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.

Journal of Molecular Structure 1240 (2021) 130534

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: www.elsevier.com/locate/molstr

In silico identification of RBD subdomain of spike protein from Pro^{322}-Thr^{581} for applications in vaccine development against SARS-CoV2

Nataraj S Pagadala[a,*], Abdolamir Landi[b], Paramahamsa Maturu[c], Jack Tuszynski[b]

[a] Carnegie Mellon University, Department of Chemistry, Mellon Institute Bldg. 4400 Fifth Avenue, Pittsburgh, PA 15213-2683, United States
[b] Li Ka Shing Applied Virology Institute, Department of Medical Microbiology and Immunology, University of Alberta, Edmonton, AB, Canada
[c] Department of Pediatrics, Baylor College of Medicine, Houston, TX, United States

ARTICLE INFO

Article history:
Received 11 January 2021
Revised 17 April 2021
Accepted 19 April 2021
Available online 30 April 2021

Keywords:
SARS-COV 2
Spike protein
ACE2

ABSTRACT

The three-dimensional hybrid structures of coronavirus spike proteins including the C-terminal sequence and receptor binding motif (RBM) was remodeled and energy minimized. Further, protein-protein docking show that Receptor Binding Domain (RBD) of SARS-CoV 2 Lys^{457}-Pro^{490} bind on the surface of ACE2 receptor near N-terminal helices to form host-pathogen attachment. In this binding interface, SARS-CoV 2 shows a tight network of hydrogen bonds than other spike proteins from BtRsRaTG13-CoV, SARS-CoV, BtRsBeta-CoV, BtRsCoV-related, Pangolin-CoV (PCoV), human-CoV (hCoV), MERS-CoV (MCoV), Avian-CoV (ACoV) and PEDV1-CoV. Further studies show that subdomains from SARS-CoV 2 RBD Pro^{322}-Thr^{581}, SARS-CoV RBD Pro^{309}-Pro^{575}, BtRsRaTG13 RBD Thr^{581}-Thr^{323}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, BtRsCoV-related Arg^{306}-Pro^{575} and PCoV RBD Gln^{319}-Ser^{589} show binding conformations with ACE2 like their full-length structures of spike proteins. In addition, the subdomains MCoV RBD Gly^{372}-Val^{616}, ACoV RBD Gly^{372}-Val^{616} and PEDV1-CoV RBD Ala^{315}-Tyr^{675} also binds on the surface of ACE2 similar to their full-length spike proteins. The B-Cell epitope mapping also identified main antigenic determinants predicting that these nine subdomains are highly useful in recombinant vaccine development in inducing cross neutralizing antibodies against SARS-CoV 2 spike protein and inhibits its attachment with ACE2.

1. Introduction

Coronaviruses (CoVs) are the largest group of RNA viruses, which belong to the genus coronavirus, the family coronaviridae, and the order nidovirales. The genome of CoV is a single stranded positive-sense RNA (+ssRNA) with a size of 27–32 kb [1,2]. There are 4 structural proteins including spike (S), envelope (E), membrane (M), and nucleocapsid (N) in coronaviruses which are embedded into envelope, contributing to the crown-like feature of viral particles [3]. This family of viruses include human coronaviruses (hCoV)−229E, hCoV-OC43, hCoV HKU-1, and hCoV NL63, causing mild upper respiratory infection, known as common cold and are constantly circulating among 70% of the human population [4]. In contrast, two fatal coronaviruses, severe acute respiratory syndrome SARS-CoV and Middle East respiratory syndrome MERS-CoV that are causing severe upper and lower respiratory diseases leading to fatal pneumonia are transmitted from animals to humans [5]. SARS-CoV, which resides in Chinese horseshoe bats as a natural reservoir, was associated with 8096 cases and 774 deaths globally in 2002–3, started in Guangdong province, China, [6]. The virus had been transmitted to humans through civet cats and raccoon dogs that were consumed as food and sold in Chinese wet markets [7]. Due to lack of specific antivirals or approved vaccines for the SARS-CoV in 2002–3, conventional measures had been taken to stop the spread of the disease, including travel restrictions and patient isolation. MERS-CoV infection that was first reported in Saudi Arabia in 2012, was mainly spread in the Middle East and later self-controlled with ~2000 infected cases with a fatality rate of ~35%. Both SARS and MERS had limited spread and are not a health concern anymore. In 2019 December, a novel form

Abbreviations: SARS-CoV 2, Severe Acute Respiratory Syndrome Coronavirus-2; BtRsRaTG13-CoV, Bat Coronavirus; SARS-CoV, Severe Acute Respiratory Syndrome Coronavirus; BtRsBeta-CoV, Bat Respiratory syndrome Beta Coronavirus; BtRsCoV-related, Bat Respiratory syndrome Coronavirus Related.; PCoV, Pangolin coronavirus; hCoV, Human Coronavirus; MCoV, MERS Coronavirus; ACoV, Avian Coronavirus; PEDV1-CoV, Porcine epidemic diarrhea virus.

* Corresponding author.
E-mail address: nattu251@gmail.com (N.S. Pagadala).

https://doi.org/10.1016/j.molstruc.2021.130534

of coronavirus named as SARS-CoV 2 emerged in China at Wuhan city, causing severe pandemic affecting the global public health resulting in progressive respiratory failure due to alveolar damage and death [8–10]. According to daily report by Centre for Systems Science and Engineering at Johns Hopkins University, as of February 25, 2021, there have been over 112 million confirmed SARS-CoV 2 infections with more than 2,5 million global fatal cases, exceeding any former epidemics by coronaviruses.

The infections by coronavirus are mainly due to the process of receptor binding by the spike membrane glycoprotein (S protein) in mediating membrane fusion [11], resulting in the high virulence of SARS-CoV 2. The mechanism of spike-mediated membrane fusion, which is similar to that of class I virus fusion proteins have been previously studied in murine coronavirus (mouse hepatitis virus; MHV) [12,13]. This mechanism of membrane fusion is due to attachment of S1 subunit of spike protein to the cellular receptor, facilitating viral attachment to the surface of target cells. Similarly, studies have shown angiotensin converting enzyme 2 (ACE2), which regulates blood pressure acts as a cellular receptor for viral entry by SARS-CoV and hCoV NL63 [14–18], where cellular serine protease TMPRSS2 is used for cleavage and conformational changes of S protein, called priming [15,19–21]. Current studies on SARS-CoV 2 have also demonstrated that ACE2 receptor is utilized as the entry point in Chinese horseshoe bats, civet, swine, but not in mouse [10]. These observations clearly reveal that ACE2 plays a key role in SARS-CoV spread. As shown for SARS-CoV, the virus binds to the peptidase domain of ACE2 and both spike and ACE2 (primarily expressed on pulmonary epithelium) are cleaved by cellular proteases such as TMPRSS2. This results in conformational change in spike and allows it to insert its S1 subunit into the membrane, facilitating virus entry. During the entry process, spike cleavage is critical for virus entry and blocking the cleavage would reduce viral entry [22].

Comparative studies on the viral sequences have demonstrated a similarity of ~80% between SARS-CoV 2 and SARS-CoV with major difference to be in three regions. These differences exist in open reading frame (ORF) 1a/b, ORF8, expressing a protein involved in immune evasion, and more importantly spike region. This similarity is even higher with BtRsRaTG13-CoV, which is 96% identical to SARS-CoV 2 in its amino acid (AA) sequence. However, since the mismatch is localized at Receptor Binding Domain (RBD) of S protein, BtRsRaTG13 CoV does not infect humans due to lack of binding to ACE2. Conversely, the RBD domain of S is highly identical to that of another Bat coronavirus detected in Pangolin; however, pangolin CoV does not infect humans either, because of significant differences in other parts of spike protein [23]. Accordingly, it has been hypothesized by other authors, that during a cross-species recombination, the RBD in BtRsRaTG13-CoV might have been substituted by that of PCoV to produce SARS-CoV 2 that can infect humans. The other unique feature of SARS-CoV 2 is the cleavage domain between S1 and S2. This domain seems to be acquired by adding a number of amino acids, making the region more susceptible to a wide range of proteases, facilitating the conformational change in S protein and insertion of its S1 subunit into membrane [24,25]. Although it is not yet known whether SARS-CoV 2 and SARS-CoV sequence similarities correlates with similar biological properties, including pandemic potential [26], the interface details for Spike/ACE2 elucidated that SARS-CoV 2 transmissibility is due to efficient use of ACE2 as a key determinant at the atomic level [27,28].

Regardless of strict health measures such as social distancing, lock down of businesses and recreation centres, flight, travel, and tourism bans in many parts of the world, the high transmissibility of the virus still results in a significant number of infected cases around the world, which makes a fatality rate of 2% a very significant loss. To tackle this crisis, scientists have started lots of efforts in two major paths to first develop a vaccine to control transmission and spread of the infection and second to manufacture antivirals to treat the infected cases. As of now, more than 10 vaccines are approved for SARS-CoV 2 while over 250 teams are still working to develop vaccines against the virus using different methods (https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines). These includes development of inactivated/weakened virus particles, nucleic acid (DNA or RNA) vaccines, non/replicating viral vectors, and protein based vaccines including recombinant subunit proteins or virus-like particles [29]. Although the non-protein developed vaccines may help with the urgent need to protect at risk population, vaccine previous experiences suggest that recombinant protein-based vaccine would be likely the most efficient and safest vaccine for long-term use as a prophylactic vaccine for public. Current evidence almost unanimously recommends spike protein as the best candidate to develop an optimal vaccine with respect to humoral and cellular immune responses. Since antibody dependent enhancement is also a potential concern for SARS-CoV 2 vaccine, it is reasonable to pick as small as possible part of spike protein that is critical target to be used as vaccine. In this study, we have studied the spike protein from 10 different coronaviruses of animals and humans, including SARS-CoV and SARS-CoV 2 to pinpoint the most critical region of S protein to be used as an antigen for vaccine development.

2. Methodology

Protein sequences of spike proteins from, SARS-CoV 2 (QHD43416.1), BtRsRaTG13-CoV (QHR63300.2), SARS-CoV (AAP13441.1 S), BtRsBeta-CoV (QDF43825.1), BtRsCoV-related (ATO98157.1), PCoV (QIQ54048.1), hCoV-HKU1 (BBA20986.1), MCoV (YP_009047204.1), ACoV (-ACV87265.1) and PEDV1-CoV (ALB35885.1) were obtained from the National Center for Biotechnology Information (NCBI) database. The initial homology models of full-length spike protein from SARS-CoV 2 was remodeled including the missing C-terminal sequence and receptor binding motif using the crystal structure of 2019-ncov chimeric receptor-binding domain (PDB ID: 6VW1) with MODELLER 9v7 on windows operating system [30]. The co-ordinates for the structurally conserved regions (SCRs) of RBD SARS-CoV 2 sequence were assigned from the template using pair wise sequence alignment, based on the Needleman-Wunsch algorithm [31,32]. In addition, BtRsRaTG13-CoV, SARS-CoV, BtRsBeta-CoV and BtRsCoV-related, PCoV, hCoV, MCoV, ACoV, and PEDV1-CoV homology models were developed with the same methodology as described above using the build homology model of SARS-CoV 2 as the template. Further, protein-protein docking studies and their interactions of the full-length SARS-CoV 2, BtRsRaTG13-CoV, SARS-CoV, BtRsBeta-CoV, BtRsCoV-related, PCoV, hCoV, MCoV, ACoV and PEDV1-CoV spike proteins with its receptor ACE2 (PDBID: 6VW1) was performed with the online server ZDOCK in which proteins were treated as rigid objects and 6-dimenional rotational and translational degrees of freedom were explored. Similarly, protein-protein docking was also performed using the spike RBD subdomains, SARS-CoV 2 RBD Pro^{322}-Thr^{581}, BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323}, SARS-CoV RBD Pro^{309}-Pro^{575}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, BtRsCoV-related Arg^{306}-Pro^{575}, PCoV RBD Gln^{319}-Ser^{589}, hCoVRBD subdomain Ala^{315}-Tyr^{675}, MCoV RBD Gly^{372}-Val^{616}, ACoV RBD -Asp^{250}-Gln^{489} and PEDV1-CoV RBM Ala^{315}-Tyr^{675} using the above method and their top ten conformations were extracted. For protein-protein docking, the residues from Lys^{417}–^{508} of the exposed loop regions of the SARS-CoV 2 RBM and Ser^{19}-Met^{81} of ACE2 receptor were specified in a filter, feature blocking all other residues to involve in the binding interface with the receptor cavity of the ACE2. Finally, ZRANK, a scoring algorithm that relies on the usage of

