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.
Antigenic sites in SARS-CoV-2 spike RBD show molecular similarity with pathogenic antigenic determinants and harbors peptides for vaccine development

Tikam Chand Dakal

Genome & Computational Biology Lab, Department of Biotechnology, Mohanlal Sukhadia University, Udaipur 313001, Rajasthan, India

ARTICLE INFO

Keywords:
SARS-CoV-2
Mycobacterium
Yersinia
Listeria
Plasmodium
Bacillus
Borrelia
Vaccine
Antigenic variations
Antigenic drift
Antigenic shift

ABSTRACT

The spike protein of coronavirus is key target for drug development and other pharmacological interventions. In current study, we performed an integrative approach to predict antigenic sites in SARS-CoV-2 spike receptor binding domain and found nine potential antigenic sites. The predicted antigenic sites were then assessed for possible molecular similarity with other known antigens in different organisms. Out of nine sites, seven sites showed molecular similarity with 54 antigenic determinants found in twelve pathogenic bacterial species (Mycobacterium tuberculosis, Mycobacterium leprae, Bacillus anthracis, Borrelia burgdorferi, Clostridium perfringens, Clostridium tetani, Helicobacter Pylori, Listeria monocytogenes, Staphylococcus aureus, Streptococcus pyogenes, Vibrio cholera and Yersinia pestis), two malarial parasites (Plasmodium falciparum and Plasmodium knowlesi) and influenza virus A. Most of the bacterial antigens that displayed molecular similarity with antigenic sites in SARS-CoV-2 RBD (receptor binding domain) were toxins and virulent factors. Antigens from Mycobacterium that showed similarity were mainly involved in modulating host cell immune response and ensuring persistence and survival of pathogen in host cells. Presence of a large number of antigenic determinants, similar to those in highly pathogenic microorganisms, not merely accounts for complex etiology of the disease but also provides an explanation for observed pathophysiological complications, such as deregulated immune response, unleashed or dysregulated cytokine secretion (cytokine storm), multiple organ failure etc., that are more evident in aged and immune-compromised patients. Over-representation of antigenic determinants from Plasmodium and Mycobacterium in all antigenic sites suggests that anti-malarial and anti-TB drugs can prove to be clinical beneficial for COVID-19 treatment. Besides this, anti-leprosy, anti-lyme, anti-plague, anti-anthrax drugs/vaccine etc. are also expected to be beneficial in COVID-19 treatment. Moreover, individuals previously immunized/vaccinated or had previous history of malaria, tuberculosis or other disease caused by fifteen microorganisms are expected to display a considerable degree of resistance against SARS-CoV-2 infection. Out of the seven antigenic sites predicted in SARS-CoV-2, a part of two antigenic sites were also predicted as potent T-cell epitopes (KVGGNYNYL\textsuperscript{444-452} and SVLYNSASF\textsuperscript{366-374}) against MHC class I and three (KRISNCVADYSVLYN\textsuperscript{356-370}, DLCFTNVYADSFVI\textsuperscript{389-402}, and YRVVVLSFELLHA\textsuperscript{508-520}) against MHC class II. All epitopes possessed significantly lower predicted IC50 value which is a prerequisite for a preferred vaccine candidate for COVID-19.

1. Introduction

The recent outbreak of novel coronavirus SARS-CoV-2 (earlier known as 2019-nCoV) in Wuhan city of China has led to serious health crisis as well as impacted socio-economic development globally. Alike previous SARS-CoV-1, the SARS-CoV-2 is also an enveloped “positive” single-stranded RNA virus. SARS-CoV-2’s genome encodes for 12 proteins that include non-structural proteins and four structural proteins such as S (spike), E (envelop), M (membrane) and N (nucleocapsid) proteins (Kuiken et al., 2003; Marra et al., 2003; Peiris et al., 2003; Rota et al., 2003). The ability of coronaviruses to enter into host cells and infect them is due to their ability to establish strong attachment with host cell receptor proteins such as integrins and others via their RGD motif present in their receptor binding domain of spike glycoproteins (Dakal, 2020; Li et al., 2005). The RGD motifs are well known in the field of cell biology, cell therapy and tissue engineering because of their...
Bivalent ions such as Ca\textsuperscript{2+} and Mg\textsuperscript{2+} play indispensable role in promoting such cell adhesion and EDTA are used as a chelating agent for disaggregating Ca\textsuperscript{2+} mediated cell bondage. The RGD-integrin mediated virus-host attachment has been found to be dependent upon the calcium and other divalent ions’ concentration (Dakal, 2020).

Initially the spike proteins bind to cellular receptor angiotensin-converting enzyme 2 (ACE2). Following the cleavage of spike protein and with the help of proteases produce 2 domains of spike protein, S1 known as the receptor binding domain, which recognizes and bind to the RBD and S2 which fuse with the membrane following the entry of the virus into the host cell (Hoffmann et al., 2020). After that the viral RNA get exposed in the cytoplasm of the host cell. Virus infection in humans results in two major types of immune response. The first is an innate immune response, which is first line of defense and involves synthesis of proteins called interferons (key regulators of viral replication) and stimulation of a number of immune cells (Theofilopoulos et al., 2005; Wei et al., 2014). The viral infection is detected by innate immune system by pattern recognition receptors (PRRs) that recognize pathogen-associated molecular patterns (PAMPs). The PRRs mainly include toll-like receptor (TLR), RIG-I-like receptor (RLR), NOD-like receptor (NLR), C-type lectin-like receptors (CLR), and free-molecule receptors in the cytoplasm, such as cGAS, IFI16, STING, DAI, and so on (Thompson et al., 2011; Schoek et al., 2017; Bottermann and James, 2018). After recognizing pathogenic features in the invading molecule, virus is delivered to macrophage and dendritic cells (DCs) that play crucial roles for viral destruction and immune response induction in mucosal-associated lymphoid tissues (MALT) (Spiegel et al., 2006; Spiegel and Weber, 2006; Frieman and Baric, 2008). Besides this, the host response is also regulated by a highly controlled network of cytokines (TNF-\(\alpha\), IL-6, IL-8, IP-10, MCP-1), chemokines (CXCL-1, CXCL-2, CCL-3 and CCL-5) and complement proteins (Friedman and Baric, 2008; Chen et al., 2010; van den Brand et al., 2014). Studies have also shown that activation of macrophages and DCs by SARS-CoV lead to excessive pro-inflammatory cytokine responses (Tseng et al., 2005; Mehta et al., 2020). On activation, an unleashed production of inflammatory cytokines and chemokines can be seen in the tissues and serum of the COVID-19 patients, a pathophysiological condition formally known as cytokine storm (Mehta et al., 2020; Huang et al., 2020). Also, the levels of IFN-\(\gamma\), IL-1\(\beta\), IL-2, IL-6, IL-7, IL-8, IL-10, IL-12, MCP-1, MIP1A, IP-10 and TNF-\(\alpha\) are generally increased in the early infection; however, subsequently get lowered in the recovery stage (Huang et al., 2020). Upregulation of the inflammatory cytokines and chemokines like IL-1\(\beta\), IFN-\(\gamma\), IP-10, and MCP-1 may lead to activated T-helper-1 (Th1) cells response (Huang et al., 2020). However, it was also observed that SARS-CoV-2 patients secreted excessive IL-4 and IL-10 that may suppress inflammation via T-helper-2 (Th2) which makes the SARS-CoV-2 different from other previous coronavirus infections (Zhao and Zhao, 2020). The induction of interferons (IFN) along with anti-viral actions of macrophages and DCs at the sites of infection effectively hinders viral tropism in lung tissues/ cells and eventually dampens virus’s efficiency to replicate and finally led to their elimination (Theofilopoulos et al., 2005; Garcia-Sastre and Biron, 2006; Seth et al., 2006; Versteeg et al., 2007).

