Biochemical and Biophysical Research Communications xxx (xxxx) xxx

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Structure analysis of the receptor binding of 2019-nCoV

Yun Chen [a, 1], Yao Guo [a, 1], Yihang Pan [a, **], Zhizhuang Joe Zhao [a, b, *]

[a] *Edmond H. Fischer Translational Medical Research Laboratory, Precision Medicine Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, 518107, Guangdong, China*
[b] *Department of Pathology, University of Oklahoma Health Sciences Center, 940 Stanton L. Young Blvd, Oklahoma City, OK, 73104, USA*

ARTICLE INFO

Article history:
Received 6 February 2020
Accepted 10 February 2020
Available online xxx

Keywords:
Coronavirus
Sequence analysis
Structure analysis
Molecular modeling

ABSTRACT

2019-nCoV is a newly identified coronavirus with high similarity to SARS-CoV. We performed a structural analysis of the receptor binding domain (RBD) of spike glycoprotein responsible for entry of coronaviruses into host cells. The RBDs from the two viruses share 72% identity in amino acid sequences, and molecular simulation reveals highly similar ternary structures. However, 2019-nCoV has a distinct loop with flexible glycyl residues replacing rigid prolyl residues in SARS-CoV. Molecular modeling revealed that 2019-nCoV RBD has a stronger interaction with angiotensin converting enzyme 2 (ACE2). A unique phenylalanine F486 in the flexible loop likely plays a major role because its penetration into a deep hydrophobic pocket in ACE2. ACE2 is widely expressed with conserved primary structures throughout the animal kingdom from fish, amphibians, reptiles, birds, to mammals. Structural analysis suggests that ACE2 from these animals can potentially bind RBD of 2019-nCoV, making them all possible natural hosts for the virus. 2019-nCoV is thought to be transmitted through respiratory droplets. However, since ACE2 is predominantly expressed in intestines, testis, and kidney, fecal-oral and other routes of transmission are also possible. Finally, antibodies and small molecular inhibitors that can block the interaction of ACE2 with RBD should be developed to combat the virus.

© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

A mysterious pneumonia illness was first reported in late December 2019 in Wuhan, China, and has rapidly spread to a dozen of countries including the United States with thousands of infected individuals and hundreds of deaths within a month [1]. Scientists in China have isolated the virus from patients and determined its genetic code. The pathogen responsible for this epidemic is a new coronavirus designated 2019-nCoV by the World Health Organization. 2019-nCoV belongs to the same family of viruses as the well-known severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), which have killed hundreds of people in the past 17 years.

Coronaviruses consist of a large diverse family of viruses. They can be classified into four genera: Alpha-, Beta-, Gamma-, and Delta coronavirus [2,3]. Representative alphacoronaviruses include human coronavirus NL63 (HCoV-NL63), while the betacoronaviruses include the best-known SARS-CoV and MERS-CoV. Based on nucleic acid sequence similarity, the newly identified 2019-nCoV is a betacoronavirus. The entry of all coronaviruses into host cells is mediated by spike glycoprotein that gives coronaviruses a crown-like appearance by forming spikes on their surface. The amino acid sequence of spike glycoprotein consists of a large ectodomain, a single-pass transmembrane anchor, and a short C-terminal intracellular tail [3]. The ectodomain contains a receptor-binding unit S1 and a membrane-fusion unit S2. Electron microscopic imaging illustrated that spike glycoprotein forms a clove-shaped spike with three S1 heads and a trimeric S2 stalk. For a virus to enter a host cell, S1 binds to a specific cell surface receptor via its receptor-binding domain (RBD), and S2 fuses the host cell and viral membranes, enabling the entry of viral genomes into host cells. Specific RBD-receptor binding determines if a cell or animal can be infected and also serves as a target for therapeutic inventions to treat diseases caused by coronaviruses. Previous studies have identified angiotensin converting enzyme 2 (ACE2) as a functional receptor for SARS-CoV [4,5]. In this study, we analyzed the structure of spike

---

* Corresponding author. Edmond H. Fischer Translational Medical Research Laboratory, Precision Medicine Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, 518107, Guangdong, China.
** Corresponding author.
E-mail addresses: yihang73@hotmail.com (Y. Pan), joe-zhao@ouhsc.edu (Z.J. Zhao).
[1] These authors contribute equally.

https://doi.org/10.1016/j.bbrc.2020.02.071
0006-291X/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071

glycoprotein RBD of 2019-nCoV and identified a unique feature that potentially allows a high affinity binding to ACE2 in human cells. We further discussed potential candidates for natural hosts of 2019-nCoV, routes of transmission, and strategies to inhibit virus entry for therapeutic applications.

2. Methods

The genomic sequence of 2019-nCoV as deposited by Wang et al. was downloaded from the GenBank database (MN908947.3). DNA and protein sequences were compared by using the BLAST program. Multiple sequence alignment was performed by using the Clustal Omega program. Three-dimensional structure was analyzed by using the Cn3D program from the NCBI. Protein structure simulation was performed by using Swiss-model based on the co-crystal structure of human ACE2 with the SARS-CoV spike glycoprotein RBD (6, PDB ID 2AJF). ACE2 and RDB interaction was analyzed by molecular docking using the PatchDock and FireDock programs.

