Structural analysis of COVID-19 spike protein in recognizing the ACE2 receptor of different mammalian species and its susceptibility to viral infection.

Tirthankar Koley
Department of Biophysics, All India Institute of Medical Sciences, New Delhi
Shivani Madaan
Department of Computer Science, Jamia Millia Islamia, New Delhi
Sanghati Roy Chowdhury
Department of Biophysics, All India Institute of Medical Sciences, New Delhi
Manoj Kumar
Department of Biophysics, All India Institute of Medical Sciences, New Delhi
Punit Kaur
Department of Biophysics, All India Institute of Medical Sciences, New Delhi
Tej Pal Singh
Department of Biophysics, All India Institute of Medical Sciences, New Delhi
Abdul S. Ethayathulla (ethayathulla@aiims.edu)
Department of Biophysics, All India Institute of Medical Sciences, New Delhi https://orcid.org/0000-0001-8958-3916

Research Article

Keywords: Human ACE2 receptor, SARS-COV-2, spike, COVID-19, protein-protein docking, Glycosylation

Posted Date: October 30th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-98738/v1

Version of Record: A version of this preprint was published at 3 Biotech on February 1st, 2021. See the published version at https://doi.org/10.1007/s13205-020-02599-2.

Abstract

The pandemic COVID-19 caused by a novel coronavirus SARS-CoV-2 spread worldwide as a new public health emergency. The SARS-CoV-2 infects humans by binding to glycosylated ACE2 receptor present in the inner lining of the lungs, heart, intestine and kidney. The COVID spike 2 protein recognizes the ACE2 receptor at the N-terminal helices of the metalloprotease domain. The residues Gln24, Thr27, Lys31, His34, Glu37, Asp38, Tyr41, Gln42 from helix α1; Leu79, Met82, Tyr83 from helix α2 and Gln325, Glu329, Asn330, Lys353 from loop connecting β4 and β5 strands form a concave surface surrounded by four glycosylation sites Asn53, Asn90, Asn103 and Asn322 form interactions with the spike protein. However, no significant data on the susceptibility of animals for infection or transmission. Therefore, we performed the comparative protein-protein docking analysis using the crystal structure of spike protein and homology models of the ACE2 receptor from 16 commonly found mammalian species to understand the potential mode of spike binding. Our comprehensive sequence and structure-based interaction analysis revealed the natural substitution of amino acid residues Gln24, His34, Phe40 and Met82 in the N-terminal α1 and α2 helices results in loss of crucial network of hydrogen-bonded and hydrophobic interactions with spike 2 RBD domain. Besides, the absence of N-linked glycosylation site Asn103 in other mammals further reduces the binding affinity between spike and ACE2 receptor. Hence, these changes explain the differences in the susceptibility and host pathogenesis in other mammalian species.

Introduction

COVID-19 is one of the most dreadful pandemic diseases of the 21st century responsible for most deaths worldwide. The highly pathogenic virus originated from Wuhan city, Hubei state of China, in December 2019 and expanded globally as a new health pandemic (Zhao et al. 2020b) (Li et al. 2020). As reported by World Health Organization till 19th October 2020, 40.1 million cases are confirmed with infection globally, and the total death is 1.1 million (World Health Organization). Belonging to the realm of Ribovaria, family Coronaviridae, suborder Coronovirineae and order Nidovirales (Contini et al. 2020) (Kannan et al. 2020)(Hasöksüz et al. 2020), the Severe Acute Respiratory Syndrome related coronaviruses are mostly zoonotic, with only a handful of species known to cause diseases in humans (Chen et al. 2020). The phylogenetic analysis showed these as four distinct groups (Bhowmik et al. 2020). Hypothesized to have originated in bats with unknown intermediate hosts, this could be the result of a “spillover event”(WHO) (Banerjee et al. 2019). Earlier studies described bats and pangolins as intermediate hosts (Zhang et al. 2020a); subsequent literature failed to reveal any such connections (Zhang et al. 2020b). Many zoonotic CoVs are sustaining in nature, constantly mutating and evolving to give rise to a new kind of infectious form (Ye et al. 2020). The person infected with this virus develops various symptoms like fever, body ache, cough, pneumonia etc. (Singhal 2020). Some of these coronaviruses can transmit between animals and humans (Cevik et al. 2020) mainly through direct physical contacts, droplet transmission or oral transmission and prevalent in all age groups. Interestingly, many animal pools with direct contact with humans have minimal disease susceptibility (Baloch et al.

2020). Moreover, the virus is mutating rapidly, and the recently emerging group of viruses indicates specific structural changes in a particular protein. The Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a positive-sense single-stranded RNA virus contains four main structural proteins (Mousavizadeh and Ghasemi 2020). The spike glycoprotein (S), NSPs, membrane glycoprotein (M), and several accessory (Jiang et al. 2020) protein with a diameter of 80 - 120 nm. It consists of a large genome of 28 to 32 kb. Coronaviruses have four genera: α, β, γ, δ and the novel coronavirus (SARS-CoV-2) (Gabutti et al. 2020) belongs to β type Coronavirus. The SARS-CoV-2 recognizes angiotensin-converting enzyme 2 (ACE2) receptors to infect humans (Turner et al. 2004). The ACE2 receptor is an integral membrane protein that belongs to a zinc metalloprotease family (Alifano et al. 2020) converts angiotensin 2 to form angiotensin 1-7 and play a vital role in the Renin-Angiotensin system (RAS) (Warner et al. 2004). The enzyme is localized in the cell's outer membrane surface in the heart, lungs, arteries, kidney, and intestine (Hamming et al. 2004). Coronavirus attaches to the receptor with (Schwarz and Aebi 2011a) spike receptor-binding domain by a hinge-like motion (Towler et al. 2004; Shang et al. 2020). The receptor N-terminal helices α1 and α2 form a concave binding surface with glycosylation sites recognized by spike protein (Zhai et al. 2020) (Han et al. 2007). The sugar moieties at the glycosylation site act as a ligand for cell surface receptors to mediate cell attachment or induce a signal transduction pathway (Ohtsubo and Marth 2006). Overall thermodynamic of the protein complex depends on the receptor and spike glycosylation (Shental-Bechor and Levy 2008). In humans, the glycosylation pattern of the ACE2 receptor is known to be altered depending on age, disease and ethnicity, probably responsible for the susceptibility of infection (Zhao et al. 2020a). The human ACE2 receptor has 7 potential N-linked glycosylation sites Asn53, Asn90, Asn103, Asn322, Asn 432, Asn546 and Asn690 (Donoghue et al. 2000a; Tipnis et al. 2000; Schwarz and Aebi 2011b). The studies have shown that the glycosylation pattern influences viral entry into the host and pathogenesis (Han et al. 2007; Almendros and Gascoigne 2020; Shi et al. 2020). Interestingly the COVID-19 pathogenesis in animals is less compared to humans. Only a few reports of COVID-19 infection with some pet animals like Pomeranian in Hong Kong, a Malayan tiger in New York City Zoo and a domestic cat in Belgium (Brooke and Prischi 2020; Hoffmann et al. 2020; Leist et al. 2020; Mallapaty 2020). But there is no significant report related to viral transmission in animals or viral transmission from animals to humans. To infect the spike protein should recognize the ACE2 receptor, which acts as a doorknob for virus pathogenesis (Walls et al. 2020). Few studies reported comparing the ACE2 receptor, Transmembrane protease serine 2 precursor and Furin proteases from different species to highlight the variation in serine proteases. Some reports have been in developing therapeutics models for vaccine development (Hoffmann et al. 2020; Leist et al. 2020). This study performed a complete sequence and detailed molecular interactions analysis of 16 ACE2 receptors from different mammals with the COVID-spike to understand the molecular difference influencing the viral pathogenesis. The mammalian species selected in this study is based on four-category viz. common domestic animals which are in daily direct contact with humans, wild animals, animals reportedly prone to the disease, and some closely related species.

Materials And Methods

Multiple Sequence Alignment

The sequences of the ACE2 receptor from human and 16 other common mammalian species (**Fig. S1**) were retrieved from the National Center for Biotechnology Information (NCBI) protein sequence database. The multiple sequence alignment was performed using the online sequence alignment tool Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers and Higgins 2014).

Homology Modelling

The homology models of the ACE2 receptor of the 16 common mammalian species were generated using the SWISS-MODEL server (https://swissmodel.expasy.org/) (Waterhouse et al. 2018). The best model for each individual species was selected based on Levitt-Gerstein (LG) score (Levitt and Gerstein 1998) using the neural network-based method and ProQ server (https://proq.bioinfo.se/ProQ/ProQ.html). The chosen model of each species were refined by energy minimization using previously reported parameters (Yadav et al. 2019; Naz et al. 2020). The stereochemical quality of the 16 models was evaluated by the PROCHECK program (Laskowski et al. 1993) and ERRAT server (Colovos and Yeates 1993).

Docking of SARS-CoV-2 spike protein with ACE-2 Protein

The protein-protein docking of the ACE2 receptor homology model with the receptor-binding domain of spike protein was performed using the HDOCK server (Yan et al. 2020b), a hybrid approach template-based modeling and ab initio free docking. Based on the docking score, the best-docked conformation was selected for each species, and energy minimized using previously reported parameters (Yadav et al. 2019). The energy minimized docked complex of individual species was superimposed on to the crystal complex of human ACE-2 and spike protein (PDB ID - 6M0J and 6M18) (Brooke and Prischi 2020; Stout et al. 2020)for structural comparison and interaction analysis.

Binding Free Energy Calculation

The binding energy of each energy minimized docked complex of SARS-CoV-2 spike protein and the ACE2 receptor was calculated by MM-GBSA (Molecular Mechanics-Generalized Born Surface Area) approach to estimate the relative binding affinity. Binding energy was calculated using OPLS-2005 molecular mechanics force field parameters (Shivakumar et al. 2010) and continuum (implicit) solvation models similar to the previously described approach (Beard et al. 2013).