a combination of three atom-based terms, i.e., Van der Waals, electrostatics, and desolvation was used to rank the structures [33–38]. Out of top 10 conformations that were generated, the top conformation of the SARS-CoV 2 spike protein RBM binds on the surface of the receptor ACE2 for viral host interaction was analysed. In addition, the binding conformations of full-length BtRsRaTG13-CoV, SARS-CoV, BtRsBeta-CoV, BtRsCoV-related, PCoV, hCoV, MCoV, ACoV and PEDV1-CoV full-length spike proteins and their subdomains SARS-CoV 2 RBD Pro^{322}-Thr^{581}, SARS-CoV RBD Pro^{309}-Pro^{575}, -BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, PCoV RBD Gln^{319}-Ser^{589}, hCoV RBD Ala^{315}-Tyr^{675}, BtRsCoV-related Arg^{306}-Pro^{575}, MCoV RBD Gly^{372}-Val^{616}, ACoV RBD Asp^{250}-Gln^{489} - and PEDV1-CoV RBD Ala^{315}-Tyr^{675} with its receptor ACE2 (PDBID: 6VW1) were also analyzed. The docking conformations of receptor binding spike subdomains that replicates the docking conformation of the full-length spike proteins, still able to bind with higher affinity and with similar hydrogen bonding network with the receptor ACE2 were extracted. Further, the B-Cell antigenic determinants or epitopic sequences of these isolated sub domains were predicted using “ElliPro”, a web-based tool for the prediction of antibody epitopes in protein antigens of a given sequence or structure [39,40].

3. Results and discussion

3.1. Phylogeny

Phylogenetic analysis of the CoV spike proteins falls under five subfamilies. Sequences from PCoV, SARS-CoV 2 and BtRsRaTG13-CoV fall under cluster I with SARS-CoV 2. -. On the other hand, MCoV, hCoV, ACoV and PEDV1-CoV are closely related falling under cluster II where MCoV and hCOV falls under subfamily-I while ACoV and PEDV1-CoV falls under subfamily-II. Finally, SARS-CoV, BtRsCoV-related and BtRSBeta-CoV falls under cluster-III where BtRSBeta-CoV is too divergent showing separate branch in the phylogenetic tree. The percentage of identity between the sequences reveals that SARS-CoV 2 has 97%, 92%, 76%, 76%, 75%, 26%, 24%, 21% and 19% identity with, BtRsRaTG13-CoV, PCoV, BtRsCoV-related, BtRsBeta-CoV, SARS-CoV, MCoV, hCoV, ACoV and PEDV1-CoV, respectively. This shows that SARS-CoV 2, BtRsRaTG13-CoV and PCoV are very closely related to each other compared to others in the evolution (Fig. 1). Furthers structural studies shows that the RMSD of the full-length SARS-CoV 2 with other species - showed a wide range of deviation from 2.6 to 17.2 Å while the super pose structures of CoV spike subdomain-ACE2 complexes show a least back bone RMSD difference with its full-length spike protein-ACE2 complexes within a range of 0.1–4.1 Å. However, the superimposition of SARS-CoV 2 RBM (Receptor Binding Motif) with others show a least RMSD’s ranging from 0.16 to 0.85 Å where both BtRsRaTG13-CoV and PCoV RBM’s are closed related to SARS CoV 2 RBM with 0.16 and 0.18 Å indicating a clear evolutionary ship between these three species (Table 1 & Fig 2A-2J).

3.2. Spike protein subdomains-ACE2 interactions

Protein-Protein docking studies of spike protein subdomains shows that Arg^{403}, Lys^{417} and Tyr^{453} from SARS-CoV 2 RBD Pro^{322}-Thr^{581} in the middle of the bridge shows strong network of hydrogen bonds and ionic interactions with His^{34}, Asp^{30}, Asp^{38}, Lys^{31} and Glu^{35} of α1. In addition, Tyr^{449}, Glu^{484}, Gly^{485}, Pro^{491}, Gln^{493}, Ser^{494} and Tyr^{495} forms another set of strong hydrogen bonding network with six hydrogen bonds and π-stacking interaction in the middle of the bridge with the residues of ACE2 α2, Asn^{64}, Ala^{71}, Lys^{74}, Glu^{75} and Lys^{68}. In this network of hydrogen bonds, the positively charged Arg^{403} forms two hydrogen bonds with His^{34} and Asp^{38} (3.4 and 3.7 Å), while Lys^{417} shows contacts with Asp^{30}, Glu^{35} and Lys^{31} with two ionic and a hydrogen bond (3.1, 3.0 and 3.8 Å). In addition, the hydroxyphenyl ring of Tyr^{453} shows contacts with Glu^{35} in the middle of ACE2 α1 (2.6 Å). Apart from these interactions at α1 of ACE2, Tyr^{449} towards the N-terminal end of spike protein α2 shows π-interaction with Asn^{64}. In the middle of the α2, the backbone oxygens of Tyr^{495} and Ser^{494} and terminal nitrogen of Gln^{493} shows contacts with Lys^{68} with three hydrogen bonds (2.9, 3.4 and 2.8 Å). Towards the C-terminal end, Pro^{491}, Glu^{484} and Gly^{485} of α2 shows another three hydrogen bonds with the positively charged Lys^{74} (2.8, 3.4 and 3.5 Å).

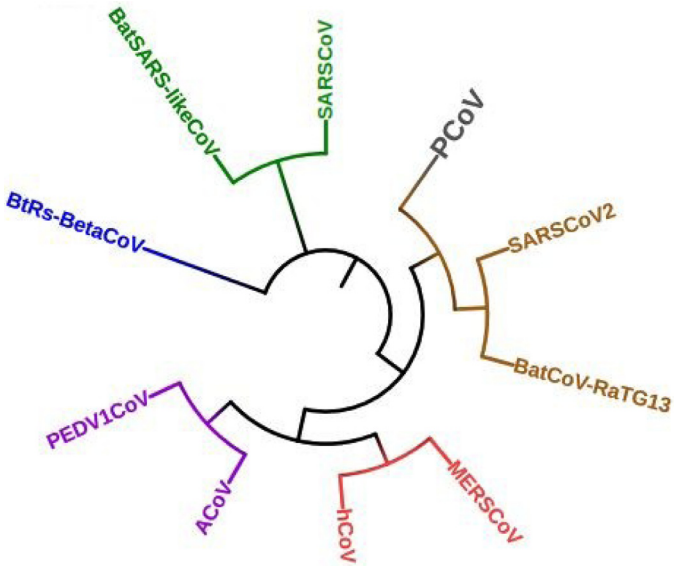

Fig. 1. Phylogenetic analysis of CoV spike proteins predicted using Tree Tol software. Four different cluster of families are indicated in Blue, Maroon, Orange and Green, Family-I is indicated in Blue, Family-II is indicated in maroon, Family-III and Family-IV are indicated as separate branches in the phylogeny indicated with orange and green colors.

Similarly, SARS-CoV 2 RBD Pro^{322}-Thr^{581} also indicates similar type of interactions with seven hydrogen bonds and two ionic interactions including one π-stacking on the surface of ACE2. At the N-terminal of α1, Leu^{455} and Tyr^{473} shows both hydrogen and π-stacking interactions with Lys^{31} (3.8 and 3.2 Å). The residues, Glu^{406} and Lys^{417} in the middle of the bridge forms one hydrogen and two ionic interactions with His^{34}, Asp^{30} and Glu^{35} (2.9, 2.6 and 3.5 Å). At the C-terminal end of α1, Asp^{501} shows a hydrogen bond with Gln^{42} (2.5 Å). In addition, the residues Glu^{445}, Tyr^{495}, Gly^{485} and Asn^{487} shows four hydrogen bonds with Glu^{57} at the N-terminal of ACE2 α2, Lys^{68} at the middle of the bridge and Thr^{78} at the C-terminal of α2 (3.0, 3.7, 2.9 and 3.2 Å).

In SARS-CoV Pro^{309}-Pro^{575} the residues, $Asp^{392,}$ Arg^{395}, Asp^{463} and Asp^{480} forms ionic interactions with Lys^{74}, Glu^{110}, Lys^{353} and Lys^{31} (3.5, 3.4, 2.5 and 3.6 Å). In addition, Tyr^{442} and Trp^{476} shows two hydrogen bonds with Glu^{75} and Glu^{35} (3.6 and 2.7 Å). The subdomain, BtRsBeta-CoV Ser^{311}-Thr^{568} forms a network of hydrogen bonds along with π-interaction on the receptor surface. The residues, Gly^{483}, Asn^{488} and Tyr^{492} in the middle of the bridge interacts with Asp^{38} and Lys^{353} (3.3, 3.0 and 3.9 Å). However, BtRsCoV-related Arg^{306}-Pro^{575} binds far from N-ter Met^{82} of ACE2 with similar orientation like BtRsBeta-CoV Ser^{311}-Thr^{568} on the receptor surface predicting different type of interactions between the protein and the receptor. This allows the subdomain to form five hydrogen bonds, including one π-stacking on the receptor surface. The residues, Cys^{475} at the C-terminal end of spike protein α2 and Asp^{481} in the middle of the bridge shows two hydrogen bonds with Thr^{78} and Lys^{68} (2.7 and 3.6 Å). In extension to these hydrogen bonds, Gly^{483}, Asn^{488} and Tyr^{492} forms another set of strong hy-

Table 1
RMSD difference in Å between the full length and its subdomains of CoV spike proteins predicted using MOE software suite (Molecular Operating Environment (MOE), 2014.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada) H3A 2R7, 2021).

Species	Full length Spike that interacts with its receptor ACE2 used in this study	RMSD (Å) with full length SARS CoV2	Spike Subdomain that interacts with its receptor ACE2 used in this study	RMSD (Å) of Spike Subdomain-ACE2 interaction with its full length-ACE2 interaction	Receptor Binding Motif (RBM)	RMSD (Å) With RBM of SARS CoV 2
SARS	SARS-CoV2	–	SARS-CoV 2 RBD Pro^{322}-Thr^{581}	0.1	SARS-CoV 2 RBM Asn^{440}-Pro^{507}	—
Bat	BtRsCoVRaTG13-CoV Met^{1}-Ser^{1143}	5.5	-BtRs RaTG13-CoV Thr^{323} -Thr^{581}	1.4	-BtRs RaTG13-CoV RBM His^{440}-Tyr^{508}	0.16
SARS	SARS-CoV Met^{1}-Ser^{1129}	3.9	SARS-CoV Pro^{309}-Pro^{575}	0.5	SARS-CoV RBM Asn^{427}-Tyr^{494}	1.22
Bat	BtRsBeta-CoV Met^{1}-Ser^{1130}	2.7	BtRsBeta-CoV Ser^{311}-Thr^{568}	1.4	BtRsBeta CoV RBM Asn^{428}-Tyr^{495}	0.21
Bat	BtRsCoV-related Met^{1}-Ser^{1129}	3.5	BtRsCoV-related Arg^{306}-Pro^{575}	–	BtRsCoV- related RBM Ser^{424}-Tyr^{494}	0.52
Pangolin	PCoV Met^{1}-Ser^{1143}	5.5	PCoV RBD Gln^{319}-Pro^{589}	4.1	PCoV RBM Lys^{440}-Phe^{508}	0.18
MERS	MCoV Met^{1}-Glu^{1237}	2.6	MCoV RBD Gly^{372}-Val^{616}	–	MCoV RBM His^{486}-Tyr^{564}	0.85
Human	hCoV Met^{1}-Asp^{1230}	17.2	hCoV RBD Ala^{315}-Tyr^{675}	–	hCoV– RBM Leu^{431}-Pro^{548}	—-
Avian	ACoV Met^{1}-Glu^{1067}	10.2	ACoV RBD Asp^{250}-Gln^{489}	–	ACoV RBM RBM Ser^{351}-Lys^{40589}	1.78
Pig	PEDV1-CoV Met^{1}-Asn^{1261}	3.0	PEDV1-CoV RBD Ala^{315}-Tyr^{675}	–	PEDV1- CoV RBM Asn^{491}-Lys^{566}	0.57

drogen bonds and π-stacking in the middle of α1 with Asp^{38} and Lys^{353} (3.3, 3.0, and 3.9 Å).