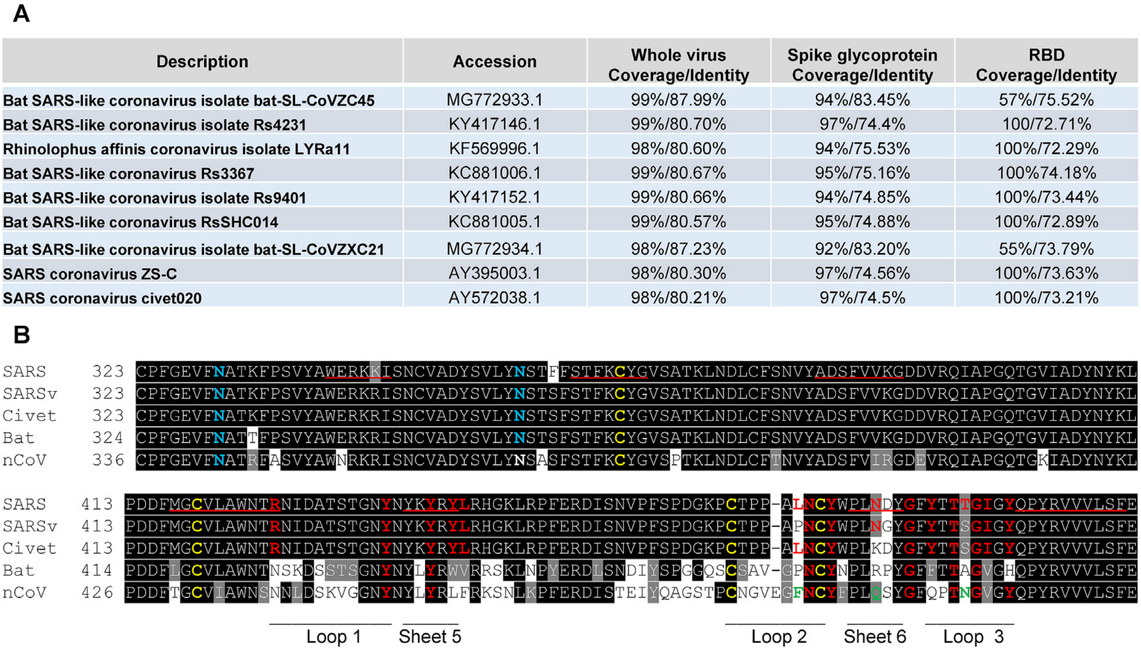
3. Results

3.1. General sequence characteristics of 2019-nCoV in comparison with SARS-CoV

By using the initially reported sequence MN908947.3, a BLAST search of the NCBI database revealed 6 inputs for the virus with essentially identical sequences (accession NC_045512.2, MN908947.3, MN975262.1, MN985325.1, MN988713.1, and MN938384.1). The closest homolog of 2019-nCoV is a SARS-like coronavirus isolated from bat (MG772933.1) with a sequence identity of 87.99% at 99% coverage (Fig. 1A). It also shows 80% sequence identity with SARS coronavirus isolated from human patients or civet with 98% coverage. Throughout the entire 29,903bp genome of 2019-nCoV, the least conserved region encodes for the spike glycoprotein with sequence identity of 74–83%. Spike glycoprotein forms spikes on the surface of coronaviruses and is responsible for entrance of the viruses into the host cells. The RBD in the spike glycoprotein molecule directly binds receptors on the surface of host cells [3]. In the case of SARS-CoV and bat/civet SARS-like CoV, the receptor is ACE2, an exopeptidase that catalyzes the conversion of angiotensin I to the nonapeptide angiotensin1-9 or the conversion of angiotensin II to angiotensin1-7 [3–7]. At the protein level, the whole spike glycoprotein and its RBD share 76% and 72% sequence identity with SARS-CoV, respectively. SARS-CoV spike glycoprotein is known to be glycosylated. A total of 22 predicated N-glycosylation sites is found in spike glycoprotein of 2019-nCoV, which are shared by SARS-CoV except that the latter contains an extra glycosylation site at N370. A detailed sequence alignment of the RBD of SARS-CoV spike glycoprotein with those from closely related coronaviruses at the protein level is shown in Fig. 1B.

3.2. Unique RBD structure in 2019-nCoV

The crystal structure of SARS-CoV RBD in complex with its receptor, human ACE2, has been solved [6]. By performing molecular simulation, we obtained a ternary structure for RBD of 2019-nCoV that is essentially superimposable with that of SARS-CoV (Fig. 2A–C), except for a noted structural variation in a loop (Loop 2). The backbone of the deduced RBD structure consists of 7 beta sheets. Peptide segments involved in the formation of this secondary structure are all highly conserved without the presence of secondary structure breakers. Four cysteinyl residues that form