In cases when the host response is not enough to prevent viral infection, adaptive immunity comes into action, especially during the later stages of viral infection in which infection has already proceeded beyond the first few rounds of viral replication. As the part of adaptive immune response, antigen presenting cells such as B-cells, macrophages and dendritic cells (DCs) recognize and present viral antigens and trigger activation and proliferation of T-helper (Th) cells. The T-helper cells are required for the generation of the humoral response (the synthesis of virus-specific antibodies by B lymphocytes) and the cell-mediated response (recognition and targeted killing of virus infected antigen-displaying altered self-cells by cytotoxic T-cells). The respiratory dendritic cells (rDCs) that resides in lung epithelium acquire the invading virus or its antigens from infected lung cells and become activated (Peebles and Graham, 2011; Tognarelli et al., 2019). The rDCs process viral antigen and subsequently migrate to the draining lymph nodes (DLN), where the presentation of the processed antigen to naïve T cells takes place in the form of MHC/peptide complex (Briciale et al., 2012; Guilliams et al., 2013; Nast and Lambrecht, 2013). Once the T cell receptor (TCR) get engaged with MHC/peptide complex and additional co-stimulatory signals, T cells get activated which in turn result into their rapid proliferation, differentiation and recruitment to the site of virus infection (Larsson et al., 2000; Norbury et al., 2002; Belz et al., 2004). At the site of virus infection or tropism, effector T cells produce antiviral pro-inflammatory cytokines and chemokines, most notably, IFN-\(\gamma\), TNF-\(\alpha\), IL-2, CXCL-9, CXCL-10, and CXCL-11 as well as some cytokote molecules such as perforin and granzyme B (Wherry and Ahmed, 2004). Under the influence of IL-2, IL-15 and other cytokines and chemokines, natural killer (NK) cells also develop and target virus-infected cells using antibody dependent cell-mediated cytotoxicity (ADCC) (Cooper et al., 2009). These effector molecules (cytokines, chemokines and cytotoxic molecules) in a multi-faceted way inhibit viral replication and enhance antigen presentation leading to recruitment of additional immune cells (such as NK-cells) of innate and adaptive system at the site of viral infection for destroying infected epithelial cells (virus-infected altered self-cells) and eliminating virus (Cerwenka et al., 1999; Cerwenka et al., 1999; Roman et al., 2002; Swain et al., 2002; Saha et al., 2010).

The spike protein is the key target for drug and vaccine development for preventing SARS-CoV-2 infection and combating COVID-19 (Du et al., 2009). In particular, the receptor binding domain (RBD), which stretches from residue 330 to 583 within the spike protein, is the most important structural module (Dakal, 2020; Wong et al., 2004). Several researchers demonstrated that spike protein of SARS-CoV-1 plays a key role in eliciting potent T-cell responses and binds with neutralizing-antibodies (Prabakaran et al., 2006; Janice Oh et al., 2012). In particular, the S1 glycoprotein of SARS-CoV-1 spike protein has been found be an important immunodominant epitope which induces a number of neutralizing antibodies (Tian et al., 2020).

Of all structural proteins in SARS-CoV-2, spike proteins are the first to interact with the host cells, and therefore, the initial host cell immune response in COVID-19 patients is expected to be against the exposed antigenic epitopes in the spike proteins. This suggested that there is a strong need for exploration and prediction of antigenic sites and potent cytotoxic T-cell epitopes in spike RBD for rapid developments of novel vaccine candidates and neutralizing antibodies against SARS-CoV-2 infection. In current work, we have employed an integrated approach that comprises identification of potential antigenic sites and antigenic determinants in SARS-CoV-2’s RBD based on its primary sequence and 3D structure. These antigenic sites/determinants can trigger B-cells, T-cells and other immune cells mediated immune response and can explain the complex etiology, pathophysiology and other clinical features of COVID-19 patients. MHC molecules, also called as human leukocyte-associated (HLA) antigens, are cell surface glycoproteins that bind peptide fragments of proteins that either have been synthesized within the cell (class I MHC molecules) or that have been ingested by the cell and proteolytically processed (class II MHC molecules). Potent CTL epitopes against MHC class I/II have also been predicted. These CTL epitopes (peptides) can be used for development of DNA and recombinant vaccines as well as neutralizing antibodies against SARS-CoV-2. We believe that, it is most probable that SARS-CoV-2’s spike RBD-based vaccines will bear fruit in the near future, as they are expected to induce neutralizing antibodies to prevent viral entry into host cells and elicit long-term immune response against COVID-19.
2. Materials and methods

2.1. Retrieval of the SARS-CoV-2 proteome sequence data, sequence alignment and structural modeling

The complete proteome sequence of SARS-CoV-2 (YP_009724389, Wuhan, 12-2019) containing more than 7000 amino acids was downloaded from NCBI virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vss/#/). The protein sequence of SARS-CoV-1 and SARS-CoV-2 coronaviruses were subjected to pair-wise sequence alignment using Clustal Omega using default setting (https://www.ebi.ac.uk/Tools/msa/clustalo/). The structural modeling was done using Chimera ver. 1.10.

2.2. Consensus antigenicity prediction

The consensus antigenicity of SARS-CoV-2 spike RBD was ascertained using AntigenPro (http://scratch.proteomics.ics.uci.edu/) (Magnan et al., 2010). The antigenic propensity was predicted using Antigenic Peptide tool of Universidad Complutense Madrid (imed.med.ucm.es/Tools/antigenic.pl).

2.3. Sequence- and Structure-based epitope prediction

We used the three-dimensional structure model of spike receptor binding domain (PB ID: 6LZG) (Wang et al., 2020) for the prediction of epitopes using ElliPro online server (http://tools.iedb.org/ellipro/) (Ponomarenko et al., 2008). We used an additional tool, namely SVMTrrip, for prediction of linear epitope in spike RBD of SARS-CoV-2 using its primary sequence (https://sysbio.unl.edu/SVMTrrip/prediction.php). The tool employs support vector machine (SVM) in which the tripeptide similarity score and the propensity scores are combined to yield improved predictions (Yao et al., 2012).

2.4. Molecular similarity match

The predicted antigenic epitope sites were matched for molecular similarity with antigens present in other organisms using AntigenDB (https://webs.iiitd.edu.in/raghava/antigendb/epiquery.html) (Ansari et al., 2010). The antigenic sites were used as input in tripeptide format and output was obtained in the form of comprehensive information about experimentally verified epitopes from a wide range of known antigens.

