

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

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

# Virus Research

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

# SARS-CoV-2 host tropism: An in silico analysis of the main cellular factors

H.R. Rangel<sup>a</sup>, J.T. Ortega<sup>b</sup>, S. Maksoud<sup>c</sup>, F.H. Pujol<sup>a</sup>, M.L. Serrano<sup>d,\*</sup>

<sup>a</sup> Laboratorio de Virología Molecular, Centro de Microbiología y Biología Celular, Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela

<sup>b</sup> Department of Pharmacology, School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA

<sup>c</sup> Department of Neurology and Experimental Therapeutics and Molecular Imaging Laboratory, Massachusetts General Hospital, MA 02129, USA

<sup>d</sup> Unidad de Química Medicinal, Facultad de Farmacia, Universidad Central de Venezuela, Caracas 1041-A, Venezuela

ARTICLE INFO

Keywords: COVID-19 SARS-CoV-2 Receptor ACE2 TMPRSS2 GRP78

## ABSTRACT

Recent reports have shown that small and big felines could be infected by SARS-CoV-2, while other animals, like swines and mice, are apparently not susceptible to this infection. These findings raise the question of the role of cell factors associated with early stages of the viral infection in host selectivity. The cellular receptor for SARS-CoV-2 is the Angiotensin Converting Enzyme (ACE2). Transmembrane protease serine 2 (TMPRSS2) has been shown to prime the viral spike for its interaction with its receptor. GRP78 has also been proposed as a possible coreceptor. In this study, we used several bioinformatics approaches to bring clues in the interaction of ACE2, TMPRSS2, and GRP78 with SARS-CoV-2. We selected several mammalian hosts that could play a key role in viral spread by acting as secondary hosts (cats, dogs, pigs, mice, and ferrets) and evaluated their predicted permissiveness by *in silico* analysis. Results showed that ionic pairs (salt bridges, N–O pair, and long-range interactions) produced between ACE2 and the viral spike has an essential function in the host interaction. On the other hand, TMPRSS2 and GRP78 are proteins with high homology in all the evaluated hosts. Thus, these proteins do not seem to play a role in host selectivity, suggesting that other factors may play a role in the non-permissivity in some of these hosts. These proteins represent however interesting cell targets that could be explored in order to control the virus replication in humans and in the intermediary hosts.

## 1. Introduction

The recent outbreak associated with Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has made a deep impact on modern science. Several research groups have been working to achieve two main goals: to obtain a vaccine and/or pharmacological targets that could help to control the pandemic (Ahn et al., 2020; Amanat and Krammer, 2020; Liu et al., 2020; Ricke and Malone, 2020; Tu et al., 2020). However, as it happened in previous outbreaks, we need to fully understand the viral biology/host interaction to reach these goals (Parvez and Parveen, 2017). In this report, we are focused on the early stages driving host selectivity for SARS-CoV-2. Similar to SARS-CoV, SARS-CoV-2 infection starts with the interaction between the viral spike protein and its counterpart, the ACE2 cell receptor (Wang et al., 2020,). Previously, we predicted that the changes in the SARS-CoV-2 viral spike protein improved its ability to bind to ACE2, compared to SARS-CoV (Ortega et al., 2020). Nevertheless, more scientific data have been published since then, helping us to better understand the main factors associated with these interactions. First, several crystals structures detailing the

interaction between the viral spike protein and ACE2 are available in the Protein Data Bank (Lan et al., 2020; Wang et al., 2020; Yan et al., 2020). Also, new receptors, such as GRP78, are described as possible co-receptors for SARS-CoV-2 (Ibrahim et al., 2020). GRP78 is a chaperone protein related to the unfolded protein response in the Endoplasmic Reticulum (ER) (Ge and Kao, 2019). Under stress conditions, GRP78 is over expressed and translocated to the cell membrane (Wang et al., 2009); there, the substrate-binding domain (SBD) serves as a receptor for viral entry (Ibrahim et al., 2020). Furthermore, these receptors have been correlated with other CoVs such as MERS, Bat-HKU9, and other viral species (Dengue and Human papillomavirus) (Nain et al., 2017; Chu et al., 2018; Jindadamrongwech et al., 2004; Elfiky, 2020).