A

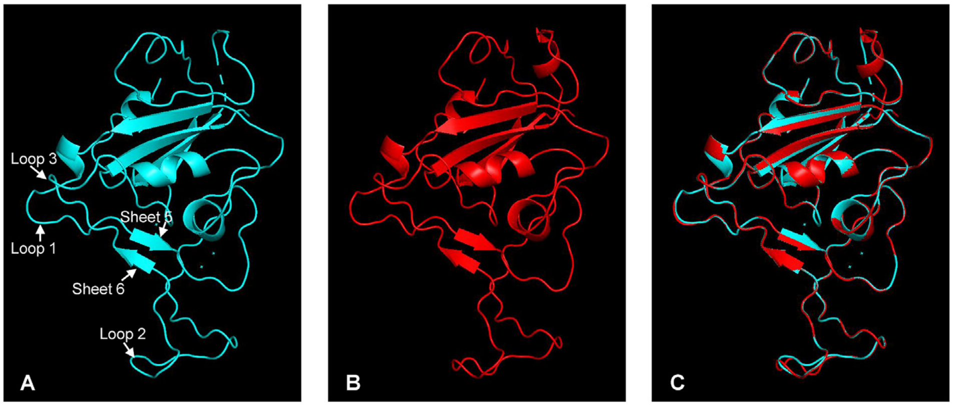
Description	Accession	Whole virus Coverage/Identity	Spike glycoprotein Coverage/Identity	RBD Coverage/Identity
Bat SARS-like coronavirus isolate bat-SL-CoVZC45	MG772933.1	99%/87.99%	94%/83.45%	57%/75.52%
Bat SARS-like coronavirus isolate Rs4231	KY417146.1	99%/80.70%	97%/74.4%	100/72.71%
Rhinolophus affinis coronavirus isolate LYRa11	KF569996.1	98%/80.60%	94%/75.53%	100%/72.29%
Bat SARS-like coronavirus Rs3367	KC881006.1	99%/80.67%	95%/75.16%	100%74.18%
Bat SARS-like coronavirus isolate Rs9401	KY417152.1	99%/80.66%	94%/74.85%	100%/73.44%
Bat SARS-like coronavirus RsSHC014	KC881005.1	99%/80.57%	95%/74.88%	100%/72.89%
Bat SARS-like coronavirus isolate bat-SL-CoVZXC21	MG772934.1	98%/87.23%	92%/83.20%	55%/73.79%
SARS coronavirus ZS-C	AY395003.1	98%/80.30%	97%/74.56%	100%/73.63%
SARS coronavirus civet020	AY572038.1	98%/80.21%	97%/74.5%	100%/73.21%

B

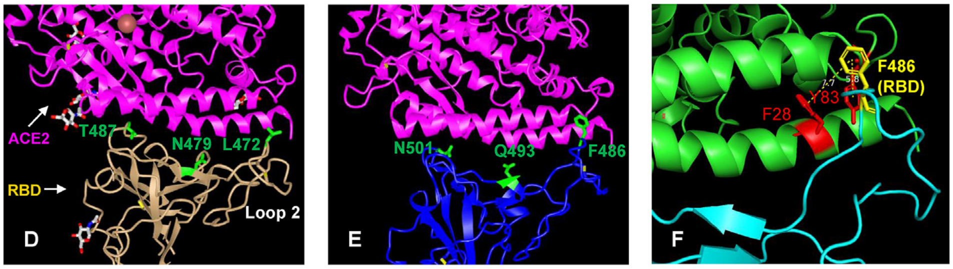
```
SARS     323 CPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKL
SARSv    323 CPFGEVFNATKFPSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKL
Civet    323 CPFGEVFNATKFPSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKL
Bat      324 CPFGEVFNATTFPSVYAWERKRISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKL
nCoV     336 CPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKL

SARS     413 PDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPP-ALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFE
SARSv    413 PDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPP-APNCYWPLNGYGFYTTSGIGYQPYRVVVLSFE
Civet    413 PDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPP-ALNCYWPLKDYGFYTTSGIGYQPYRVVVLSFE
Bat      414 PDDFLGCVLAWNTNSKDSSTSGNYNYLYRWVRRSKLNPYERDLSNDIYSPGGQSCSAV-GPNCYNPLRPYGFFTTAGVGHQPYRVVVLSFE
nCoV     426 PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFE
```

Fig. 1. Comparison of 2019-nCoV with closely related coronaviruses. A. Identity of 2019-nCoV with closely related coronaviruses in nucleotide sequences. Data are given for the percentage of coverage and sequence identity for the whole genome and the coding region of the entire spike glycoprotein and its receptor-binding domain (RBD). **B.** Amino acid sequence alignment for RBD of spike glycoprotein from five closely related coronaviruses. GenBank accession numbers are AY274119 for SARS (SARS-CoV isolated during the initial 2002–2003 SARS outbreak, ref. 8), AY525636 for SARSv (SARS-CoV isolated during the mild 2003–2004 outbreak, ref. 9), AY304486 for civet (SARS-CoV-like virus isolated from palm civets, ref. 10), AGZ48806.1 for bat (SARS-CoV-like coronavirus RsSHC014 isolated from bat, ref. 11), and MN908947.3 for nCoV (2019-nCoV). Amino acid residues interacting with ACE2 in the SARS-CoV RBD and conserved ones in other viruses are highlighted in red. Major secondary structures are underlined (double underlines for β-sheet). Major altered amino acids in 2019-nCoV are highlighted in green. Cysteine residues forming disulfide bonds and putative N-linked glycosylation sites are highlighted in yellow and cyan, respectively.

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071



SARS-CoV RBD 2019-nCov RBD Superimpose



SARS-CoV RBD/ACE2 2019-nCov RBD/ACE2 2019-nCov RBD/ACE2

Fig. 2. Comparison of the known structures SARS-CoV RBD and its ACE2 complex with the deduced molecular models of 2019-nCoV RBD. Some structure segments and key amino acid residues are indicated.

disulfide bonds (corresponding to C366/C418 and C467/C474 of SARS-CoV) are also conserved (see also Fig. 1B). We furthermore performed molecular docking to examine the binding of RBD with ACE2. The deduced complex structure reveals similar mode of extensive interaction as seen with SARS-CoV with a more favorable binding energy (−21.82 21 vs. −13.38 kcal/mol) (Fig. 2D–F). The contact between ACE2 and RBD involves two β-sheets and three loops (see Figs. 1B and 2A). There are 16 amino acid residues in SARS-CoV RBD that are directly in contact with ACE2, of which 8 are conserved in 2019-nCoV (see Fig. 1B). Presumably, the substituted amino acids can either reduce or enhance the interaction.