2.5. Vaccine epitope prediction

The predicted antigenic sites in SARS-CoV-2 RBD were assessed for their ability to be processed and presented to cytotoxic T-cells by MHC class I/II proteins. For this EpIJen (http://www.ddg-pharmfac.net/epije n/EpiJen/EpiJen.htm) was used for determining potential T-cell epitopes in SARS-CoV-2 spike RBD domain (Doychimova et al., 2006). The prediction is based on known antigens that have been shown to be presented by MHC class I proteins in various experimental studies and are capable of inducing potent CTL immune responses. In addition, we also performed an analysis for prediction of peptides against the MHC class II alleles using NetMHCpan available in IEDB Analysis Resource of National Institute of Allergy and Infectious Diseases (NIAID), USA (http://tools.iedb.org/mhcii/) (Andreatta et al., 2015). Alleles of this HLA class are normally express on professional antigen-presenting cells (APCs) such as B lymphocytes, dendritic cells, mononuclear phagocytes, endothelial cells and thymic epithelial cells that are crucial in exacerbating immune responses in multi-faceted manner.

3. Results

3.1. SARS-CoV-2 spike RBD domain is highly antigenic

We performed pairwise sequence alignment of spike protein from SARS-CoV-1 and SARS-CoV-2 and found that the similarity between SARS-CoV-1 and SARS-CoV-2 spike receptor binding domain is approximately 70–80% (Dakal, 2020). The SARS-CoV-2 spike RBD was found to be highly antigenic with a predicted probability of antigenicity equal to 0.85, which is even more than the predicted probability of antigenicity of the complete SARS-CoV-2 spike protein (0.72). The average antigenic propensity of the SARS-CoV-2 is 1.0416 (Fig. 1). This showed that SARS-CoV-2 spike RBD possess highly antigenic sites that may activate immune cells and trigger immune response in multiple ways.

3.2. SARS-CoV-2 spike RBD domain possess nine potential antigenic sites

We subjected the 3D structure (PDB file) of SARS-CoV-2 spike RBD to epitope prediction using ElliPro and found that the RBD domain contains nine potential antigenic sites (Table 1). The length of these antigenic sites stretched from four amino acids to twenty three amino acids. The predicted antigenic sites were TNLCS33-336, RISNCVADYSLVNASPG57-374, GVSPTKLDLCLFTYJN381-394, PDDF426-429, SNNLDSKVGNY84-450, LFRRKSNLKFDERD455-470, STEIYQAGSTPCNGVEFNCYFP469-491, YGFQPTNGGVOG595-506 and ELLHAPATVCGBP516-527 (Table 1, Fig. 2). The antigenic sites prediction using sequence-based method (SVMTriP) also resulted in almost similar predictions with prediction of an additional antigenic site LFRRKSNLKFDERD455-470 (Table 2). We performed structural analysis of spike RBD domain using Chimera and found that all the predicted antigenic sites are surface exposed loops of the spike RBD validating the ElliPro prediction (Fig. 2). The positions of predicted antigenic sites using ElliPro and SVMTriP in SARS-CoV-2 spike RBD domain have been also been represented in its sequence aligned with the spike proteins sequence from SARS-CoV-1 (Fig. 3).

3.3. Predicted antigenic sites show molecular similarity with antigens from twelve pathogenic bacterial species, two Plasmodium parasites and influenza virus A

The predicted antigenic sites were then submitted to AntigenDB database for assessing their molecular similarity with other known antigens or antigenic determinants in different microorganisms. While using the whole antigenic site, we could not predict any similar antigenic determinant and as such the antigenic sites predicted in SARS-CoV-2 spike RBD appeared to be unique. However, when the same antigenic sequence was subjected for similarity search in small tripeptide fragments (with three amino acid residues at a time), for example the RISNCVADYSLVNASPG57 antigenic sequence was used as RIS, ISN, SNC, NCV, CVA and so on, we could successfully predict antigenic sites’ similarity with other known antigens in different organisms.

We subjected all nine antigenic sites’ sequence for similarity search in the form of tripeptide fragment and listed all the similar known antigens (Table 3). TNLCS33-336 and PDDF426-429 showed no similarity with any known antigens from other organisms. Rest all other seven predicted antigenic sites showed molecular similarity with known antigens/antigenic determinants. To our surprise, the predicted antigenic sites showed similarity with antigenic determinants from twelve pathogenic bacterial species (Mycobacterium tuberculosis, Mycobacterium leprae, Bacillus anthracis, Borrelia burgdorferi, Clostridium perfringens, Clostridium tetani, Helicobacter pylori, Listeria monocytogenes, Staphylococcus aureus, Streptococcus pyogenes, Vibrio cholera and Yersinia pestis), two malaria parasites (Plasmodium falciparum and Plasmodium knowlesi) and Influenza virus A. The antigenic determinants from both gram+ and gram− bacteria were predicted in spike RBD antigenic sites. All these
microorganisms are well known for their pathogenesis and are associated with life threatening diseases in humans and in some animals as well. In total, out of nine sites, seven sites showed molecular similarity with 54 antigens from twelve pathogenic bacterial species, two malarial parasites and influenza virus A. The predicted molecular similarity between antigenic sites in SARS-CoV-2 and antigenic determinants from pathogenic organisms has three major implications: 1) the human immune system will recognize and activate the immune cells and trigger response in response to SARS-CoV-2 infection in the similar manner as the antigenic determinants could arouse after infection by these fifteen microorganisms (Huang et al., 2020; Qiu et al., 2020; Wang et al., 2020), and 3) the SARS-CoV-2 infection (due to unique antigenic features) is expected to have complex etiology and multiple pathophysiological features (symptoms), possibly multiple organ failure also.

### 3.4. All antigenic sequences in SARS-CoV-2 RBD had similarity with Mycobacterium tuberculosis and Plasmodium falciparum

All seven predicted antigenic sites in SARS-CoV-2 RBD had at least molecular similarity with antigenic determinants from *Mycobacterium tuberculosis* and *Plasmodium falciparum* (Fig. 4) suggesting that COVID-19 patients should have at least symptoms of these two diseases and the same can be useful for identification of SARS-CoV-2 infection. Antigenic sites in SARS-CoV-2 showed similarity with a number of proteins, enzymes, toxins and virulent factors with diverse functional role in pathogens. The role of proteins, enzymes, toxins and virulence factors has been presented here. Diacylglycerolacyltransferase fbpB from *M. Tuberculosis* have high affinity for extracellular matrix protein, fibronectin, which facilitates strong binding of *M. tuberculosis* to macrophages. The 10-kDa chaperonin (groS) of *M. tuberculosis* as found in *M. Leprae* is implicated for its virulence (Roberts et al., 2003). An in vitro study using recombinant *M. tuberculosis* showed that 10-kDa co-chaperonin (cpn10) led to bone weakness and fragility (Corrado et al., 2013). The 6 kDa early secretory antigenic target of *M. tuberculosis* is expected to have complex etiology and multiple pathophysiological features (symptoms), possibly multiple organ failure also.