Comparison of glycosylation pattern between human and other mammalian species

The ACE2 receptor glycosylation sites of all 16 mammalian species were predicted and analyzed using the NetNGlyc1.0 server.

Results

Multiple Sequence Alignment

The multiple sequence alignment of the Human ACE2 receptor with other mammalian species (**Table 1**) showed the highest sequence identity with *Equus caballus* (Horse) (86.78%) and lowest identity with the *Rattus rattus* (Black rat) (81.01%). Based on the structure complex of human ACE2 and spike protein (PDB ID:6M0J), the residues majorly involved in hydrogen-bonded and hydrophobic interactions with spike 2 RBD are Gln24, Asp30, Lys31, His34, Glu35, Asp38, Tyr41, Gln42, Lys353 and Asp355 (**Fig. 1)**. The substitution of these residues leads to the loss of affinity. Out of these 10 residues, Lys31, Glu35, Asp38, Gln42, and Tyr41 were conserved in all species. Asp30 is mostly replaced by Glu30, maintaining the residual charge integrity and Tyr41 is conserved in most species except in *Equus asinus and Equus caballus,* where Histidine replaces it. Similarly, Tyr83 and Lys353 are conserved in most species except for *Mus musculus*, *Rattus rattus* and *Loxodonta africana,* where the residues are replaced by phenylalanine and histidine, respectively. The crucial substitutions observed in the sequence alignment were Gln24 and Met82 to Leu and Thr, respectively. Besides, the absence of a glycosylation site at position Asn103 in all other mammals also plays an essential role in spike recognition of the ACE2 receptor. In human, the ACE2 receptor contains seven N-linked glycosylation sites (Asn53, Asn90, Asn103, Asn322, Asn432, Asn546 and Asn690) and three O-linked glycosylation sites (S155, S496 and S730) (Shajahan et al. 2020a). Out of seven N-linked glycosylation sites, four (Asn53, Asn90, Asn103 and Asn322) are aligned close to the spike 2 recognition site increasing the binding affinity. Interestingly, in almost all species, the position Asn103 is replaced by Ser103 (**Fig. 1**), making this position exclusive for the human ACE2 receptor.

Homology modeling

The homology model of ACE2 receptors from different species were generated using PDB 6M0J (Lan et al. 2020). The Levitt-Gerstein (LG) score of each model calculated by the ProQ server showed that the quality parameters of the structures are within the acceptable range (>5.0) (**Table 1**). The PROCHECK analysis showed that greater than 99% of the residues fall in the allowed, additionally allowed and generously allowed regions of the Ramachandran plot (**Table S1**, **Fig. S2, S3**). The ERRAT value of all the structure models was found to be greater than 93%. The generated models have the right geometry; hence the models were further used for protein-protein docking.

Protein-protein docking

The protein-protein docking of ACE2 receptor model with spike 2 RBD was performed using the HDock server. The human ACE2 receptor and Spike 2 RBD domain complex (6M0J) was used as a docking reference (Lan et al. 2020). The top model (out of 10) was used for molecular interaction analysis using the Schrodinger Maestro MMGBSA method (Beard et al. 2013). The Gibbs free energy for each complex was calculated and compared with human (-128.67 Kcal/mol with 949.8 $Å^2$ buried surface area between two proteins) **(Table 2)**. The highest Gibbs free energy ΔG next to human was observed for *Equus caballus* (-97.24 Kcal/mol) followed by *Felis catus* (-98.04 Kcal/mol), *Bos tauras* (-92.32 Kcal/mol) and *Ovis aries* (-90.20 Kcal/mol). Similarly, the lowest ΔG was found for the complexes of *Canis lupus familiaris* (-65.12 Kcal/mol), *Mus musculus* (-62.12 Kcal/mol), *Ailuropoda melanoluca* (-62.46 Kcal/mol), *Panthera tigris altaica* (-52.41 Kcal/mol) and *Rattus rattus* (-43.21 Kcal/mol) (**Table 2**). The free binding energy indicates that there is the loss of some crucial interactions result in lower affinity between spike protein and ACE2 receptor from other species

Glycosylation of ACE2 receptor

Glycosylation plays a significant role in host-parasite interaction. The pattern of glycosylation contributes to the recognition of the ACE2 receptor by COVID spike 2 protein. The metalloprotease domain of ACE2 receptor is highly glycosylated with seven N-linked and three O-linked glycosylation sites (Shajahan et al. 2020a). In the reported crystal structures of the human ACE2 receptor (PDB ID: 6M0J and 6M18) the N-linked glycosylation sites Asn53, Asn90, Asn103, Asn322, Asn432, Asn546 and Asn690 found with sugar moieties. The structure analysis of human ACE2 receptors in complex with spike 2 RBD domain showed four glycosylation sites (Asn53, Asn90, Asn103 and Asn322) situated in the close vicinity of the spike 2 interacting interface (**Fig. 2**). The four glycosylation sites form a pillar-like structure around the interacting interface (**Fig. 2**). These four N-linked glycosylation sites play an essential role in spike 2 recognition. Based on the previously reported study on the human ACE2 receptor glycosylation pattern (Shajahan et al. 2020a). We modeled the sugar molecules at these N-linked glycosylation sites as shown in **Fig. 2b**. According to previous study (Shajahan et al. 2020b), the N-linked glycosylation sites in human ACE2 receptor majorly contain N-acetylglucosamine, mannose, fucose, galactose and N-acetylneuraminic acid sugar moieties with different branching length. The sugar branching is an essential aspect of COVID spike 2 recognition. Upon sequence and structure comparison of all selected mammalian species, the Asn103 N-linked glycosylation is either absent or replaced by Ser. The NetOGlyc 4.0 server prediction also showed no possibility for glycosylation. Except for *Ailuropoda melanoleuca, Equus asinus, Equus caballus, Felis catus, and Panthera tigris altaica* all other species lack more than one glycosylation site. In *Loxodonta africana, Mus musculus* and *Rattus rattus*, three glycosylation sites were absent. The complete list of glycosylation sites of all the species is shown in **Table 4**. Hence, we can conclude that Asn103 glycosylation is also an essential factor influencing the susceptibility to viral infection and pathogenesis.

Structural comparison of ACE2 receptor and spike 2 complexes

Spike protein juxtaposes with human ACE2 receptor predominantly by hydrogen bonds, hydrophobic interactions and ionic interactions. Additionally, glycosylation plays a significant role in specifically recognizing the human ACE2 receptor by SARS-COV-2 spike protein. Spike 2 recognizes the concave surface of the ACE2 receptor. The concave surface can be divided into four interacting regions, majorly involves N-terminal α1, α2 helices and β4, β5 strands. The N-terminal part of α1 helix forms the interacting region 1; the middle part of the α1 helix forms region 2 and the end part of α1 helix and β4, β5 strands form as region 3 and the 4^{th} region is formed by end part of α2 helix (Lan et al. 2020)(Yan et al. 2020). Comparison of the sequence of ACE2 receptor from different mammals with the human receptor revealed much notable natural amino acid substitution. A detailed structural analysis of each species interacting region compared to the human ACE2 receptor was performed to understand the spike 2 protein recognition and affinity towards other species ACE2 receptor.

N-terminal α1 helix interacting region1

Upon structural analysis of the human ACE2 receptor and spike 2 RBD complex, the interacting region1 consists of three stable hydrogen-bonded networks formed between Gln24 - Asn487, Tyr83 - Asn487, Tyr83 - Tyr489 in humans (**Fig. 2**). The Gln24 residue (OE1) forms strong hydrogen bond interaction with ND2 of spike Asn487 (2.7Å) (**Table 3**). Consecutively OD1 of Asn487 of spike form another hydrogen-bonded interaction with OH group of human Tyr83 (3.5 Å). Consequently, the spike protein becomes more invasive in the left-hand side of the α1 helix and makes a strong interaction with the ACE2 receptor. In other mammalian species, Gln24 is mostly replaced by hydrophobic residue Leu24, in *Mus musculus* by Asn24 and in *Rattus rattus* by Lys24 (**Fig. 3**). The substitution of Gln leads to loss of network of interaction with Asn487 of spike protein. As a result, Tyr83 is displaced from forming a hydrogen bond with Tyr489. Hence, the complete interchain hydrogen bonding network of Gln24 - N487 - Tyr83 - Tyr489 observed in the human ACE2 receptor was not observed in other species models. Another interesting feature observed in the human ACE2 receptor is the presence of Asn103 N-linked glycosylation to this α1 helix interacting interface region 1 (**Fig. 2**). The α1 helix gets appropriately sandwiched with the interacting surface of the spike protein. But in the case of the ACE2 receptor of other species, the α1 helix differently acclimatizes to another direction (**Fig. 3-6**). The Gln24 hydrogen-bonding network, N-linked glycosylation site Asn90 at the proximal to interacting interface and the N-linked glycosylation site Asn103 at the distal position from the interface region uniquely form an intense binding surface, which is absent in other species. Interestingly, this interacting region is also more specific to SARS-COV-2. In other SARS-CoV, the Phe486 is mutated to Leu472 due to which the hydrophobic interactions formed between Phe486 and Gln24, Leu79, Met82 and Tyr83 is lost result in reduced affinity towards host receptor.

