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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Epitope based vaccine prediction for SARS-COV-2 by deploying immuno-informatics approach

Amit Joshi a, Bhuwan Chandra Joshi b, M. Amin-ul Mannan a, Vikas Kaushik a,∗

a Domain of Bioinformatics, School of Bio-Engineering and Bio-Sciences, Lovely Professional University, Punjab, India
b Department of Biological Sciences and Biotechnology, Surajmal Agarwal Private Girls PG Degree College Kiccha, Kumaun University, Nainital, Uttarakhand, India

1. Introduction

The outbreak of COVID-19 in the Hubei region of the Chinese city of Wuhan [37] has resulted in a difficult situation for the global populace and for the World Health Organization (WHO). On 30 Jan 2020, WHO declared an emergency regarding COVID-19 spread, prevention, and control [28,29]. As of 20 March 2020, there are over 240,000 cases and over 10,000 confirmed deaths, affecting 181 countries [9]. Present updates indicate that there are over 1 million confirmed cases and over 100,000 confirmed deaths due to the virus. Hence, prophylactic and therapeutic strategies are promptly needed. In this study we report an epitope, ITLCFTLKR, which is biochemically fit to HLA allelic proteins. We propose that this could be used as a potential vaccine candidate against SARS-COV-2. A selected putative epitope and HLA-allelic complexes show not only better binding scores, but also RMSD values in the range of 0–1 Å. This epitope was found to have a 99.8% structural favorability as per Ramachandran-plot analysis. Similarly, a suitable range of IC50 values and population coverage was obtained to represent greater validation of T-cell epitope analysis. Stability analysis using MDWeb and half-life analysis using the ProtParam tool has confirmed that this epitope is well-selected. This new methodology of epitope-based vaccine prediction is fundamental and fast in application, ad can be economically beneficial and viable.

ARTICLE INFO

A B S T R A C T

A new virus termed SARS-COV-2 (causing COVID-19 disease) can exhibit a progressive, fatal impact on individuals. The World Health Organization (WHO) has declared the spread of the virus to be a global pandemic. Currently, there are over 1 million cases and over 100,000 confirmed deaths due to the virus. Hence, prophylactic and therapeutic strategies are promptly needed. In this study we report an epitope, ITLCFTLKR, which is biochemically fit to HLA allelic proteins. We propose that this could be used as a potential vaccine candidate against SARS-COV-2. A selected putative epitope and HLA-allelic complexes show not only better binding scores, but also RMSD values in the range of 0–1 Å. This epitope was found to have a 99.8% structural favorability as per Ramachandran-plot analysis. Similarly, a suitable range of IC50 values and population coverage was obtained to represent greater validation of T-cell epitope analysis. Stability analysis using MDWeb and half-life analysis using the ProtParam tool has confirmed that this epitope is well-selected. This new methodology of epitope-based vaccine prediction is fundamental and fast in application, ad can be economically beneficial and viable.