### Table 1

Structure-based antigenic sites prediction in SARS-CoV-2 spike RBD using ElliPro.

| S. No. | Start | End | Peptide | Number of residues | Score |
|-------|-------|-----|---------|--------------------|-------|
| 1     | 469   | 491 | STEIYQAGSTPCNGVEGFNCYFP | 23     | 0.821 |
| 2     | 516   | 527 | ELHHPATVCGP | 12     | 0.783 |
| 3     | 333   | 336 | TNLC    | 4      | 0.773 |
| 4     | 438   | 450 | SNNLDSKVGNNYN | 13     | 0.689 |
| 5     | 495   | 506 | YGFOPTNGVGYQ | 12     | 0.674 |
| 6     | 381   | 394 | GVSPTKNLDCFTN | 14     | 0.652 |
| 7     | 357   | 374 | RISNCAVADYSVLNNSAF | 18     | 0.57  |
| 8     | 426   | 429 | PDDF    | 4      | 0.505 |
| 9     | 455   | 470 | LFRRSNIKPFERDIST | 16     | 0.50  |

Fig. 1. Antigenic plot for SARS-CoV-2 spike protein showing antigenic propensity for different antigenic sites predicted by Antigenic Peptide tool of Universidad Complutense Madrid (imed.med.ucm.es/Tools/antigenic.pl).

Fig. 2. Structure model of SARS-CoV-2 receptor binding domain showing potential antigenic sites as predicted by ElliPro. All predicted sites are marked with their sequence and position.
involved in persistence and survival of pathogen in host cells and tissues (Kruh-Garcia et al., 2014). Sequence of DNA polymerase 1 of M. tuberculosis was also predicted similar to antigenic sites in SARS-CoV-2 RBD. In nutshell, the antigenic sites in SARS-CoV-2 were similar to those found in pathogenic proteins in Mycobacterium having role in modulating the immune response for escaping immune recognition and ensuring persistence in host body, possibly in latent stage.

The spike RBD was also found to possess antigenic determinants as found in proteins from malarial parasites such as Plasmodium falciparum and Plasmodium knowlesi. The related proteins from malarial parasites were circumsporozoite protein (CSP), circumsporozoite protein-related antigen, malaria protein EXP-1, thrombospondin-related anonymous protein (TRAP), liver stage antigen-1 (LSA-1), liver stage antigen-3 (LSA-3), merozoite surface protein 1 (MSP1), erythrocyte-binding antigen, malaria protein EXP-1, thrombospondin-related anonymous protein, liver stage antigen-3, Plasmodium knowlesi, Plasmodium vivax, and Plasmodium falciparum having role in Plasmodium falciparum. The RESA is released by Plasmodium falciparum inside the RBCs on entry. The RESA migrates to the host cell membrane, where it binds to spectrin; however, the mechanism and type of binding and its pathological consequences are largely unknown yet (Pei et al., 2007).

### 3.5. Influenza virus A having role in virus assembly were predicted in SARS-CoV-2 RBD

Antigenic determinants of two proteins from influenza virus A were predicted similar to SARS-CoV-2 RBD antigenic sites. The proteins are hemagglutinin (HA) (one of the three transmembrane protein of virus) and nucleoprotein (NP) (encapsulated in viral nucleocapsid). While HA is involved in viral assembly at host cell membrane, the viral nucleoproteins helps in incorporation of viral genetic material into newly budded virions (Zheng et al., 2000; Leser and Lamb, 2005). The nucleo-proteins of influenza virus have RNA binding and protein interaction sites that may be important for their host cell functions such as stability and nuclear export of mRNA (Krug, 1993; Qian et al., 1994; Qiu and Krug, 1994) and activation of IRF3, STAT1 and NF-κB dependent pathways (Chien et al., 2004; Krug et al., 2003; Min et al., 2007).

### 3.6. Predicted similar antigenic sequences from bacteria were mainly toxins, virulent factors or other proteins that help invasion and survival of pathogens in host cells and tissues

Two antigenic determinants predicted from Bacillus anthracis were of protective antigen (pagA) and lethal factor. These proteins are two of the three proteins (PA, EF, and LF) that form anthrax toxin (Friebe et al., 2016). Five antigenic determinants predicted from spirochete Borrelia burgdorferi were of flagellar filament core protein (Fla1) and outer surface proteins A (OspA) and C (OspC). Fla1 protein is an immunodominant141Da antigen present in the sera of Lyme-disease patients that allows bacteria to efficiently bore into the host cells for colonization and survival using outer surface proteins (Steere et al., 2004; Neelakanta

---

**Table 2**

Sequence-based antigenic sites prediction in SARS-CoV-2 spike RBD using SVMTriP. The potential sites with score >0.5 have been marked with asterisk * sign.

| S.No. | Site | Sequence | Score (0.5) |
|-------|------|----------|-------------|
| 1     | 571-590 | DTTDAVDRPQ7E1L1D1TPC | 0.466 |
| 2     | 595-614 | VSVT1GTSNSQVAVLYQD | 0.457 |
| 3     | 386-403 | KLNDLDFTNVYADFSVIR | 0.781* |
| 4     | 357-374 | RISNCVADSVLYNSASF | 0.711* |
| 5     | 459-476 | SNNLPERDISTGVAG | 0.581* |
| 6     | 357-372 | RISNCVADSVLSNA | 0.69* |
| 7     | 455-470 | LFRKSNLKPFFERDIST | 0.605* |
| 8     | 388-403 | NDLCFTNVYADSFVIR | 0.544* |
| 9     | 356-369 | KRISNCVADSVLY | 0.622* |
| 10    | 298-311 | DSFVIRGDEVRQIA | 0.496 |
| 11    | 457-470 | RSKNLKFPERDIST | 0.426 |
| 12    | 357-368 | RISNCVADSVLY | 0.727* |
| 13    | 356-365 | KRISNCVADY | 0.496 |

---

**Fig. 3.** Pairwise sequence alignment of SARS-CoV-1 and SARS-CoV-2 showing the position of antigenic sites as predicted using sequence-based (SVMTriP) and structure-based (ElliPro) antigenic sites prediction tools. The site TNLCC533-536 and PDDF36-40 sites are similar in both SARS-CoV-1 and SARS-CoV-2 while other six antigenic sites have sequence divergence. The antigenic site RISNCVADSVLYNSASF357-374 of SARS-CoV-2 has been found to be quite unique as it is not present in SARS-CoV-1 as can be seen in alignment.
Table 3
List of antigenic proteins from different pathogenic microorganisms and virus showing sequence similarity with predicted antigenic sites’ sequence in SARS-CoV-2 spike RBD.