The middle portion of α1 helix region

The middle portion of α1 helix form interacting region 2 formed by Asp30 - Lys417, His34 - Tyr453 and Glu35 - Gln493 with the intra-network salt bridge (**Fig. 2**). The OD2 atom of Asp30 (human) forms an interchain H-bond with NZ atom of Lys417 (2.9 Å) and the OD1 atom form another intrachain H-bond with His34 NE2. Sequentially the network completes upon the formation of another interchain H-bond

between ND1 of His34 with OH group of Tyr453 (2.8 Å), establishing a tight interaction between human ACE2 receptor and SARS-COV-2 spike protein (**Table 4, Fig. 2**). The His34 residue being the heart of the network formation is mostly replaced by Tyr34 in *Ailuropoda melanoleuca* and *Canis lupus familiaris* (Fig. **4**); Ser34 in *Equus caballus* and *Equus asinus* (**Fig. 3**); Gln34 in *Loxodonta africana*, *Mus musculus*, *Oryctolagus cuniculus* and *Rattus rattus* and finally Leu34 in *Sus scrofa* (**Fig. 3,5**). Interestingly, in the human ACE2 receptor, the combination of His34 and Asp30 residues is concertedly forming Lys417 (spike) - Asp30 (human) - His34 (human) - Tyr453 (spike), leading to an intricate H-bonding network. However, Asp30 is replaced by Glu30 in some species with the same charge distribution. The residue His34 is the crucial player in this network, mostly substituted with Ser, Leu or Gln. Exceptionally in *Bubalus bubalis*, *Panthera tigris altaica* and *Felis catus,* the combination of His34 and Glu30 forms the whole interchain bonding network (**Fig. 4-6**) than other species. All the interchain and intrachain hydrogen bonding interactions and hydrophobic and ionic bonds assist the spike protein in recognizing the human ACE2 receptor concave interface.

The end region of α1 helix

The interacting region 3 is formed by the end portion of α1 helix. The interaction Asp38 - Tyr449 (2.7 Å) is fully conserved despite D38E substitution observed in *Equus asinus*, *Equus caballus*, *Felis catus* and *Panthera tigris altaica*. The interaction Tyr41- Thr500 is also conserved except in *Equus asinus* and *Equus caballus* (Y41H replacement) (**Fig. 3-5**). Gln42 residue gets into the vicinity of Gly446 and Tyr449 residues and consecutively forming two interactions, Gln42 - Gly446 and Gln42 - Tyr449 in humans. In human, these two interactions are stabilized further by the hydrophobic residue Phe40. Interestingly in *Sus scrofa* (**Fig. 5**) both Gln42 - Gly446 and Gln42 - Tyr449 interactions are conserved, but Phe40 is replaced by small chain hydrophobic residue Ala40. In contrast, F40Y (in *Ailuropoda melanoleuca* and *Camelus dromedaries*) (**Fig. 4**) and F40S (in rest of the species) leads to complete disruption of Gln42 - Gly446 interaction in all species and partial disruption of Gln42 - Tyr449 in *Bubalus bubalis*, *Capra hircus*, *Felis catus*, *Loxodonta africana* and *Panthera tigris altaica* (**Table 3**).

The region 3 interacting surface also involves the residues from the loop connecting β4 and β5 strands. The three hydrogen bond interactions Lys353 - Gly496, Lys353 - Gly502 and Asp355 - Thr500 clustered together, forming a tight grip with the spike 2 protein (**Fig. 2**). The Lys353 - G502 and Asp355 - Thr500 hydrogen bonds are conserved in all species. While in *Canis lupus*, *Camelus dromedarius* and *Felis catus,* the Lys353 and Gly496 are situated far away from each other (**Fig. 4**). In Mus musculus and *Rattus rattus* Lys353 is replaced by His353 and not making any considerable interaction (**Table 3**).

In the vicinity of the region 3 interacting interface two glycosylation sites, Asn53 (site1) and Asn322 (site4) are present. The hydrogen-bonding network supported with two glycosylations at Asn322 and Asn53 helps spike 2 protein to recognize the human receptor (**Fig. 2**). The residue Asn53 is conserved in all the species. In contrast, Asn322 residue is conserved only in *Ailuropoda melanoluca*, *Canis lupus familiaris Equus caballus*, *Equus asinus*, *Felis catus*, *Sus scrofa*, *Camelus dromedaries and Panthera tigris altaica* (**Table 4**).

The end region of α2 helix

The region 4 interacting surface is formed by the end portion of α2 helix, where Met82 is crucial in terms of both hydrophobic interaction and holding the glycosylation site. The Phe486 of spike 2 is sandwiched between Met82 and Tyr83 to form strong hydrophobic interaction to hold the helix (**Fig. 2**). The Met82 and Gln81 hold the sugar moieties at the N-linked glycosylated site Asn103. This entire network of hydrophobic interaction and N-linked glycosylation is absent in all species as Thr or Asn replaces Met82 (**Fig. 1, 3-6**). As the N-linked glycosylation Asn103 site is not present in other mammalian species, it might be necessary for spike 2 protein recognition.

Other significant hydrophobic interactions

The hydrophobic residues in the spike recognizing site of the ACE2 receptor are evenly distributed from left to right of the interface. Overall the hydrophobic residues give considerable stability to the spike-receptor complex. Some noteworthy hydrophobic interactions are Thr27 - Phe456, His34 - Gln493, Tyr41 - Gln498, Leu79 - Phe486, Met82 - Phe486, Tyr83 - Phe486 and Lys353 - Tyr505. The Left-hand portion of the concave interface is supported by Thr27 - Phe456, Leu79 - Phe486, Met82 - Phe486, Tyr83 - Phe486. The middle part and the right-hand side of the interface is held up by His34 - Gln493, Tyr41- Gln498 and Lys353 - Tyr505 sequentially (**Table 3**). Notably, in other mammalian species, many hydrophobic residues are naturally replaced by polar/charged residues leading to reduce affinity. All the residue substitutions are shown in (**Table 3**).

Discussion

The recognition of the ACE2 receptor by spike protein is essential for host infection. The spike protein recognizes the human ACE2 receptor with more affinity than other mammalian species (Lam et al. 2020). Based on our sequence and structural analysis of four interacting sites, the difference observed in region 4 and region 1 explains the reduced affinity towards other species. The significant difference is the substitution of Asn103; the N-linked glycosylation site leads to loss of hydrogen-bonding network formed by Gln24 and the loss of stacking and hydrophobic interaction created by Met82. The spike protein residue Phe486 gets sandwiched between Tyr83 and Met82, sequentially forming hydrophobic interaction with the glycan chains at the Asn103 site. Besides, the Glycan chain at Asn103, through the interaction with Gln81 form an intrachain hydrogen bonding network with the Asn90 glycosylation site. Similarly, the glycan chain of Asn90 forms another hydrogen-bonded network with Lys26 and Thr27. The Thr27, in turn, creates hydrophobic interactions with Tyr489 and Phe456. Collectively all these interactions form an inverted U-shaped H-bonding, hydrophobic and stacking interactions network stretched from Asn103 to Asn90 crucial for holding spike protein at one end of the concave interface (**Fig. 2**). The glycosylations site Asn90 and Asn103 push the helix α1 and α2 interface upwards for the tight spike binding.

Similarly, the sugar moieties at these glycosylation sites are held by nearby supporting residues (**Table 4**). The Glycan chain at Asn53 is supported by Thr55, Glu57, Asn58 and Gln340. While the glycan chains at Asn90, Asn103, Asn322 and Asn546 are supported by Lys26 and Gln81, His194, Asn195 and Glu312 and

Ser420, respectively. Overall, four glycosylations and supporting residual hydrogen bonding and hydrophobic interactions make the human ACE2 receptor more specific than other mammalian species. Upon comparing the surface potential of human ACE2 receptor with other mammalian species, it clearly showed that substitution of residues Gln24, His34, Phe40 and Met82 at the N-terminal α1 helix and α2 helix results in a change in charge distribution at the interacting site (**Fig. 7**). Besides, the absence of glycosylation at the Asn103 makes the human ACE2 receptor more specific than other species. The residual interactions and the approximate free energies of binding also indicate significantly higher ΔG in humans than other species. The buried hydrophobic surface area of spike receptor assembly was 949.8 $Å^2$ in humans, while in other species, it averaged around approx 850 $Å^2$. Cumulatively all these interactions constructed a favorable environment for spike protein invasion in the left-hand side of the human ACE2 receptor with enhanced affinity over all other selected mammalian species. Besides these, the four N-linked glycosylation sites increase the specificity of the receptor. The ACE2 receptor with symmetrical glycosylation sites makes it more accessible to SARS-COV-2 compared to other mammalian species. Our study shows that natural substitution at crucial interacting regions in the ACE2 receptors makes animals less susceptible to infection, with no primary role in transmission.

Conclusion

Understanding the mechanism of host-pathogen interactions and cross-species infections is essential to control the pandemic disease COVID-19 (Donoghue et al. 2000b; Ayati et al. 2020; Das and Choudhuri 2020; Tay et al. 2020). In general, mammals are the most susceptible group concerning Coronavirus infection (Mallapaty 2020). The ACE2 receptor is the access point for SARS-CoV-2 entry and further infection in the healthy human cell (Hofmann and Pöhlmann 2004). A hinge-like movement happens when the virus's receptor-binding domain attaches with the ACE2 receptor increasing interchain surface accessibility (Lan et al. 2020; Roy et al. 2020; Yan et al. 2020). Multiple sequence alignment of the ACE2 receptor revealed high sequence similarity/identity among different mammalian species (Letko et al. 2020). So it is evident that the minor changes are imparting a significant contribution in host-parasite interaction. Among all the interactions, two unique interchain bonding networks Gln24 - Asn487 - Tyr83 - Tyr489 and Lys417 - Asp30 - His34 - Tyr453 are crucial to grip the ACE2 receptor by spike protein tightly. Gln24Leu substitution in most of the species disrupts the network except common cattle (*Bos tauras*), goat (*Capra hircus*), sheep (*Ovis aries*) and domestic water buffalo (*Bubalus bubalis*). The second interaction network is observed only in Siberian tiger and domestic cat (Brooke and Prischi 2020), corroborating with the recent COVID-19 infection report of these two species. The network of interaction supported by mainly two glycosylations at positions Asn90 and Asn103 push the α1-α2 helices upwards, making it easily accessible for recognition by COVID spike 2 protein. To further validate our study, the glycosylation pattern in the ACE2 receptor of other mammalian species is required to understand the complete COVID-19 infection mechanism.

Declarations

Acknowledgments

The authors thank AIIMS Intramural funding for providing the software and hardware for the modeling studies.