After the virus interacts with its receptor, several changes occur driving viral internalization *via* endocytosis (Letko et al., 2020). Nonetheless, the proteolytic processing of the viral spike protein is a key step to expose the fusion motif and release the virus to the cytoplasm (Hoffmann et al., 2020); this process is well known for other coronaviruses (SARS and MERS) (Belouzard et al., 2009; Glowacka et al., 2011; Kleine-Weber et al., 2018). Recently, Hoffman et al. (2020) described that SARS-CoV-2 priming

https://doi.org/10.1016/j.virusres.2020.198154

Received 29 May 2020; Received in revised form 1 September 2020; Accepted 3 September 2020 Available online 9 September 2020 0168-1702/© 2020 Elsevier B.V. All rights reserved.





<sup>\*</sup> Corresponding author. *E-mail address:* maria.serrano@ucv.ve (M.L. Serrano).

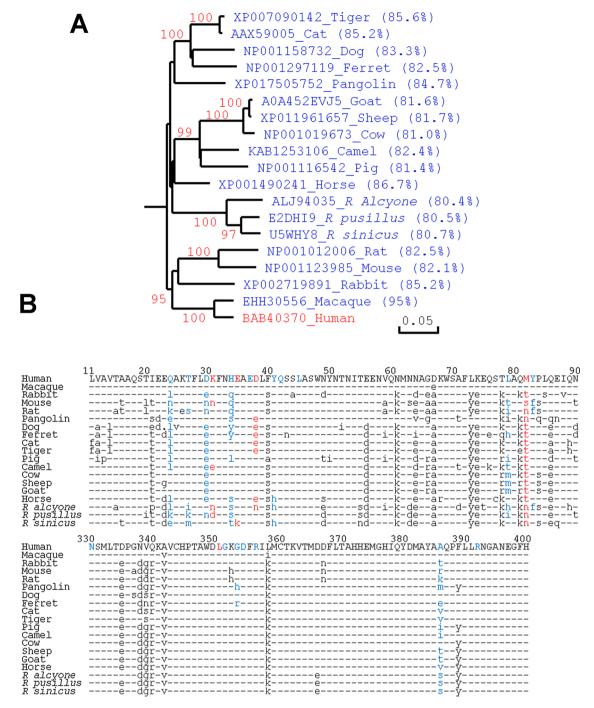


Fig. 1. Sequence analysis for ACE2 receptor from several mammalian species, including some SARS-CoV-2 putative host. A) Phylogenetic analysis of ACE2 proteins. Distance was calculated with Poisson correction. Each protein is named with its accession number and percent homology with human protein is shown. B) Protein alignment of ACE2 regions showing the important (blue) and critical (red) residues involved in the interaction with SARS-CoV-2 RBD.

through proteolytic processing is driven by Transmembrane protein 2 (TMPRSS2). However, there is no clear evidence of whether or not ACE2, GRP78, or TMPRSS2 are crucial to determine the host selectivity for SARS-CoV-2. Interestingly, recent reports showed that SARS-CoV-2 could infect domestic cats and ferrets, partially dogs but not pigs (Shi et al., 2020). Thus, there is a need to know the possible reservoirs or alternative viral hosts to fully control the pandemic. Moreover, a better understanding of the viral infection in these hosts could provide some clues to control the virus in humans. Based on this, we established a comparative *in silico* analysis of the main factors (ACE2, GRP78, and TMPRSS2) associated with SARS-CoV-2 entry to understand the structural mechanism related to the host selectivity.

