92:418-423. 10.1002/jmv.25681

De Gregorio E, and Rappuoli R. 2012. Vaccines for the future: learning from human immunology. *Microb Biotechnol* 5:149-155. 10.1111/j.1751-7915.2011.00276.x

de Wilde AH, Snijder EJ, Kikkert M, and van Hemert MJ. 2018. Host Factors in Coronavirus Replication. *Curr Top Microbiol Immunol* 419:1-42. 10.1007/82_2017_25

Dimitrov I, Naneva L, Doytchinova I, and Bangov I. 2013. AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatics* 30:846-851. 10.1093/bioinformatics/btt619

Dong N, Yang X, Ye L, Chen K, Chan E, and Chen S. 2020. Genomic and protein structure modelling analysis depicts the origin and pathogenicity of 2019-nCoV, a new coronavirus which caused a pneumonia outbreak in Wuhan, China [version 2; peer review: 1 not approved]. *F1000Research* 9. 10.12688/f1000research.22357.2

Dowling JK, and Mansell A. 2016. Toll-like receptors: the swiss army knife of immunity and vaccine development. *Clinical & translational immunology* 5:e85-e85. 10.1038/cti.2016.22

Doytchinova IA, and Flower DR. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8:4. 10.1186/1471-2105-8-4

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, and Bairoch A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31:3784-3788. 10.1093/nar/gkg563

Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW, Penzar D, Perlman S, Poon LLM, Samborskiy D, Sidorov IA, Sola I, and Ziebuhr J. 2020. Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group. *bioRxiv*:2020.2002.2007.937862. 10.1101/2020.02.07.937862

Gralinski LE, and Menachery VD. 2020. Return of the Coronavirus: 2019-nCoV. 12. 10.3390/v12020135

Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, and Raghava GP. 2013. In silico approach for predicting toxicity of peptides and proteins. *PLoS One* 8:e73957. 10.1371/journal.pone.0073957

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, and Cao B. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497-506. 10.1016/s0140-6736(20)30183-5

Jensen KK, and Andreatta M. 2018. Improved methods for predicting peptide binding affinity to MHC class II molecules. 154:394-406. 10.1111/imm.12889

Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, Sette A, Peters B, and Nielsen M. 2018. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154:394-406. 10.1111/imm.12889

Jespersen MC, Peters B, Nielsen M, and Marcatili P. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 45:W24-w29. 10.1093/nar/gkx346

Jiayuan C, Jinsong S, Tungon Y, Chang L, Xin L, Qiang Z, Jishou R, and Gao S. 2020. Bioinformatics analysis of the 2019 novel coronavirus genome. *Chinese Journal of Bioinformatics* 18:96-102. 10.12113/202001007

Kelley LA, Mezulis S, Yates CM, Wass MN, and Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10:845-858. 10.1038/nprot.2015.053

Khan A, Junaid M, Kaushik AC, Ali A, Ali SS, Mehmood A, and Wei DQ. 2018. Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PLoS One* 13:e0196484. 10.1371/journal.pone.0196484

Krogh A, Larsson B, von Heijne G, and Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567-580. 10.1006/jmbi.2000.4315

Laimer J, Hiebl-Flach J, Lengauer D, and Lackner P. 2016. MAESTROweb: a web server for structure-based protein stability prediction. *Bioinformatics* 32:1414-1416. 10.1093/bioinformatics/btv769

Li G, and De Clercq E. 2020. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). *Nat Rev Drug Discov* 19:149-150. 10.1038/d41573-020-00016-0

Li X, Lamothe PA, Ng R, Xu S, Teng M, Walker BD, and Wang J-H. 2016. Crystal structure of HLA-B*5801, a protective HLA allele for HIV-1 infection. *Protein & cell* 7:761-765. 10.1007/s13238-016-0309-y

Lu C, and Meng S. 2017. A novel multi-epitope vaccine from MMSA-1 and DKK1 for multiple myeloma immunotherapy. 178:413-426. 10.1111/bjh.14686

Mirza MU, Rafique S, Ali A, Munir M, Ikram N, Manan A, Salo-Ahen OM, and Idrees M. 2016. Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. *Sci Rep* 6:37313. 10.1038/srep37313

Nain Z, Karim MM, Sen MK, and Adhikari UK. 2020. Structural basis and designing of peptide vaccine using PE-PGRS family protein of Mycobacterium ulcerans-An integrated vaccinomics approach. *Mol Immunol* 120:146-163. 10.1016/j.molimm.2020.02.009

Olejnik J, Hume AJ, and Muhlberger E. 2018. Toll-like receptor 4 in acute viral infection: Too much of a good thing. 14:e1007390. 10.1371/journal.ppat.1007390

Omasits U, Ahrens CH, Muller S, and Wollscheid B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30:884-886. 10.1093/bioinformatics/btt607

Patronov A, and Doytchinova I. 2013. T-cell epitope vaccine design by immunoinformatics. *Open Biol* 3:120139. 10.1098/rsob.120139

Perlman S. 2020. Another Decade, Another Coronavirus. *N Engl J Med* 382:760-762. 10.1056/NEJMe2001126

Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, and Deal C. 2009. The NIH human microbiome project. *Genome research* 19:2317-2323.

Purcell AW, McCluskey J, and Rossjohn J. 2007. More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov* 6:404-414. 10.1038/nrd2224

Qamar MTu, Shahid F, Ashfaq UA, Aslam S, Fatima I, Fareed MM, Zohaib A, and Chen L-L. 2020. Structural modeling and conserved epitopes prediction against SARS-COV-2 structural proteins for vaccine development. Research Square.

Ramos PIP, Fernández Do Porto D, Lanzarotti E, Sosa EJ, Burguener G, Pardo AM, Klein CC, Sagot M-F, de Vasconcelos ATR, Gales AC, Marti M, Turjanski AG, and Nicolás MF. 2018. An integrative, multi-omics approach towards the prioritization of Klebsiella pneumoniae drug targets. *Sci Rep* 8:10755. 10.1038/s41598-018-28916-7

Saadi M, Karkhah A, and Nouri HR. 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. 10.1016/j.meegid.2017.04.009

Saha S, and Raghava GP. 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65:40-48. 10.1002/prot.21078

Seema M. 2020. *T Cell Epitope-Based Vaccine Design for Pandemic Novel Coronavirus 2019-nCoV*.