Keywords:
Epitope
Immuno-informatics
Vaccine
HLA-Alleles
Simulation
SARS-COV-2

1. Introduction

The outbreak of COVID-19 in the Hubei region of the Chinese city of Wuhan [37] has resulted in a difficult situation for the global populace and for the World Health Organization (WHO). On 30 Jan 2020, WHO declared an emergency regarding COVID-19 spread, prevention, and control [28,29]. As of 20 March 2020, there are over 240,000 cases and over 10,000 confirmed deaths, affecting 181 countries [9]. Present updates indicate that there are over 1 million confirmed cases and over 100,000 confirmed deaths worldwide [39] (information is obtained from https://www.worldometers.info/coronavirus/). The coronavirus is an RNA type virus with positive sense strand feature and is associated with the Coronaviridae family under order Nidovirales, and is found to be dispersed among Primate order, members of class Mammalia, and specifically in humans [31]. Severe acute respiratory syndrome coronavirus (SARS-CoV) [11,21,22] and Middle East respiratory disorder coronavirus (MERS-CoV) [7,40], classified as β-coronaviruses, are found to be related to the novel SARS-COV-2. Until now, there has been no efficacious therapy to regulate its spread [24]. Investigating its spread across continents and the vulnerability, there is an urgent need to craft vaccines for reinforcing immune defense against the SARS-COV-2 virus. One of the strategies to combat it is the development of vaccines that can initiate an adaptive immune response in humans. Here an attempt has been made to design an epitope based vaccine directed at SARS-COV-2, by analyzing the proteome of the virus by using Immuno-informatics tools. In this study, we deployed the use of various bioinformatics servers and Immuno-informatics tools for identifying and recognizing the T-cell epitopes from the intensive study of available protein sequences and structures that are related to SARS-COV-2. These epitope stretches can interact with MHC Class I and Class II HLA alleles; further validation of epitopes was analyzed by Ramachandran Plot analysis, Antigenicity parameters evaluation, Toxicity analysis, Population coverage, Molecular dynamics, and ProtParam analysis [19]. This approach is an excellent method in modern vaccine design, as it provides a lead over classical trial and error methods of wet labs [23]. We tried to identify T-cell epitopes that can elicit a robust immune response in the global human population and act as potential vaccine candidates. However, the ability of these epitopes to act as a vaccine candidate needs to be analyzed in Molecular biology lab studies. Our investigation can open new dimensions in crafting peptide-based vaccine regimens for Novel SARS-COV2. The greatest decline in virus expansion was noted following ORF3a removal. ORF7a encodes a 122-amino-acid type I
transmembrane protein and structural studies disclose a packed seven-stranded β sandwich comparable in fold and topology to members of the immunoglobulin super family. SARS-CoV is an enveloped, positive-stranded RNA virus with a genome of around 29,700 bases. The genome incorporates at least 14 open reading frames (ORFs) that encode 28 proteins in three distinct classes: two large polyproteins P1a and P1ab that are cleaved into 16 non-structural proteins (nsp1–nsp16) during viral RNA synthesis; four structural proteins (S, E, M and N) that are necessary for viral entrance and gathering; and eight accessory proteins – nucleocapsid phosphoprotein, ORF7a protein, and membrane glycoprotein, which are crucial for the structural integrity and functionality of the virus [38].

**2. Methodology**

**2.1. Protein retrieval and allergenicity analysis**

Five protein sequences were selected from NCBI-GenBank, listed in Table 1, for SARS-COV-2 based on their Allergenicity that relies on the Tanimoto similarity index score produced by AllergenFP 1.0 [8]. The selected proteins were the envelope protein, ORF3a protein, nucleocapsid phosphoprotein, ORF7a protein, and membrane glycoprotein.

**2.2. T-cell epitope prediction for MHC HLA alleles**

IEDB (Immune epitope database) [20] along with NetMHCII PAN 3.2 and NetMHC 4.0 servers [18] were effectively used for finding putative peptide sequences that were aimed to interact with the MHC Class II and I HLA alleles, respectively (because of efficient algorithms based on artificial neural networks). The VaxJen score is determined for screening the best antigenic epitopes using the VaxJen online tool [10] with a threshold ≥ 1.0 for the viruses’ domain.

**2.3. Structural prediction: Putative epitopes and MHC HLA alleles**

The epitopes 3D structural findings were conducted by using the PEP-FOLD-3.5 server [25,33,34] and MHC HLA Allelic peptides tertiary structure predictions with the IEDB (Immune epitope database) [20] along with NetMHCII PAN 3.2 and PatchDock [32].

**2.4. Molecular docking analysis**

The selected epitopes and HLA complexes were docked for calculating refined interactions and binding energies, along with atomic contact energy (ACE), by using two docking web-servers: DINC 2.0 [2] and PatchDock [32].

**2.5. Molecular dynamics-simulation analysis of docked complex**

Molecular dynamics study was conducted to analyze RMSD values and atomic fluctuations for all amino acids under the 100 ps time frame by deploying the MDWeb server. MDWeb server was deployed to analyze Coarse grained MD Brownian dynamics (C-alpha) with specifications → Time: 100 ps, output frequency (steps) = 10, force constant (kcal/mol Å²) = 40, distance between alpha carbon atoms = 3.8 for both the interacting epitopes, and it was based on a GROMACS MD setup with solvation using an Amber-99sb* force-field [17].

**2.6. Toxicity, Ramachandran-plot, and population coverage analysis**

The ToxinPred server [16] is utilized for determining the toxicity
scoring of Epitopes for selecting non-toxic ones; also, the Ramachandran plot analysis was deployed by using the MolProbity 4.2 server [6] to analyze the quantitative presence of residues in the favorable region. The Immune Epitope Database (IED) resource web-server of population coverage was used to predict population coverage of the MHC II and MHC I alleles that interact with screened out epitopes based on their restriction database [5]. The MHCPred web-server was effectively used in quantitative prediction of sorted out epitopes interacting with HLA alleles of MHC II and MHC I [15]. Thereafter the ProtParam tool [42] of the ExPASy server was used to screen final stable epitopes based on the instability index and half-life.