#### 2. Methods

## 2.1. Sequence analysis

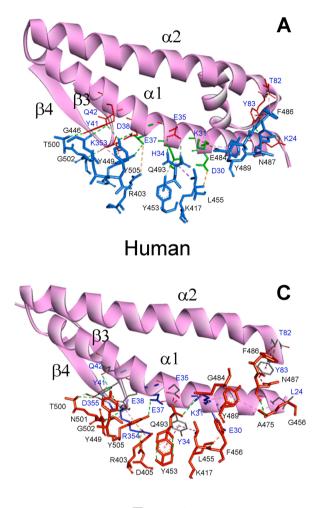
Sequences for ACE2 and TMPRSS2 were retrieved from GenBank; accession numbers are shown in the sequence alignments. Sequences used for GRP78 are shown in the respective alignment. Multiple sequence alignments and phylogenetic analysis were generated using DNAMAN with default parameters.

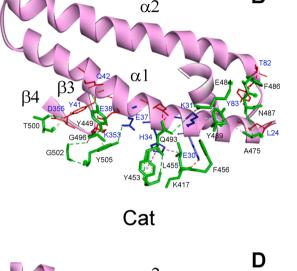
Β

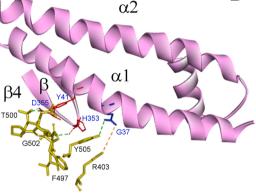
## 2.2. Protein modeling

The crystal structures of the SARS-CoV-2 viral spike protein bound to human ACE2, PDB code 6M0J, 6LZG, 6M17, were retrieved and

analyzed. Homology structural models of mice, cat, pig, ferret, and dog for ACE2 were determined. The data obtained for the crystal structure 6M0J were selected for the comparative analyses between the different viral spike protein-ACE2 interactions. The homology models for the non-







Mice

Ferret

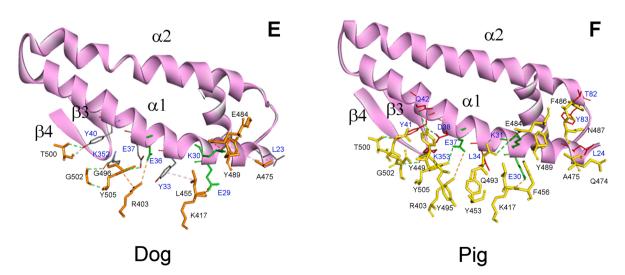


Fig. 2. Docking between the SARS-CoV-2 spike protein and ACE2 for some mammalian hosts. The best-docked structure between the viral spike protein and the hosts ACE2 is presented. Also, the main interacting atoms are shown for each protein. Residues involved in the formation of ionic pairs are colored differently. The chains alpha and beta as the main interaction regions on ACE2 are showed.

human ACE2 enzymes were generated using the crystal structure of human ACE2 as a template (PDB code 6M17). On the other hand, the structural modeling of TMPRSS2 was developed using the crystal structure of Serine protease Hepsin (PDB code 5CE1) as a template (Ortega et al., 2020c). All models were obtained with the SWISS-MODEL modeling server and the DeepView/Swiss-PdbViewer 4.01 software (Arnold et al., 2006). Hydrogen atoms on each structure were added and partial charges were assigned for energy refinement. The obtained models were subjected to a MD simulation using NAMD 2.9 (Phillips et al., 2005), as described by Ortega et al. (2019) using the CHARMM 22 force field (Vanommeslaeghe et al., 2010) and Gasteiger charges. The obtained structures represent the lowest energy frame of the MD simulations. The quality of the models was established *via* ProSA (Wiederstein and Sippl, 2007) and PROCHECK programs (Laskowski et al., 1993).

# 2.3. Protein-protein docking

The crystal structure for SARS-CoV-2 viral spike protein bound to human ACE2 (PDB code 6M0J), human ACE2 (PDB code 6M17) and for human GRP78 (PDB code 5E84) were downloaded from Protein Data Bank. Protein preparation was carried out as described above. Human ACE2 was used as a template for the different ACE2 hosts. These models were also evaluated against the viral spike protein through molecular docking. Then, the obtained binding patterns and affinity estimations were analyzed and compared. The docking was developed in two steps: first, the structure of the complex between the ligand (viral spike protein) and receptor (ACE2 or GRP78), delimiting the Receptor Binding Domain (RBD) to that described for human ACE2 6M0J or the Substrate Binding Domain (SBD)  $\beta$  for GRP78, was performed using Z-dock 3.0.2 software (Pierce et al., 2014). Second, the resulting docking data were processed and analyzed employing the tools from PRODIGY software (Xue et al., 2016). Finally, these results were processed, clustered, and analyzed considering binding energies and interacting residues for each viral spike protein-ACE2 or GRP78 complex.