Oppositely, PCoV RBD Gln^{319}-Ser^{589} shows similar binding site and orientation compared to SARS-CoV 2 with six hydrogen bonds on the surface of the receptor ACE2. The residues, Tyr^{419} and Tyr^{471} at the C-terminal and in the middle of the α1 contacts with two hydrogen bonds (3.6 and 3.4 Å). However, Ala^{473} shows contacts with Glu^{75} and Thr^{78} at the middle of the α2 (3.2 and 2.9 Å) while both Gly^{444} and Glu^{491} also show two hydrogen bonds with both negatively and positively charged Glu^{57} and Lys^{68} (3.4 Å). In MCoV RBD Gly^{372}-Val^{616},out of a total of six hydrogen bonds and an ionic interaction with the receptor ACE2, two of them show contacts with N-acetyl-D-glucosamine through Thr^{492} and Lys^{493} (1.5 and 3.3 Å). The residue, Lys^{493} also shows contacts with Glu^{57} at the N-terminal of α2 (3.3 Å), while Ser^{524} at the C-terminal end of the α2 contacts with Glu^{75} (2.9 Å). However, Tyr^{541} and Lys^{543} in the middle of the α1 shows contacts with Asp^{38} through both hydrogen and ionic interactions (2.7 and 3.7 Å). Surprisingly, hCoV RBD Ala^{315}-Tyr^{675} shows contacts with N-acetyl-D-glucosamine with eleven hydrogen bonds with the receptor ACE2. The residue, Asn^{452} shows contacts with Asn^{49} and Thr^{52} at the C-terminal end of α1 through two hydrogen bonds (3.8 and 3.0 Å). At the N-terminal end of α2, Pro^{490} and Pro^{491} shows two hydrogen bonds with Glu^{56} (2.8 and 3.7 Å). In addition, the residue Ser^{494} show hydrogen bond with Thr^{125} at the C-terminal of α5 (2.9 Å). Similar to SARS-CoV 2 subdomain, ACoV RBD Asp^{250}-Gln^{489} - also shows similar binding orientation with six hydrogen bonds and an ionic interaction with ACE2. The residue, Ser^{382} at the C-terminal end of α2 shows contacts with Glu^{75} (3.1 Å) while Tyr^{390} and Val^{392} in the middle of the helix α2 also shows contacts with Glu^{75} and Lys^{68} (3.0 Å). However, Arg^{477} at the N-terminal end of the α2 contacts with Glu^{56} with an ionic interaction (3.9 Å). In extension to these contacts, Gln^{385} and Cys^{388} shows contacts with the same Lys^{31} towards the N-terminal end of the α1 (3.0 Å). Finally, PEDV1-CoV RBD Ala^{315}-Tyr^{675} shows similar binding site and orientation compared to SARS-CoV 2 subdomain on the ACE2 receptor surface with three hydrogen bonds. The residues, Asn^{508} and Thr^{558} at the C-terminal and in the middle of the α1 contacts with two hydrogen bonds (3.5 and 2.8 Å) while Ile^{551} shows contacts with Thr^{78} at the middle of the α2 with the distance (3.5 Å).

These protein-protein docking studies of spike protein subdomains reveal that the residues Arg^{403}, Lys^{417}, $Tyr^{449,}$ Tyr^{453}, Glu^{484}, Gly^{485}, Pro^{491}, Gln^{493}, Ser^{494} and Tyr^{495} of SARS-CoV 2 RBD Pro^{322}-Thr^{581} (Fig. 3A); Glu^{406}, Lys^{417}, Glu^{445}, Leu^{455}, Tyr^{473}, Gly^{485}, Asn^{487}, Tyr^{495} and Asp^{501} of BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323} (Fig. 3B); Asp^{392}, Arg^{395}, Tyr^{442}, Asp^{463}, Trp^{476}, and Asp^{480} of SARS-CoV RBD Pro^{309}-Pro^{575} (Fig. 3C); Cys^{475}, Asp^{481}, Gly^{483}, Asn^{488}, and Tyr^{492} of BetaCoV RBD Ser^{311}-Thr^{568} (Fig. 3D); Cys^{475}, Asp^{481}, Gly^{483}, Asn^{488}, and Tyr^{492} of BtRsCoV-related RBD Arg^{306}-Pro^{575} (Fig. 3E); Tyr^{491}, Gly^{444}, Tyr^{471}, Ala^{473} and Glu^{491} of PCoV RBD Gln^{319}-Ser^{589} (Fig. 3F)); Thr^{492}, Lys^{493}, Ser^{524}, Tyr^{541} and Lys^{543} of MCoV RBD Gly^{372}-Val^{616} (Fig. 3G); Ser^{382}, Gln^{385}, $Cys^{388,}$ Tyr^{390}, Val^{392}, and Arg^{477} of ACoV RBD Asp^{250}-Gln^{489} (Fig. 3H) and the residues Asn^{508}, Ile^{551} and Thr^{558} of PEDV1-CoV RBD Ala^{315}-Tyr^{675} (Fig. 3J) acts as common pharmacophores in viral host interactions with stronger hydrogen bonds with its receptor ACE2 (Table 2). However, the results reveal that N-acetyl-D-glucosamine from ACE2 plays an important role in viral host interactions with stronger hydrogen bonds in hCOV (Fig. 3I).

3.3. Full-length spike protein -ACE2 interactions

In comparison to the spike subdomains, the residues of the full-length SARS-CoV 2 RBD, $Gln^{474,}$ Gln^{498}, Thr^{500}, and Asn^{501} at the N and C terminus of α1 form a network of hydrogen bonds with Gln^{24}, Tyr^{41}, Gln^{42}, Met^{82}, Lys^{353} and Arg^{357} of ACE2 receptor (Fig. 4A). The residues of the -BtRs RaTG13-CoV RBD, Lys^{417}, Tyr^{453}, Arg^{494}, Tyr^{498}, Asp^{501} and His^{505} shows contacts through eight hydrogen bonds, while BtRsBeta-CoV RBD shows only four hydrogen bonds with ACE2 receptor surface (Fig. 4B and Fig. 4D). However, the residues of SARS-CoV RBD, Thr^{433} Tyr^{475}, Pro^{477} and Tyr^{481} at the N and C-terminus of α1 makes contacts with Asp^{38}, $Lys^{68,}$ Glu^{57} and Glu^{75}. In the middle of the bridge, Ser^{461} and Leu^{472} interacts with Met^{82} and $Lys^{74,}$ respectively. (Fig. 4C). Both Trp^{442} and Arg^{479} of BtRsCoV-related forms an hydrogen and ionic interaction with the same His^{34} while Arg^{479} also forms ionic interaction with Asp^{30}. Apart from these interactions, Gly^{471}, Asn^{473} and Tyr^{475} forms hydrogen bonds and π-interaction with Thr^{78}, Gln^{24} and Lys^{31} with ACE2, respectively (Fig. 4E). The substitution of interface residues of PCoV RBD allow Tyr^{488} and Glu^{491} to form

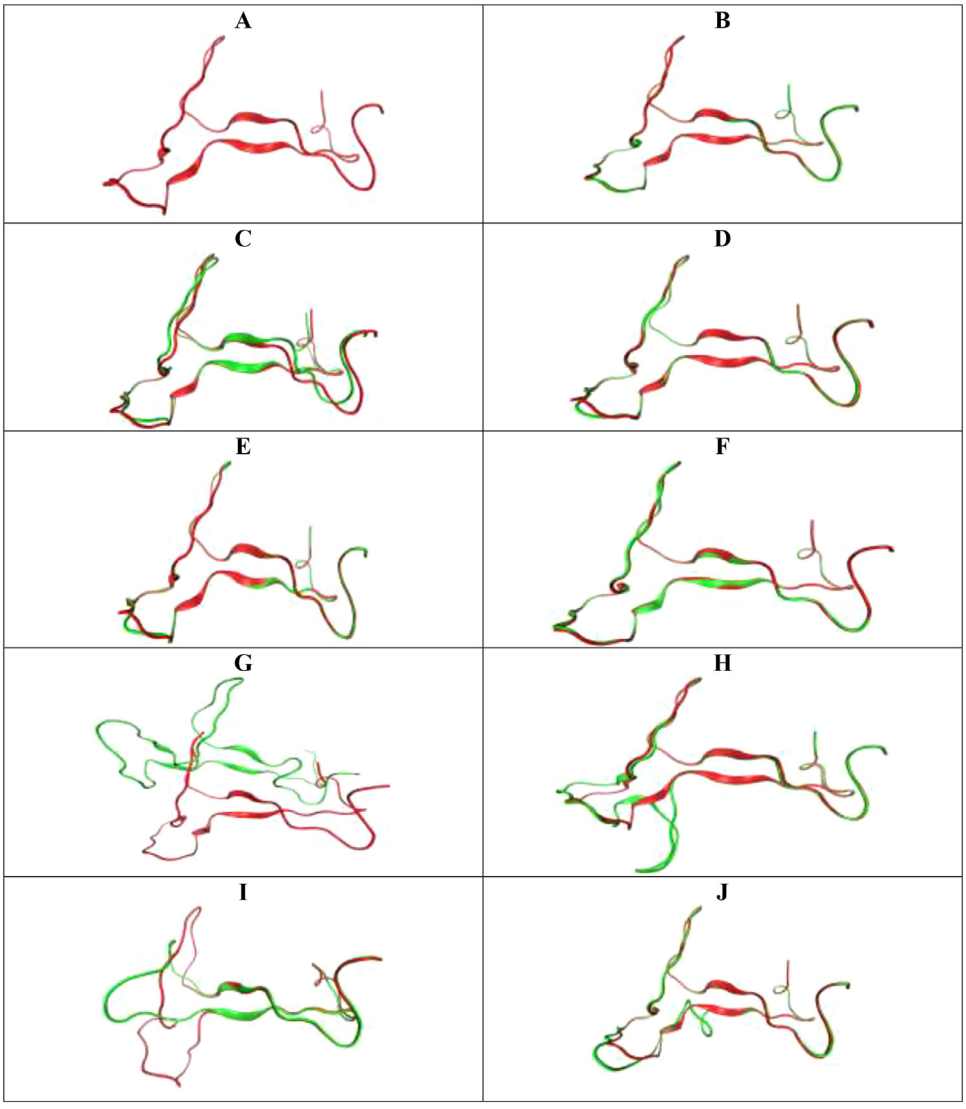

Fig. 2. Superimposition of SARS-CoV 2 RBM **(Fig 2A)** with BtRsRaTG13-CoV RBM **(Fig 2B)**, SARS-CoV RBM **(Fig 2C),** BtRsBeta-CoV RBM **(Fig 2D)**, BtRsCoV-related RBM **(Fig 2E),** PCoV RBM (**Fig 2F**), hCoV RBM **(Fig 2 G)**, MCoV RBM **(Fig 2H)**, ACoV RBM **(Fig 2I)** and with PEDV1-CoV RBM **(Fig 2 J)** with its receptor ACE2 predicted using MOE software suite (Molecular Operating Environment (MOE), 2014.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada) H3A 2R7, 2021). SARS-CoV 2 RBM is represented in red color while the RBM of the other species is shown in green color.