### 3. Results

#### 3.1. T-cell epitopes prediction and VaxiJen scoring

Non-allergen proteins selected (Table 1) based on allergenicity scores (by deploying AllergenFP 1.0) New to NetMHCII PAN 3.2 and NETMHC 4.0 servers for determining 1-log50k value and affinity values for selecting the best possible pair of epitopes with their corresponding HLA alleles. In Table 2 and Table 3, results for MHC Class I HLA and MHC Class II HLA alleles paired epitopes along with their VaxiJen scores were represented respectively, to obtain putative epitopes. The results are self-explanatory in Fig. 2. Graphical representation of selected peptides for docking based on their interaction with MHC Class I and Class II HLA alleles along with their Antigenicity.

**Table 2**

| NCBI-GenBank ID | MHC I Allele | POS | Core peptide | 1-LOG50K | Affinity(nM) | VaxiJen score | Antigenicity |
|-----------------|--------------|-----|--------------|----------|--------------|---------------|--------------|
| QHD43417.1     | HLA-A*68:01  | 7   | FTIGTVTLK    | 0.853    | 4.9          | 2.0317        | ANTI GEN     |
| QHD43417.1     | HLA-A*31:01  | 125 | RLWLCWKCR    | 0.74     | 16.59        | 1.1604        | ANTI GEN     |
| QHD43419.1     | HLA-A*11:01  | 5   | GTYTVEELK    | 0.719    | 20.98        | 1.0976        | ANTI GEN     |
| QHD43421.1     | HLA-A*11:01  | 109 | ITLCFTLKR    | 0.71     | 22.97        | 2.0208        | ANTI GEN     |
| QHD43421.1     | HLA-A*68:01  | 109 | ITLCFTLKR    | 0.695    | 27.24        | 2.0208        | ANTI GEN     |
| QHD43423.2     | HLA-A*23:01  | 107 | VFTLCFTL     | 0.621    | 60.42        | 1.2490        | ANTI GEN     |
| QHD43423.2     | –            | –   | –            | –        | –            | –             | NO INTERACTION |

**Table 3**

| NCBI-GenBank ID | MHC II Allele | POS | Core peptide | 1-LOG50K | Affinity(nM) | VaxiJen score | Antigenicity |
|-----------------|--------------|-----|--------------|----------|--------------|---------------|--------------|
| QHD43417.1     | –            | –   | –            | –        | –            | –             | NO INTERACTION |
| QHD43418.1     | –            | –   | –            | –        | –            | –             | NO INTERACTION |
| QHD43419.1     | HLA-DRB1*04:01 | 55  | WLIWPVTA     | 0.282    | 2375.78      | 1.0631        | ANTI GEN     |
| QHD43421.1     | HLA-DRB1*01:01 | 74  | VYQLRARSV    | 0.484    | 267.05       | 1.3108        | ANTI GEN     |
| QHD43421.1     | HLA-DRB1*07:01 | 74  | VYQLRARSV    | 0.342    | 1235.08      | 1.3108        | ANTI GEN     |
| QHD43423.2     | –            | –   | –            | –        | –            | –             | NO INTERACTION |

**Table 4**

| Allele Name     | Template structure (PDB-ID) | Crystal structure/Model |
|-----------------|-----------------------------|-------------------------|
| HLA-A*11:01     | 2HN7                        | CRYSTAL STRUCTURE       |
| HLA-A*23:01     | 36L                         | CRYSTAL STRUCTURE       |
| HLA-A*31:01     | 3RL1                        | CRYSTAL STRUCTURE       |
| HLA-A*68:01     | 6PBH                        | CRYSTAL STRUCTURE       |
| HLA-DRB1*01:01  | 4AJ2                        | CRYSTAL STRUCTURE       |
| HLA-DRB1*04:01  | 5LAX                        | CRYSTAL STRUCTURE       |
| HLA-DRB1*07:01  | 6BLJ                        | CRYSTAL STRUCTURE       |
4

In Table 4 the crystal structure/model structure details with server; the HLA allele Class II HLA alleles along with their antigenicity. - peptides for docking are based on their interaction with MHC Class I and II HLA alleles along with their antigenicity.