To define the contribution of the variant amino acids to the RBD/ACE2 interaction, we compare the sequences of RBD from three other SARS-CoV-associated viruses (Fig. 1B). These include coronaviruses isolated from patients during a short, weak SARS outbreak in 2003–2004 (denoted SARSv here) and from palm civets and bats, possible sources of SARS-CoV found in humans [8–11]. Recombinant proteins containing RBD of these 3 viruses are all known to bind to human ACE2 [11,12]. In comparison with SARS-CoV (responsible for the major SRAS outbreak during the 2002–2003), binding with SARSv and civet RBDs is substantially weaker [12], while quantitative binding affinity with bat RBD has been not been determined [11]. Amino acid residues in the RBD/ACE2 binding interface plays a crucial role in determining the binding affinity. Among the 16 amino acid residues in RBD of SARS that are in contact with ACE2, 14, 14, 7, and 8 are shared by SARSv, civet, bat, and 2019-nCoV, respectively (Fig. 1B). N479 found in both SARS viruses isolated from human patients is changed to K and R in civet and bat, respectively. An earlier study demonstrated that an N479 to K substitution resulted in significantly lower affinity (30-fold increase in Kd values) [12]. Interestingly, this amino acid is substituted by a similar amino acid glutamine (Q493) in 2019-nCoV, which also contains an amide group but at an extended position, which can potentially carry out similar functions. In comparison with SARS-CoV, T487 is changed to asparagine (N501) in 2019-nCoV but alanine or serine in the other viruses. It has been shown that a T487 to S substitution increased Kd by 20-fold, suggesting the methyl group rather than the hydroxyl group in this threonine residue is more important for the interaction [12]. It is hard to predict if N501 with an amide group can confer a better interaction. Hydrophobic amino acid L472 is also important for interaction between RBD and ACE2. Interestingly, it is substituted by proline in SARSv and phenylalanine in 2019-nCoV (corresponding to F486). L472 is located in a loop formed by disulfide bond C467/C474. Interestingly, this loop with CTPPALNC in SARS-CoV is replaced by CNGVEGFNC in 2019-nCoV containing one extra amino acid residue and totally different amino acid compositions. The replacement of two proline residues by two flexible glycine residues converts a rigid structure to a very flexible one. Further examination of the deduced RBD/ACE2 complex structure reveals that this unique phenylalanine F486 in the flexible loop can penetrate deep into a hydrophobic pocket in ACE2 formed by F28, L79, Y83, and L97 (Fig. 2F). The presence of two aromatic amino acids in the pocket may provide additional binding force via π-stacking interactions [13]. Taken together, 2019-nCoV likely has a stronger binding to ACE2 via its spike glycoprotein.

Glycosylation may also affect the interaction of RBD with ACE2. Among the 23 glycosylation sites on spike glycoprotein, two are in RBD (Fig. 1B). Glycosylation has been detected on one of these residues, Asn330 [6]. N330 corresponds to N343 in the spike glycoprotein of 2019-nCoV and is a conserved glycosylation site. Since it is well separated from the RBD/ACE2 interaction interphase, glycosylation at this site is unlikely to interfere with the interaction [6]. It should be noted that another potential glycosylation site corresponding to N357 in SARS-CoV is not conserved in 2019-nCoV because of substitution of T by A in the +2 position. Lack

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071

of this glycosylation is not expected to affect the receptor binding.

3.3. Distribution of ACE2 in the animal kingdom and implications for virus reservoirs

The 2019-nCoV outbreak is thought to be initiated from a seafood market that also carried many other wild live animals including snakes, birds, and various mammals. Interestingly, a study by Ji et al. suggests that snakes might serve as a likely reservoir for the novel nCoV-2019 based on the observation that the codon usage of nCoV-2019 was more similar to snakes than other potential hosts they investigated [14]. While the data and premise are being debated, we sought to address the problem by analyzing the structure of ACE2 in different animals. ACE2 is widely expressed in the animal kingdom from fish, amphibians, reptiles, birds, to mammals. Remarkably, its structure is highly conserved. Comparison of human ACE2 with that of a civet (Paguma larvata, AAX63775.1), a bat (Rhinolophus sinicus, ADN93475.1), a bird (Nipponia nippon, KFQ92425.1), a snake (Protobothrops mucrosquamatus, XP_029140508.1), a frog (*Xenopus laevis*, XP_018104311.1), and a fish (*Callorhinchus milii*, XP_007889845.1) revealed amino acid sequence identity of 83%, 81%, 83%, 61%, 60%, and 59%, respectively. Fig. 3 aligns parts of ACE2 sequences that contain all the interaction sites in contact with SARS-CoV RBD according to the published co-crystal structure [6]. The interaction involved mainly two α-helices of ACE2. Out of 20 amino acid residues involved in the direct interaction, 4 of them are shared by all seven species of animals analyzed in the study, including F28 that supposedly interacts with F486 of spike glycoprotein from 2019-nCoV (Fig. 2F). Many of the remaining resides in the contact are conserved or replaced by amino acids of similar chemical properties. It is interesting to note that bird ACE2 shares as many conserved contacting amino acid residues as bat and civet ACE2. ACE2 molecules from any of these has the potential to interact with RBD of 2019-nCoV with high affinity. Therefore, it would not be a surprise if any of these wild animals is found to be a primary or secondary host of 2019-nCoV. SARS-CoV-like coronaviruses have been found in many bats that are considered as natural reservoirs for the viruses. They may well be the host for 2019-nCoV. However, the possibility that cold-blooded animals like snakes can serve as a host cannot be ruled out. The flexible interacting loop identified in our study may allow the virus to adapt to both the cold-blooded and warm-blooded hosts.