Conflict of Interest

The authors declare no conflict of interest with the current work or its publication.

Ethical approval

This article does not involve any study with human participants or animals to be performed by any authors.

Authors contribution

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Tirthankar Koley, Shivani Madaan and Abdul S. Ethayathulla. The first draft of the manuscript was written by Tirthankar Koley, Shivani Madaan and Abdul S. Ethayathulla and all authors gave their inputs in the previous versions of the manuscript. All authors read and approved the final manuscript.

References

Alifano M, Alifano P, Forgez P, Iannelli A (2020) Renin-angiotensin system at the heart of COVID-19 pandemic. Biochimie. https://doi.org/10.1016/j.biochi.2020.04.008

Almendros A, Gascoigne E (2020) Can companion animals become infected with Covid-19? Veterinary Record

Ayati N, Saiyarsarai P, Nikfar S (2020) Short and long term impacts of COVID-19 on the pharmaceutical sector. DARU, Journal of Pharmaceutical Sciences. https://doi.org/10.1007/s40199-020-00358-5

Baloch S, Baloch MA, Zheng T, Pei X (2020) The coronavirus disease 2019 (COVID-19) pandemic. Tohoku Journal of Experimental Medicine

Banerjee A, Kulcsar K, Misra V, et al (2019) Bats and coronaviruses. Viruses

Beard H, Cholleti A, Pearlman D, et al (2013) Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. PLoS One 8:e82849. https://doi.org/10.1371/journal.pone.0082849

Bhowmik D, Pal S, Lahiri A, et al (2020) Emergence of multiple variants of SARS-CoV-2 with signature structural changes. bioRxiv. https://doi.org/10.1101/2020.04.26.062471

Brooke GN, Prischi F (2020) Structural and functional modelling of SARS-CoV-2 entry in animal models. Sci Rep 10:15917. https://doi.org/10.1038/s41598-020-72528-z

Cevik M, Bamford CGG, Ho A (2020) COVID-19 pandemic—a focused review for clinicians. Clinical Microbiology and Infection

Chen Y, Liu Q, Guo D (2020) Emerging coronaviruses: Genome structure, replication, and pathogenesis. Journal of Medical Virology

Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2:1511–1519. https://doi.org/10.1002/pro.5560020916

Contini C, Nuzzo M Di, Barp N, et al (2020) The novel zoonotic COVID-19 pandemic: An expected global health concern. Journal of Infection in Developing Countries. https://doi.org/10.3855/jidc.12671

Das P, Choudhuri T (2020) Decoding the global outbreak of COVID-19: the nature is behind the scene. VirusDisease. https://doi.org/10.1007/s13337-020-00605-y

Donoghue M, Hsieh F, Baronas E, et al (2000a) A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circulation research. https://doi.org/10.1161/01.res.87.5.e1

Donoghue M, Hsieh F, Baronas E, et al (2000b) A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circulation research. https://doi.org/10.1161/01.res.87.5.e1

Gabutti G, d'Anchera E, Sandri F, et al (2020) Coronavirus: Update Related to the Current Outbreak of COVID-19. Infectious Diseases and Therapy

Hamming I, Timens W, Bulthuis MLC, et al (2004) Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. Journal of Pathology. https://doi.org/10.1002/path.1570

Han DP, Lohani M, Cho MW (2007) Specific Asparagine-Linked Glycosylation Sites Are Critical for DC-SIGN- and L-SIGN-Mediated Severe Acute Respiratory Syndrome Coronavirus Entry. Journal of Virology. https://doi.org/10.1128/jvi.00315-07

Hasöksüz M, Kiliç S, Saraç F (2020) Coronaviruses and sars-cov-2. Turkish Journal of Medical Sciences

Hoffmann M, Kleine-Weber H, Schroeder S, et al (2020) SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181:271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052

Hofmann H, Pöhlmann S (2004) Cellular entry of the SARS coronavirus. Trends in Microbiology

Jiang S, Hillyer C, Du L (2020) Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. Trends in Immunology

Kannan S, Shaik Syed Ali P, Sheeza A, Hemalatha K (2020) COVID-19 (Novel Coronavirus 2019) - recent trends. European Review for Medical and Pharmacological Sciences. https://doi.org/10.26355/eurrev_202002_20378

Lam SD, Bordin N, Waman VP, et al (2020) SARS-CoV-2 spike protein predicted to form stable complexes with host receptor protein orthologues from mammals. bioRxiv

Lan J, Ge J, Yu J, et al (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. https://doi.org/10.1038/s41586-020-2180-5

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 26:283–291. https://doi.org/10.1107/S0021889892009944

Leist SR, Schäfer A, Martinez DR (2020) Cell and animal models of SARS-CoV-2 pathogenesis and immunity. Dis Model Mech 13:. https://doi.org/10.1242/dmm.046581

Letko M, Marzi A, Munster V (2020) Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nature Microbiology. https://doi.org/10.1038/s41564-020-0688-y

Levitt M, Gerstein M (1998) A unified statistical framework for sequence comparison and structure comparison. Proc Natl Acad Sci U S A 95:5913–5920. https://doi.org/10.1073/pnas.95.11.5913

Li R, Pei S, Chen B, et al (2020) Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. https://doi.org/10.1126/science.abb3221

Mallapaty S (2020) What's the risk that animals will spread the coronavirus? Nature. https://doi.org/10.1038/d41586-020-01574-4

Mousavizadeh L, Ghasemi S (2020) Genotype and phenotype of COVID-19: Their roles in pathogenesis. Journal of Microbiology, Immunology and Infection

Naz F, Mashkoor M, Sharma P, et al (2020) Drug Repurposing Approach to Target FtsZ Cell Division Protein From Salmonella Typhi. In: International journal of biological macromolecules. https://pubmed.ncbi.nlm.nih.gov/32417543/. Accessed 24 Jun 2020

Ohtsubo K, Marth JD (2006) Glycosylation in Cellular Mechanisms of Health and Disease. Cell

Roy S, Jaiswar A, Sarkar R (2020) Dynamic Asymmetry Exposes 2019-nCoV Prefusion Spike. The journal of physical chemistry letters. https://doi.org/10.1021/acs.jpclett.0c01431

Schwarz F, Aebi M (2011a) Mechanisms and principles of N-linked protein glycosylation. Current Opinion in Structural Biology

Schwarz F, Aebi M (2011b) Mechanisms and principles of N-linked protein glycosylation. Current Opinion in Structural Biology. https://doi.org/10.1016/j.sbi.2011.08.005

Shajahan A, Archer-Hartmann S, Supekar NT, et al (2020a) Comprehensive characterization of N- and O-glycosylation of SARS-CoV-2 human receptor angiotensin converting enzyme 2. bioRxiv. https://doi.org/10.1101/2020.05.01.071688

Shajahan A, Archer-Hartmann S, Supekar NT, et al (2020b) Comprehensive characterization of N- and O-glycosylation of SARS-CoV-2 human receptor angiotensin converting enzyme 2. bioRxiv. https://doi.org/10.1101/2020.05.01.071688

Shang J, Wan Y, Luo C, et al (2020) Cell entry mechanisms of SARS-CoV-2. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.2003138117

Shental-Bechor D, Levy Y (2008) Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.0801340105

Shi J, Wen Z, Zhong G, et al (2020) Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. Science. https://doi.org/10.1126/science.abb7015

Shivakumar D, Williams J, Wu Y, et al (2010) Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. J Chem Theory Comput 6:1509–1519. https://doi.org/10.1021/ct900587b

Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large numbers of sequences. Methods Mol Biol 1079:105–116. https://doi.org/10.1007/978-1-62703-646-7_6

Singhal T (2020) A Review of Coronavirus Disease-2019 (COVID-19). Indian Journal of Pediatrics

Stout AE, André NM, Jaimes JA, et al (2020) Coronaviruses in cats and other companion animals: Where does SARS-CoV-2/COVID-19 fit? Vet Microbiol 247:108777. https://doi.org/10.1016/j.vetmic.2020.108777

Tay MZ, Poh CM, Rénia L, et al (2020) The trinity of COVID-19: immunity, inflammation and intervention. Nature Reviews Immunology. https://doi.org/10.1038/s41577-020-0311-8

Tipnis SR, Hooper NM, Hyde R, et al (2000) A human homolog of angiotensin-converting enzyme: Cloning and functional expression as a captopril-insensitive carboxypeptidase. Journal of Biological Chemistry. https://doi.org/10.1074/jbc.M002615200

Towler P, Staker B, Prasad SG, et al (2004) ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. J Biol Chem 279:17996–18007. https://doi.org/10.1074/jbc.M311191200

Turner AJ, Hiscox JA, Hooper NM (2004) ACE2: From vasopeptidase to SARS virus receptor. Trends in Pharmacological Sciences

Walls AC, Park YJ, Tortorici MA, et al (2020) Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. https://doi.org/10.1016/j.cell.2020.02.058

Warner FJ, Smith AI, Hooper NM, Turner AJ (2004) Angiotensin-converting enzyme-2: A molecular and cellular perspective. Cellular and Molecular Life Sciences

Waterhouse A, Bertoni M, Bienert S, et al (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303. https://doi.org/10.1093/nar/gky427

Yadav P, Kumar M, Bansal R, et al (2019) Structure model of ferrochelatase from Salmonella Typhi elucidating metalation mechanism. Int J Biol Macromol 127:585–593. https://doi.org/10.1016/j.ijbiomac.2019.01.066

Yan R, Zhang Y, Li Y, et al (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science. https://doi.org/10.1126/science.abb2762

Ye ZW, Yuan S, Yuen KS, et al (2020) Zoonotic origins of human coronaviruses. International Journal of Biological Sciences

Zhai X, Sun J, Yan Z, et al (2020) Comparison of Severe Acute Respiratory Syndrome Coronavirus 2 Spike Protein Binding to ACE2 Receptors from Human, Pets, Farm Animals, and Putative Intermediate Hosts. Journal of Virology. https://doi.org/10.1128/jvi.00831-20

Zhang T, Wu Q, Zhang Z (2020a) Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. Current Biology. https://doi.org/10.1016/j.cub.2020.03.022

Zhang T, Wu Q, Zhang Z (2020b) Probable Pangolin Origin of 2019-nCoV Associated with Outbreak of COVID-19. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3542586

Zhao P, Praissman JL, Grant OC, et al (2020a) Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. Cell Host and Microbe. https://doi.org/10.1016/j.chom.2020.08.004

Zhao WM, Song SH, Chen ML, et al (2020b) The 2019 novel coronavirus resource. Yi chuan Hereditas. https://doi.org/10.16288/j.yczz.20-030

Tables

Table 1 Sequence identity of the human ACE2 receptor with other mammalian species. The calculated LG score from ProQ online server of different species ACE2 receptor homology models are also shown.