Table 5

| Epitope     | HLA Allele | Binding score (kcal/mol) | Patch dock score | ACE | Selection |
|-------------|------------|--------------------------|-----------------|-----|-----------|
| FTIGTVTLK   | HLA-A*68:01| -8.00                    | 8066            | 178.91| Selected |
| RLWLOWKCR   | HLA-A*31:01| -4.80                    | 8916            | 117.53| Rejected |
| GITVEELK    | HLA-A*11:01| -3.80                    | 8040            | 78.78 | Rejected |
| ITLCFTLKR   | HLA-A*11:01| -3.70                    | 8206            | 25.88 | Selected |
| ITLCFTLKR   | HLA-A*68:01| -7.60                    | 8136            | 184.55| Selected |
| VITLFCFTL   | HLA-A*23:01| -4.40                    | 7706            | 134.91| Rejected |
| WLLKVVLTA   | HLA-DRB1*04:01| -8.60                  | 9432            | 150.96| Rejected |
| VYQLRARSV   | HLA-DRB1*01:01| -6.20                | 6874            | 168.74| Selected |
| VYQLRARSV   | HLA-DRB1*07:01| -6.20                | 6842            | 262.74| Selected |

MHC class I HLA Alleles, and VYQLRARSV epitope interacted with MHC Class II HLA alleles with a perfect binding score and ACE values as shown in Table 5. The ITLCFTLKR Epitope of the ORF-7A protein exhibits binding with 2 HLA alleles (HLA-A*11:01, HLA-A*68:01) of MHC Class I, while FTIGTVTLK Epitope of ORF-3a protein interact with 1 HLA Allele (HLA-A*68:01) of MHC Class I. The VYQLRARSV Epitope of ORF-7a protein interacts clearly with 2 HLA Alleles (HLA-DRB1 *01:01, HLA-DRB1 *07:01) of the MHC Class II domain.

In Figs. 3, 4, and 5, interactions between a selected three T-Cell epitopes with respective MHC Class I and II HLA-Alleles via hydrogen bond formation and van der Waals interactions is depicted. After positive docking results, these epitopes were subjected to further Molecular dynamic simulation and biochemical parameters assessment. Fig. 6 represents a graphical plot of binding scores for epitopes interacting with HLA-Alleles.

3.4. Molecular dynamics and simulation analysis

RMSD values and Atomic fluctuation per amino acid residue were obtained for Epitopes interacting with the HLA-Allele structure; this analysis allows a perfect pair selection and validation. Moreover, only two Epitope pairs, i.e., ITLCFTLKR and VYQLRARSV, were identified as probable T-cell epitopes and as putative vaccine specimens. Fig. 7 shows the RMSD Plot and Atomic fluctuation per residue for the ITLCFTLKR-HLA-A*68:01 complex, the RMSD Plot and Atomic fluctuation per residue for the VYQLRARSV-HLA-DRB1*07:01 complex. Both results were positive as best interactions, for protein-ligand docked complexes must possess RMSD values from 0 to 1.0 Å as a preferred range [13].

3.5. Toxicity analysis, Ramachandran Plot analysis, and population coverage results

ToxinPred 4.0 server results (in Table 6.) represent Finalized T-cell Epitopes that were nontoxic from the biochemical perspective.

Ramachandran plot analysis, Fig. 8A and B suggest that most of the residues are allowed in a favored region; this gives more confidence in the structural conformation for targeted T-Cell Epitopes.

MHCpred results (Table 7) indicate quantitative estimation of IC50 values for both MHC I and MHC II alleles for respective Epitopes shows elicitation of an immune response when this data is deployed in a population coverage analysis.

IEDB population coverage analysis suggests that ITLCFTLKR and VYQLRARSV epitopes exhibit a suitable population coverage, as depicted in the graphical representation of Fig. 9A and B. This allows only two probable Epitopes for the final selection of vaccine crafting.

In Table 8, ProtParam analysis further reveals the stability of the considered epitopes and final revelation of one epitope ITLCFTLKR is screened out. This particular Epitope exhibits an instability index of...
35.68, with a grand average of hydropathicity (GRAVY) calculated was 0.844, and the estimated half-life for this peptide was determined to be 20 h for mammalian reticulocytes.