3.4. Expression of ACE2 in human tissues and implications for virus transmission

By performing immunostaining, earlies studies have demonstrated the expression of ACE2 in lung alveolar epithelial cells as well as arterial and venous endothelial cells, arterial smooth muscle

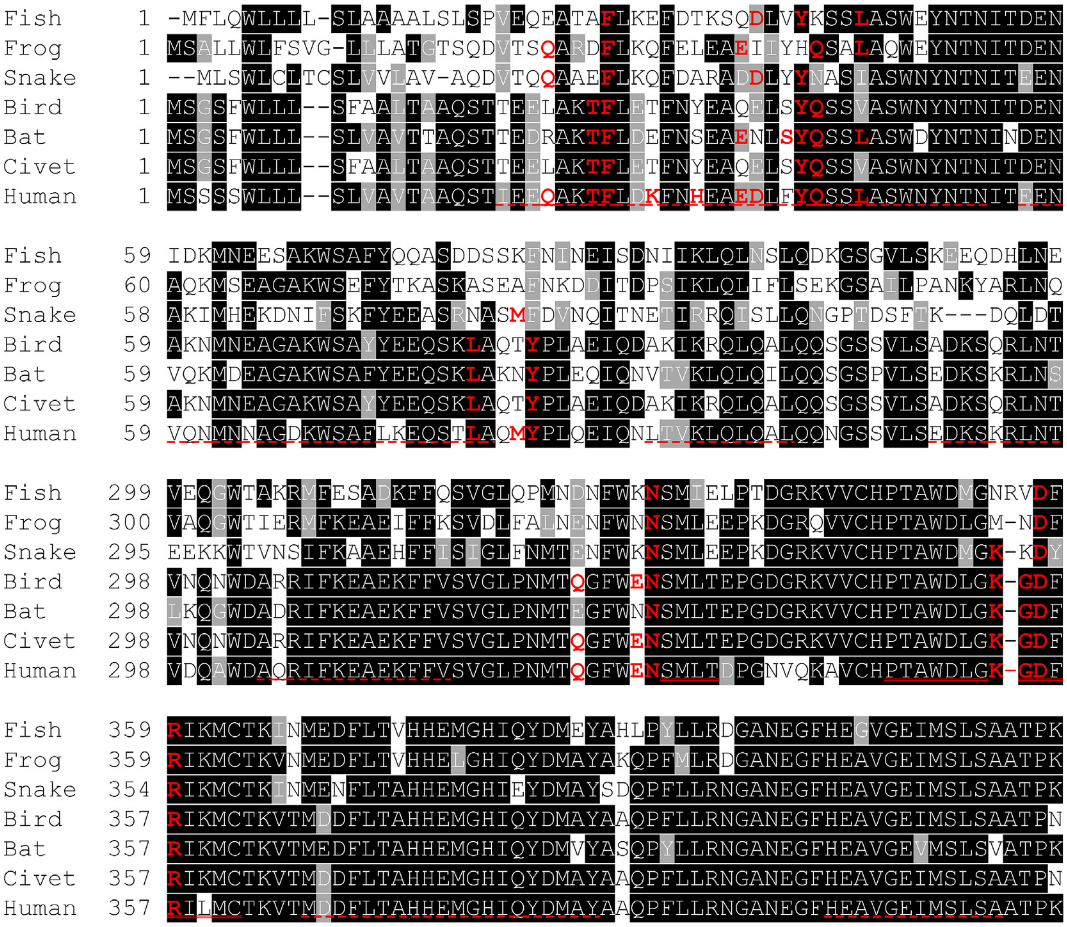

Fish 1 -MFLQWLLLL-SLAAAALSLSPVEQEATAFLKEFDTKSQDLVYKSSLASWEYNTNITDEN
Frog 1 MSALLWLFSVG-LLLATGTSQDVTSQARDFLKQFELEAEIIYHQSALAQWEYNTNITDEN
Snake 1 --MLSWLCLTCSLVVLAV-AQDVTQQAAEFLKQFDARADDLYYNASIASWNYNTNITEEN
Bird 1 MSGSFWLLL--SFAALTAAQSTTEELAKTFLETFNYEAQELSYQSSVASWNYNTNITDEN
Bat 1 MSGSFWLLL--SLVAVTTAQSTTEDRAKTFLDEFNSEAENLSYQSSLASWDYNTNINDEN
Civet 1 MSGSFWLLL--SFAALTAAQSTTEELAKTFLETFNYEAQELSYQSSVASWNYNTNITDEN
Human 1 MSSSSWLLL--SLVAVTAAQSTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEEN
Fish 59 IDKMNEESAKWSAFYQQASDDSSKFNINEISDNIIKLQLNSLQDKGSGVLSKEEQDHLNE
Frog 60 AQKMSEAGAKWSEFYTKASKASEAFNKDDITDPSIKLQLIFLSEKGSAILPANKYARLNQ
Snake 58 AKIMHEKDNIFSKFYEEASRNASMFDVNQITNETIRRQISLLQNGPTDSFTK---DQLDT
Bird 59 AKNMNEAGAKWSAYYEEQSKLAQTYPLAEIQDAKIKRQLQALQQSGSSVLSADKSQRLNT
Bat 59 VQKMDEAGAKWSAFYEEQSKLAKNYPLEQIQNVTVKLQLQILQQSGSPVLSEDKSKRLNS
Civet 59 AKNMNEAGAKWSAYYEEQSKLAQTYPLAEIQDAKIKRQLQALQQSGSSVLSADKSQRLNT
Human 59 VQNMNNAGDKWSAFLKEQSTLAQMYPLQEIQNLTVKLQLQALQQNGSSVLSEDKSKRLNT
Fish 299 VEQGWTAKRMFESADKFFQSVGLQPMNDNFWKNSMIELPTDGRKVVCHPTAWDMGNRVDF
Frog 300 VAQGWTIERMFKEAEIFFKSVDLFALNENFWNNSMLEEPKDGRQVVCHPTAWDLGM-NDF
Snake 295 EEKKWTVNSIFKAAEHFFISIGLFNMTENFWKNSMLEEPKDGRKVVCHPTAWDMGK-KDY
Bird 298 VNQNWDARRIFKEAEKFFVSVGLPNMTQGFWENSMLTEPGDGRKVVCHPTAWDLGK-GDF
Bat 298 LKQGWDADRIFKEAEKFFVSVGLPNMTEGFWNNSMLTEPGDGRKVVCHPTAWDLGK-GDF
Civet 298 VNQNWDARRIFKEAEKFFVSVGLPNMTQGFWENSMLTEPGDGRKVVCHPTAWDLGK-GDF
Human 298 VDQAWDAQRIFKEAEKFFVSVGLPNMTQGFWENSMLTDPGNVQKAVCHPTAWDLGK-GDF
Fish 359 RIKMCTKINMEDFLTVHHEMGHIQYDMEYAHLPYLLRDGANEGFHEGVGEIMSLSAATPK
Frog 359 RIKMCTKVNMEDFLTVHHELGHIQYDMAYAKQPFMLRDGANEGFHEAVGEIMSLSAATPK
Snake 354 RIKMCTKINMENFLTAHHEMGHIEYDMAYSDQPFLLRNGANEGFHEAVGEIMSLSAATPK
Bird 357 RIKMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPN
Bat 357 RIKMCTKVTMEDFLTAHHEMGHIQYDMVYASQPYLLRNGANEGFHEAVGEVMSLSVATPK
Civet 357 RIKMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPN
Human 357 RILMCTKVTMDDFLTAHHEMGHIQYDMAYAAQPFLLRNGANEGFHEAVGEIMSLSAATPK