Table 2
Amino acids of CoV spike protein subdomains involved in binding to its receptor ACE2 predicted using MOE software suite ((Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada) H3A 2R7, 2021).

Species	Spike Subdomain	Residues of Spike Subdomain involved in binding with the receptor ACE2.	Residues of ACE2 receptor involved in binding with the spike subdomains.
SARS	SARS-CoV 2 RBD Pro^{322}-Thr^{581}	Arg^{403}, Lys^{417}, $Tyr^{449.}$ Tyr^{453}, Glu^{484}, Gly^{485}, Pro^{491}, Gln^{493}, Ser^{494} and Tyr^{495}	Asp^{30}, Lys^{31}, His^{34}, Glu^{35}, Asp^{38}, Asn^{64}, Lys^{68}, Ala^{71}, Lys^{74} and Glu^{75}
Bat	BtRs RaTG13-CoV RBD Thr^{581}-Thr^{323}	Glu^{406}, Lys^{417}, Glu^{445}, Leu^{455}, Tyr^{473}, Gly^{485}, Asn^{487}, Tyr^{495} and Asp^{501}	Asp^{30}, Lys^{31}, His^{34}, Glu^{35}, Gln^{42}, Glu^{57}, Lys^{68}, Thr^{78}
SARS	SARS-CoV RBD Pro^{309}-Pro^{575}	Asp^{392}, Arg^{395}, Tyr^{442}, Asp^{463}, Trp^{476}, and Asp^{480}	Asp^{392}, Arg^{395}, Tyr^{442}, Asp^{463}, Trp^{476}, and Asp^{480}
Bat	BtRsBeta-CoV RBD Ser^{311}-Thr^{568}	Cys^{475}, Asp^{481}, Gly^{483}, Asn^{488}, and Tyr^{492}	Asp^{38}, Lys^{68} Thr^{78} and Lys^{353}
Bat	BtRsCoV-related RBD Arg^{306}-Pro^{575}	Cys^{475}, Asp^{481}, Gly^{483}, Asn^{488}, and Tyr^{492}	Asp^{38}, Lys^{68}, Thr^{78}, and Lys^{353}
Pangolin	PCoV RBD Gln^{319}-Ser^{589}	Tyr^{491}, Gly^{444}, Tyr^{471}, Ala^{473} and Glu^{491}	Lys^{31}, Glu^{57}, Glu^{75}, Thr^{78}, and Lys^{68}
MERS	MCoV RBD Gly^{372}-Val^{616}	Thr^{492}, Lys^{493}, Ser^{524}, Tyr^{541} and Lys^{543}	Asp^{38}, Glu^{57} and Glu^{75}
Human	hCoV- RBD Ala^{315}-Tyr^{675}	Ala^{315}-Tyr^{675}, Asn^{452}, Pro^{490}, Pro^{491} and Ser^{494}	Asn^{49}, Thr^{52}, Thr^{52}, Thr^{125}
Avian	ACoV RBD Asp^{250}-Gln^{489}	Ser^{382}, Gln^{385}, $Cys^{388.}$ Tyr^{390}, Val^{392}, and Arg^{477}	Lys^{31}, Glu^{56}, Lys^{68}, Glu^{75}
Pig	PEDV1-CoV RBM RBD Ala^{315}-Tyr^{675}	Asn^{508}, Ile^{551} and Thr^{558}	Ser^{19}, Asp^{38} and Thr^{78}

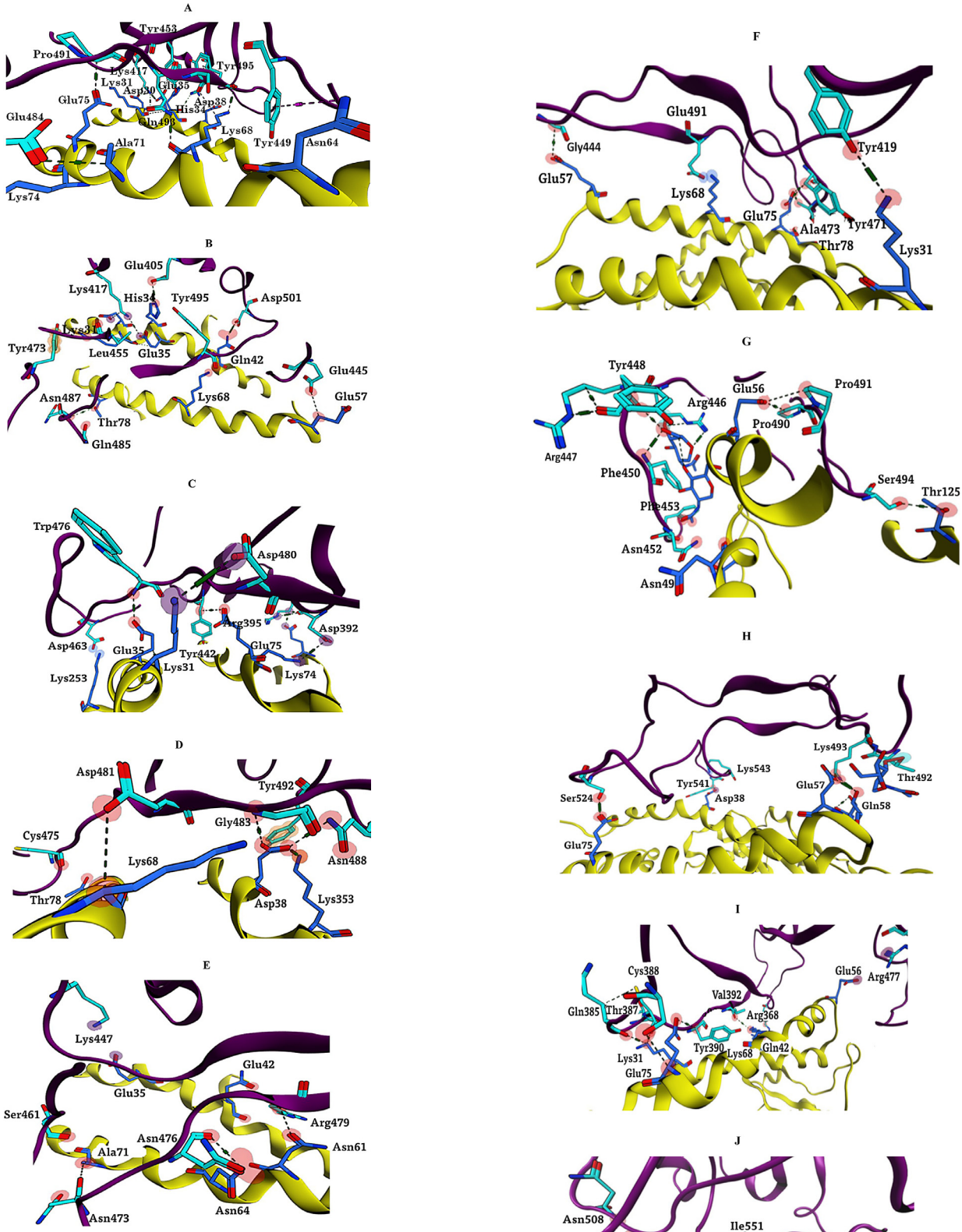

Fig. 3. Protein-Protein interactions of SARS-CoV 2 RBD Pro^{322}-Thr^{581} **(A)**, BtRsRaTG13-CoV RBD Thr^{323}-Thr^{581} (**B**), SARS-CoV RBD Pro^{309}-Pro^{575} **(C),** BtRsBeta-CoV RBD Ser^{311}-Thr^{568} **(D)**, BtRsCoV-related RBD Arg^{306}-Pro^{575} **(E),** PCoV RBD Gln^{319}-Ser^{589} (F), hCoV RBD subdomain Ala^{315}-Tyr^{675}, **(G)**, MCoV RBD Gly^{372}-Val^{616} **(H)**, ACoV RBD Asp^{250}-Gln^{489} **(I)**, PEDV1-CoV RBD Ala^{315}-Tyr^{675} **(J)** with its receptor ACE2 predicted using MOE software suite (Molecular Operating Environment (MOE), 2014.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada) H3A 2R7, 2021). Spike protein is represented in maroon ribbons with residues in cyan color while the ACE2 receptor is represented in yellow ribbons with residues in blue colors.

Fig. 3. Continued

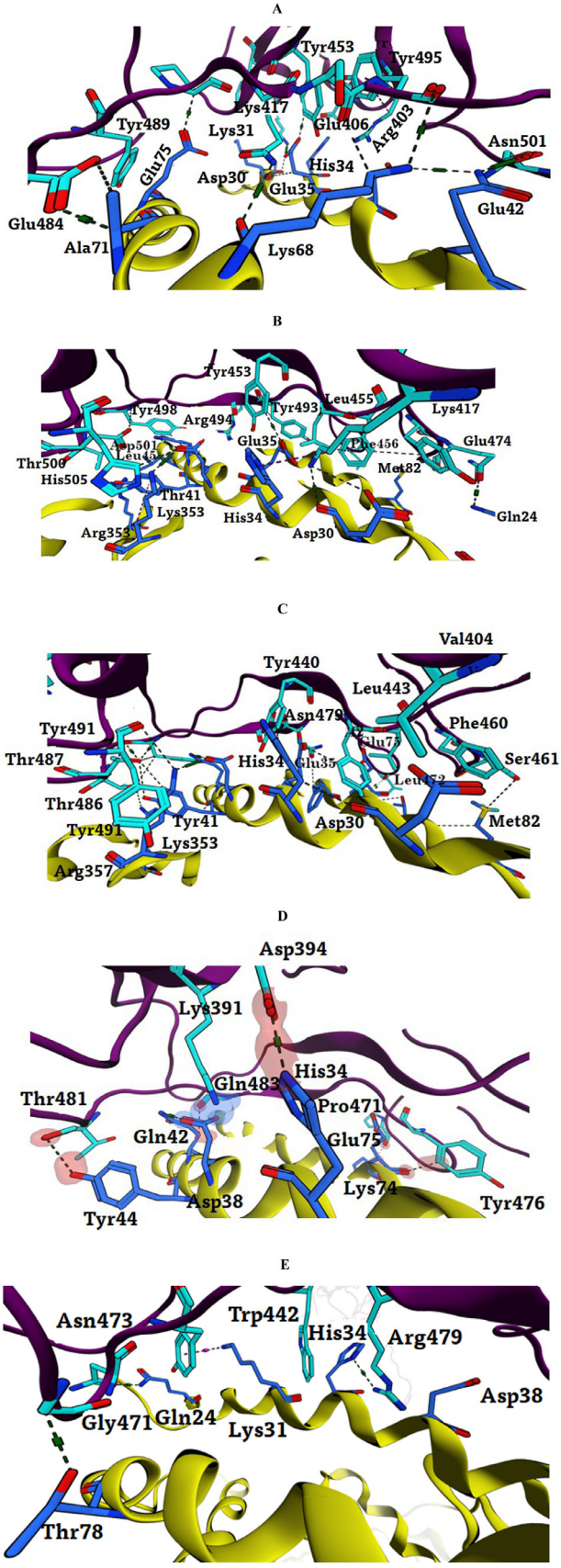

Fig. 4. Protein-Protein interactions of SARS-CoV 2 (A), BtRsRaTG13-CoV (B), SARS-CoV (C), BtRsBeta-CoV (D), BtRsCoV-related (E), PCoV (F), hCoV (G), MCoV (H), ACoV (I), PEDV1- CoV (J), with its receptor ACE2 predicted using MOE software suite. Spike protein is represented in maroon ribbons with residues in cyan color while the ACE2 receptor is represented in yellow ribbons with residues in blue colors.