# 2.4. Molecular docking

TMPRSS2 homology models from different hosts were used to evaluate the binding of TMPRSS2 human inhibitors. The initial models were optimized through addition of hydrogen atoms and partial charges. The obtained structures were further submitted to restrained molecular mechanic refinement via NAMD software (Phillips et al., 2005), using the CHARMM22 force field (Vanommeslaeghe et al., 2010). Next, a structural analysis of the binding pocket was performed using the Swiss PDB viewer server (Guex and Peitsch, 1997). The ligand-binding pocket located in the catalytic site was acquired manually and then verified by a priori docking approach with camostat mesylate through the Achilles Blind Docking server (Sánchez-Linares et al., 2012). The 3D structure of each inhibitor was taken from PubChem and ZINC databases. Molecular docking was performed with VINA/Vega ZZ 3.1.0.21 (Pedretti et al., 2004) and 30 runs conducted for each compound (Ortega et al., 2020b). Results were prioritized according to the predicted free energy of binding energy in kcal/mol. Docking results were visualized via the Biovia Discovery Studio Visualizer 17.2.0 software.

### 3. Results

Within this study, we developed a bioinformatics analysis of three proteins (ACE2, TMPRRS2, and GRP78) described as main factors associated with the early steps of SARS-CoV-2 infection in several hosts. First, sequence analysis was carried out for a broad range of hosts, emphasizing our interest in the main interacting regions related to SARS-CoV-2 spike protein recognition and establishing comparison to the well-known susceptible hosts. Then, proteins of hosts with known susceptiblity to SARS-CoV-2 were evaluated for homology modeling

and docking studies. Altogether, these bioinformatics approaches help to gain deeper understanding of the molecular interactions between the viral spike and these proteins and their relation to host susceptibility.