Fig. 3. Amino acid sequence alignment of ACE2 molecules from 7 animals. Only the part containing amino acid residues (highlighted in red) in human ACE2 that directly interact with RBD of SARS-CoV is shown. Correspondent amino acids for other animals are also highlighted in red if they are shared by human ACE2. Major secondary structures are underlined (dash-lines for α-helix and double lines for β-sheet). GenBank accession numbers are AAX63775.1 for civet (Paguma larvata), ADN93475.1 for bat (Rhinolophus sinicus), KFQ92425.1 for bird (Nipponia nippon), XP_029140508.1 for snake (Protobothrops mucrosquamatus), XP_018104311.1 for frog (*Xenopus laevis*), and XP_007889845.1 for fish (*Callorhinchus milii*).

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071

cells, renal tubular epithelium, and epithelia of the small intestine [15,16]. The lung expression provides strong support for infection of SARS-CoV and 2019-nCoV through the airways of the lung. However, by searching the Human Protein Atlas database, we found that ACE2 mRNA is mainly detected in small intestine, colon, duodenum, kidney, testis, and gallbladder. Its expression level in the lung is minimal (Fig. 4). Furthermore, by examining data from two single-cell RNA-seq studies [17,18], we only identified 2 out of 4599 and 13 out of 540 lung epithelial cells expressed a detectable level of ACE2 (www.ebi.ac.uk/gxa/sc). This confirms that the overall expression of ACE2 in the lung is low and may also suggest the presence of selected cells with upregulated ACE2 expression under certain conditions.

The tissue expression pattern of ACE2 suggests other modes of virus transmission that may involve the functions of intestine, kidney, testis, and other tissues. Particular attention should be paid to the intestines which expressed the highest level of ACE2. Earlier studies have demonstrated that diarrhea was present in up to 70% of patients infected with SARS-CoV [19]. More importantly, a recent case report demonstrated the presence of 2019-nCoV in feces of a patient with an initial diarrhea episode [20]. While this finding has been noted in other reports, tests of feces and urine samples for the presence of 2019-nCoV is warranted, which may help to reveal alternative routes of virus transmission.