Shahid F, Ashfaq UA, Javaid A, and Khalid H. 2020. Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. *Infect Genet Evol* 80:104199. 10.1016/j.meegid.2020.104199

Solanki V, and Tiwari V. 2018. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against Acinetobacter baumannii. *Sci Rep* 8:9044. 10.1038/s41598-018-26689-7

Stranzl T, Larsen MV, Lundegaard C, and Nielsen M. 2010. NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* 62:357-368. 10.1007/s00251-010-0441-4

Tahir Ul Qamar M, Bari A, Adeel MM, Maryam A, Ashfaq UA, Du X, Muneer I, Ahmad HI, and Wang J. 2018. Peptide vaccine against chikungunya virus: immuno-informatics combined with molecular docking approach. *J Transl Med* 16:298. 10.1186/s12967-018-1672-7

Tahir Ul Qamar M, Saleem S, Ashfaq UA, Bari A, Anwar F, and Alqahtani S. 2019. Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study. *J Transl Med* 17:362. 10.1186/s12967-019-2116-8

Tian X, Li C, Huang A, Xia S, Lu S, and Shi Z. 2020. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody. 9:382-385. 10.1080/22221751.2020.1729069

Tiwari V. 2020a. Denovo designing, retrosynthetic analysis, and combinatorial synthesis of a hybrid antiviral (VTAR-01) to inhibit the interaction of SARS-CoV2 spike glycoprotein with human angiotensin-converting enzyme 2. *Biol Open*. 10.1242/bio.054056

Tiwari V. 2020b. Novel hybrid antiviral VTRRT-13V2.1 against SARS-CoV2 main protease: retro-combinatorial synthesis and molecular dynamics analysis. *Heliyon* 6:e05122-e05122. 10.1016/j.heliyon.2020.e05122

Wang P, Sidney J, Dow C, Mothe B, Sette A, and Peters B. 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol* 4:e1000048. 10.1371/journal.pcbi.1000048

World Health O. 2020. Surveillance case definitions for human infection with novel coronavirus (nCoV): interim guidance v1, January 2020. Geneva: World Health Organization.

Wrapp D, and Wang N. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. 367:1260-1263. 10.1126/science.abb2507

Wright WC, Chenge J, Wang J, Girvan HM, Yang L, Chai SC, Huber AD, Wu J, Oladimeji PO, Munro AW, and Chen T. 2020. Clobetasol Propionate Is a Heme-Mediated Selective Inhibitor of Human Cytochrome P450 3A5. *Journal of Medicinal Chemistry* 63:1415-1433. 10.1021/acs.jmedchem.9b02067

WU C, Zheng S, Chen Y, and Zheng M. 2020a. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCoV, in the nasal tissue. *medRxiv*:2020.2002.2011.20022228. 10.1101/2020.02.11.20022228

Wu JT, Leung K, and Leung GM. 2020b. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 395:689-697. 10.1016/s0140-6736(20)30260-9

Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D, Chen J, Luo Y, Guo H, Jiang R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Zheng X-S, Zhao K, Chen Q-J, Deng F, Liu L-L, Yan B, Zhan F-X, Wang Y-Y, Xiao G, and Shi Z-L. 2020. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. *bioRxiv*:2020.2001.2022.914952. 10.1101/2020.01.22.914952

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 GF, and Tan W. 2020. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 382:727-733. 10.1056/NEJMoa2001017

Figure Legends

Figure 1. Brief workflow of combinational chimeric multi-epitope vaccine designing with predicted immune cell response.

Figure 2. Tertiary structure of modelled VTC3 construct (A) and Ramachandran plot of the modelled proteins (B).

Figure 3. Docked pose of VTC3-TLR1/2 complex (A), and VTC3-TLR4 complex (B).

Figure 4. Root-mean-square deviation and Root mean square fluctuations during molecular dynamics simulation analysis of VTC3-TLR1/2 complex (A and B), and VTC3-TLR4 complex (C and D).

Figure 5. Root-mean-square deviation during molecular dynamics simulation analysis of VTC3-HLA complex. MDS was performed till 50ns.

Figure 6. Diagrammatic presentation of proposed combinational chimeric multi-epitope vaccine VTC3 showing the position of different epitopes and linker in the vaccine.

Table 1(on next page)

Table 1 to Table 8

Table 1. MHC I binder epitopes positive antigenicity, allergenicity, toxicity and immunogenicity analysis.

S.No	Protein name	Start	epitope	Antigenicity	Allergenicity	Toxicity	Immunogenicity
1	ORF3a	4	MRIFTIGTV	0.69	Non-allergenic	Non-toxic	0.37
2	surface glycoprotein	242	WTAGAAAYY	0.63	Non-allergenic	Non-toxic	0.15
		702	FTISVTTEI	0.85	Non-allergenic	Non-toxic	-0.18
3	orf1ab polyprotein	3449	LSFKELLVY	0.72	Non-allergenic	Non-toxic	-0.07
		1890	EIDPKLDNY	1.61	Non-allergenic	Non-toxic	-0.2
		1502	ETISLAGSY	0.59	Non-allergenic	Non-toxic	-0.16
		3841	LSDDAVVCF	0.58	Non-allergenic	Non-toxic	0.1
		5467	YTEISFMLW	1.21	Non-allergenic	Non-toxic	-0.03
		2413	VVTTFDSEY	0.45	Non-allergenic	Non-toxic	0.1
		295	FMGRIRSVY	0.52	Non-allergenic	Non-toxic	0.125
		3471	LLDKRTTCF	1.76	Non-allergenic	Non-toxic	-0.12
		4672	SMMGFKMNY	1.3	Non-allergenic	Non-toxic	-0.26
		4615	LQAENVTGL	0.82	Non-allergenic	Non-toxic	0.19
		2166	NYMPYFFTL	1	Non-allergenic	Non-toxic	0.15
		5726	IQLSSYSLF	0.75	Non-allergenic	Non-toxic	-0.48
		524	EQKSILSPL	0.55	Non-allergenic	Non-toxic	-0.26
		5406	FELEDFIPM	1.26	Non-allergenic	Non-toxic	0.33
		724	EETGLLMPL	0.48	Non-allergenic	Non-toxic	-0.12
		4055	KLVLSVNPY	0.54	Non-allergenic	Non-toxic	-0.13

Table 2. MHC II peptide percentile rank with antigenicity score.