Fig. 4. Continued

hydrogen and ionic interactions with the positively charged Lys^{68}. However, another bulky residue Tyr^{471} shows hydrogen bond with the negatively charged Glu^{75}. Apart from these interactions, both Arg^{455} and Arg^{492} forms hydrogen bonds with Lys^{31} and Asn^{61}. In addition, both Phe^{454} and Ala^{473} also forms hydrogen bonds with Glu^{35} and Thr^{78} with ACE2 receptor surface (Fig. 4F).

Conversely, the residues of the hCoV forms hydrogen bonds with N-acetyl-D-glucosamine around the surface of the receptor. The positively charged Arg^{446} and Arg^{447} and the backbone of Phe^{450} shows hydrogen bonds with NAG711 oxygens. In addition, both the Tyr^{448} and Gly^{449} also shows hydrogen bonds with NAG710 oxygens (Fig. 4G). Similar to hCoV, MCoV also shows a different mode of binding and allows the residues Pro^{471}, Gly^{483}, Thr^{487} to form hydrogen and π-interactions at the attachment site with the residues of the receptor ACE2. The phenyl ring of Phe^{418} and Asn^{421} at the middle of the bridge shows both π-stack and a hydrogen bond with Gln^{24}. However, His^{486} at the C terminus of $\alpha 1$ forms a hydrogen bond with Asn^{103} (Fig. 4H). Moreover, ACoV-ACE2 also shows a different mode of binding and no alignment was seen from Phe^{490} -Pro^{499} of SARS-CoV 2-RBD with ACoV-RBD. The positively charged Arg^{207}, shows ionic interaction with Asp^{111} while Ser^{298} and Thr^{432} of RBD forms hydrogen bonds with Glu^{110} and Met^{82} of ACE2 respectively (Fig. 4I). Likewise, PEDV1-CoV with different mode of binding allows Asn^{508}, Ile^{551}, and Thr^{558} to form three hydrogen bonds with Ser^{19}, Thr^{78} and Asp^{38} of ACE2 respectively (Fig. 4J). These network of hydrogen bonds with different amino acids in different species with the receptor ACE2 is due to amino acid variations of RBM in comparison to SARS CoV 2 RBM (Table 3).

Over all, the critical residues at receptor binding motif of spike proteins shows that both positively charged Arg^{403}, Lys^{417}, Lys^{444} and negatively charged Glu^{406} play an important role in formation of ionic interactions/salt bridges in attaching to the host receptor ACE2 and is only discussed further. The residue Arg^{403} of SARS-CoV 2 and Lys^{390} of SARS-CoV shows an ionic interaction with the nearby residue Asp^{38}. However, Lys^{391} of BtRsBeta-CoV deviates and forms interaction with His^{34} instead of Asp^{38} at the interface of the receptor ACE2. In comparison, ε-amine of Lys^{390} in BatBtRs-CoV related contacts with internal residue Asp^{405} making less stable on the surface of the receptor ACE2. On the other hand, SARS-CoV 2 RBD Pro^{322}-Thr^{581} also shows contacts with $Asp^{30,}$ Glu^{35} and Lys^{31} through ionic interactions and a hydrogen bond with −8.2, −17.9 and −1.2 kcal/mol higher than the energies obtained with full length protein. On the other side, BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323} also contacts with Asp^{30} and Glu^{35} similar to full length protein with −9.2 and −10.3 Kcal/mol. However, no contacts were seen with Lys^{304} of ACoV and ACoV RBD Asp^{250}-Gln^{489} due to its positioning far away from the receptor surface similar to full length protein.

In addition, the mutants Arg^{403}Thr in both full-length and subdomains of BtRsRaTG13-CoV and BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323}, ACoV and ACoV RBD Asp^{250}-Gln^{489} , Arg^{403}Ser, Arg^{403}Pro and Arg^{403}Tyr in MCoV and MCoV RBD Gly^{372}-Val^{616}, hCoV and hCoV-RBD Ala^{315}-Tyr^{675}, PEDV1-CoV and PEDV1-CoV RBD Ala^{315}-Tyr^{675} resulted in a loss of this ionic interaction due to their smaller side chain and lack of positive charge on the receptor surface ACE2. This confirms that higher stability of Arg^{403} in SARS-CoV 2 is due to its internal network of hydrogen bonds with Ile^{402}, Gly^{404} and Asp^{405} that allows to orient and makes stronger interaction with the receptor ACE2 are not seen with other viral species.

Furthermore, another ionic interaction/salt bridge is formed at the nearby residues between Glu^{406}-His^{34} of SARS-CoV 2 with the binding energy of −1.4 kcal/mol. The residue also contacts internally with Asp^{405}, Val^{407}, Arg^{408} and Gln^{409} through hydrogen bonds predicting to be highly stable in this orientation. Although the residue Glu^{406} is highly conserved in BtRsRaTG13-CoV and shows internal contacts with Arg^{403}, Arg^{408} and Gln^{409}, no hydrogen bonds are seen with receptor surface. This predicts that Glu^{406} to show lesser contribution in receptor binding compared to SARS-CoV 2. Similar type of internal contacts was also seen with mutant Glu^{406}Asp, where two hydrogen bonds are seen in BtRsBeta-CoV and BtRsCoV-related with Arg^{493} and Gln^{409} while two hydrogen bonds are seen only with Arg^{404} through terminal oxygens in SARS-CoV. In extension to these mutant, Glu^{406}Met, Glu^{406}Arg, Glu^{406}Gly and Glu^{406}Gln mutants also shows similar effect in MCoV, hCoV, PEDV1-CoV and ACoV without any internal and external hydrogen bonds predicting to show lesser contribution in binding affinity with the receptor surface.

On the other hand, the subdomains of BtRsBeta-CoV RBD Ser^{311}-Thr^{568} and SARS-CoV RBD Pro^{309}-Pro^{575} also maintains the ionic interaction/salt bridge with Lys^{353} with −5.1 and −4.2 kcal/mol showing no sign of formation in BtRsCoV-related. Similarly, the Glu^{406}Met, Glu^{406}Arg, Glu^{406}Gly and Glu^{406}Gln in other viral species like MCoV RBD Gly^{372}-Val^{616}, hCoV-RBD Ala^{315}-Tyr^{675}, PEDV1-CoV RBD Ala^{315}-Tyr^{675} and ACoV RBD Gly^{372}-Val^{616} has not shown any sign of ionic interaction/salt bridge with surface receptor showing an order SARS-CoV RBD Pro^{309}-Pro^{575} > BtRsBeta-CoV RBD Ser^{311}-Thr^{568} > BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323} > SARS-CoV 2 RBD Pro^{322}-$Thr^{581,}$ respectively.

In extension to these charge interactions, the residue Lys^{417} of SARS-CoV 2 which is conserved in BtRsRaTG13-CoV and ACoV result in tighter association because of the two ionic interactions/salt bridges formed between the terminal amino group of Lys^{317}, Asp^{30} and Glu^{35} of ACE2 with a total energy of −14.3 kcal/mol. In addition, Asp^{30} also accepts electrons with the terminal carbons of Lys^{417} to form additional bonds with −1.6 and −0.5 kcal/mol. In this orientation, the backbone oxygen of Lys^{417} also forms internal hydrogen bond by accepting electrons with the near by residue Tyr^{421} for higher stability. In comparison, Lys^{417} of BtRsRaTG13-CoV donates electrons to both Asp^{30} and Glu^{35} to form two hydrogen bonds with lesser binding energies of −4.2 and −9.5 kcal/mol less than the energies formed in SARS-CoV 2. However, the terminal carbons donate electrons to leu^{455} while backbone oxygen accepts electrons with the same Tyr^{421} to form two internal hydrogen bonds. On the other hand, SARS-CoV 2 RBD Pro^{322}-Thr^{581} also contacts with $Asp^{30,}$ Glu^{35} and Lys^{31} with −8.2, −17.9 and −1.2 kcal/mol through ionic interactions and a hydrogen bond higher than the energies obtained with full length protein. In addition, BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323} also contacts with Asp^{30} and Glu^{35} similar to full length protein with −9.2 and −10.3 kcal/mol. However, no contacts were seen with Lys^{304} of ACoV and ACoV RBD Gly^{372}-Val^{616} due to its positioning far away from the receptor surface similar to full length protein. Replacing with hydrophobic Valine (Lys^{417}Val) shows no contacts between the protein and the receptor both in BtRsBeta-CoV and BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, BtRsCoV- related and BtRsCoV-related Arg^{306}-Pro^{575}, SARS and SARS-CoV RBD Pro^{309}-Pro^{575} due to its hydrophobic environment and inducing structural changes to stay way from the charged residues on the receptor surface. In addition, elimination of the positively charged group and introducing the bulky side chain of phenylalanine at position Lys^{417} restrict the conformational changes which in turn greatly decreases the affinity of RBD showing no direct contact with the receptor ACE2 both in hCoV and hCoV RBD subdomain Ala^{315}-Tyr^{675}. However, mutations of Lys^{417}Pro result in the appearance of the fastest phase in folding at RBM of MCoV and disrupt the local structure near the binding interface. This causes the shift at the RBM to bind away from the surface showing internal hydrogen bond between the α-amino group and the nearby residue Gly^{462}, both in MCoV and MCoV RBD Gly^{372}-Val^{616}. Similarly, replacing the uncharged

Table 3
Amino acids of CoV full-length spike proteins involved in binding to its receptor ACE2 predicted using MOE software suite (Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada) H3A 2R7, 2021).