4. Discussion

Since its initial outbreak, the 2019-nCoV infection is much more contagious than it was originally thought. We know that the virus is capable of spreading quickly from human to human and that people can spread the virus even before they become symptomatic [1]. This makes it harder to contain the virus, and many are concerned about the possibility of a new pandemic. Our study suggests unique structural features of the spike glycoprotein RBD of 2019-nCoV that confers potentially higher affinity binding for its receptor than found with SARS-CoV. With a higher affinity binding capability, the number of viruses required to infect a cell is much reduced. This partly explains why 2019-nCoV appears to be more aggressive than SARS-CoV. This also reminds us of a lesser-known coronavirus HCoV-NL63 that also uses ACE2 also as a receptor. HCoV-NL63 was initially isolated from a child with bronchiolitis in the Netherlands [21]. It belongs to the alphacoronavirus subfamily. The RBD of SARS-CoV shares no structural homology with that of SARS-CoV but recognizes the same region in ACE2. However, co-crystal structure reveals that RBD of NL63-CoV has a narrower contact with ACE2, involving fewer amino acids [22]. This presumably results in a weaker interaction. Evidently, NL63-CoV does not spread aggressively and only causes mild to moderate respiratory infections [23].

The exact mode of transmission for 2019-nCoV has not been firmly established. SARS-CoV is thought to be transmitted by respiratory droplets produced when an infected person coughs or sneezes [19]. The respiratory droplets spread can occur only through direct person-to-person contact or at a close distance. Presumably, 2019-nCoV can be transmitted through respiratory droplets. It may also be transmitted more effectively through the air over a long distance (airborne spread) or by other ways. Considering the predominant expression of ACE2 in intestines and kidney, 2019-nCoV may infect cells in these tissues and find its way into feces and urine. This makes transmission through the fecal-oral route and bodyfluids (urine) possible. The presence of 2019-nCoV in feces supports such a notion [20].

Specific RBD-receptor binding determines if a cell or animal can be infected and also serves as a target for therapeutic inventions to treat diseases caused by coronaviruses. By binding directly to ACE2 on the surface of host cells, spike glycoprotein plays an essential role in virus infection. An obvious way to stop the virus infection is to block the RBD and ACE2 interaction. This can be achieved by using antibodies or small molecular inhibitors. Naturally, antibodies and inhibitors that can disrupt the interaction of RBD with ACE2 is of therapeutic importance. By using a molecular docking approach, an earlier study identified N-(2-aminoethyl)-1 aziridine-ethanamine as a novel ACE2 inhibitor that effectively blocks the SARS-CoV RBD-mediated cell fusion [24]. This has provided a potential candidate and lead compound for further therapeutic drug development. Meanwhile, biochemical and cell-based assays can be established to screen chemical compound libraries to identify novel inhibitors. On the other hand, many ACE inhibitors are currently used to treat hypertension and other cardiovascular diseases [25]. Among them are captopril, perindopril, ramipril, lisinopril, benazepril, and moexipril. Although these drugs primarily target ACE, a homolog of ACE2 with 42% sequence identity and 61% sequence similarity in the catalytic domain, they may be effective toward ACE2 as well [26]. It should be noted that ACE inhibitors bind to the catalytic center rather than RBD binding site. Nonetheless, these enzymatic inhibitors may indirectly alter conformation of the RBD binding site and thereby affect the interaction of ACE2 with RBD. It is certainly worthwhile to test these drugs for their ability to block the RBD/ACE2 interaction.

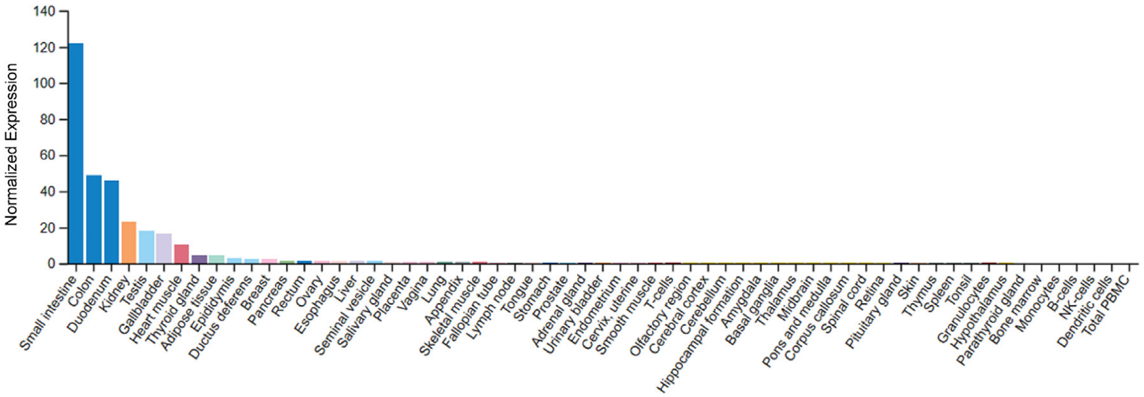


Fig. 4. Expression of ACE2 in human tissues. Data represent Consensus Normalized eXpression levels for 55 tissue types and 6 blood cell types obtained from The Human Protein Atlas (www.proteinatlas.org).

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported in part by the 100 Top Talents Program of Sun Yat-sen University, National Natural Science Foundation of China (NSFC, Grant No. 31871400), and the Sanming Project of Medicine in Shenzhen (No. SZSM201911004).

References

[1] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet (2020 Jan 29). Epub ahead of print, https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30185-9/fulltext.

[2] A.R. Fehr, S. Perlman, Coronaviruses: an overview of their replication and pathogenesis, Methods Mol. Biol. 1282 (2015) 1–23. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4369385/.

[3] F. Li, Structure, function, and evolution of coronavirus spike proteins, Annu Rev Virol 3 (2016) 237–261. https://www.annualreviews.org/doi/abs/10.1146/annurev-virology-110615-042301?rfr_dat=cr_pub%3Dpubmed&url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&journalCode=virology.