S.N	Protein Name	Allele	Start	Peptide	Percentile_rank	Antigenicity
1	ORF8 protein	HLA-DRB3*01:01	14	QPYVVDDPCPIHFYS	0.07	0.4574
		HLA-DRB3*01:01	10	CTQHQPYVVDDPCPI	0.08	0.5165
		HLA-DRB3*01:01	13	HQPYVVDDPCPIHFY	0.08	0.55
		HLA-DRB3*01:01	12	QHQPYVVDDPCPIHF	0.08	0.8637
		HLA-DRB3*01:01	11	TQHQPYVVDDPCPIH	0.08	0.6706
2	surface glycoprotein	HLA-DRB1*13:02	98	KTQSLLIVNNATNVV	0.01	0.63
		HLA-DRB1*13:02	102	LLIVNNATNVVIKVC	0.01	0.09
		HLA-DRB1*13:02	100	QSLLIVNNATNVVIK	0.01	0.43
		HLA-DRB1*13:02	101	SLLIVNNATNVVIKV	0.01	0.47
		HLA-DRB1*13:02	99	TQSLLIVNNATNVVI	0.01	0.43
		HLA-DRB3*02:02	100	QSLLIVNNATNVVIK	0.02	0.43
		HLA-DPA1*01:03/DPB1*04:01	323	FGEVFNATRFASVYA	0.03	0.04
		HLA-DRB1*01:01	498	LSFELLHAPATVCGP	0.03	0.5
		HLA-DRB1*01:01	497	VLSFELLHAPATVCG	0.03	0.47
		HLA-DRB1*01:01	496	VVLSFELLHAPATVC	0.03	0.86
		HLA-DRB1*13:02	103	LIVNNATNVVIKVCE	0.03	-0.11
		HLA-DRB1*13:02	97	SKTQSLLIVNNATNV	0.03	0.62
		HLA-DRB3*02:02	101	SLLIVNNATNVVIKV	0.03	0.47
		HLA-DRB3*02:02	99	TQSLLIVNNATNVVI	0.06	0.433
		HLA-DPA1*01:03/DPB1*04:01	324	GEVFNATRFASVYAW	0.07	-0.12
		HLA-DPA1*01:03/DPB1*04:01	322	PFGEVFNATRFASVY	0.07	0.03
		HLA-DRB1*01:01	499	SFELLHAPATVCGPK	0.09	0.2
		HLA-DRB1*01:01	495	VVVLSFELLHAPATV	0.09	0.8

		HLA-DRB3*02:02	102	LLIVNNATNVVIKVC	0.09	0.09
		HLA-DRB1*13:02	98	KTQSLLIVNNATNVV	0.01	0.63
		HLA-DRB1*13:02	102	LLIVNNATNVVIKVC	0.01	0.09
3	orf1ab polyprotein	HLA-DRB1*09:01	54	AIILASFSASTSAFV	0.01	0.23
		HLA-DQA1*01:02/DQB1*06:02	45	AFASEAARVVRSIFS	0.08	-0.02
		HLA-DRB1*07:01	10	CTFTRSTNSRIKASM	0.12	-0.02
		HLA-DPA1*03:01/DPB1*04:02	1	YFFTLLLQLCTFTRS	0.16	-0.37
		HLA-DRB5*01:01	33	LGRYMSALNHTKKWK	0.17	0.04
		HLA-DRB1*01:01	20	KSAFYILPSIISNEK	0.27	0.71
		HLA-DPA1*03:01/DPB1*04:02	2165	CTNYMPYFFTLLLQL	0.01	0.45
		HLA-DRB1*01:01	1801	ESPFVMMSAPPAQYE	0.01	0.54
		HLA-DRB1*09:01	474	AIILASFSASTSAFV	0.01	0.23
		HLA-DPA1*03:01/DPB1*04:02	1244	EETKFLTENLLLYID	0.03	-0.11
		HLA-DRB3*01:01	903	ATYYLFDESGEFKLA	0.04	0.23
		HLA-DQA1*01:02/DQB1*06:02	535	AFASEAARVVRSIFS	0.08	-0.02
		HLA-DRB1*15:01	747	AMPNMLRIMASLVLA	0.01	0.09
		HLA-DRB3*02:02	2720	AFVTNVNASSSEAFL	0.01	0.15
		HLA-DRB1*11:01	865	NEFYAYLRKHFSMMI	0.02	0.22
		HLA-DQA1*05:01/DQB1*02:01	2390	QMEIDFLELAMDEFI	0.03	0.61
		HLA-DRB3*02:02	2755	NYIFWRNTNPIQLSS	0.04	0.92
		HLA-DQA1*05:01/DQB1*02:01	2393	IDFLELAMDEFIERY	0.05	0.25
		HLA-DPA1*02:01/DPB1*14:01	663	QMNLKYAISAKNRAR	0.07	1.5

5

6 **Table 3. Analysis of MHC II epitopes allergenicity and toxicity.**

S.No	Protein Name	start	Peptide	Allergenicity	Toxicity
1	ORF8 protein	14	QPYVVDDPCPIHFYS	Allergen	-
		10	CTQHQPYVVDDPCPI	Non-allergenic	Non-toxic
		13	HQPYVVDDPCPIHFY	Non- allergenic	Non-toxic
		12	QHQPYVVDDPCPIHF	Allergen	-
		11	TQHQPYVVDDPCPIH	Non-allergenic	Non-toxic
2	Surface Glycoprotein	98	KTQSLLIVNNATNVV	Non-allergenic	Non-toxic
		100	QSLLIVNNATNVVIK	Non- allergenic	Non-toxic
		101	SLLIVNNATNVVIKV	Non-allergenic	Non-toxic
		99	TQSLLIVNNATNVVI	Non- allergenic	Non-toxic
		100	QSLLIVNNATNVVIK	Non-allergenic	Non-toxic
		498	LSFELLHAPATVCGP	Non- allergenic	Non-toxic
		497	VLSFELLHAPATVCG	Non-allergenic	Non-toxic
		496	VVLSFELLHAPATVC	Non- allergenic	Non-toxic
		97	SKTQSLLIVNNATNV	allergenic	-
		101	SLLIVNNATNVVIKV	Non-allergenic	Non-toxic
		99	TQSLLIVNNATNVVI	Non- allergenic	Non-toxic
		495	VVVLSFELLHAPATV	Non-allergenic	Non-toxic
3	ORF1ab polyprotein	20	KSAFYILPSIISNEK	Non-allergenic	Non-toxic
		2165	CTNYMPYFFTLLLQL	allergenic	-
		1801	ESPFVMMSAPPAQYE	Non-allergenic	Non-toxic

	2390	QMEIDFLELAMDEFI	Non- allergenic	Non-toxic
	2755	NYIFWRNTNPIQLSS	Non-allergenic	Non-toxic
	663	QMNLKYAISAKNRAR	Non-allergenic	Non-toxic

Table 4. Comparative analysis of B cell epitope using IEDB and ABCpred server and antigenicity analysis.