S. No.	Species Name	NCBI ID	Sequence Identity (%)	LG score
1.	*Homo sapiens* (Human)	NP_001358344.1	100	6.472
2.	*Ailuropoda melanoleuca* (Giant Panda)	XP_034505781.1	83.55	5.826
3.	*Bos taurus* (Exotic Cattle)	NP_001019673.2	81.21	5.76
4.	*Bubalus bubalis* (Water buffalo)	XP_006041602.1	81.21	5.677
5.	*Canis lupus familiaris* (Dog)	XP_013966804.1	84.18	5.90
6.	*Capra hircus* (Goat)	NP_001277036.1	81.97	5.721
7.	*Camelus dromedarius* (Arabian camel)	XP_010991717.1	83.25	5.657
8.	*Equus asinus* (Ass)	XP_014713133.1	85.99	5.071
9.	*Equus caballas* (Domestic horse)	XP_001490241.1	86.78	5.72
10.	*Felis catus* (Domestic cat)	NP_001034545.1	85.39	5.99
11.	*Loxodonta africana* (African elephant)	XP_023410960.1	81.12	5.688
12.	*Mus musculus* (House Mouse)	NP_001123985.1	81.86	5.40
13.	*Ovis aries* (Sheep)	XP_011961657.1	81.97	5.829
14.	*Oryctolagus cuniculus* (European rabbit)	XP_002719891.1	85.14	5.80
15.	*Panthera tigris altaica* (Siberian tiger)	XP_007090142.1	85.77	5.693
16.	*Rattus rattus* (Black rat)	XP_032746145.1	81.01	5.125
17.	*Sus scrofa* (Wild Boar)	NP_001116542.1	81.08	5.93

Table 2 Buried surface area and binding free energies calculated between SARS-CoV-2 spike protein and the ACE2 receptors.

S. No.	Species	Buried surface Area Å^2	ΔG Kcal/mol (MMGBSA)
1	*Homo sapiens*	949.8	-128.67
2	*Ailuropoda melanoluca*	819.7	-62.46
3	*Bos tauras*	882.9	-92.32
4	*Bubalus bubalis*	829.0	-66.81
5	*Canis lupus familiaris*	898.7	-65.12
6	*Capra hircus*	850.5	-81.49
7	*Camelus dromedarius*	919.4	-126.33
8	*Equus asinus*	842.2	-77.07
9.	*Equus caballus*	871.1	-97.24
10.	*Felis catus*	945.1	-98.04
11.	*Loxodonta africana*	850.0	-78.01
12.	*Mus musculus*	880.2	-62.12
13.	*Ovis aries*	819.6	-90.20
14.	*Oryctolagus cuniculus*	945.3	-93.07
15.	*Panthera tigris altaica*	822.6	-52.41
16.	*Rattus rattus*	800.8	-43.21
17.	*Sus scrofa*	873.8	-46.31

Table 3 Comparison of the hydrogen bonds, ionic bonds and hydrophobic interactions observed between SARS-CoV-2 spike protein RBD and ACE2 receptor from different species.

Types of interactions	*Homo sapiens*	*Ailuropoda melanoluca*	*Bos tauras*	*Bubalus bubalis*	*Canis lupus familiaris*	*Capra hircus*	*Camelus dromedarius*	*Equus asinus*	*Equus caballus*
Hydrogen bond	Q24 - N487 K31 - Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500	L24 x N487 K31 - Q493 Y34 x Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K354 - G496 K354 - G502 D356 - T500	Q24 - N487 K31 - Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K352- G496 K352 - G502 D354 - T500	Q23 - N487 K3O - Q493 H33 - Y453 E34 - Q493 D37 - Y449 Y40 - T500 Q41 - G446 Q41 - Y449 Y82 - Y489 Y82 - N487 K351 - G496 K351 - G502 D353 - T500	L23 x N487 K30 - Q493 Y33 - Y453 E34 - Q493 E37 - Y449 Y40 - T500 Q41 - G446 Q41 - Y449 Y82 - Y489 Y82 - N487 K352 - G496 K352 - G502 D354 - T500	Q24 - N487 K31 - Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K352 - G496 K352 - G502 D354 - T500	L24 x N487 E31 x Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500	L24 x N487 K31 - Q493 S34 x Y453 E35 - Q493 E38 - Y449 H41 x T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K331 - G496 K331 - G502 D333 - T500	L24 x N487 K31 - Q493 S34 x Y453 E35 - Q493 E38 - Y449 H41 x T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500
Hydrophobic	T27- F456 H34 - Q493 Y41 - Q498 L79 - F486 M82 - F486 Y83 - F486 K353 - Y505	- - Y41 - Q498 H79 - F486 - – K354 - Y505	T27 - F456 - Y41 - Q498 M79 - F486 T82 - F486 Y83 - F486 K352 - Y505	- - - - T81 - F486 - K351 - Y505	- - - L78 - F486 T81 - F486 - K351 -Y505	- - - M79 - F486 - - K352 -Y505	T27 - F456 - Y41 - Q498 T79 - F486 T82 - F486 Y83 - F486 K353 -Y505	- - - - T82 - F486 - K331 -Y505	- - H41 - Q498 - T82 - F486 Y83 - F486 K353-Y505
Ionic	D30 - K417	E30 - K417	Absent	E29 - K417	E29 - K417	E30 - K417	E30 - K417	E30 - K417	E30 - K417

Types of interactions	*Homo sapiens*	*Felis catus*	*Loxodonta africana*	*Mus musculus*	*Ovis aries*	*Oryctolagus cuniculus*	*Panthera tigris altaica*	*Rattus rattus*	*Sus scrofa*
Hydrogen bond	Q24 - N487 K31 - Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500	L24 x N487 K31- Q493 H34 - Y453 E35 - Q493 E38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 - Y489 Y83 - N487 K353 - G502 D355 - T500	L24 x N487 T31 x Q493 Q34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 x G446 Q42 x Y449 F83 x Y489 F83 x N487 K348 - G496 K348 - G502 D350 - T500	N24 x N487 N31 - Q493 Q34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 F83 x Y489 F83 x N487 H353 x G496 H353 - G502 D355 - T500	Q24 -N487 K31 - Q493 H34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 x Y489 Y83 - N487 K352 - G496 K352 - G502 D354 - T500	L24 x N487 K31 - Q493 Q34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 x G446 Q42 - Y449 Y83 x Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500	L16 x N487 K23 - Q493 H26 - Y453 E27 - Q493 E30 x Y449 Y33 - T500 Q34 - G446 Q34 - Y449 Y75 - Y489 Y75 - N487 K345 - G496 K345 - G502 D347 - T500	K24 - N487 K31 - Q493 Q34 - Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 x G446 Q42 - Y449 F83 x Y489 F83 x N487 H353 x G496 H353 - G502 D355 - T500	L24 x N487 K31 - Q493 L34 x Y453 E35 - Q493 D38 - Y449 Y41 - T500 Q42 - G446 Q42 - Y449 Y83 x Y489 Y83 - N487 K353 - G496 K353 - G502 D355 - T500
Hydrophobic	T27 - F456 H34 - Q493 Y41 - Q498 L79 - F486 M82 - F486 Y83 - F486 K353 - Y505	- - - L79 – F486 T82 – F486 - K353 -Y505	T27 - F456 - Y41 - Q498 L79 - F486 D82 - F486 Y83 - F486 K348 - Y505	- - Y41 - F498 T79 - F486 S82 - F486 F83 - F486 H353- Y505	T27 - F456 - - M79 - F486 T82 - F486 - K352 - Y505	- - - - - Y83 - F486 K353 - Y505	- - - - T74 - F486 Y75 - F486 K345 - Y505	- - Y41 - Q498 I79 - F486 D82 - F486 F83 - F486 H353 - Y505	- L34 - Q493 Y41 - Q493 I79 - F486 - Y83 - F486 K353 - Y505
Ionic	D30 - K417	E30 - K417	D30 - K417	Absent	E30 - K417	E30 - K417	E22 - K417	Absent	E30 - K417

x - indicates the absence of interaction due to residue substitution

Table 4 Comparison of N-linked glycosylation sites and the supporting residues of the ACE2 receptor from different species.