4. Discussion

In this study, SARS-COV-2 virus proteins were analyzed by using In-silico methods, and can be further utilized for vaccine trials as per earlier successes in the case of similar SARS-COV studies, and later observed in the development of polyclonal antibodies [27]. Here we obtained two epitopes ITLCFTLKR and VYQLRARSV after successful docking and molecular dynamics simulation; furthermore, these two epitopes were subjected to population coverage and toxicity analysis. Similarly, in another study, for MERS-COV, nucleocapsid peptides were used for T-cell epitope prediction, and found to be successful [36]. The IEDB and NCBI-GenBank database were fully deployed to analyze sequence homology, to predict targets for COV-2 in case of viral protein identification as per the related studies [14], as VIPR (Virus Pathogen database

Fig. 6. Binding energy graphical plot for selected Epitope and HLA-Allelic pair.

Fig. 7. A. RMSD Plot for ITLCFTLKR- HLA-A*68:01 complex, for each amino acid residue by Molecular dynamics analysis, B. B-Factor (atomic fluctuation) values per amino acid residue for Epitope ITLCFTLKR- HLA-A*68:01 docked complex, C. RMSD Plot for VYQLRARSV- HLA-DRB1*01:01, for each amino acid residue by Molecular dynamics analysis, D. B-Factor (atomic fluctuation) values per amino acid residue for Epitope VYQLRARSV- HLA-DRB1*07:01 docked complex.
analysis resource) are also dependent on IEDB and GenBank primarily [30]. We analyzed five different proteins in SARS-COV-2 for the present study (because of their availability in the NCBI-GenBank database and importance in a structural role in SARS-COV-2 [14] and finally revealed T–Cell epitopes that can be used for wet lab considerations and time savings. In a very recent study, different epitopes were found for SARS-COV-2, based on In-silico approaches and focused on only surface glycoprotein [3], but in our research study there are many differences as we analyzed a different group of proteins from SARS-COV-2 to sort out short length T-Cell epitopes specific to MHC I as well as MHC II diversified HLA-Alleles.

It is reported for SARS-CoV HLA-B*4601, HLA-B*0703, HLA-DR B1*1202 are activated [26], interaction with different MHC I and II allelic forms namely HLA-A*11:01, HLA-A*68:01, HLA-DRB1 *01:01 and HLA-DRB1 *07:01. CD4+ and CD8+ memory T cells. Based on prior literature, it is anticipated that it can persist for four years as in the case of SARS-CoV recovered individuals, show T-cell proliferation, DTH response, and production of IFN-γ [12]. We surmise that our screen can be more effective and useful. Primarily molecular docking reveals three Epitopes, but as we proceed to Molecular dynamic simulations, it reveals best interactions for two epitopes i.e., ITLCFTLKR and VYQLRARSV, with acceptable stability analyzed with the help of MDWeb and identified by using best available tools with easy-to-apply methods. One recent study was found to be focused on developing monoclonal

| Peptide/Probable antigen | SVM score | Hydrophilicity | Molecular weight | Toxicity |
|--------------------------|-----------|----------------|------------------|----------|
| FTITGTVTLK               | −1.36     | −1.23          | 979.32           | NON-TOXIN|
| GIVITVEELK               | −0.98     | 0.34           | 989.27           | NON-TOXIN|
| ITLCFTLKR                | −1.32     | −0.41          | 1094.51          | NON-TOXIN|
| VFTILCFITL               | −1.21     | −1.52          | 1056.46          | NON-TOXIN|
| WLLWPVTLA                | −1.18     | −1.62          | 1098.49          | NON-TOXIN|
| VYQLRARSV                | −1.07     | −0.12          | 1091.40          | NON-TOXIN|

Table 6
Results of ToxinPred on probable antigens.