S. N.	Protein name	Start	Epitope IEDB BepiPred	Antigenicity	ABCpred	Start	Antigenicity
1	ORF8 protein	12	QHQPYVVDDP	0.4127	PYVVDDPCPIHFYSKW	15	0.56
		49	EAGSKSPI	0.2081	ELCVDEAGSKSPIQYI	44	-0.18
2	ORF7a protein	18	EPCSSGTYEGNSPFHPLAD	0.39	SGTYEGNSPFHPLADN	22	0.29
		58	HVYQLRARSVSPKLFIRQE	0.59	HVYQLRARSVSPKLFI	58	0.43
3	membrane glycoprotein	7	TITVEELKK	0.57	DSNGTITVEELKKLLE	3	0.07
4	envelope protein	23	YSRVKNLNSSRVP	0.47	YVYSRVKNLNSSRVPD	21	0.54
5	ORF3a protein	15	LKQGEIKDATPSDFVR	0.81	QGEIKDATPSDFVRAT	17	0.91
					IGTVTLKQGEIKDATP	10	1.22
6	surface glycoprotein	55	VSGTNGTKRF	0.53	HRSYLTPGDSSSGWTA	230	0.6
		123	DPFLGVYYHKNNKSWMESEFRVYSSA	0.49	TVEKGIYQTSNFRVQP	292	0.67
		234	LTPGDSSSGWTA	0.68	GCLIGAEHVNNSYECD	633	0.84
		298	YQTSNFRVQP	1.18	LQSYGFQPTNGVGYQP	477	0.52
		315	PNITNLCPFGEVFNATRFASVYAWNRKRISNC	0.47	TEIYQAGSTPCNGVEG	455	-0.01
		389	GDEVRQIAPGQTGKIAD	1.06	KQIYKTPPIKDFGGFN	771	-0.22
		441	FRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPT	0.39	CGPKKSTNLVKNKCVN	510	0.2
		501	ELLHAPATVCGPKKSTNLVKN	0.0029	FERDISTEIYQAGSTP	449	-0.29
		619	RVYSTGSNVFQ	-0.1	SWMESEFRVYSSANNC	136	0.17
		641	VNNSYECDIPI	0.6124	EVRQIAPGQTGKIADY	391	1.38
		657	ASYQTQTNSPRRARSVASQ	0.2556	TPTWRVYSTGSNVFQT	615	0.18
		680	YTMSLGAENSVAYSNN	0.6434	VIGIVNNTVYDPLQPE	1114	0.71
		771	KQIYKTPPIKDFGGF	-0.3896	SQSIIAYTMSLGAENS	674	0.56
		792	PDPSKPSKR	0.478	VSGTNGTKRFDNPVLP	55	0.51
		1093	NFYEPQIITTD	0.36	FPNITNLCPFGEVFNA	314	0.6
		1118	VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISG	0.13	SQILPDPSKPSKRSFI	788	0.26
					YQTQTNSPRRARSVAS	659	0.192
7	orf1ab polyprotein	373	CHNSEVGPEH	1.14	VVKIYCPACHNSEVGP	365	0.76
		763	LQPLEQPTSEAVEAP	0.05	TLKGGAPTKVTFGDDT	814	0.98
		813	FTLKGGAP	0.88	TSRYWEPEFYEAMYTP	3996	0.4
		1458	NLEEAAR	-0.14	AVTAYNGYLTSSSKTP	1482	0.35
		1482	AVTAYNGYLTSSSKTPEE	0.5	LNLEEAARYMRSLKVP	1457	0.33
		2241	FSSEIIGYKAI	0.26	SEAVEAPLVGTPVCIN	771	0.74
		3072	GCSCDQLREPMLQSADAQS	0.92	CGMWKGYGCSCDQLRE	3065	0.17
		3993	NDNTSRYWEP	0.27	SSEIIGYKAIDGGVTR	2242	0.74

Table 5: Analysis of antigenic epitope allergenicity and toxicity

S. No	Protein name	Start	epitope IEDB bepipred	Allergenicity	Toxicity	ABCpred	Start	Allergenicity	Toxicity
1	ORF8 protein	12	QHQPYVVDDP	Allergenic	-	PYVVDDPCPIHFYSKW	15	Non-allergenic	
2	ORF7a protein	58	HVYQLRARSVSPKLFIRQE	Allergenic	-	HVYQLRARSVSPKLFI	58	Allergen	-
3	membrane glycoprotein	7	TITVEELKK	Allergenic	-				
4	envelope protein	23	YSRVKNLNSSRVP	Non-Allergenic	Non-toxic	YVYSRVKNLNSSRVPD	21	Allergen	-
5	ORF3a protein	15	LKQGEIKDATPSDFVR	Non-Allergenic	Non-toxic	QGEIKDATPSDFVRAT	17	Allergen	-
						IGTVTLKQGEIKDATP	10	Allergen	-
6	surface glycoprotein	55	VSGTNGTKRF	Non-Allergenic	Non-toxic	HRSYLTPGDSSSGWTA	230	Non-allergenic	Non-toxic
		123	DPFLGVYYHKNNKSWMESEFRVYSSA	Non-Allergenic	Non-toxic	TVEKGIYQTSNFRVQP	292	Allergen	-
		234	LTPGDSSSGWTA	Non-Allergenic	Non-toxic	GCLIGAEHVNNSYECD	633	Non-allergenic	Toxic
		298	YQTSNFRVQP	Allergenic	-	LQSYGFQPTNGVGYQP	477	Non-allergenic	Non-toxic
		315	PNITNLCPFGEVFNATRFASVYAWNRKRISNC	Allergenic	-	EVRQIAPGQTGKIADY	391	Non-allergenic	Non-toxic
		389	GDEVRQIAPGQTGKIAD	Non-Allergenic	Non-toxic	TPTWRVYSTGSNVFQT	615	Non-allergenic	Non-toxic
		641	VNNSYECDIPI	Non-Allergenic	Non-toxic	VIGIVNNTVYDPLQPE	1114	Non-allergenic	Non-toxic
		680	YTMSLGAENSVAYSNN	Non-Allergenic	Non-toxic	SQSIIAYTMSLGAENS	674	Non-allergenic	Non-toxic
		792	PDPSKPSKR	Non-Allergenic	Non-toxic	VSGTNGTKRFDNPVLP	55	Allergen	-
						FPNITNLCPFGEVFNA	314	Allergen	-
7	orf1ab polyprotein	373	CHNSEVGPEH	Allergenic	-	VVKIYCPACHNSEVGP	365	Non-allergenic	Non-toxic
		813	FTLKGGAP	Non-Allergenic	Non-toxic	TLKGGAPTKVTFGDDT	814	Allergen	-
		1482	AVTAYNGYLTSSSKTPEE	Non-Allergenic	Non-toxic	TSRYWEPEFYEAMYTP	3996	Allergen	-
		3072	GCSCDQLREPMLQSADAQS	Non-Allergenic	Non-toxic	SEAVEAPLVGTPVCIN	771	Non-allergenic	Non-toxic
						SSEIIGYKAIDGGVTR	2242	Non-allergenic	Non-toxic