| S. No. | UniProt ID | Protein Name | Protein Description | Organisms |
|--------|------------|---------------|---------------------|-----------|
| 1      | O33084     | ESXB_MYCLE   | ESAT-6-like protein ExsB | Mycobacterium leprae (strain TN) |
| 2      | P01556     | CHTB_VIBCH   | Cholera enterotoxin subunit B | Vibrio cholerae serotype O1 (strain ATCC 39315/El Tor Inaba N16961) |
| 3      | P01558     | ELTB_CLOP    | Heat-labile enterotoxin B chain | Clostridium perfringens |
| 4      | P02893     | CSP_PLAFA    | Circumporozoite protein | Plasmodium falciparum |
| 5      | P02894     | CSP_PLAKH    | Circumporozoite protein | Plasmodium knowlesi (strain H) |
| 6      | P03438     | HEMA_I000X   | Hemagglutinin | Influenza A virus (strain A/X-31 H3N2) |
| 7      | P03449     | HEMA_J71A1   | Hemagglutinin | Influenza A virus (strain A/Memphis/1/1971 H3N2) |
| 8      | P03452     | HEMA_J34A1   | Hemagglutinin | Influenza A virus (strain A/Puerto Rico/8/1934 H1N1) |
| 9      | P03453     | HEMA_J77AB   | Hemagglutinin | Influenza A virus (strain A/USSR/90/1977 H1N1) |
| 10     | P03459     | HEMA_J34A0   | Hemagglutinin | Influenza A virus (strain A/Fowl plague virus/Rostock/B/1934 H7N1) |
| 11     | P03466     | NCAP_J34A1   | Nucleoprotein | Influenza A virus (strain A/Puerto Rico/8/1934 H1N1) |
| 12     | P04923     | CRA_PLAFA    | Circumporozoite protein-related antigen | Plasmodium falciparum |
| 13     | P04926     | EXP1_PLAFA   | Malaria protein EXP-1 | Plasmodium falciparum |
| 14     | P04934     | MSP1_PLAFC   | Merosome surface protein 1 | Plasmodium falciparum (isolate Camp/Malaysia) |
| 15     | P04958     | TETX_CLOTE  | Tetanus toxin | Clostridium tetani (strain Massachusetts/B8) |
| 16     | P08089     | M68_STRPY    | M protein, serotype 6 | Streptococcus pyogenes |
| 17     | P09621     | CH10_MYCTU   | 10 kDa chaperonin | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 18     | P09621     | CH10_MYCTO   | 10 kDa chaperonin | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 19     | P0A0L2     | ETA1_STAAU   | Enterotoxin type A | Staphylococcus aureus |
| 20     | P0A4V2     | ABSA_MYCTU   | Diacylglycerolacyltransferase | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 21     | P0A4V2     | ABSA_MYCTO   | Diacylglycerolacyltransferase | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 22     | P0A550     | DPO1_MYCTU   | DNA polymerase I | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 23     | P0A550     | DPO1_MYCTO   | DNA polymerase I | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 24     | P0A564     | ESXA_MYCTU   | 6 kDa early secretory antigen target | Mycobacterium tuberculosis (strain ATCC 25,618 / H37Rv) |
| 25     | P0A564     | ESXA_MYCTO   | 6 kDa early secretory antigen tar | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 26     | P0A5J0     | LPQH_MYCTU   | Lipoprotein LpqH | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 27     | P0A5J0     | LPQH_MYCTO   | Lipoprotein LpqH | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 28     | P0A5Q4     | MP64_MYCTU   | Immunogenic protein MPT64 | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 29     | P0A5Q4     | MP64_MYCTO   | Immunogenic protein MPT64 | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 30     | P11089     | FLA1_BORBU   | Flagellar filament 41 kDa core protein | Borrelia burgdorferi (strain ATCC 35210/B31/CIP 102532/DSM 4680) |
| 31     | P13423     | PAG_BACAN    | Protective antigen | Bacillus anthracis |
| 32     | P13830     | RESA_PLAFF   | Ring-Infected erythrocyte surface antigen | Plasmodium falciparum (isolate FC27/Papua New Guinea) |
| 33     | P14013     | OSPA_BORBU   | Outer surface protein A | Borrelia burgdorferi (strain ATCC 35210/B31/CIP 102532/DSM 4680) |
| 34     | P14013     | OSPA_BORBN   | Outer surface protein A | Borrelia burgdorferi (strain N40) |
| 35     | P14013     | OSPA_BOBZ    | Outer surface protein A | Borrelia burgdorferi (strain 257) |
| 36     | P14916     | URE23_HELPC  | Urease subunit alpha | Helicobacter pylori (strain ATCC 700929/26695) (Campylobacter pylori) |
| 37     | P15712     | PSTS1_MYCTU  | Phosphate-binding protein PsTs 1 | Mycobacterium tuberculosis (strain ATCC 25618/H37Rv) |
| 38     | P15712     | PSTS1_MYCTO  | Phosphate-binding protein PsTs 1 | Mycobacterium tuberculosis (strain CDC 1551/Oshkosh) |
| 39     | P15917     | LEP_BACAN    | Lethal factor | Bacillus anthracis |
| 40     | P16893     | TRAP_PLAFA   | Thrombospondin-related anonymous protein | Plasmodium falciparum |
| 41     | P19214     | EBA1_PLAFC   | Erythrocyte-binding antigen 175 | Plasmodium falciparum (isolate Camp/Malaysia) |
| 42     | P21177     | P60_LISMO    | Probable endopeptidase p60 | Listeria monocytogenes serovar 1/2a (strain ATCC BAA-679/EGD-e) |
| 43     | P23024     | TCPA_VIBCL   | Toxin coregulated pilin | Vibrio cholera |
| 44     | P24301     | CH10_MYCLE   | 10 kDa chaperonin | Mycobacterium leprae (strain TN) |
| 45     | P25893     | CAF1_YERPE   | F1 capsule antigen | Yersinia pestis |
| 46     | P26948     | CAF1_YERPE   | F1 capsule antigen | Mycobacterium leprae (strain TN) |
| 47     | P31951     | ABSB_MYCLE   | Diacylglycerolacyltransferase | Mycobacterium leprae (strain TN) |
| 48     | P31952     | ABSB_MYCTA   | Diacylglycerolacyltransferase | (continued on next page)
The antigenic sequences predicted from *Clostridium perfringens* and *Clostridium tetani* were also of the toxin proteins such as heat-labile enterotoxin B chain and tetanus toxin. Similarly, some other predicted antigenic determinants were also of well known toxins and virulent factor such as enterotoxin A (from *Staphylococcus aureus*), M protein (from *Streptococcus pyogenes*), cholera enterotoxin subunit B and toxin coregulated pilin (from *Vibrio cholerae*). The 30kDa urease subunit alpha (ureA) from *Helicobacter pylori* acts as virulence factor and cause infection in stomach via host-pathogen interaction (Schoep et al., 2010) and mediate immune response in humans (Schoep et al., 2010). The endopeptidase p60 (Lm-p60) of *Listeria monocytogenes* whose sequence was also predicted in similar search is a highly conserved carbohydrate binding module which can be engineered for binding to peptidoglycans with high affinity (Yu et al., 2016).