![Fig. 8. A. 99.8% residues of the ITLCFTLKR Epitope were in the allowed and favored region under Ramachandran Plot analysis. B. 99.8% residues of the VYQLRARSV Epitope were in the allowed and favored region under Ramachandran Plot analysis.](image)

Table 7
MHCpred results depict IC50 Values for HLA Alleles and confidence of the prediction.

| HLA Alleles | Amino acid groups | Predicted -logIC50 (M) | Predicted IC50 Value (nM) | Confidence of prediction (Max = 1) |
|-------------|-------------------|------------------------|---------------------------|-----------------------------------|
| HLA-A*68:01 | FTITGTVTLK        | 7.116                  | 76.56                     | 1.00                              |
| HLA-A*11:01 | ITLCFTLKR         | 7.028                  | 93.76                     | 1.00                              |
| HLA-A*68:01 | ITLCFTLKR         | 6.282                  | 522.40                    | 0.78                              |
| HLA-DRB1*01:01 | VYQLRARSV   | 7.624                  | 23.77                     | 0.89                              |
| HLA-DRB1*07:01 | VYQLRARSV | 6.734                  | 184.50                    | 0.89                              |
antibodies like CR-3022 against the Spike protein of SARS-COV-2 that also exhibits interaction with ACE (Angiotensin Converting Enzyme) enzyme of the Human respiratory epithelium and requires complex neutralizing mechanisms for several binding domains [35], whereas in our study the putative T-cell epitopes can directly interact with MHC-Allelic sets that can be useful for developing immunization against SARS-COV-2. ProtParam [42] analysis further reveals the stability of the considered epitopes, and final revelation of one epitope ITLCFTLKLR is screened out. This particular Epitope shows an instability index of 35.68 with a grand average of hydrophaticity (GRAVY) calculated as 0.844, and the estimated half-life for this peptide was determined to be 20 h for mammalian reticulocytes.

Satisfactory population coverage was observed for targeted epitopes - HLA allelic complexes at the worldwide, South Asia, and India level. The biochemical integrity in epitope structure was further evident by deploying Ramachandran plot analysis. Both epitopes were non-toxic.

Table 8
ProtParam analysis for selected epitopes.

| Selected Epitope | GRAVY Score | Instability Index (Indication) | Estimated Half-Life (Mammalian reticulocytes) | Theoretical pl | Aliphatic Index |
|------------------|-------------|-------------------------------|---------------------------------------------|---------------|----------------|
| ITLCFTLKLR       | 0.844       | 35.68 (Stable)                | 20 Hours                                    | 9.51          | 130.00         |
| VYQLRARSV        | -0.067      | 70.73 (Unstable)              | 100 Hours                                   | 10.83         | 118.89         |

Fig. 9. A. Graphical representation of population conservancy analysis of ITLCFTLKLR Epitope. B. Graphical representation of population conservancy analysis of VYQLRARSV Epitope.
non-allergenic, and possess good antigenicity. In a similar study of the
preliminary analysis of COVID-19 vaccine targets [1] the investigators
tried to use the spike protein and nucleo-capid protein sequences of
SARS-CoV that are homologous to some extent with SARS-CoV-2 pro-
teins to determine multiple different epitopes for Vaccine prediction, but
in our study out of five two proteins namely ORF-3a and ORF-7a specific
to SARS-CoV-2 were found to be putative T-cell epitope determinants
that create useful information; these proteins are also important for viral
replication [41]. Both biochemical parameters, as well as an advanced
HMm and ANN based algorithm in selected Immunoinformatics tools,
were very useful to present a clear picture of predicted epitopes for
crafting vaccine against SARS-CoV-2. The only limitation that can be
considered as future scope is that these easily synthesized peptides
should be tested with In-vitro study for more practical validation.

5. Conclusion

ITLCTFLKR epitope was selected for crafting and designing a vaccine
against SARS-CoV-2. This particular epitope has good antigenicity, ex-
hibits active binding with MHC HLA-Alleles, and has maximum popu-
lation coverage for different geographical regions. Therefore, this
can be further used in vaccine design against SARS-CoV-2 after
wet lab verification. This novel approach can also assist life science
research groups to reduce time, monetary expenditures, as well as
physical hit-trial efforts.

Ethical approval

I confirm that authors did not perform any experiments on human or
animals.

Declaration of competing interest

I confirm that the authors hereby declare they that have no conflict of
interest.

Acknowledgement

We acknowledge the support provided by Computational lab of
Lovely professional University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.imu.2020.100338.

References

[1] Ahmed SF, Quadeer AA, McKay M. Preliminary identification of potential vaccine
targets for 2019-nCoV based on SARS-CoV immunological studies. Viruses 2020;12
(3):254. https://doi.org/10.3390/v12030254.

[2] Antunes DA, Moli M, DeVauers D, Jackson KH, Lizzie G, Ravrlesi LE. DINC-2.0: a new
protein-peptide docking webserver using an incremental approach. Canc Res 2017;
77(21). e55–57.