Table 6. Comparative analysis of Final MHC I, MHC II and B cell epitopes.

S. No	Protein Name	Start	B cell Epitopes (IFNscore)	Cross-reactivity with human	START	MHC I Epitopes (IFNscore)	Cross-reactivity with human	Start	MHC II Epitopes (IFNscore)	Cross-reactivity with human
1	ORF8 protein	15	PYVVDDPCPIHFYSKW(1)	NO				10	CTQHQPYVVDDPCPI(1)	NO
2	envelope protein	23	YSRVKNLNSSRVP(-0.49)	NO						
3	ORF3a protein	15	LKQGEIKDATPSDFVR(0.62)	NO	4	MRIFTIGTV(-0.43)	NO			
4	surface glycoprotein	55	VSGTNGTKRF(-0.75)	NO	242	WTAGAAAYY (0.57)	NO	98	KTQSLLIVNNATNVV(-0.37)	NO
		123	DPFLGVYYHKNNKSWMESEFRVYSSA(3.7)	NO				498	LSFELLHAPATVCGP(-0.51)	NO
		234	LTPGDSSSGWTA(-0.14)	NO	-	-	-	-	-	-
		641	VNNSYECDIPI(1)	NO	-	-	-	-	-	-
		680	YTMSLGAENSVAYSNN(-0.04)	NO	-	-	-		-	-
		792	PDPSKPSKR(0.12)	NO	-	-	-	-	-	-
		477	LQSYGFQPTNGVGYQP(14)	NO						
		391	EVRQIAPGQTGKIADY(1)	NO						
		1114	VIGIVNNTVYDPLQPE(12)	NO						
5	orf1ab polyprotein	365	VVKIYCPACHNSEVGP(2)	NO	3841	LSDDAVVCF(-0.30)	NO	20	KSAFYILPSIISNEK(0.33)	NO
		771	SEAVEAPLVGTPVCIN(2)	NO	2413	VVTTFDSEY(-0.51)	NO	1801	ESPFVMMSAPPAQYE(-0.21)	NO
		2242	SSEIIGYKAIDGGVTR(2)	NO	295	FMGRIRSVY(-0.25)	NO	2390	QMEIDFLELAMDEFI(0.31)	NO

	813	FTLKGGAP(-0.22)	NO	4615	LQAENVTGL(-0.21)	NO	2755	NYIFWRNTNPIQLSS(-0.27)	NO
	1482	AVTAYNGYLTSSSKTPEE(-0.04)	NO	2166	NYMPYFFTL(-0.58)	NO	663	QMNLKYAISAKNRAR(1)	NO
	3072	GCSCDQLREPMLQSADAQS(0.41)	NO	5406	FELEDFIPM(-0.28)	NO			

Table 7. designing of multi-epitope Vaccine constructs.

S.No	Name	Vaccine construct	Antigenicity	Allergenicity
1	VTC1	EAAAKCTQHQPYVVDDPCPIHEYGAEALERAGYSRVKNLNSSRVPGGGSMRIFTIGTVHEYGAEALERAGKTQSLLIVNNATNVVGGGSLKQGEIKDATPSDFVRHEYGAEALERAGWTAGAAAYYGGGSLSFELLHAPATVCGPHEYGAEALERAGVSGTNGTKRFGGGSLSDDAVVCFHEYGAEALERAGKSAFYILPSIISNEKGGGS VVKIYCPACHNSEVGPEAAAK	0.51	Non-allergenic
2	VTC2	EAAAKCTQHQPYVVDDPCPIHEYGAEALERAGYSRVKNLNSSRVPEAAAKLKQGEIKDATPSDFVRHEYGAEALERAGMRIFTIGTVEAAAKKTQSLLIVNNATNVVHEYGAEALERAGDPFLGVYYHKNNKSWMESEFRVYSSAEAAAKWTAGAAAYYHEYGAEALERAGPDPSKPSKREAAAKNYIFWRNTNPIQLSSHEYGAEALERAG SSEIIGYKAIDGGVTREAAAKLQAENVTGL	0.47	Non-allergenic
3	VTC3	EAAAKGCSCDQLREPMLQSADAQSHEYGAEALERAGFELEDFIPMEAAAKQMNLKYAISAKNRARHEYGAEALERAGEVRQIAPGQTGKIADYEAAAKLSFELLHAPATVCGPHEYGAEALERAGWTAGAAAYYEAAAKLKQGEIKDATPSDFVRHEYGAEALERAGMRIFTIGTVEAAAKYSRVKNLNSSRVPHEYGAEALERAG PYVVDDPCPIHFYSKWEAAAK	0.61	Non-allergenic

Table 8. Molecular docking of vaccine construct with different TLRs and HLA alleles.

S.No.	Receptor	Receptor PDB	Global Energy
1	TLR4	2Z62	-1.64
		3UL8	2.16
		2Z63	-28.76
		3UL9	-1.43

		2Z65	-18.55
		2Z66	-23.62
		3FXI	-6.93
		3ULA	-1.75
		4G8A	-2.65
		5NAM	-6.18
		2Z64	1.23
2	TLR1-2 hetero dimer	2Z81	-0.88
		2Z80	-31.20
		2Z82	0.22
		2Z7X	3.75
3	TLR1	6NIH	-1.95
		1FYV	-6.49
4	TLR2	6NIG	10.81
5	TLR6	3A79	-13.05
6	TLR3	2A0Z	-8.85
		1ZIW	6.11
7	HLA Alleles	5IM7	-24.66
		1XR8	-1.95
		1SYS	-17
		1A6A	-5.59
		1A1M	-3.24
		1BX2	6.09
		1H15	-9.32
		1ZSD	5.51
		3C5J	5.12
		4O2E	9.68

62

Figure 1

Figure 1

Figure 1. Brief workflow of combinational chimeric multi-epitope vaccine designing with predicted immune cell response.