# 3.1. ACE2

The viral spike protein mediates coronavirus entry into host cells. SARS-CoV-2 spike protein contains a receptor-binding domain (RBD) that specifically recognizes ACE2, which appears to be the key functional receptor for the SARS-CoV-2. The receptor recognition of the spike protein is an important determinant of viral infectivity, pathogenesis, and host range. In the present study, we developed homology models of a series of ACE2 proteins from different hosts (mice, cat, pig, ferret, and dog) in order to gain insight about the structural mechanism related to host selectivity. Then, a computational protein-protein docking approach was performed to predict the structure of complexes with the SARS-CoV-2 RBD and compared it with that of the human ACE2 complex. Phylogenetic analysis of the ACE2 from different mammalian hosts reveals a high sequence identity among them (with more than 80 % identity in comparison to the human protein) (Fig. 1A and Table S1). Comparison of the amino acid sequences of the contact residues directly involved in the recognition of the SARS-CoV-2 RBD domain are shown in Fig. 1 B. The ACE2 enzyme showed substitution of specific amino acids residues of the contact region and particularly of those involved in the recognition of the SARS-CoV-2 RBD domain. The amino acid residues Lys31, Glu35, Asp38, Met82, and Lys353 from the human cell receptor have been linked to recognition of viral spike protein (Shang et al., 2020). Based on the homology observed between the different proteins and the known degree of susceptibility of each host previously reported (Shi et al., 2020), the structural models of ACE2 enzymes (mice, cat, pig, ferret, and dog) were established using the human enzyme as a template (PDB code 6M17). Comparison of homology models of ACE2 from selected mammalian hosts with that of human suggested that the structures are very similar, with a root mean square deviation (RMSD) ranging from 0.09 to 0.77 (708 C- $\alpha$  atoms). In addition, an examination of the Ramachandran plot (not shown) indicated a good overall geometry for ACE2 models. The protein folding energy from the protein models exhibited typical features of native structures (Table S1). These resulting ACE2 models (peptidase domain (PD) of ACE2) were employed to collect the complexes with the SARS-Co-V-2 RBD domain. The hACE2 receptor recognition of the SARS-CoV-2 RBD domain has been extensively studied and these studies showed that the SARS-Co-V-2 RBM forms a large binding interface with 20 amino acids of hACE2 (Lan et al., 2020; Shang et al., 2020; Yan et al., 2020). The overall structures of ACE2 from different mammalian hosts in conjunction with the RBD domain are similar to those of the hACE2/SARS-CoV-2 RBD complex (Table S2). These results show that ACE2 enzymes recognize the RBD domain of the viral spike protein with the putative binding domain, including mainly the arch-shaped  $\alpha$ 1 helix. The  $\alpha 2$  helix and the loop that connects the  $\beta 3$  and  $\beta 4$  antiparallel strands also make limited contributions to the coordination of RBD (Fig. 2 and Table S3). Despite the overall similarity, some non-synonymous sequence variations are found on their respective interface. In the ACE2  $\alpha$ 1 helix: Gln24 $\rightarrow$ Leu24 (dogs, pigs, cats, and ferrets), Asp30→Asn30 (mice) and His34→Tyr34 (dogs and ferrets)/Leu34 (pigs). In the ACE2  $\alpha$ 2 helix: Leu79 $\rightarrow$ Thr79 (mice)/His79 (ferrets), Met82→Ser82 (mice)/Thr82 (dogs, pigs, cats and ferrets) and Tyr83 $\rightarrow$ Phe83 (mice). Variations found in the  $\alpha$ 1 helix between the different mammalian hosts reduce the hydrogen bonds formed with the RBD domain of the viral spike protein in mice and dogs complexes (Fig. 2D and 2E). The variations found at the end of the  $\alpha$ 2 helix leaves this region without any hydrogen bond in the mice ACE2/SARS-CoV-2 RBD complex (Fig. 2D). The formation of ionic pairs (salt bridges, N-O pair, and long-range interactions) (Kumar and Nussinov, 2002) between ACE2/SARS-CoV-2 RBD complexes was also examined (Table 1). Linking the arch-shaped  $\alpha$ 1 helix of ACE2 with the RBD

#### Table 1

Differences based on key residues describes as fundamental for host selectivity.

Δ

Host	Persimitivity for SARS-CoV-2	Number of ionic pairs	Residues ACE2 and Spike (length in Å)				
Human	Infected	4	Asp30-Lys417	Lys31-Glu484	His34-Glu406	Glu37-Arg403	
			3.14	5.1	6.67	6.4	
Mice	Non-infected	1				Glu37-Arg403	
						6.1	
Dog	Partially infected	3	Glu29-Lys417	Lys30-Glu484		Glu36-Arg403	
			1.79	5.19		5.46	
Pig	Non-infected	3	Glu30-Lys417	Lys31-Glu484		Glu37-Arg403	
			3.93	4.06		6.03	
Cat	Infected	4	Glu30-Lys417	Lys31-Glu484	His34-Glu406	Glu37-Arg403	
			3.36	4.1	6.58	6.0	
Ferret	Infected	4	Glu30-Lys417	Lys31-Glu484		Glu37-Arg403	Arg354-Asp405
			3.3	4.1		6.09	3.4