**3.7 Five peptide sequences predicted as novel vaccine candidates for COVID-19 treatment**

The spike RBD of SARS-CoV-2 was subjected for T-cell epitope prediction using EpIen which predicts peptides molecules that can be recognized by human T-cells after presentation by MHC class I proteins. In total fourteen peptide sequences were recognized as possible epitopes for T-cells (predicted against MHC class I) and only eight peptide sequences were predicted to be overlapping with the antigenic sites predicted (Table 4). However, the predicted IC50 value for six peptides was very high and thus these peptides failed to qualify the criteria of becoming a potent vaccine candidate. Finally, two peptides (KVGGNYNYL<sup>444-452</sup> and SVLYNSASF<sup>366-374</sup>) with low predicted IC50 value of 29.38 and 31.26 nM, respectively were considered suitable for designing DNA vaccine and recombinant vaccine and other vaccine type against COVID-19. Besides this, the receptor binding domain of the spike protein in SARS-CoV-2 has a stretch of sequence (TEIYQAGSTPCNGVEGF<sup>470-486</sup>) showing mismatch with SARS-CoV-1 and the same stretch has been found to be one of the antigenic sites and harboring peptide sequence for a potent vaccine candidate identified in the current study (Table 4). We have done an additional analysis in which prediction was done against the HLA class II alleles using NetMHCpan available in IEDB Analysis Resource of National Institute of Allergy and Infectious Diseases (NIH), USA (http://tools.iedb.org/mhcII/). Alleles of this HLA class are normally found on professional antigen-presenting cells (APCs) such as B lymphocytes, dendritic cells, mononuclear phagocytes, endothelial cells and thymic epithelial cells. These cells are important in exacerbating immune responses in different ways. Herein, we predicted three potential peptides in SARS-CoV-2 RBD against MHC class II alleles (Table 5). These peptides are KRISNCVADYSVLYN<sup>356-370</sup> (IC50 = 2.6 nM), DLCFTNVYADSFVI<sup>389-402</sup> (IC50 = 3.8 nM), and YRVVVSFELLHA<sup>508-520</sup> (IC50 = 3.8 nM).
context to auto-immune pathogenesis, molecular similarity has been well presented and similarities (mainly in antigens) between two (or more) organisms. The molecular similarity of spike RBD with antigenic determinants (predicted as antigenic determinants in current study) can activate T-cells and can induce immune response (Hemmer et al., 2000). Additionally, other researchers demonstrated how an antigenic tripeptide motif (CVADYSVL, 356-374) can be traced only as small tripeptide motifs and this makes their antigenic patches from pathogenic microorganisms in SARS-CoV-2 specially, antigenic determinants those are unique to SARS-CoV-2. The more pathogenic than any other previously known coronavirus. Especially, antigenic determinants those are unique to SARS-CoV-2. The antigenic patches from pathogenic microorganisms in SARS-CoV-2 could be traced only as small tripeptide motifs and this makes their antigenic determination as found in different pathogenic microorganisms is expected to induce exuberant innate and adaptive immune responses leading to excessive secretion of cytokines that is expected to adversely affect several vital organs leading to multi-organ failure in COVID-19 patients (Fig. 5). Such an exacerbated innate and adaptive immune response is attributed to hyper-activation of B-cells, T-cells (both CTL and Th-cells), DCs, NK-cells and macrophage/monocyte lineage cells (Fig. 5). Since, T-cell play critical role in controlling immune response and one can expect deregulated immune response and dysregulated cytokine secretion in aged and immune-compromised patients who have less number of T-cells in their body (Kim et al., 2007; Palm and Medzhitov, 2007). There is also a correlation between the secretion of some cytokines and chemokines, such as IL-6, IL-8, and MCP-1 and IP-10, with higher mortality and severity of disease (Reghunathan et al., 2005).

We predicted nine potential surface exposed antigenic sites in SARS-CoV-2 spike RBD. We also predicted CTL epitopes that can be presented with MHC class I and class II and found five peptides in spike RBD of SARS-CoV-2 (KVVGNNYL, 344-345 and SYLVNNSF, 366-374) against MHC class I; KRISNCVADYSVLYN, 356-370, DLFCTNYYADSVFI, 389-407, and YRVVVLFSFELLA, 508-520 against MHC class II) with low predicted IC50 values, which confirms their suitability for vaccine design against COVID-19. Some researchers used RVDFCGKGY peptide (CTL epitope) to design a vaccine against SARS-CoV-1 and experimental trials showed the vaccine to be effective as well on different animal models (Choy et al., 2004). Finally, the results obtained in current study are also in congruence with previous experimental studies in which SARS-CoV-specific CD4+ and CD8+ T cell epitopes were found in C57BL/6 and BALB/C mice (Zhi et al., 2005; Huang et al., 2007; Zhao et al., 2010).

Presence of antigenic sites in SARS-CoV-2 with similarity with antigenic determinants found in pathogenic bacterial, malarial and viral species is seriously alarming as their presence makes the SARS-CoV-2 more pathogenic than any other previously known coronavirus. Especially, antigenic determinants those are unique to SARS-CoV-2. The antigenic patches from pathogenic microorganisms in SARS-CoV-2 could be traced only as small tripeptide motifs and this makes their antigenic origin uncertain. However, most of the antigenic determinants found in pathogenic microorganisms were related to antigenic proteins having established function as toxin and proven role in virulence and pathogenicity. Besides this, other antigenic determinants found similar to Mycobacterium antigens having role in persistence and survival of pathogen in host cells. Some antigenic determinants from HA and NP of influenza virus A were also in antigenic sites of SARS-CoV-2 and these antigens from influenza virus have role in assembly of newly budded virions in case of influenza virus A. Due to the presence of antigenic sites with molecular similarity with antigens from different pathogenic organisms, parasites and virus, it can be ruled out that COVID-19 patients may possibly suffer multi-organ failure as these pathogenic antigens are well known to cause serious damages to different vital organs of human body such as liver, kidney, blood, stomach, heart and bones. This is also clearly evident from complex etiology and pathophysiological outcomes (clinical features or symptoms) as observed in COVID-19 patients. Common in all seven antigenic sites predicted in SARS-CoV-2 was the presence of antigenic determinants from Mycobacterium and Plasmodium antigens suggesting that antimalarial and anti-TB drugs and vaccines could be a good treatment options for COVID-19. Besides this, anti-leprosy, anti-lyme, anti-plague, anti-anthrax drugs/vaccine etc are also expected to be beneficial in COVID-19 treatment. Moreover, individuals previously immunized/vaccinated or had previous history of malaria, do not experience severe complications due to COVID-19.

4. Discussion

In brief, in current work, we have made an attempt to ascertain molecular similarity of antigenic sites predicted in SARS-CoV-2 spike protein with other proteins/antigens in other organisms. Molecular similarity can be defined as the theoretical explanation for sequence similarities (mainly in antigens) between two (or more) organisms. The concept of molecular or antigen similarity has been well presented and entrusted in literature in context to autoimmune disorders (Fujinami et al., 2006; Cusick et al., 2012; Pontes-de-Carvalho et al., 2013). In context to auto-immune pathogenesis, molecular similarity has been implicated between pathogenic organisms and humans self-antigens which results in generation of cross-reactive T-cells that targets human’s self-cells (Chodisetti et al., 2012). In context to current manuscript, the molecular similarity has been observed between antigenic sites in SARS-CoV-2 spike protein and similar antigenic determinants in fifteen microorganisms, including bacteria, parasites and viruses. We believe that the presence of tripeptide sequence(s) (displaying molecular similarity with highly potent antigenic determinants as present in fifteen pathogenic microorganisms) in the antigenic sites of SARS-CoV-2 spike proteins are sufficient to activate host immune cells (T or B cells or other immune cells) in similar as the fifteen microorganisms can do. Some researchers showed that even a tripeptide motif (predicted as antigenic determinants in current study) can activate CD4+ T-cells and can induce immune response (Hemmer et al., 2000). Additionally, other researchers demonstrated how an antigenic tripeptide motif in the amino acid sequence of a protein in pathogens can also differentially activate one of the two immune cells, either B-cell or T-cell (Yao et al., 2012). The molecular similarity of spike RBD with antigenic determinants as found in different pathogenic microorganisms is expected to induce exuberant innate and adaptive immune responses leading to excessive secretion of cytokines that is expected to adversely affect several vital organs leading to multi-organ failure in COVID-19 patients (Fig. 5). Such an exacerbated innate and adaptive immune response is attributed to hyper-activation of B-cells, T-cells (both CTL and Th-cells), DCs, NK-cells and macrophage/monocyte lineage cells (Fig. 5). Since, T-cell play critical role in controlling immune response and one can expect deregulated immune response and dysregulated cytokine secretion in aged and immune-compromised patients who have less number of T-cells in their body (Kim et al., 2007; Palm and Medzhitov, 2007). There is also a correlation between the secretion of some cytokines and chemokines, such as IL-6, IL-8, and MCP-1 and IP-10, with higher mortality and severity of disease (Reghunathan et al., 2005).