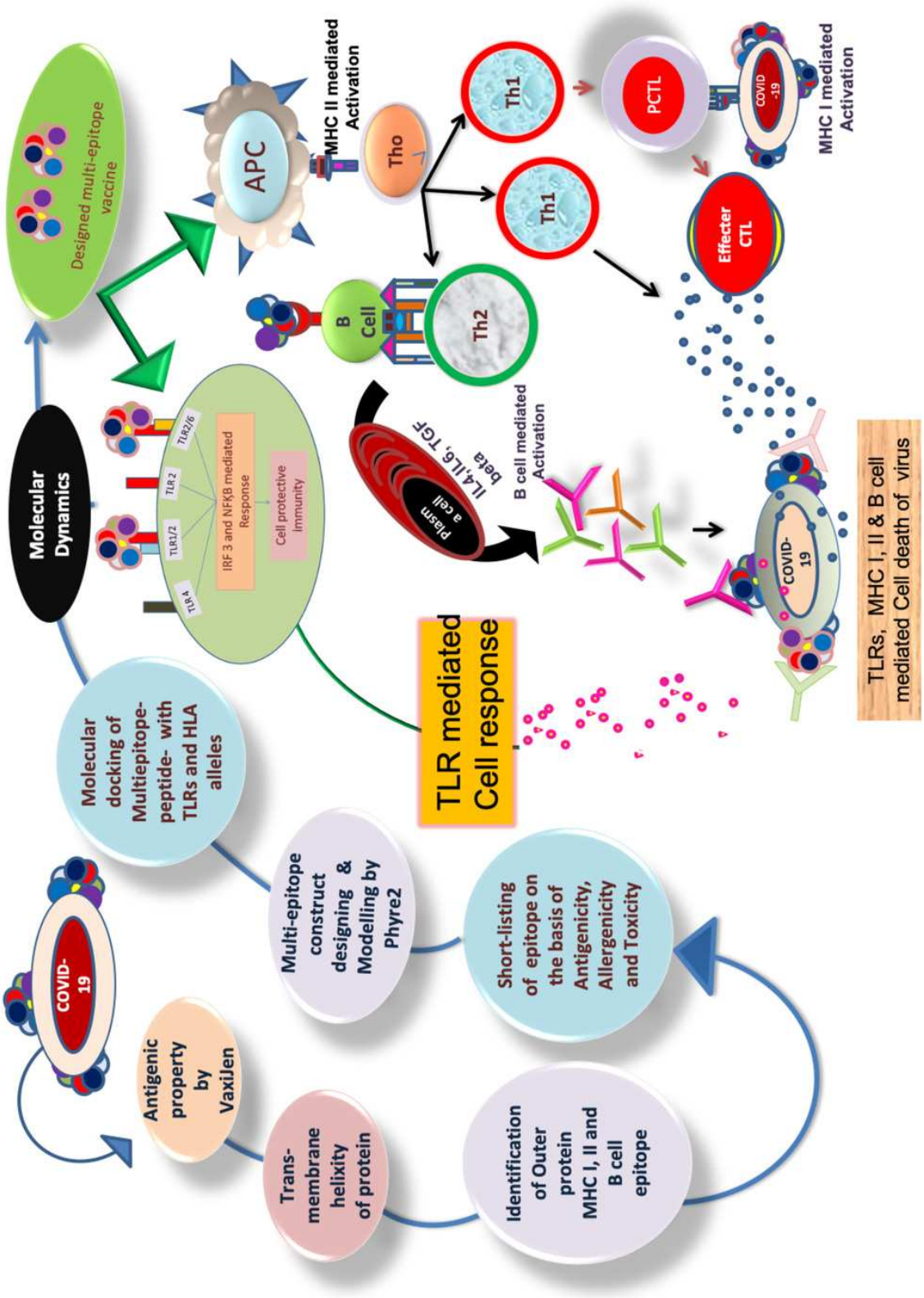
COVID-19
Antigenic property by VaxiJen
Trans-membrane helixity of protein
Identification of Outer protein MHC I, II and B cell epitope
Short-listing of epitope on the basis of Antigenicity, Allergenicity and Toxicity
Multi-epitope construct designing & Modelling by Phyre2
Molecular docking of Multiepitope-peptide- with TLRs and HLA alleles
Molecular Dynamics
Designed multi-epitope vaccine
TLR mediated Cell response
TLR4
TLR1/2
TLR 2
TLR2/6
IRF 3 and NFkB mediated Response
Cell protective Immunity
APC
MHC II mediated Activation
Tho
Th1
Th1
Th2
B Cell
PCTL
COVID -19
MHC I mediated Activation
Effecter CTL
IL4,IL6, TGF beta
Plasm a cell
B cell mediated Activation
COVID-19
TLRs, MHC I, II & B cell mediated Cell death of virus

Figure 2

Figure 2

Figure 2. Tertiary structure of modelled VTC3 construct (A) and Ramachandran plot of the modelled proteins (B).