100 \_\_\_\_\_ NP005647\_Human

		L XP0287	01149_Macaque	(87.8%)		
	1	NP056	590 Mouse (78.4	1 % )		
	μ. μ.					
		100 BAB70	683_Rat (76.98	5)		
	F	XP008	250697 Rabbit	(75.8%)		
	Iſ					
		100 C NP00107	5054_Cow (78.6	58)		
			.6505 Sheep (79			
	U	1 4				
	d.		322124_Goat (7			
		BAF76	737_Pig (77.2%)	)		
		XP0109	93167 Camel (7	77.6)		
	L			0.0000.00		
		BBD3386.	1_Dog (79.9%)			
			- XP012916721_F	erret (70.3%)		
	10	. XP0230	94479 Cat (79.	7%)		
	10					
В		100	<u> </u>	Tiger (74.7%)		
					0.05	
XP028701149_Macaque	RIVGGESqn	·····	i	sqv	AGILRQSFMFYGAGYQVEKVISHP -tsek-hr	ĥt
NP005647 Human XP028701149 Macaque NP056590 Mouse BaB70683 Bat	qn	l		sqv ess-ry	/-tsek-hr	ĥt ta
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit	qn. ln. st.	l sd sd	i g giqq	ss-qv ess-ry ress-ry 1-esary-m	7-tsek-hr srhsrh k-lsrh sfs-l-psa	ht 
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023485975_Horse	qn. ln. st. sh.	1 sd sd ae	~_i	v	/-tsek-hr srh	ñt ta ta t it ea
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023485975_Horse NP001075054_Cow	qn. ln. st. sh. sn.		~_i g giq giq dgi	 	/-tsek-hr srhsrh s-fs-l-psa s-len-hr-ga k-ysr-a	
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023485975_Horse NP001075054_Cow BBD33861_Dog	qn ln st sh sh sn sn		i		/-tsek-hr shsh s-fs-1-psa s-1en-hr-ga k-ys-r-a	ht 
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023485975_Horse NP001075054_cow BBD33861_Dog XP027816505_Sheep	qn ln st sh sh sn sn		i- giq	8q 8	/-tsek-hr srhsrh s-fs-l-psa s-len-hr-ga k-ysr-a	 
XP028701149_Macaque NP056550_Mouse BAB70683_Rat XP008250597_Rabbit XP03485375_Horse NP001075054_Cow BBD33861_Dog XP037816505_Sheep XP013822124_Coat XP013816721_Ferret	qn. 		i		r-tsek-hr- 	ht
XP028701149_Macaque NP056590_MOUSe BAB70683_Rat XP008250597_Rabbit XP023485975_Horse NP001075054_Cow BBD33861_Dog XP013861505_Sheep XP013822124_Goat XP012916721_Ferret XP012916721_Ferret	qn. 		i		r-tsek-hr- srh	
XP028701149_Macaque NP056500_MOUSe BAB70683_Rat XP008250637_Rabbit XP003465975_Horse BBD33861_Dog XP027816505_sheep XP013822124_Coat XP012916721_Perret XP015336688_Tiger	qn. 		i gi gi 		r-tsek-hr- srh	
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023465975_Horse NP001075054_Cow BBD33861_Dog XP01386505_Sheep XP013822124_Coat XP01291672_Ferret XP023094479_Cat XP01395467_Came1	qn. 				r-tsek-hr 	ht
XP028701149_Macaque NP056590_Mouse BAB70683_Rat XP008250697_Rabbit XP023485975_Horse BB03861_Dog XP027816505_Sheep XP013822124_Goat XP012916721_Ferret XP01393668_Tiger XP0139967_Came1 BAF76737_Pig	qn. 		     		r-tsek-hr- srhsrh s1en-hr-g 	nt           ta           ta           ta           ta           ea           ta           ea           t
XP028701149_Macaque NP065500/MoUse BAB70603_Rat XP008250697_Rabbit XP003485975_Horse NP001075054_Cow BB033861_Dog XP027816505_Sheep XP013822124_Goat XP013822124_Goat XP01392124_Goat XP01392124_Goat XP015396688_Tiger XP015993167_Camel BAF76737_Pig NP005647_Human	qn st. st. sh. 				r-t	
XP028701149_Macaque NP056550_Mouse BAB70683_Rat XP008250597_Rabbit XP023465975_Horse BB03861_Dog XP027816505_sheep XP013822124_Goat XP0138216721_Perret XP01390479_cat XP01593668_Tiger XP010993167_Came1 BAF76737_Pig NP005647_Human XP028701149_Macaque	qn. 				r-tsek-hr- srh	