Species	Full length Spike that interacts with its receptor ACE2 used in this study	Amino acid Variations in comparison to SARS-CoV2	Residues of Spike involved in binding with the receptor ACE2.	Residues of receptor ACE2 involved in binding with the spike subdomains.
SARS	SARS-CoV2	—	Gln^{498}, Thr^{500}, and Asn^{501} $Gln^{474,}$ Phe^{486}	Tyr^{41}, Gln^{42}, Lys^{353}, $Arg^{357,}$ $Gln^{24,}$ and Met^{82}
Bat	BtRsRaTG13-CoV Met^{1}-Ser^{1143}	Phe^{485} → Gln^{498} and Ser^{443} → Leu^{455}	Lys^{417}, Tyr^{453}, Arg^{494}, Tyr^{498}, Asp^{501} and His^{505}	$Asp^{30,}$ Glu^{35}, Asp^{38}, Gln^{42}, Lys^{68},
SARS	SARS-CoV Met^{1}-Ser^{1129}	Arg^{426} →Asn^{439}, Tyr^{484} → Gln^{498}, and Thr^{487} → Asn^{501}, Val^{404}→ Lys^{417}, Tyr^{442} → Leu^{455}, Leu^{443} → Phe^{456}, Phe^{460} → Tyr^{473}, Asn^{479} → Gln^{493}, and Leu^{472}→Phe^{486}	Thr^{433} $Tyr^{475,}$ Pro477 and Tyr481	$Glu^{57,}$ Met^{82} and $Lys^{74,}$ Glu^{75}, Asp^{38} and Lys^{68}
Bat	BtRsBeta -CoV Met^{1}-Ser^{1130}	Phe^{485} → Gln^{498} and Ser^{443} → Leu^{455}	$Tyr^{476,}$ Thr^{487} Gly^{483} Pro^{471}	Tyr^{41}, Gln^{42}, Lys^{74}, Glu^{75}
Bat	BtRsCoV- related Met^{1}-Ser^{1129}	Val^{404}→Lys^{417}, Trp^{442} → Leu^{455}, Ile^{525} → $Phe^{486,}$ Phe^{456} →Val^{443}, Gln^{474}→Ser^{461}, Phe^{486}→Pro^{472}, Ala^{487}→Asn^{501}, His^{491}→$Tyr^{505,}$ Phe^{484} → $Gln^{498,}$ Arg^{479}→Gln^{493} Ile^{525}→ Phe^{486}	Arg479 Ile525, Gly471, Asn473 and Tyr475	His^{34}, Thr^{78}, Gln^{24} and Lys^{31}, Asp^{30}
Pangolin	PCoV Met^{1}-Ser^{1143}	Val^{415}→Lys^{417}, Leu^{484}-Phe^{486}, Asp^{491}-Gln^{493}, Arg^{492}-Ser^{494}, Thr^{499}-Asn^{501}, Val^{437}-Asn^{439}, His^{496}-$Gln^{498,}$ Asp^{491}-Gln^{493} and Arg^{492}-Ser^{494}	Phe^{454}, Arg^{455}, Tyr^{471}, Ala^{473}, Tyr^{488}, Glu491 and Arg^{492}	Glu^{35}, Lys^{31}, Asn^{61}, Lys^{68} Glu^{75} &Thr^{78},
MERS	MCoV Met^{1}-Glu^{1237}	Phe^{473} → Leu^{455}, Gly^{538} → $Phe^{486,}$ Glu^{549}→Gln^{493}, Gly^{550} →Ser^{550}, Ser^{557}→Asn^{501}, Val^{561}→Tyr^{505}, Leu^{441} →Asn^{439}, Leu^{554} → $Gln^{498,}$ Ser^{474} → Phe^{456}, Pro^{515} → Tyr^{473}, and Glu^{549} → Gln^{493}, Gly^{538}→ Phe^{486}	Pro^{471} Tyr^{476} Gly^{483} Thr^{487} Phe^{418} Asn^{421}	Gln^{24} Asn^{103}
Human	hCoV Met^{1}-Asp^{1230}	Phe^{408}→Lys^{417}, Cys^{466} → Leu^{455}, Ile^{525} → $Phe^{486,}$ Ser532→Gln^{493}, Val^{540}→Asn^{501}, Glu^{544}→Tyr^{505}, Ser^{442} →Asn^{439}, Lys^{537} → Gln^{498}, Pro^{490} → Tyr^{473}, and Ser^{532} → $Gln^{493,}$ Ile^{525}→ Phe^{486}	Arg^{446} $Arg^{447,}$ Tyr^{448}, $Gly^{449,}$ Phe^{450}	——
Avian	ACoV Met^{1}-Glu^{1067}	Glu^{358}-Asn^{439}, Val^{371}-Leu^{452}, Leu^{386}-Phe^{456}, Tyr^{386}-Phe^{486} and Ser^{303}-Gln^{493}	Arg^{207} & $Ser^{298,}$	Glu^{110} & Asp^{111}
Pig	PEDV1 Met^{1}-Asn^{1261}	Val^{468}→Lys^{417}, Ser^{506} → Leu^{455}, Ile551→Gln^{493}, Asn^{556} →Gln^{498}, Thr^{490} →Asn^{439}, Asn^{556} → Gln^{498}, His^{524} → Tyr^{473}, and Ile^{551} → Gln^{493}	Ile^{551} Thr^{558}	Ser^{19}, Thr^{78} and Asp^{38}

amino acid Thr^{423} with Lys^{417} (Lys^{417}Thr) decreases the net charge near the negatively charged amino acids on the receptor surface showing no sign of contacts with the viral protein both in PEDV1-CoV and PEDV1-CoV RBD Ala^{315}-Tyr^{675}. This indicates that introduction of a positive charge of the ε-NH2 of Lys^{417} may be generating charge attraction with spatially close to negatively charged residues to form hydrogen bond with Asp^{30} and Glu^{35} of ACE2 predicting to be the hot spot for protein-protein interactions. Previous structural studies also show that Lys^{317} of the RBD in the middle of the "bridge may result in tight hydrogen bonds with Asp^{30} of ACE2 [41]. In addition, Lys^{444} shows another salt bridge with Glu^{57} with −0.5 and −2.0 Kcal/mol indicating lesser contribution in SARS-CoV 2 than in BtRsRaTG13-CoV. However, no salt bridge was seen in ACoV although the positively charged Lys^{423} (Lys^{444} in SARS CoV 2) is highly conserved due to its positioning far way from the surface receptor ACE2. In turn, the salt bridge between the carboxylic group of Glu^{57} and the Lys^{444} amine is disrupted in the mutants Lys^{444}Thr in BtRsBeta-CoV, SARS-CoV and MCoV, where the amino group of threonine probably cannot make such an ionic interaction with the receptor due to its longer distance with Glu^{57}.

Protein-protein docking studies of SARS-CoV 2 RBD Pro^{322}-Thr^{581}, BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323}, SARS-CoV RBD Pro^{309}-Pro^{575}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, BtRsCoV-related Arg^{306}-Pro^{575} and PCoV RBD Gln^{319}-Ser^{589} subdomains binds similar to full length spike proteins of SARS-CoV 2, BtRsRaTG13-CoV, SARS-CoV, BtRsBeta-CoV, BtRsCoV-related and PCoV to the N-terminal helices of ACE2 receptor predicting to be the important subdomains that may induce antibodies to cross reactive against SARS-CoVs spike protein attachment. Protein-protein interactions also show that the mutants Lys^{417}Val, Lys^{417}Phe, Lys^{417}Pro, Lys^{417}Thr and Lys^{444}Thr in which the basic group is removed, support the importance of the binding of the carboxyl group of the Asp^{30}, Glu^{35} and Asp^{54} with the basic Lys^{417} and Lys^{444} for the proper positioning of the RBD on the receptor surface exhibiting a very low affinity with ACE2. The results indicate that these positive and negative charges of all the nine subdomains expect hCoV-RBD subdomain Ala^{315}-Tyr^{675} are directly involved in the formation of salt bridges in stabilizing ACE2-spike protein interactions which is higher in SARS-CoV 2 compared to other viral species used in this study. This may be the reason why only SARS-CoV 2 and SARS-CoV RBDs were recognized by SARS-CoV RBD-specific, but not MCoV RBD-specific, polyclonal antibodies, whereas only MCoV RBD was recognized by MCoV RBD-immunized polyclonal antibodies, suggesting the cross-reactivity of SARS-CoV RBD-specific antibodies with SARS-CoV 2 RBD protein [42]. On the other side both full length and hCoV-RBD subdomain Ala^{315}-Tyr^{675}, binds away from the N-terminal helices of ACE2 which is in correlation with previous study demonstrating that hCoV uses certain types of O-acetylated sialic acid residues on glycoproteins to initiate the infection of host cells. Studies also reveal that HKU1 is only one of the six hCoVs identified with an unidentified cellular receptor [43]. The same way, MCoV also shows S-mediated attachment to sialo-

Table 4
B-cell epitope sequences of CoV spike protein subdomains predicted using ElliPro web server.

Species	Subdomain	Epitope	Score
SARS	SARS-CoV 2 RBD Pro^{322}-Thr^{581}	^{545}GLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT^{581}	0.78
Bat	Bat-CoV RaTG13 Thr^{581}-Thr^{323}	^{532}GLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKT^{568}	0.78
SARS	SARS-CoV Pro^{309}-Pro^{575}	^{454} DISNVPFSPDGKPCTPPALNCYWP^{477}	0.78
Bat	BtRs-Beta CoV Ser^{311}-Thr^{568}	^{532}GLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKT^{568}	0.78
Bat	BtRs-CoV related Arg^{306}-Pro^{575}	^{532}GLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKT^{568}	0.78
Pangolin	PCoV RBD Gln^{319}-Ser^{589}	^{319}QPTIS^{323}	0.78
MERS	MCoV RBD Gly^{372}-Val^{616}	^{593}DTKIASQLGNCVEYSLYGVSGRGV^{616}	0.78
Human	hCoV-HKU1 RBD Ala^{315}-Tyr^{675}	^{557}GVLDGSYNVSCLCSTDAFLG^{576}	0.84
Avian	ACoV RBD Asp^{250}-Gln^{489}	^{414}DFGTAMYSVKSA^{425}	0.77
Pig	PEDV1 CoV RBM Ala^{315}-Tyr^{675}	^{605}GYPEFGSGVKFTSLYFQFTKGELITGTPKPLEGVTDVSFMTLDVC^{649}	0.78

sides and entry into human airway epithelial cells [44]. In addition, studies also shown that corona viruses that belong to group-I namely, human coronavirus-229E (HCoV-229E), feline infectious peritonitis virus (FIPV), canine coronavirus (CCoV), transmissible gastroenteritis virus (TGEV), and porcine epidemic diarrhea virus (PEDV), are known to commonly use the aminopeptidase N (APN) of their natural host species as a functional receptor for virus entry [45–49]. This shows that these four viral species hCoV, MCoV, ACoV and PEDV1-CoV spike proteins may not prefer ACE2 receptor as a viral host attachment. However, these subdomains hCoV RBD subdomain Ala^{315}-Tyr^{675}, MCoV RBD Gly^{372}-Val^{616}, ACoV RBD Asp^{250}-Gln^{489} and PEDV1-CoV RBD Ala^{315}-Tyr^{675} may also show cross neutralization against SARS-CoV 2 viral infection since they bind on the surface of N-terminal alpha helices of ACE2 receptor. This may be supported by the previous data showing SARS-CoV RBD and MERS-CoV RBD efficiently induce production of neutralizing antibodies [50,51].

3.4. Sequence analysis of epitopic regions

Further sequence analysis of all epitopic sequences shows that Leu^{552} of SARS-CoV 2 which is highly conserved in BtRsRaTG13-CoV, BtRsBeta-CoV, BtRsCoV-related and hCoV is replaced with phenylalanine in PEDV1-CoV and isoleucine in both, MCoV and SARS-CoV. The residue Gln^{564} of SARS-CoV 2 which is also highly conserved in BtRsRaTG13-CoV, BtRsBeta-CoV, BtRsCoV-related and PEDV1-CoV is replaced by glutamic acid, serine, and proline in ACoV, hCoV and SARS-CoV. Another critical residue Phe^{565} which is highly conserved in BtRsRaTG13-CoV, BtRsBeta-CoV, BtRsCoV-related, ACoV, PEDV1-CoV and SARS-CoV is replaced by tyrosine and leucine in hCoV and MCoV. The residue Thr^{573} of SARS-CoV 2 which is conserved in BtRsRaTG13-CoV, BtRsBeta-CoV, BtRsCoV-related, hCoV and SARS-CoV is replaced with serine, proline, and glycine in ACoV, PEDV1-CoV and MCoV. Finally, the residue Val^{576} of SARS-CoV 2 which is conserved in BtRsRaTG13-CoV, BtRsBeta-CoV, BtRsCoV-related, ACoV is replaced with phenylalanine in hCoV, leucine in PEDV1-CoV, glycine in MCoV and alanine in SARS-CoV respectively (Table. 4). These conserved mutations along with variable amino acids may show impact on the cross neutralization of antibodies against SARS-CoV 2 showing unique structural features of the spike glycoprotein RBD of SARS-CoV 2 that confers potentially higher affinity binding for its receptor than found with other CoV viral species. These results show that the epitopic region from Gln^{319}-Ser^{323} of PCoV RBD Gln^{319}-Ser^{589} is only five residues with only Pro^{320} conserved with respect to SARS-CoV 2 RBD Pro^{322}-Thr^{581} epitope. On the other hand, SARS-CoV 2 RBD Pro^{322}-Thr^{581} epitope show high variations with other four epitopes of hCoV RBD subdomain Ala^{315}-Tyr^{675} from Gly^{262}-Gly^{280}, MCoV RBD Gly^{372}-Val^{616} from Asp^{322}-Val^{345}, ACoV RBD Asp^{250}-Gln^{489} from Asp^{414}-Ala^{425}, and PEDV1-CoV RBD Ala^{315}-Tyr^{675} from Gly^{283}-Cys^{327} used in the study (Fig. 5B). This predicts that the epitopes of these four subdomains along the epitope of PCoV RBD Gln^{319}-Ser^{589} might be effective in a wide range in inducing antibodies for cross neutralization against SARS-CoV 2 spike protein attachment with its receptor ACE2. Previous data also confirms two-way antigenic cross reactivity between SARS-CoV and porcine group 1 CoVs through group 1 CoV N proteins and not the S protein [52]. More importantly, the epitopic sequences of six subdomains from SARS-CoV 2 RBD Pro^{322}-Thr^{581}, BtRsRaTG13-CoV RBD Thr^{581}-Thr^{323}, SARS-CoV RBD Pro^{309}-Pro^{575}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568} and BtRsCoV-related Arg^{306}-Pro^{575} might play an important role as antigenic determinants against SARS-CoV 2 viral infection and cross neutralization.