We predicted nine potential surface exposed antigenic sites in SARS-CoV-2 spike RBD. We also predicted CTL epitopes that can be presented with MHC class I and class II and found five peptides in spike RBD of SARS-CoV-2 (KVVGNNYL, 344-345 and SYLVNNSF, 366-374) against MHC class I; KRISNCVADYSVLYN, 356-370, DLFCTNYYADSVFI, 389-407, and YRVVVLFSFELLA, 508-520 against MHC class II) with low predicted IC50 values, which confirms their suitability for vaccine design against COVID-19. Some researchers used RVDFCGKGY peptide (CTL epitope) to design a vaccine against SARS-CoV-1 and experimental trials showed the vaccine to be effective as well on different animal models (Choy et al., 2004). Finally, the results obtained in current study are also in congruence with previous experimental studies in which SARS-CoV-specific CD4+ and CD8+ T cell epitopes were found in C57BL/6 and BALB/C mice (Zhi et al., 2005; Huang et al., 2007; Zhao et al., 2010).

Presence of antigenic sites in SARS-CoV-2 with similarity with antigenic determinants found in pathogenic bacterial, malarial and viral species is seriously alarming as their presence makes the SARS-CoV-2 more pathogenic than any other previously known coronavirus. Especially, antigenic determinants those are unique to SARS-CoV-2. The antigenic patches from pathogenic microorganisms in SARS-CoV-2 could be traced only as small tripeptide motifs and this makes their antigenic origin uncertain. However, most of the antigenic determinants found in pathogenic microorganisms were related to antigenic proteins having established function as toxin and proven role in virulence and pathogenicity. Besides this, other antigenic determinants found similar to Mycobacterium antigens having role in persistence and survival of pathogen in host cells. Some antigenic determinants from HA and NP of influenza virus A were also in antigenic sites of SARS-CoV-2 and these antigens from influenza virus have role in assembly of newly budded virions in case of influenza virus A. Due to the presence of antigenic sites with molecular similarity with antigens from different pathogenic organisms, parasites and virus, it can be ruled out that COVID-19 patients may possibly suffer multi-organ failure as these pathogenic antigens are well known to cause serious damages to different vital organs of human body such as liver, kidney, blood, stomach, heart and bones. This is also clearly evident from complex etiology and pathophysiological outcomes (clinical features or symptoms) as observed in COVID-19 patients. Common in all seven antigenic sites predicted in SARS-CoV-2 was the presence of antigenic determinants from Mycobacterium and Plasmodium antigens suggesting that antimalarial and anti-TB drugs and vaccines could be a good treatment options for COVID-19. Besides this, anti-leprosy, anti-lyme, anti-plague, anti-anthrax drugs/vaccine etc are also expected to be beneficial in COVID-19 treatment. Moreover, individuals previously immunized/vaccinated or had previous history of malaria, do not experience severe complications due to COVID-19.
tuberculosis or other disease caused by fifteen microorganisms are expected to display a considerable degree of resistance against SARS-CoV-2 infection. The possible explanation is that the memory B or T cells previously generated by the microorganisms (in context here) would get activated again upon SARS-CoV-2 infection (long-lived immunity to reinfection) because of the similar antigenic specificity and due to presence of common antigenic determinants in both. For some of the pathogenic antigens, whose antigenic determinants were predicted similar to SARS-CoV-2’s antigenic sites, very limited information was available in literature regarding their function.

As such, nothing can be explicitly stated regarding the presence of antigenic determinants as found in this current study. We speculate that viruses may be doing this as a genetic trick for tackling genetic constraints imposed by host cells to ensure their persistence and survival. There are two explanations for high divergence between spike RBD in SARS-CoV-1 and SARS-CoV-2. The SARS-CoV-2 might have changed or switched the molecular composition of its spike protein using a process called as antigenic variation (Smith, 2004; Chibo and Birch, 2006; de Jong et al., 2007). Antigenic variations are genetic tricks using which some viruses, such as influenza virus and coronavirus etc., evade protective immune system by altering their immunodominant epitopes that otherwise could be recognized by the host adaptive immune system leading to their elimination (Bidokhti et al., 2013; Lewis et al., 2014). Antigenic variations are brought up by periodic and random genetic mutations, generally in viral surface antigens, so as to temporarily camouflage host immune cells and to prevent clearance (Smith, 2004; de Jong et al., 2007). Antigenic variation that occurs in the HA and NA antigens of influenza virus A are known. The antigenic alteration can occur in two different ways: 1) genetic drift that causes subtle changes in amino acids, and 2) antigenic shift that causes major alterations in the antigenic properties of the protein (Smith, 2004; de Jong et al., 2007; Ren et al., 2015). In latter case, the acquisition of new antigenic determinants in virus would render virus no longer identifiable (similarly as “eclipsed antigens” do) by host immune cells and neutralizing antibodies (Smith, 2004; de Jong et al., 2007). Such antigenic-shifted strains of virus generate periodically when genes encoding structural proteins are acquired from viruses that infect animal hosts. These antigenic-shifted strains are known to cause global pandemic and recurring acute infections (Smith, 2004; de Jong et al., 2007).

Currently, computational biology approaches for epitope prediction, identification and analysis are well-developed and have been proved highly successful to predict & identify both weak and strong antibody epitopes, some of them are often experimentally ignored (Jespersen et al., 2017). Antigenic properties of spike glycoprotein of SARS-CoV-2, especially of the receptor binding domain (RBD), were well appreciated by many experimental researchers including myself in current study (Baruah and Bose, 2020; Lucchese, 2020). Most of the vaccines such as virus vector and protein subunit vaccine developed or under clinical trails have been developed using full length spike protein or RBD of spike protein of SARS-CoV-2. The identified sequence in current study are part of all vaccine under clinical or pre-clinical development. (Baruah and Bose, 2020). The strategy to find antigenic sites in spike protein of SARS-CoV-2 is also known (Ren et al., 2003). Some new antibody epitopes have also been found in SARS-CoV-2 spike protein that dominate the antigenicity of spike protein in SARS-CoV-2 as compared to other coronaviruses (Zheng and Song, 2020). Potential antigenic cross-reactivity of SARS-CoV-2 with dengue virus has also been observed (Lustig et al., 2020).