XP028701149_Macaque NP056500_Mouse BAB70663_Rat XP008250637_Rabbit XP003458575_Horse BR030861_Dog XP027816505_sheep XP01382124_coat XP01382124_coat XP01392124_coat XP01392124_coat XP013936688_Tiger XP01393167_camel BAF76737_Pig NP005647_Human XP025550_Mouse	qn. 				r-t	ht
XP028701149_Macaque NP056500_Mouse BAB70683_Rat XP008250575_Horse NP001075054_Cow BBD33861_Dog XP027816505_Sheep XP01382214_Goat XP01382214_Goat XP0139216721_Perret XP01392674_Goat XP01396687_Camel BAF76737_Pig NP005647_Human XP02580701149_Macaque NP056550_Mouse BAB70663_Rat	qn 				r-tsek-hr- srh	ht
XP028701149_Macaque NP056509_MoUse BAB70603_Rat XP008250697_Rabbit XP003465975_Horse BBD33861_Dog XP027816505_sheep XP01362124_Goat XP013822124_Goat XP013822124_Goat XP013822124_Goat XP01382124_Goat XP01382124_Goat XP01382167_Camel BAF76737_Pig NP005647_Human XP028701149_Macaque NP056509_MoUse BAB70663_Rat XP003463975_Horse					r-t	
XP028701149_Macaque NP056590_MoUse BAB70683_Rabt XP00250657_Rabbit XP00250657_Rabbit D0050504_Cow BBD3861_Dog XP027816505_Rep XP01382212_Goat XP01382212_Goat XP01382212_Goat XP01382682_Tiger XP01395167_Camel BAF76737_Pig BAF76737_Pig NP005647_Human XP028701149_Macaque NP056550_Mouse BAB70663_Rabt XP00825067_Rabbit XP00825067_Rabbit					r-tsek-hr- srh	
XP028701149_Macaque NP065500 MoUse BAB70603_Rat XP008250637_Rabbit XP003485975_Horse BBD33861_Dog XP027816505_Sheep XP013822124_Goat XP013822124_Goat XP013822124_Goat XP013822124_Goat XP013822124_Goat XP013822124_Goat XP01382124_Goat XP01382147_Gaat XP01539668_Tiger XP005647_Human XP0256470_Human XP0256470_Human XP0256590_MoUse BAB70663_Rat XP003485975_Horse NP00175054_Cow BBD3861_Dog					r-t	
XP028701149_Macaque					r-tsek-hr sth	ht
XP028701149_Macaque XP02870163_Habit XP065509_MoUse BAB70663_Rabit XP023405975_Horse NP001075054_Cow BB033061_Dog XP027816505_Sheep XP013822124_Goat XP013822124_Goat XP013820149_Cat XP013920147_Cat BAF76737_Pig NP005647_Human XP02870419_Macaque NP056590_MOUse BAB7063_Rat XP0028067_Rabbit XP0230475_Horse NP001075054_Cow BB03861_Dog XP01281650_Sheep XP013822124_Goat	qn 				r-tsek-hr- srh	
XP028701149_Macaque NP056500_Mouse BAB70683_Rat XP008250697_Rabbit XP008250697_Rabbit XP003485975_Horse NP001075054_Cow BB033861_Dog XP027816505_sheep XP01382124_Goat XP01382124_Goat XP01392124_Goat XP01392124_Goat XP013936688_Tiger XP013936688_Tiger XP0056550_Mouse BAB70683_Rat XP025500_Mouse BAB70683_Rat XP0025056_Cow BD33861_Dog XP0038160_Sheep XP013822124_Goat XP013822124_Goat XP013822124_Goat XP013822124_Goat	qn 				r-tssek-hr-        srh        srh        srh         s1srh        srh         sssh-hr-ga-        sh-hr-ga-        s	ht           Image: Second Sec
XP028701149_Macaque XP02870163_Habit XP065509_MoUse BAB70663_Rabit XP023405975_Horse NP001075054_Cow BB033061_Dog XP027816505_Sheep XP013822124_Goat XP013822124_Goat XP013820149_Cat XP013920147_Cat BAF76737_Pig NP005647_Human XP02870419_Macaque NP056590_MOUse BAB7063_Rat XP0028067_Rabbit XP0230475_Horse NP001075054_Cow BB03861_Dog XP01281650_Sheep XP013822124_Goat	qn 				r-tsek-hr- srh	