Name of the species	N-linked glycosylation sites and their supporting residues								
	N53				N90	N103			N322
Homo sapiens	T55	E57	N58	Q340	K26	Q81	H194	N195	E312
Ailuropoda melanoleuca	T	E	N	Q	E	Absent			E
Bos taurus	T	E	N	R	K	Absent			Absent
Bubalus bubalis	T	E	N	R	K	Absent			Absent
Canis lupus familiaris	T	E	N	R	Absent	Absent			E
Capra hircus	T	E	N	R	K	Absent			Absent
Camelus dromedarius	T	E	N	R	K	Absent			E
Equus asinus	T	E	N	R	K	Absent			E
Equus caballus	T	E	N	R	K	Absent			E
Felis catus	T	E	N	R	K	Absent			E
Loxodonta africana	T	E	N	R	Absent	Absent			Absent
Mus musculus	T	E	N	R	Absent	Absent			Absent
Ovis aries	T	E	N	R	K	Absent			Absent
Oryctolagus cuniculus	T	E	N	R	K	Absent			Absent
Panthera tigris altaica	T	E	N	Q	K	Absent			E
Rattus rattus	T	E	N	R	Absent	Absent			Absent
Sus scrofa	T	E	N	R	Absent	Absent			E

Figures

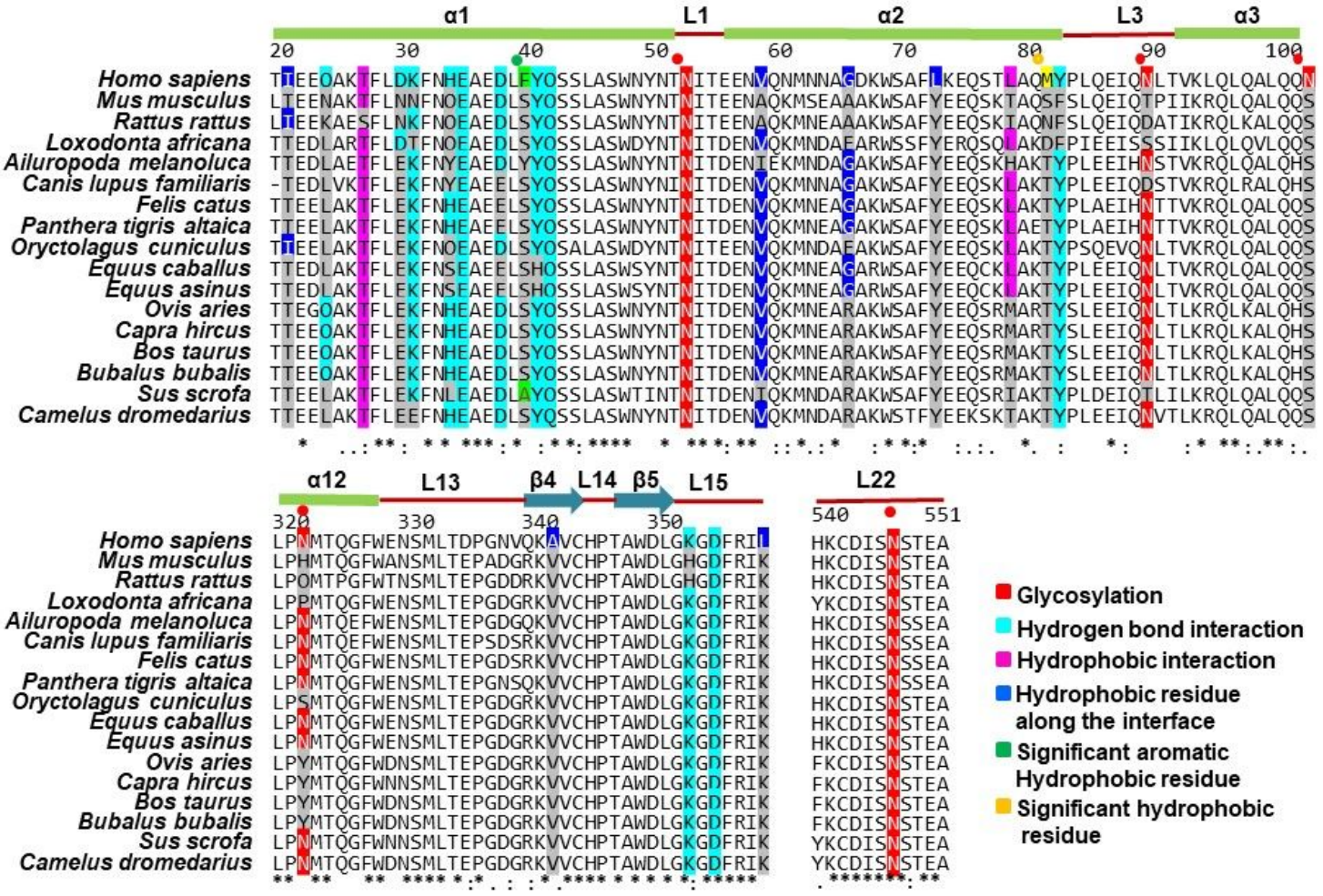

Fig.1

Figure 1

Multiple Sequence alignment of the ACE2 receptor from selected mammalian species. The key residues involved in hydrogen-bonded (cyan), hydrophobic (pink) and glycosylation sites (red) are highlighted. The hydrophobic residues lined in the concave interacting surface have also been highlighted in blue color. Hydrophobic residues having a significant role in interaction with spike protein are highlighted in green and yellow color.

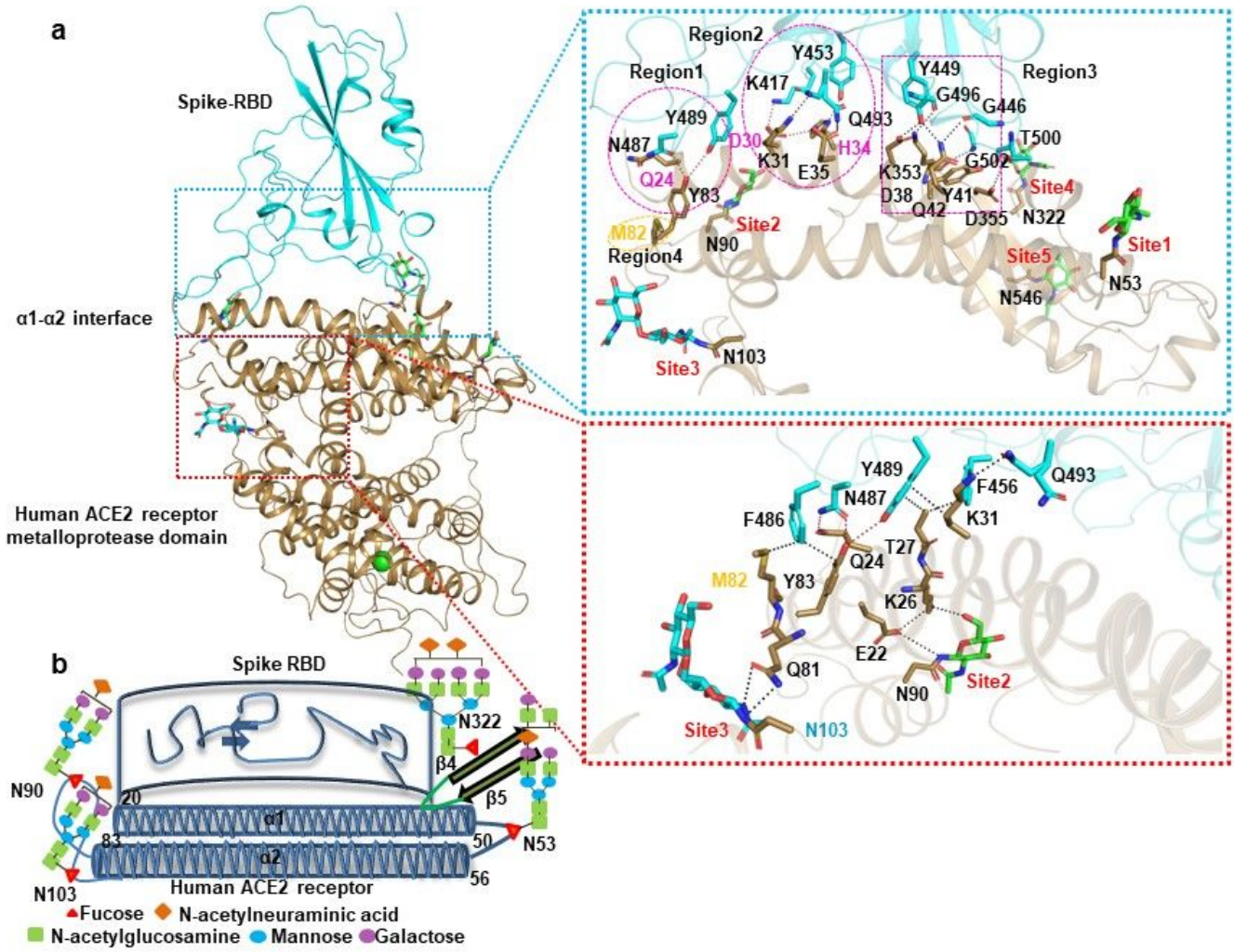

Fig.2

Figure 2

The cartoon representation of human ACE2 receptor in complex with COVID-19 spike-RBD domain. a The complex of human ACE2 receptor metalloprotease domain and SARS-CoV-2 spike receptor-binding domain. A close view of residues involved in interactions at four different interacting regions with the four glycosylation sites; b Model representation of four important glycosylation sites with sugar moieties of human ACE2 receptor required for spike recognition. The sugar moieties are made based on the study done by Varki, A., Cummings, R.D., at al. 2015.