This study provides the first evidence in favor of antigenic variations in spike RBD of SARS-CoV-2 that succour virus to undergo adaptive evolution in order to infect humans for their survival and persistence. These antigenic variations not only explain the complex immunological aspects, etiology and pathophysiology of the disease but also suggest different therapeutics (anti-malarial, anti-TB, anti-leprosy, anti-plague, anti-lyme, anti-anthrax, anti-cholera etc.), including drugs, medicines, antibodies and vaccines, for their promising role in inactivating SARS-CoV-2, which is thought to mutate quickly. Besides this, our body also requires B-cells, CD8+ and CD4+ T-helper cells and other immune cells based innate and adaptive immunity and antibodies for specifically targeting infectious cells (Janice Oh et al., 2012; Choy et al., 2004). We envisage that several lines of research endeavors are still required towards understanding the multifaceted mechanisms of immunomodulation, host-virus interaction and SARS-CoV-2 infection.

Author contributions
TCD conceived the idea and designed the experimental strategy and designed methodology, performed the analyses, wrote the manuscript, revised the manuscript.

Funding
The work was conducted in absence of any funding source.

![Fig. 5. The study explains the complex etiology and pathophysiology of COVID-19 and sheds light on implications and future prospects.](image-url)
Saha, B., Jyothi Prasanna, S., Chandrasekar, S., Nandi, D., 2010. Gene modulation and Roberts, M.M., Coker, A.R., Fossati, G., Mascagni, P., Coates, A.R.M., Wood, S.P., 2003. Ren, Y., Zhou, Z., Liu, J., Lin, L., Li, S., Wang, H., Xia, J., Zhao, Z., Wen, J., Zhou, C., Ren, L., Zhang, Y., Li, J., Xiao, Y., Zhang, J., Wang, Y., Chen, L., Paranhos-Baccal Qian, X.Y., Alonso-Caplen, F., Krug, R.M., 1994. Two functional domains of the influenza Palm, N.W., Medzhitov, R., 2007. Not so fast: adaptive suppression of innate immunity. Peebles, R.S., Graham, B.S., 2011. Viruses, dendritic cells and the lung. Respir. Res. 2, Peirs, P., Lefevre, P., Boari, S., Wang, X.-M., Denis, O., Brailant, M., et al., 2005. Mycobacterium tuberculosis with disruption in genes encoding the phosphate binding proteins PstS1 and PstS2 is deficient in phosphate uptake and demonstrates reduced in vivo virulence. J. Infect. Immun. 73, 1898–1902. https://doi.org/10.1128/IAI.73.1898-1902.2005.
Ponnomarenko, J., Bui, H.-H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. EliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Biol. 6, 27. https://doi.org/10.1186/1741-7007-6-27.
Pontes-de-Carvalho, L., Mengel, J., Figueredo C.A., Social Changes, Asthma and Allergy in Latin America (SCAALA’s) Study Group, Alcantara-Neves N.M. (2013). Antigen mimicry between infectious agents and natural or environmental antigens may lead to long-term regulation of inflammation. Front. Immunol. 4:10.3141/131008.

Prabakaran, P., Gan, J., Feng, Y., Zhu, Z., Choudhry, V., Xiao, X., Ji, X., Dimitrov, D.S., Seth, R.B., Sun, L., Chen, Z.J., 2006. Antiviral innate immunity pathways. Cell Res. 16 (4), 169–179. 10.1038/sj.crr.1s06985.

Qiao, H., Wu, J., Hong, L., Luo, Y., Song, Q., Chen, D., 2020. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. Lancet 20 (6), 689–696. https://doi.org/10.1016/S0140-6736(20)30644-3.

Regnathun, R., Jayapal, M., Huu L-Y., Chng H-H., Tai D, Leong B.P., et al. (2005). Expression profile of immune response genes in patients with severe acute respiratory syndrome. BMC Immunol. 6:2.
Ren, L., Yang, Z., Yi, L., Xiao, X., Zhang, J., Wang, Y., Chen, L., Paranhas-Baccala, G., Wang, J., 2015. Genetic drift of human coronavirus OC43 spike gene during adaptive evolution. Sci. Rep. 5 (1) https://doi.org/10.1038/srep11451.
Ren, Y., Zhou, Z., Liu, J., Lin, L., Li, S., Wang, H., Xin, J., Ji, Zhao, Z., Wen, J., Zhou, C., Wang, Y., Jin, J., Xu, N., Liu, S., 2003. A strategy for searching antigenic regions in the SARS-CoV spike protein. Genomics Proteomics Bioinformatics. 1 (3), 207–215. https://doi.org/10.11612/gpb.2003.101026-X.
Roberts, M.M., Coker, A.R., Fossati, G., Mascagni, P., Coates, A.R.M., Wood, S.P., 2003. Mycobacterium tuberculosis tozanii is a new pathovar and tuberculose-like pathogenicity index. J. Bacteriol. 185, 4172–4175. 10.1128/JB.185.11.4172-4175.2003.

Schook, S.N., Chandra, N.V., Sun, Y., Irie, T., Kitagawa, Y., Gotob, B., et al., 2017. Induction of necroptotic cell death by viral activation of the RIG-I or STING pathway. Cell Death Differ. 24, 615–625. https://doi.org/10.1038/cdd.2016.153.
Schoep, T.D., Furiljula, A., Good, F., Lu, W., Himbeck, R.P., Schwan, C., Choi, S.S., Berg, D.E., Mittl, P.R.E., Benghezal, M., Marshall, B.J., Ahmed, N., 2010. Surface properties of Helicobacter pylori urease complex are essential for persistence. PLoS Pathog. 6 (11), e1001542.
Seth, R.B., Sun, L., Chen, Z.J., 2006. Antiviral innate immunity pathways. Cell Res. 16 (2), 143–147.
Smith, D.J., 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305 (5682), 371–376. 10.1126/science.1097211.
Spiegel, M., Schneider, K., Weber, F., Weidmann, M., Huffer, F.T., 2006. Interaction of severe acute respiratory syndrome-associated coronavirus with dendritic cells. J. Gen. Virol. 87, 1953–1960.
Spiegel, M., Weber, F., 2006. Inhibition of cytokine gene expression and induction of chemokine genes in non-lymphatic cells infected with SARS coronavirus. Virol. J. 3, 17.
Steele, A.C., Coburn, J., Glickstein, L., 2004. The emergence of Lyme disease. J. Clin. Invest. 113 (8), 1093–1101.
Swain, S.L., Agerwala, J.N., Brown, D.M., Roman, E., 2002. Regulation of memory CD4 T cell generation, localization and persistence. Adv. Exp. Med. Biol. 512, 113–120. https://doi.org/10.1016/S0065-2154(02)00317-4.

Wang, J., Yin, J., Xu, N., Liu, S., 2003. A strategy for searching antigenic regions in the SARS-CoV spike protein. Genomics Proteomics Bioinformatics. 1 (3), 207–215. 10.1006/gpb.2003.101026-X.
Western, B., Glass, S., Norbury, C.C., Malide, D., Gibbs, J.S., Bennink, J.R., Haldar, K., Gratzer, W., et al., 2007. The novel coronavirus SARS-CoV-2. Int. J. Biol. Sci. 16 (10), 1718–1723.