[3] Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell
epitopes in the surface glycoprotein of 2019-nCoV. J Med Virol 2020:1–6. https://
doi.org/10.1002/jmv.25698.

[4] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weisig H, Shindyalov IN,
Boume PE. The protein data bank. Nucleic Acids Res 2000;28:235–42.

[5] Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting
replication [41]. Both biochemical parameters, as well as an advanced
HMm and ANN based algorithm in selected Immunoinformatics tools,
were very useful to present a clear picture of predicted epitopes for
crafting vaccine against SARS-CoV-2. The only limitation that can be
considered as future scope is that these easily synthesized peptides
should be tested with In-vitro study for more practical validation.

5. Conclusion

ITLCTFLKR epitope was selected for crafting and designing a vaccine
against SARS-CoV-2. This particular epitope has good antigenicity, ex-
hibits active binding with MHC HLA-Alleles, and has maximum popu-
lation coverage for different geographical regions. Therefore, this
can be further used in vaccine design against SARS-CoV-2 after
wet lab verification. This novel approach can also assist life science
research groups to reduce time, monetary expenditures, as well as
physical hit-trial efforts.

Ethical approval

I confirm that authors did not perform any experiments on human or
animals.

Declaration of competing interest

I confirm that the authors hereby declare they that have no conflict of
interest.

Acknowledgement

We acknowledge the support provided by Computational lab of
Lovely professional University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.
org/10.1016/j.imu.2020.100338.

References

[1] Ahmed SF, Quadeer AA, McKay M. Preliminary identification of potential vaccine
targets for 2019-nCoV based on SARS-CoV immunological studies. Viruses 2020;12
(3):254. https://doi.org/10.3390/v12030254.

[2] Antunes DA, Moli M, DeVauers D, Jackson KH, Lizzie G, Ravrlesi LE. DINC-2.0: a new
protein-peptide docking webserver using an incremental approach. Canc Res 2017;
77(21). e55–57.

[3] Baruah V, Bose S. Immunoinformatics-aided identification of T cell and B cell
epitopes in the surface glycoprotein of 2019-nCoV. J Med Virol 2020:1–6. https://
doi.org/10.1002/jmv.25698.

[4] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weisig H, Shindyalov IN,
Boume PE. The protein data bank. Nucleic Acids Res 2000;28:235–42.

[5] Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting
replication [41]. Both biochemical parameters, as well as an advanced
HMm and ANN based algorithm in selected Immunoinformatics tools,
were very useful to present a clear picture of predicted epitopes for
crafting vaccine against SARS-CoV-2. The only limitation that can be
considered as future scope is that these easily synthesized peptides
should be tested with In-vitro study for more practical validation.
Veit S, Jany S, Fux R, Sutter G, Volz A. CD8+ T cells responding to the Middle East respiratory syndrome coronavirus nucleocapsid protein delivered by vaccinia virus MVA in mice. Viruses 2018;10(12).

Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020;395:470–3.

Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev 2005;69 (4):635–64.

Worldometersinfo. Rapidly updating coronavirus cases information and graphical representation website. http://www.worldometers.info/coronavirus-cases/. [Accessed 12 April 2020].

Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814–20.

Hasan M, Ghosh PP, Azim KF, Mukta S, Abir RA, Nahar J, Khan MM. Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. Microb Pathog 2019 May 1;130:19–37.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Humana press; 2005. p. 571–607.

Neuman BW, Joseph JS, Saikatendu KS, Serrano P, Chatterjee A, Johnson MA, Liao L, Klaus JP, Yates JR, Wüthrich K, Stevens RC. Proteomics analysis unravels the functional repertoire of coronavirus nonstructural protein 3. J Virol 2008 Jan 1;82(11):5279–94.

Tang X, Li G, Vasilakis N, Zhang Y, Shi Z, Zhong Y, Wang LF, Zhang S. Differential stepwise evolution of SARS coronavirus functional proteins in different host species. BMC Evol Biol 2009 Dec;9(1):52.

Sanchez-Trincado JL, Gomez-Personan M, Reche PA. Fundamentals and methods for T-and B-cell epitope prediction. J Immunol Res 2017:2017.

Bhattacharya M, Sharma AR, Patra P, Ghosh P, Sharma G, Patra BC, Lee SS, Chakraborty C. Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immuninformatics approach. J Med Virol 2020 Jun;92(6):618–31.