The results show that the epitopic regions from Gly^{545}-Thr^{581} of SARS-CoV 2 and BtRsRaTG13-CoV RBD Thr^{323}-Thr^{581} are highly conserved with four variations in Leu^{560}, Ala^{570}, Ala^{575} and Gln^{580} in comparison to other two epitopic regions from Gly^{532}-Thr^{568} in BtRsBeta-CoV RBD Ser^{311}-Thr^{568} and from Gly^{531}-Thr^{567} in BtRsCoV-related Arg^{306}-Pro^{575} respectively. Expect Pro^{469} that is highly conserved, the epitopic region between Asp^{454}-Pro^{477} that was predicted in SARS-CoV RBD Pro^{309}-Pro^{575} show high variations with other five epitopic regions that were predicted in other five subdomains (Fig. 5A). Also, the epitopic region of SARS-CoV RBD Pro^{309}-Pro^{575} show a total of five prolines at Pro^{459}, Pro^{462}, Pro^{466}, Pro^{469}, Pro^{470} and Pro^{477} which might be responsible for increasing the flexibility of SARS-CoV RBM. This allows lesser binding interaction than SARS-CoV 2 with hACE2 and shows distinct epitopic features in cross neutralization studies. This may be the reason why HEK293T cells when transfected with pCAGGS plasmids containing Flag-tagged SARS-CoV S or SARS-CoV 2 S show difference in the electrostatic surface potential maps leading to different immunogenic properties of the RBD subdomains [53]. Previous results also demonstrate that most SARS-CoV RBD-specific antibodies could cross-neutralize SARS like-CoV strain WIV1 from Bat [54]. Studies also identified human mAb S309 with broad neutralizing activity binding to N343-glycan (N330 in SARS-CoV S) epitope in the RBD domain which is in correlation with our SARS-CoV 2 RBD subdomain from Pro^{322}-Thr^{581} [55]. This shows that the predicted subdomains might induce antibodies that binds to spike epitopes and shows cross neutralization. This is also supported by previous findings showing cross neutralization of antibody binding to the epitopes of SARS-CoV 2 spike protein 10-fold greater that was isolated from hyperimmune horse anti-SARS-CoV serum. Even SARS patient sera or rabbit hyperimmune sera also show cross neutralization on SARS-CoV 2 pseudo virus carrying spike protein in a limited level [10,56–58]. Recent cross-neutralizing data have also indicated that only one out of 15 SARS-CoV 2-infected patients was able to show cross reactive response weekly between SARS-CoV 2 and SARS-CoV viruses [59]. These results based on both computational and experimental clearly indicate that these five-proline shown above play an important role in both binding affinity and immunogenic properties in cross neutralization studies between SARS-CoV and SARS-CoV2.

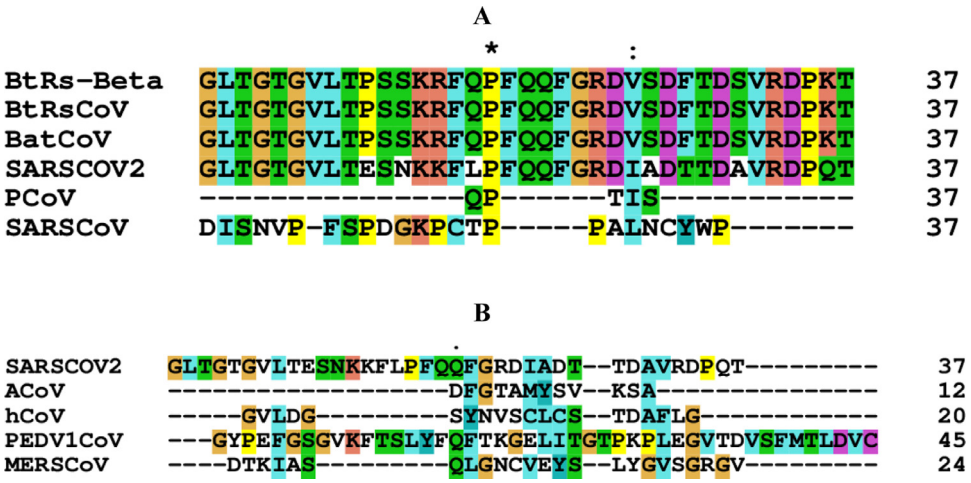

Fig. 5. Sequence analysis of the epitopic regions of BtRsRaTG13-CoV RBD Thr^{323}-Thr^{581}, BtRsBeta-CoV RBD Ser^{311}-Thr^{568}, BtRsCoV-related Arg^{306}-Pro^{575}, PCoV RBD Gln^{319}-Ser^{589} and SARS-CoV RBD Pro^{309}-Pro^{575} in comparison to SARS-CoV 2 **(Fig. A)** and with ACoV RBD Asp^{250}-Gln^{489} , hCoV-RBD subdomain Ala^{315}-Tyr^{675}, PEDV1 CoV RBD Ala^{315}-Tyr^{675} and MCoV RBD Gly^{372}-Val^{616} in comparison to SARS-CoV 2 **(Fig. B)** predicted using Clustal X software suite.

4. Conclusion

There is a health and medical emergency to control the rapid and global ever-growing SARS-CoV 2 transmission and infection. Since we are at the beginning of understanding the immune responses to the virus and due to lack of knowledge, we may need to use our previous experiences with coronaviruses along with *in silico* approaches to design vaccines as the ultimate way to protect healthy individuals. In this study, we have comprehensively compared the sequences of spike protein from 10 different coronaviruses in the context of their interaction with ACE2 to identify the best subdomain of spike protein to be used for vaccine development. Although a full-length S protein may be a better candidate to induce immunity, a more focused immune induction based on an immunogenic part of S protein may warrant a stronger and more efficient vaccination outcome, while it significantly reduces the chance of development of antibody dependent enhancement. In addition, industrial concerns support that the use of a shorter version of target antigen may be easier, faster, and more cost-efficient to be manufactured at the speed and large scale that is urgently required for the present pandemic. Although several vaccines have already been developed based on a full-length spike protein, this study suggests a shorter version of spike protein as a vaccine candidate with the same or even better immunogenicity because of its shorter length. In fact, vaccines that are designed based on shorter peptides have several advantages over longer peptides. First, a focused immune response against an essential component of a virus is much more favorable since it reduces the diversion or extension of the immune response toward less immunodominant segment of a target protein. Second, shorter peptides may reduce the chance of producing non-neutralizing or weakly-neutralizing antibodies, which can potentially facilitate viral entry through cellular FC receptor, even in cells without ACE2. This could result in a serious vaccine side effect, antibody dependent enhancement, which has been reported for respiratory syncytial virus in 1960s [60]). Third, shorter peptide can be easily scaled up and are less costly to manufacture compared to longer peptides. This is a critical industrial concern when large quantities of vaccine doses are required as such in the current SARS-CoV 2 pandemic.

This is the first *in silico* study that comprehensively compares the RBD subdomain of spike protein from ten closely related coronaviruses and their interaction with ACE2. Our protein-protein docking study identifies a short RBD subdomain of SARS-CoV 2 spike protein from Pro^{322} to Thr^{581} as the main binding site, interacting with ACE2. The current results in comparison to previous studies also indicate that SARS-CoV 2 RBD amino acids both in the full-length and suddomain Arg^{403}, Glu^{406}, Lys^{417}, Lys^{444}, Tyr^{453}, Gln^{474}, Gln^{498}, Thr^{500}, Asn^{501}, and Tyr^{505} from SARS-CoV 2 spike and Gln^{24,} Asp^{30}, Glu^{35}, His^{34}, Tyr^{41,} Asn^{49} and Lys^{353} from ACE2 acts as common pharmacophores with stronger hydrogen bonds [61]. This 260aa peptide has very high potential to be used as an efficient vaccine candidate for SARS-CoV 2. Our study demonstrates that both RBD subdomain and full-length spike protein of SARS-CoV 2 binds to ACE2 with a similar but higher affinity in comparison to that of other coronaviruses including BtRsRaTG13-CoV, BtRsBeta-CoV, PCoV, MCoV, ACoV, and PEDV1-CoV. This suggest that we might be able to design a universal vaccine that could induce cross-reactive neutralizing antibodies, which are capable of inhibiting entry of several closely related coronaviruses. These antibodies can also be produced *ex vivo* to be used as therapeutics in coronavirus infection such as COVID19. In addition, such a detailed study empowers us for an efficient and quick design or re-design of vaccine candidates to prevent future pandemic that might be caused by emerging or remerging coronaviruses infection. Taken together, this study provides an essential foundation for the design and development of SARS-CoV 2 RBD Pro^{322}-Thr^{581}-based vaccines and therapeutics while it may also be beneficial for infections caused by other coronaviruses.

Author's contributions

Nataraj Sekhar Pagadala performed the complete study, processed information, interpreted results and written the manuscript. Dr. Amir Landi interpreted the results and written the manuscript. Dr. Paramahamsa Maturu interpreted the results and written the manuscript. Prof. Jack Tuszynski interpreted the results and written the manuscript.

Declaration of Competing Interest

No potential conflict of interest was reported by the authors.

References

[1] J. Cui, F. Li, Z.L. Shi, Origin and evolution of pathogenic coronaviruses, Nat. Rev. Microbiol. 17 (3) (2019) 181–192.
[2] P.C. Woo, Y. Huang, S.K. Lau, K.Y. Yuen, Coronavirus genomics and bioinformatics analysis, Viruses 2 (8) (2010) 1804–1820.