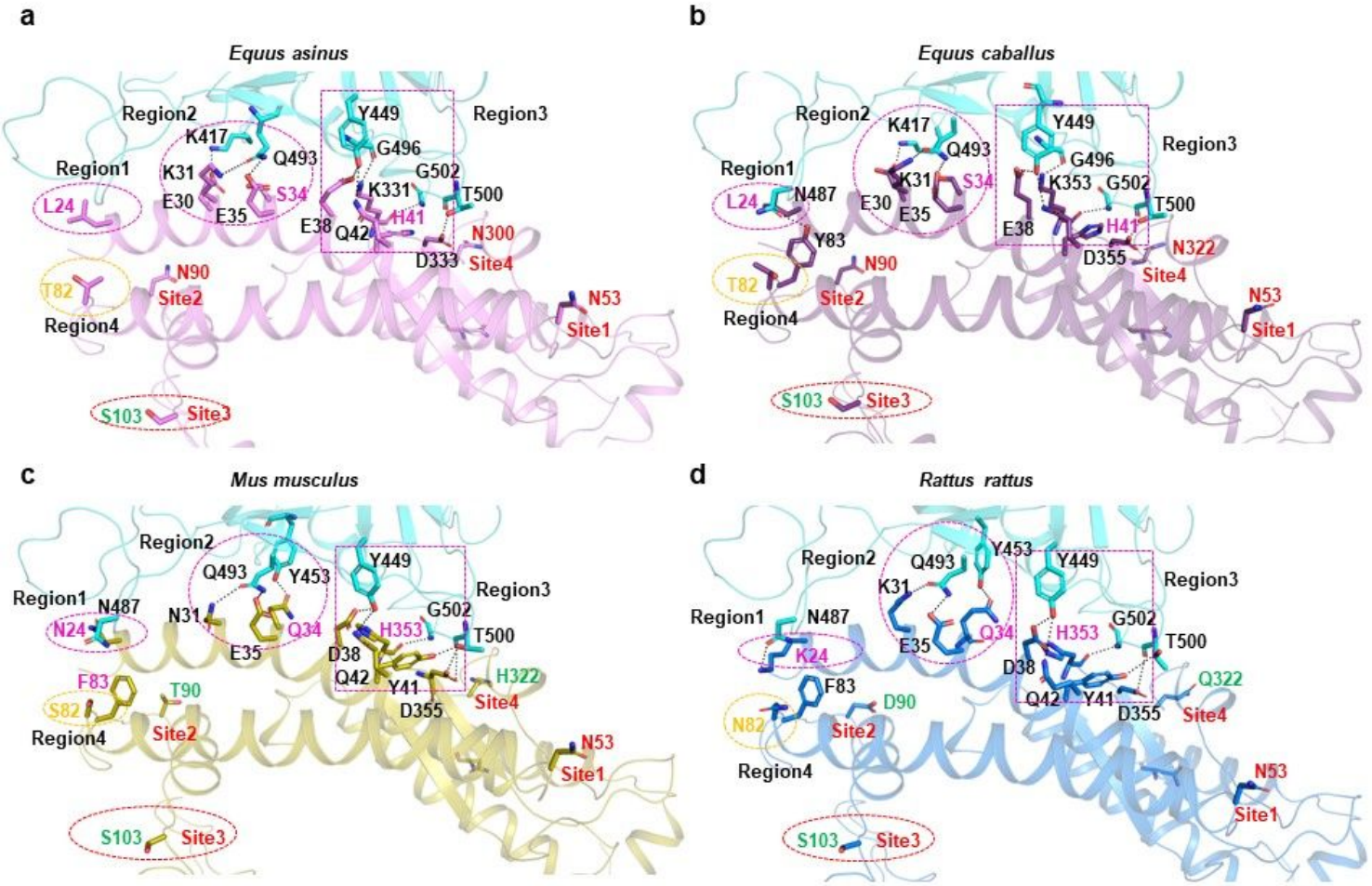

Fig.3

Figure 3

The ACE2 receptor-interacting surface of a Equus asinus; b Equus caballus; c Mus musculus; d Rattus rattus with spike 2 protein. Interactions are shown from region 1 to region 4 and the natural substitution of residues are colored as pink and yellow. Glycosylation sites are marked in red and substitution of glycosylating residue colored in green.

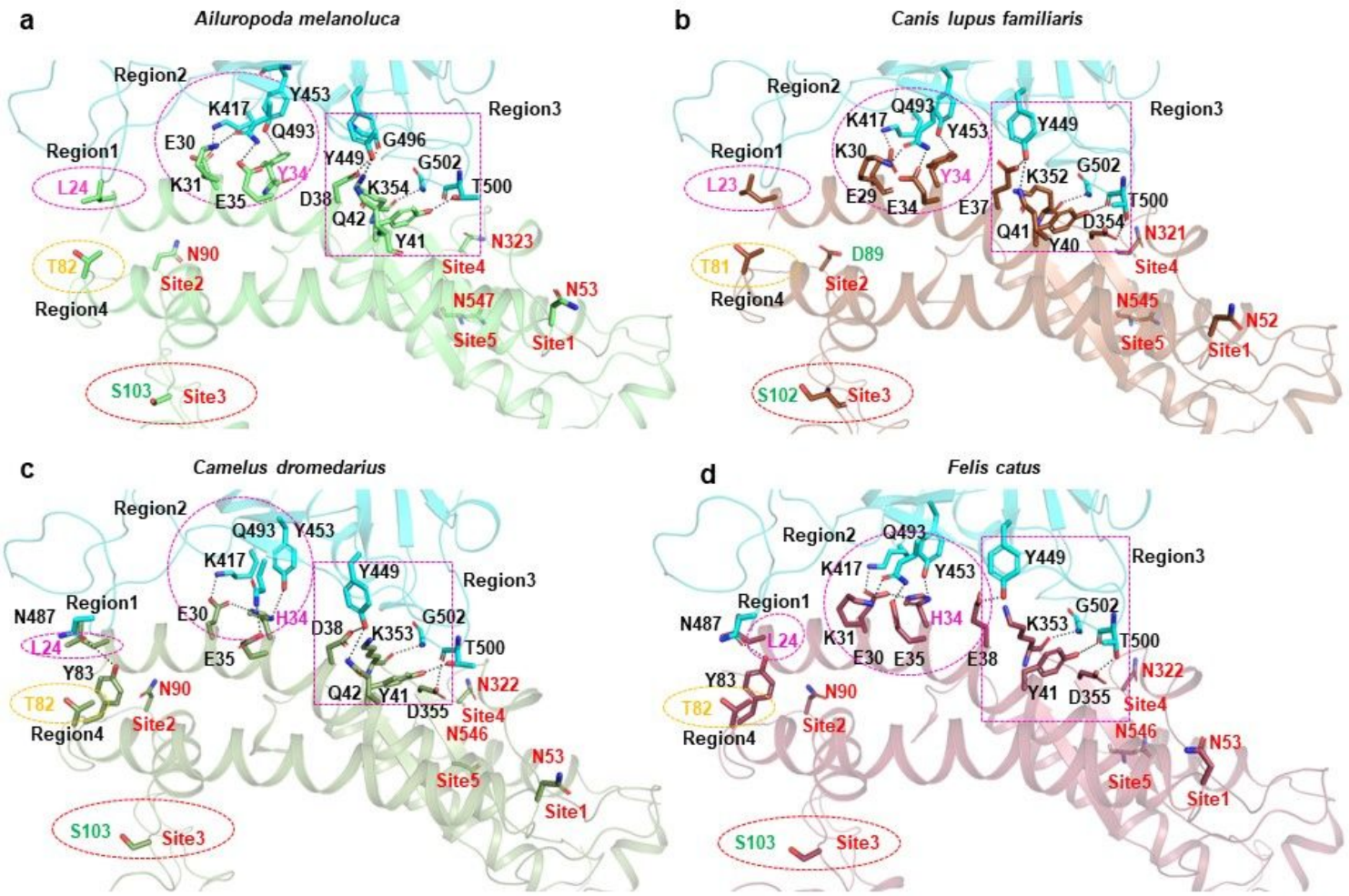

Fig.4

Figure 4

The ACE2 receptor-interacting surface of a Ailuropoda melanoluca; b Canis lupus familiaris; c Camelus dromedarius; d Felis catus. Interactions are shown from region 1 to region 4 and the natural substitution of residues are colored as pink and yellow. Glycosylation sites are marked in red and substitution of glycosylating residue colored in green.