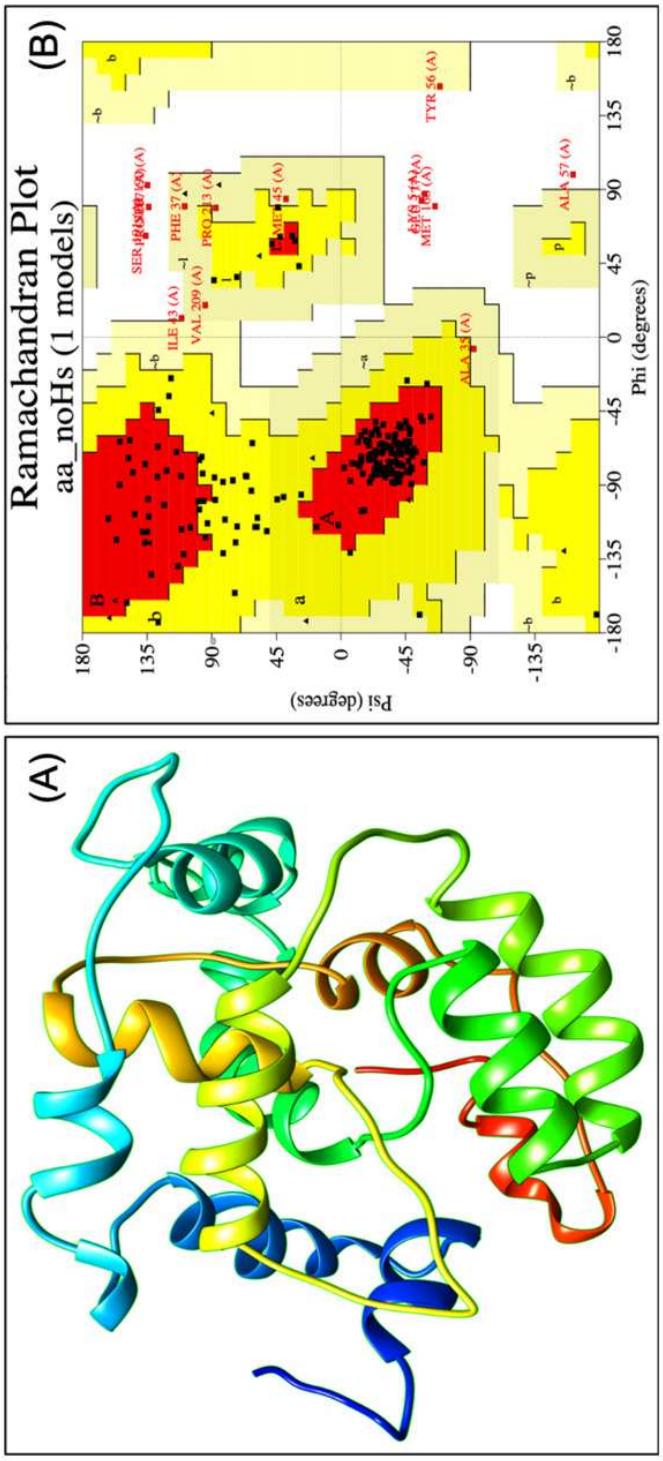
(A)
(B)
Ramachandran Plot
aa_noHs (1 models)
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
TYR 56 (A)
ALA 57 (A)
PHE 37 (A)
PRO 213 (A)
VAL 209 (A)
ALA 35 (A)
A
B
a
b
p
l

Figure 3

Figure 3

Figure 3. Docked pose of VTC3-TLR1/2 complex (A), and VTC3-TLR4 complex (B).

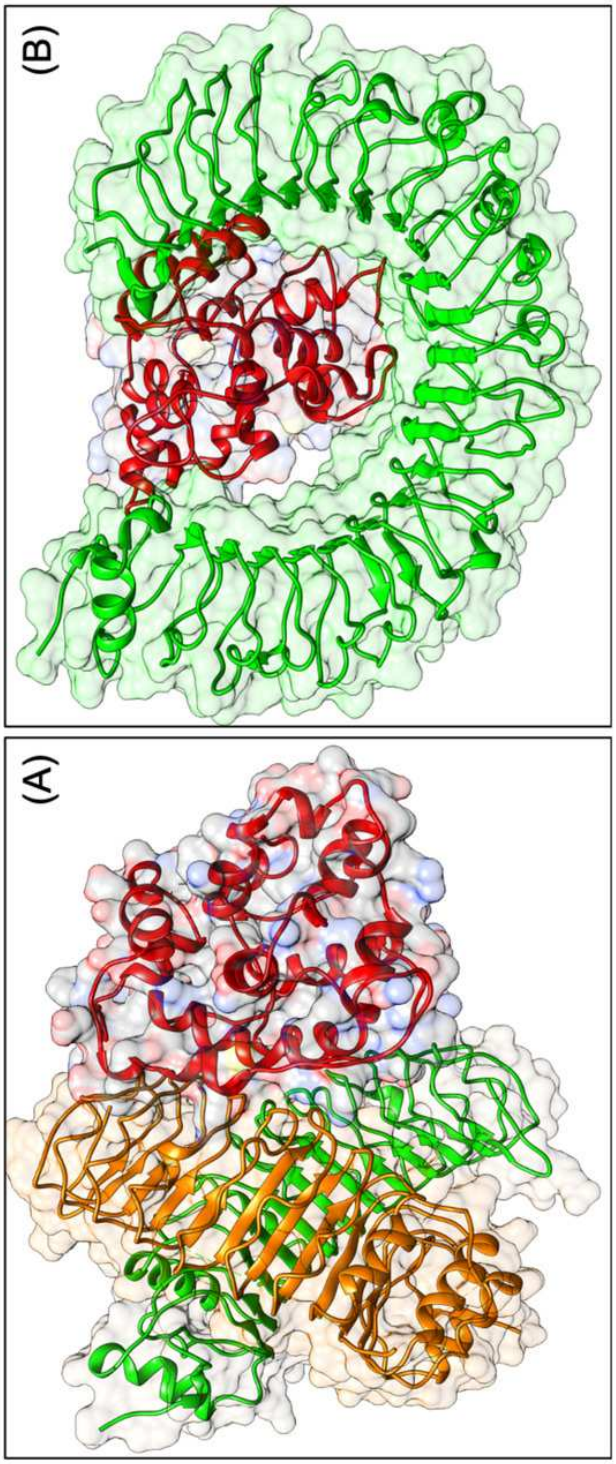
(B)
(A)

Figure 4

Figure 4

Figure 4. Root-mean-square deviation and Root mean square fluctuations during molecular dynamics simulation analysis of VTC3-TLR1/2 complex (A and B), and VTC3-TLR4 complex (C and D).