[3] Y. Wu, X. Xu, Z. Chen, J. Duan, K. Hashimoto, L. Yang, C. Liu, C. Yang, Nervous system involvement after infection with COVID-19 and other coronaviruses, Brain Behav. Immun. 87 (2020) 18–22.
[4] V.M. Corman, J. Lienau, M. Witzenrath, Coronaviruses as the cause of respiratory infections, Internist 60 (11) (2019) 1136–1145.
[5] A.R. Fehr, R. Channappanavar, S. Perlman, Middle East respiratory syndrome: emergence of a pathogenic human coronavirus, Ann. Rev. Med. 68 (2017) 387–399.
[6] E. de Wit, N. van Doremalen, D. Falzarano, V.J. Munster, SARS and MERS: recent insights into emerging coronaviruses, Nat. Rev. Microbiol. 14 (8) (2016) 523–534.
[7] Y. Guan, B.J. Zheng, Y.Q. He, X.L. Liu, Z.X. Zhuang, C.L. Cheung, S.W. Luo, P.H. Li, L.J. Zhang, Y.J. Guan, K.M. Butt, K.L. Wong, K.W. Chan, W. Lim, K.F. Shortridge, K.Y. Yuen, J.S. Peiris, L.L. Poon, Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China, Science 302 (5643) (2003) 276–278.
[8] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.
[9] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (10223) (2020) 470–473.
[10] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (7798) (2020) 270–273.
[11] S. Belouzard, J.K. Millet, B.N. Licitra, G.R. Whittaker, Mechanisms of coronavirus cell entry mediated by the viral spike protein, Viruses 4 (6) (2012) 1011–1033.
[12] Y. Xu, Y. Liu, Z. Lou, L. Qin, X. Li, Z. Bai, H. Pang, P. Tien, G.F. Gao, Z. Rao, Structural basis for coronavirus-mediated membrane fusion. Crystal structure of mouse hepatitis virus spike protein fusion core, J. Biol. Chem. 279 (29) (2004) 30514–30522.
[13] B.J. Bosch, R. van der Zee, C.A. de Haan, P.J. Rottier, The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, J. Virol. 77 (16) (2003) 8801–8811.
[14] H. Hofmann, K. Pyrc, L. van der Hoek, M. Geier, B. Berkhout, S. Pohlmann, Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry, Proc. Natl. Acad. Sci. U.S.A. 102 (22) (2005) 7988–7993.
[15] W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe, M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, Nature 426 (6965) (2003) 450–454.
[16] P. Wang, J. Chen, A. Zheng, Y. Nie, X. Shi, W. Wang, G. Wang, M. Luo, H. Liu, L. Tan, X. Song, Z. Wang, X. Yin, X. Qu, X. Wang, T. Qing, M. Ding, H. Deng, Expression cloning of functional receptor used by SARS coronavirus, Biochem. Biophys. Res. Commun. 315 (2) (2004) 439–444.
[17] M. Iwai, M. Horiuchi, Devil and angel in the renin-angiotensin system: ACE-angiotensin II-AT1 receptor axis vs. ACE2-angiotensin-(1-7)-Mas receptor axis, Hypertens. Res. 32 (7) (2009) 533–536.
[18] D.W. Lambert, N.M. Hooper, A.J. Turner, Angiotensin-converting enzyme 2 and new insights into the renin-angiotensin system, Biochem. Pharmacol. 75 (4) (2008) 781–786.
[19] I. Glowacka, S. Bertram, M.A. Muller, P. Allen, E. Soilleux, S. Pfefferle, I. Steffen, T.S. Tsegaye, Y. He, K. Gnirss, D. Niemeyer, H. Schneider, C. Drosten, S. Pohlmann, Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response, J. Virol. 85 (9) (2011) 4122–4134.
[20] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, F. Taguchi, Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2, J. Virol. 84 (24) (2010) 12658–12664.
[21] A. Shulla, T. Heald-Sargent, G. Subramanya, J. Zhao, S. Perlman, T. Gallagher, A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry, J. Virol. 85 (2) (2011) 873–882.
[22] F.A. Rabi, M.S. Al Zoubi, G.A. Kasasbeh, D.M. Salameh, A.D. Al-Nasser, SARS–CoV-2 and Coronavirus disease 2019: what we know so far, Pathogens 9 (3) (2020).
[23] X. Li, E.E. Giorgi, M.H. Marichann, B. Foley, C. Xiao, X.P. Kong, Y. Chen, B. Korber, F. Gao, Emergence of SARS-CoV-2 through Recombination and Strong Purifying Selection, bioRxiv 2 (2020) Preprint.
[24] K.G. Andersen, A. Rambaut, W.I. Lipkin, E.C. Holmes, R.F. Garry, The proximal origin of SARS-CoV-2, Nat. Med. 26 (4) (2020) 450–452.
[25] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2, Science 367 (6485) (2020) 1444–1448.
[26] V.J. Munster, M. Koopmans, N. van Doremalen, D. van Riel, E. de Wit, A Novel Coronavirus emerging in China - key questions for impact assessment, N. Engl. J. Med. 382 (8) (2020) 692–694.
[27] F. Li, W. Li, M. Farzan, S.C. Harrison, Structure of SARS coronavirus spike receptor-binding domain complexed with receptor, Science 309 (5742) (2005) 1864–1868.
[28] W. Li, C. Zhang, J. Sui, J.H. Kuhn, M.J. Moore, S. Luo, S.K. Wong, I.C. Huang, K. Xu, N. Vasilieva, A. Murakami, Y. He, W.A. Marasco, Y. Guan, H. Choe, M. Farzan, Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2, EMBO J. 24 (8) (2005) 1634–1643.
[29] E. Callaway, The race for coronavirus vaccines: a graphical guide, Nature 580 (7805) (2020) 576–577.
[30] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (3) (1993) 779–815.
[31] S.B. Needleman, C.D. Wunsch, A general method applicable to the search for similarities in the amino acid sequence of two proteins, J. Mol. Biol. 48 (3) (1970) 443–453.
[32] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucl. Acids. Res. 22 (22) (1994) 4673–4680.
[33] B.G. Pierce, K. Wiehe, H. Hwang, B.H. Kim, T. Vreven, Z. Weng, ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers, Bioinformatics 30 (12) (2014) 1771–1773.
[34] B.G. Pierce, Y. Hourai, Z. Weng, Accelerating protein docking in ZDOCK using an advanced 3D convolution library, PLoS One 6 (9) (2011) e24657.
[35] J. Mintseris, B. Pierce, K. Wiehe, R. Anderson, R. Chen, Z. Weng, Integrating statistical pair potentials into protein complex prediction, Proteins 69 (3) (2007) 511–520.
[36] R. Chen, L. Li, Z. Weng, ZDOCK: an initial-stage protein-docking algorithm, Proteins 52 (1) (2003) 80–87.
[37] R. Mosca, C. Pons, J. Fernandez-Recio, P. Aloy, Pushing structural information into the yeast interactome by high-throughput protein docking experiments, PLoS Comput. Biol. 5 (8) (2009) e1000490.
[38] B. Pierce, Z. Weng, ZRANK: reranking protein docking predictions with an optimized energy function, Proteins 67 (4) (2007) 1078–1086.
[39] S. Saha, G.P. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins 65 (1) (2006) 40–48.
[40] M.H.V. Van Regenmortel, Mapping epitope structure and activity: from one-dimensional prediction to four-dimensional description of antigenic specificity, Methods 9 (3) (1996) 465–472.
[41] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, Nature 581 (7807) (2020) 215–220.
[42] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine, Cell Mol. Immunol. 17 (6) (2020) 613–620.
[43] X. Huang, W. Dong, A. Milewska, A. Golda, Y. Qi, Q.K. Zhu, W.A. Marasco, R.S. Baric, A.C. Sims, K. Pyrc, W. Li, J. Sui, Human Coronavirus HKU1 spike protein uses O-Acetylated Sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme, J. Virol. 89 (14) (2015) 7202–7213.
[44] Y.J. Park, A.C. Walls, Z. Wang, M.M. Sauer, W. Li, M.A. Tortorici, B.J. Bosch, F. DiMaio, D. Veesler, Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors, Nat. Struct. Mol. Biol. 26 (12) (2019) 1151–1157.
[45] B. Delmas, J. Gelfi, R. L'Haridon, L.K. Vogel, H. Sjostrom, O. Noren, H. Laude, Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV, Nature 357 (6377) (1992) 417–420.
[46] A.F. Kolb, A. Hegyi, J. Maile, A. Heister, M. Hagemann, S.G. Siddell, Molecular analysis of the coronavirus-receptor function of aminopeptidase N, Adv. Exp. Med. Biol. 440 (1998) 61–67.
[47] B.X. Li, J.W. Ge, Y.J. Li, Porcine aminopeptidase N is a functional receptor for the PEDV coronavirus, Virology 365 (1) (2007) 166–172.
[48] D.B. Tresnan, K.V. Holmes, Feline aminopeptidase N is a receptor for all group I coronaviruses, Adv. Exp. Med. Biol. 440 (1998) 69–75.
[49] S.K. Wong, W. Li, M.J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2, J. Biol. Chem. 279 (5) (2004) 3197–3201.
[50] L. Du, Y. He, Y. Zhou, S. Liu, B.J. Zheng, S. Jiang, The spike protein of SARS-CoV–a target for vaccine and therapeutic development, Nat. Rev. Microbiol. 7 (3) (2009) 226–236.
[51] L. Wang, W. Shi, M.G. Joyce, K. Modjarrad, Y. Zhang, K. Leung, C.R. Lees, T. Zhou, H.M. Yassine, M. Kanekiyo, Z.Y. Yang, X. Chen, M.M. Becker, M. Freeman, L. Vogel, J.C. Johnson, G. Olinger, J.P. Todd, U. Bagci, J. Solomon, D.J. Mollura, L. Hensley, P. Jahrling, M.R. Denison, S.S. Rao, K. Subbarao, P.D. Kwong, J.R. Mascola, W.P. Kong, B.S. Graham, Evaluation of candidate vaccine approaches for MERS-CoV, Nat. Commun. 6 (2015) 7712.
[52] A.N. Vlasova, X. Zhang, M. Hasoksuz, H.S. Nagesha, L.M. Haynes, Y. Fang, S. Lu, L.J. Saif, Two-way antigenic cross-reactivity between severe acute respiratory syndrome coronavirus (SARS-CoV) and group 1 animal CoVs is mediated through an antigenic site in the N-terminal region of the SARS-CoV nucleoprotein, J. Virol. 81 (24) (2007) 13365–13377.
[53] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K.Y. Yuen, Q. Wang, H. Zhou, J. Yan, J. Qi, Structural and functional basis of SARS-CoV-2 entry by using human ACE2, Cell 181 (4) (2020) 894–904 e9.
[54] L.P. Zeng, X.Y. Ge, C. Peng, W. Tai, S. Jiang, L. Du, Z.L. Shi, Cross-neutralization of SARS coronavirus-specific antibodies against bat SARS-like coronaviruses, Sci. China Life Sci. 60 (12) (2017) 1399–1402.
[55] D. Pinto, Y.J. Park, M. Beltramello, A.C. Walls, M.A. Tortorici, S. Bianchi, S. Jaconi, K. Culap, F. Zatta, A. De Marco, A. Peter, B. Guarino, R. Spreafico, E. Cameroni,

J.B. Case, R.E. Chen, C. Havenar-Daughton, G. Snell, A. Telenti, H.W. Virgin, A. Lanzavecchia, M.S. Diamond, K. Fink, D. Veesler, D. Corti, Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody, Nature (2020).

[56] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Muller, C. Drosten, S. Pohlmann, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, Cell 181 (2) (2020) 271–280 e8.

[57] M. Yu, V. Stevens, J.D. Berry, G. Crameri, J. McEachern, C. Tu, Z. Shi, G. Liang, H. Weingartl, J. Cardosa, B.T. Eaton, L.F. Wang, Determination and application of immunodominant regions of SARS coronavirus spike and nucleocapsid proteins recognized by sera from different animal species, J. Immunol. Methods 331 (1–2) (2008) 1–12.

[58] D.E. Anderson, C.W. Tan, W.N. Chia, B.E. Young, M. Linster, J.H. Low, Y.J. Tan, M.I. Chen, G.J.D. Smith, Y.S. Leo, D.C. Lye, L.F. Wang, Lack of cross-neutralization by SARS patient sera towards SARS-CoV-2, Emerg. Microbes Infect. 9 (1) (2020) 900–902.

[59] H. Lv, N.C. Wu, O.T. Tsang, M. Yuan, R. Perera, W.S. Leung, R.T.Y. So, J.M.C. Chan, G.K. Yip, T.S.H. Chik, Y. Wang, C.Y.C. Choi, Y. Lin, W.W. Ng, J. Zhao, L.L.M. Poon, J.S.M. Peiris, I.A. Wilson, C.K.P. Mok, Cross-reactive antibody response between SARS-CoV-2 and SARS-CoV infections, Cell Rep. 31 (9) (2020) 107725.

[60] M.F. Delgado, S. Coviello, A.C. Monsalvo, G.A. Melendi, J.Z. Hernandez, J.P. Batalle, L. Diaz, A. Trento, H.Y. Chang, W. Mitzner, J. Ravetch, J.A. Melero, P.M. Irusta, F.P. Polack, Lack of antibody affinity maturation due to poor Toll–like receptor stimulation leads to enhanced respiratory syncytial virus disease, Nat. Med. 15 (1) (2009) 34–41.

[61] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus, J. Virol. 94 (7) (2020) e00127-20.