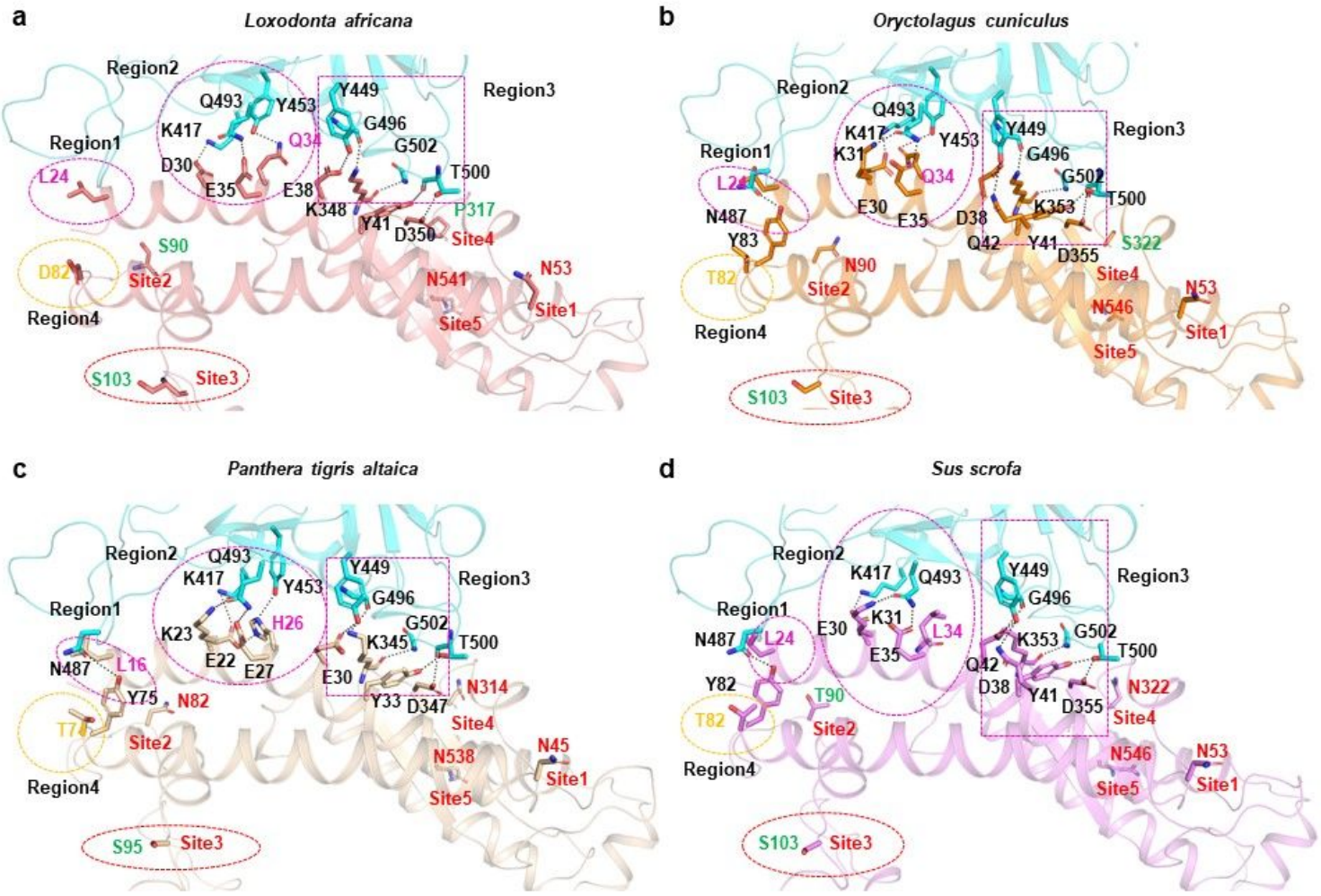

Fig.5

Figure 5

The ACE2 receptor-interacting surface of a Loxodonta africana; b Oryctolagus cuniculus; c Panthera tigris altaica; d Sus scrofa. Interactions are shown from region 1 to region 4 and the natural substitution of residues are colored as pink and yellow. Glycosylation sites are marked in red and substitution of glycosylating residue colored in green.

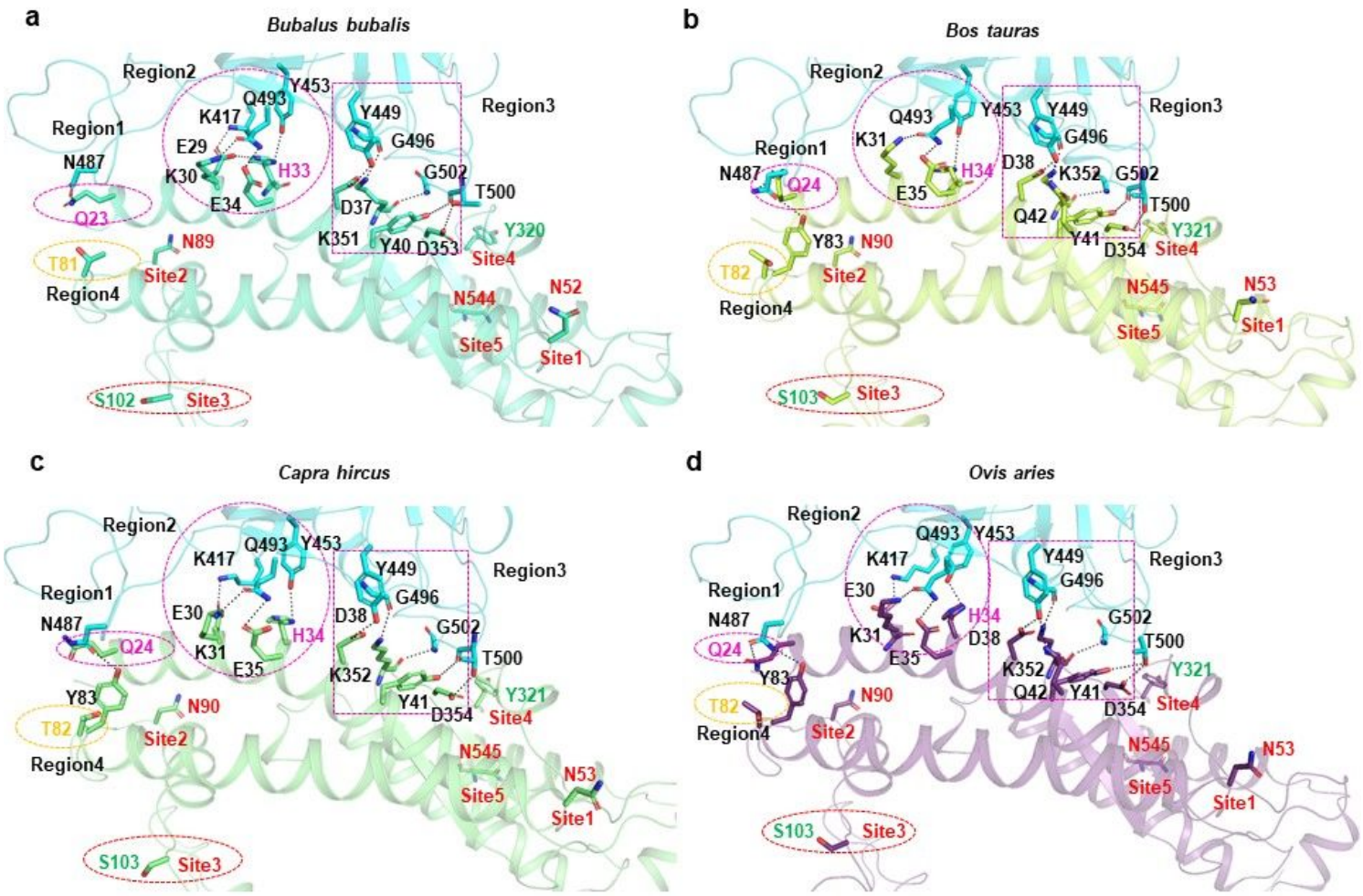

Fig.6

Figure 6

The ACE2 receptor-interacting surface of a Bubalus bubalis; b Bos tauras; c Capra hircus; d Ovis aries. Interactions are shown from region 1 to region 4 and the natural substitution of residues are colored as pink and yellow. Glycosylation sites are marked in red and substitution of glycosylating residue colored in green.

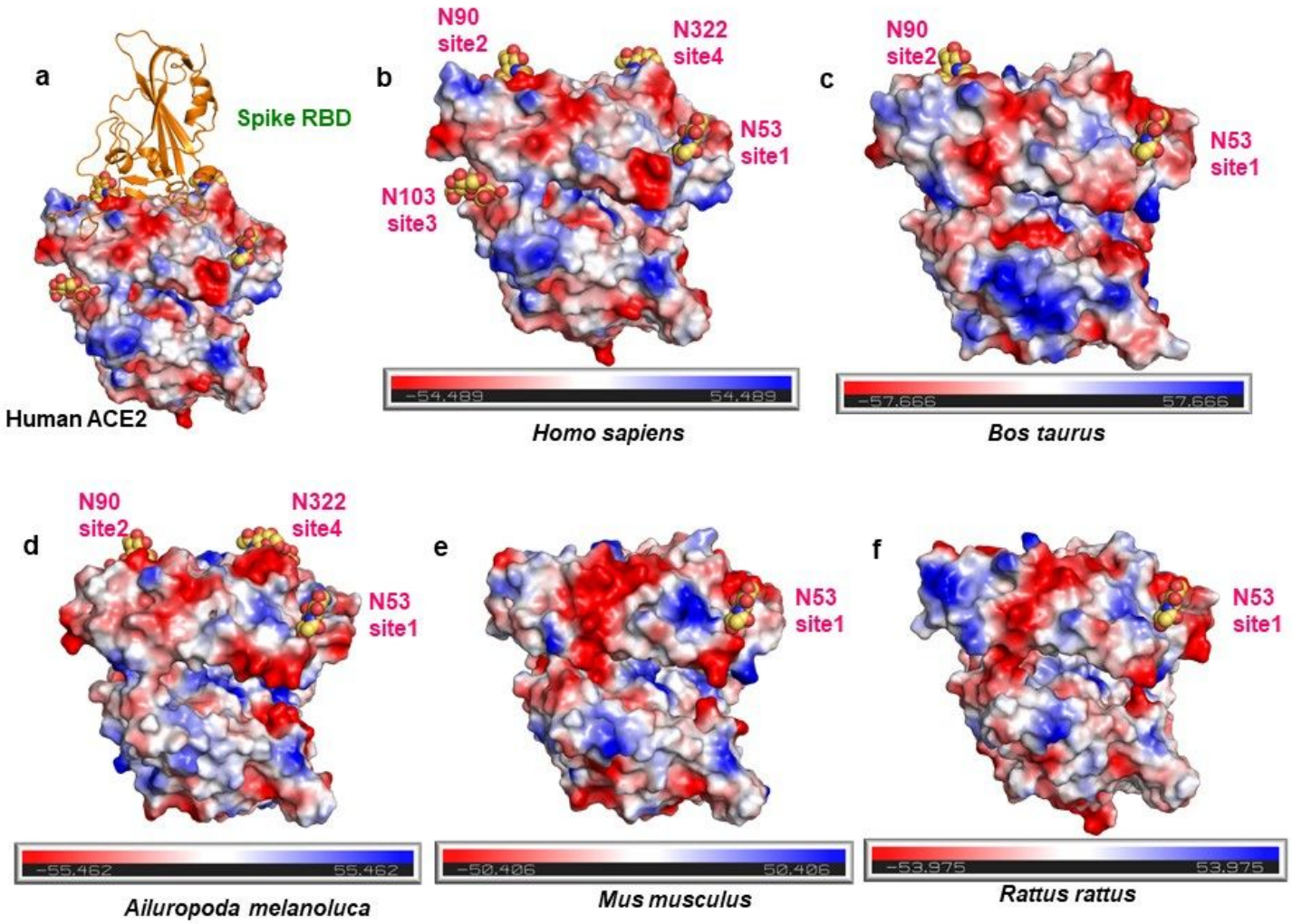

Fig.7

Figure 7

The electrostatic surface potential of the human ACE2 receptor and four other mammalian species. The surface potential are grouped based on conserved and substitution of the amino acid residue Gln24. a-b human ACE2 receptor; c Bos taurus (Q24; d Ailuropoda melanoluca (Q24L); e Mus musculus (Q24N); f Rattus rattus (Q24K).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryfiguresTables.pdf