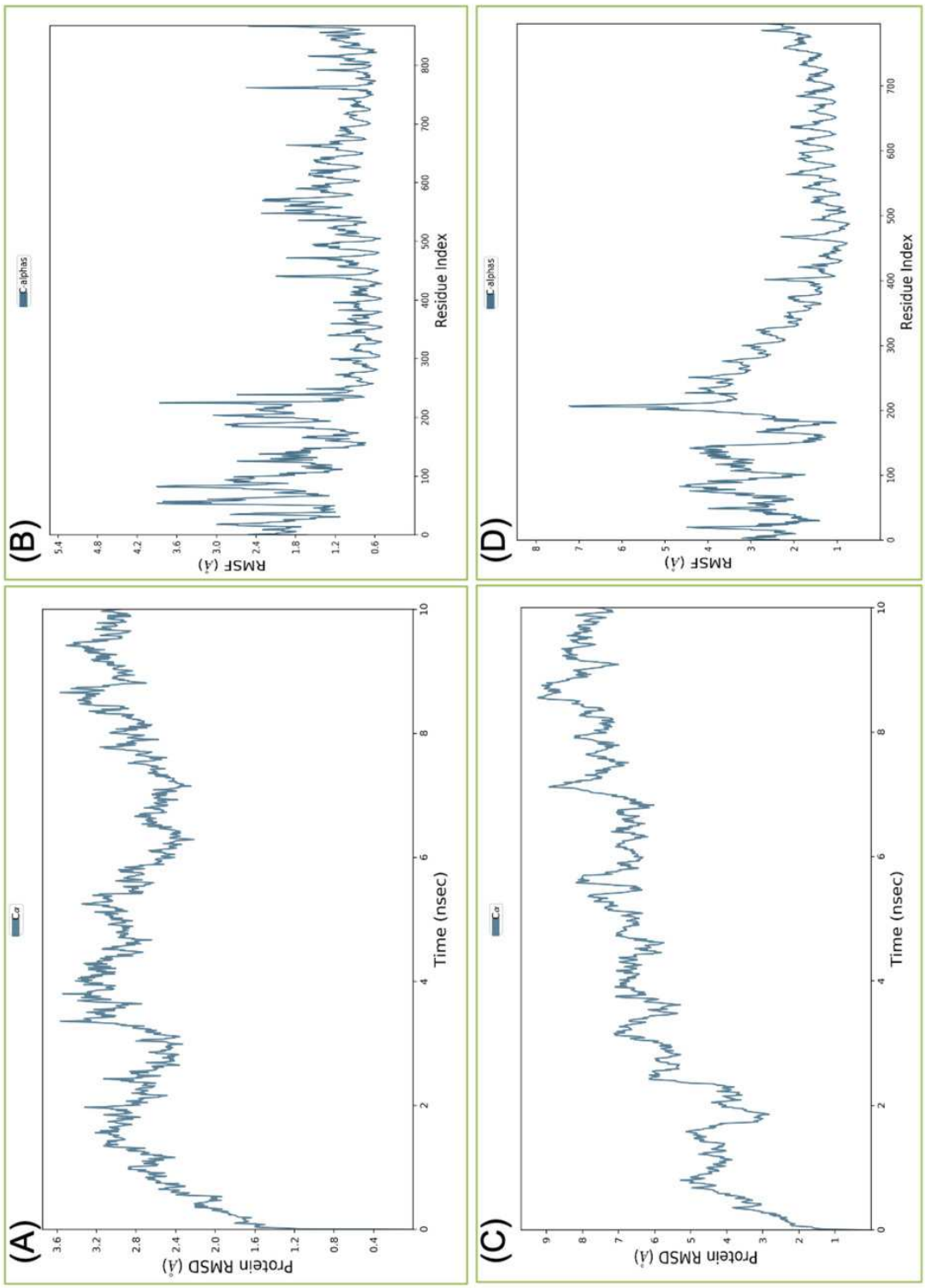
(A)
Protein RMSD (Å)
Time (nsec)
Cα
(B)
RMSF (Å)
Residue Index
C alphas
(C)
Protein RMSD (Å)
Time (nsec)
Cα
(D)
RMSF (Å)
Residue Index
C alphas

Figure 5

Figure 5

Figure 5. Root-mean-square deviation during molecular dynamics simulation analysis of VTC3-HLA complex. MDS was performed till 50ns.

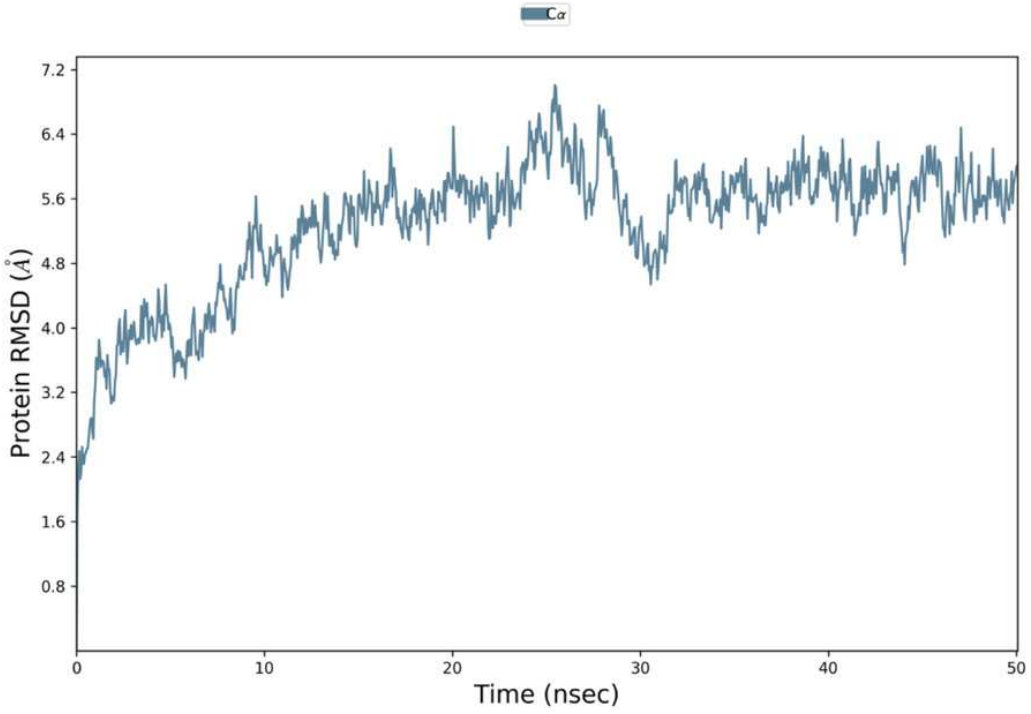
Cα
Protein RMSD (Å)
7.2
6.4
5.6
4.8
4.0
3.2
2.4
1.6
0.8
0
10
20
30
40
50
Time (nsec)

Figure 6

Figure 6

Figure 6. Diagrammatic presentation of proposed combinational chimeric multi-epitope vaccine VTC3 showing the position of different epitopes and linker in the vaccine.