[4] W. Li, M.J. Moore, N. Vasilieva, et al., Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, Nature 426 (2003) 450–454. https://www.nature.com/articles/nature02145.

[5] X. Xiao, S. Chakraborti, A.S. Dimitrov, et al., The SARS-CoV S glycoprotein: expression and functional characterization, Biochem. Biophys. Res. Commun. 312 (2013) 1159–1164. https://www.sciencedirect.com/science/article/abs/pii/S0006291X03024136?via%3Dihub.

[6] F. Li, W. Li W, M. Farzan, S.C. Harrison, Structure of SARS coronavirus spike receptor-binding domain complexed with receptor, Science 309 (2005) 1864–1868. https://science.sciencemag.org/content/309/5742/1864.full.

[7] S. Keidar, M. Kaplan, A. Gamliel-Lazarovich, ACE2 of the heart: from angiotensin I to angiotensin (1-7), Cardiovasc. Res. 73 (2007) 463–469. https://academic.oup.com/cardiovascres/article/73/3/463/367423.

[8] R. He, F. Dobie, M. Ballantine, et al., Analysis of multimerization of the SARS coronavirus nucleocapsid protein, Biochem. Biophys. Res. Commun. 316 (2004) 476–483. https://www.sciencedirect.com/science/article/abs/pii/S0006291X04003250?via%3Dihub.

[9] H.D. Song, C.C. Tu, G.W. Zhang, et al., Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 2430–2435. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC548959/.

[10] Y. Guan, B.J. Zheng, Y.Q. He, et al., Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China, Science 302 (2003) 276–278. https://science.sciencemag.org/content/302/5643/276.long.

[11] X.Y. Ge, J.L. Li, X.L. Yang, et al., Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor, Nature 503 (2013) 535–538. https://www.nature.com/articles/nature12711.

[12] W. Li, C. Zhang, J. Sui, et al., Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2, EMBO J. 24 (2005) 1634–1643. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1142572/.

[13] G.B. McGaughey, M. Gagné, A.K. Rappé, pi-Stacking interactions. Alive and well in proteins, J. Biol. Chem. 273 (1998) 15458–15463. https://www.jbc.org/content/273/25/15458.long.

[14] W. Ji, W. Wang, X. Zhao, et al., Homologous recombination within the spike glycoprotein of the newly identified coronavirus may boost cross-species transmission from snake to human, J. Med. Virol. (2020 Jan 22) [Epub ahead of print], https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25682.

[15] I. Hamming, W. Timens, M.L. Bulthuis, et al., Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis, J. Pathol. 203 (2004) 631–637. https://onlinelibrary.wiley.com/doi/abs/10.1002/path.1570.

[16] A.C. Sims, R.S. Baric, B. Yount, et al., Severe acute respiratory syndrome coronavirus infection of human ciliated airway epithelia: role of ciliated cells in viral spread in the conducting airways of the lungs, J. Virol. 79 (2005) 15511–15524. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1316022/.

[17] F.A. Vieira Braga, G. Kar, M. Berg, et al., A cellular census of human lungs identifies novel cell states in health and in asthma, Nat. Med. 25 (2019) 1153–1163. https://www.nature.com/articles/s41591-019-0468-5.

[18] Y. Xu, T. Mizuno, A. Sridharan, et al., Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis, JCI Insight 1 (2016), e90558. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5135277/.

[19] D.S.C. Hui, A. Zumla, Severe acute respiratory syndrome: historical, epidemiologic, and clinical features, Infect. Dis. Clin. 33 (2019) 869–889. https://www.sciencedirect.com/science/article/pii/S0891552019300571?via%3Dihub.

[20] M.L. Holshue, C. DeBolt, S. Lindquist, et al., First case of 2019 novel coronavirus in the United States, N Engl J Med, 2020 Jan 31. Epub ahead of print, https://www.nejm.org/doi/full/10.1056/NEJMoa2001191?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed.

[21] L. van der Hoek, K. Pyrc, M.F. Jebbink, et al., Identification of a new human coronavirus, Nat. Med. 10 (2004) 368–373. https://www.nature.com/articles/nm1024.

[22] K. Wu, W. Li, G. Peng, F. Li, Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 19970–19974. https://www.pnas.org/content/106/47/19970.long.

[23] B.C. Fielding, Human coronavirus NL63: a clinically important virus? Future Microbiol. 6 (2011) 153–159. https://www.futuremedicine.com/doi/full/10.2217/fmb.10.166?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed&.

[24] M.J. Huentelman, J. Zubcevic, J.A. Hernández Prada Ja, et al., Structure-based discovery of a novel angiotensin-converting enzyme 2 inhibitor, Hypertension 44 (2004) 903–906.

[25] S. Culver, C. Li, H.M. Siragy, Intrarenal angiotensin-converting enzyme: the old and the new, Curr. Hypertens. Rep. 19 (2017) 80. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5913745/.

[26] J.L. Guy, R.M. Jackson, K.R. Acharya, Angiotensin-converting enzyme-2 (ACE2): comparative modeling of the active site, specificity requirements, and chloride dependence, Biochemistry 42 (2003) 13185–13192. https://www.nature.com/articles/nature02145.

Please cite this article as: Y. Chen et al., Structure analysis of the receptor binding of 2019-nCoV, Biochemical and Biophysical Research Communications, https://doi.org/10.1016/j.bbrc.2020.02.071