**Fig. 3.** Sequence analysis of TMPRSS2 protein in different mammalian species, including some SARS-CoV-2 putative host. A) Phylogenetic analysis of TMPRSS2 proteins. Distance was calculated with Poisson correction. Each protein is named with its accession number and percent homology with human protein is shown. B) Protein alignment of the Serine protease domain of TMPRSS2 showing the residues of the binding site (blue) and the catalytic (red) residues.

domain in the human complex generates a strong salt bridge interaction between Asp30 and Lys417 (Lan et al., 2020; Yan et al., 2020), and three long-range interactions between Lys31-Glu484, His34-Glu406, and Glu37-Arg403. Less ionic pairs are present for mice, dogs, and pigs complexes with RBD; moreover, for the mice complex only one long-range interaction is seen. In cats and ferrets ACE2/SARS-CoV-2 RBD complexes we detected four ionic pairs; although, in the ferret complex the third ion pair is absent, due to a non-synonymous mutation, a salt bridge interaction is observed between Arg354 and Asp405.

Altogether, the results showed here suggest that the ionic pairs produced between the host ACE2 and the viral spike play a pivotal role in the complex stabilization, then, allowing the virus to attach into the target cell. However, other cellular factors could play a secondary role that allows them to finally establishing a successful infection. Some of these factors are discussed in the following sections.

# 3.2. TMPRSS2

The proteolytic processing of the viral spike protein is a key step during the infection process of SARS-CoV-2. Several reports showed that this reaction could be driven by the transmembrane serine protease TMPRSS2. This protein had been described with an intracellular domain (residues 1–84), transmembrane spanning domain (residues 84–106), low-density lipoprotein receptor domain (LDLRA: residues 133–147)

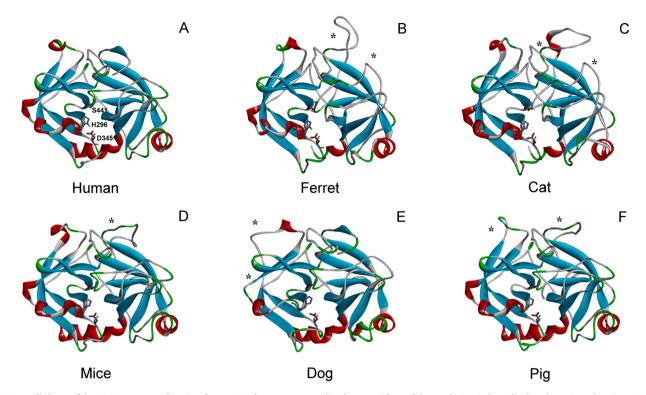


Fig. 4. Overall shape of the Serine protease domain of TMPRSS2 for some mammalian host. Residues of the catalytic triad are displayed. Regions showing significant differences are marked (\*).

and two extracellular domains. The homology models of TMPRSS2 from different hosts (mice, cat, pig, ferret, and dog) showed the two extracellular domains; the cysteine-rich domain (residues 148-242) and the serine protease domain (residues 255-489) with the presence of the residue Ser441 as a catalytic residue. Here, we performed multiple sequence alignments and phylogenetic analysis of TMPRSS2 for several mammalian hosts. These proteins display a certain degree of homology (more than 70 %) with the human protein (Fig. 3A). Of note, the catalytic triad is conserved among all the proteins and in the amino acid residues of the active site (Fig. 3B), no major differences were observed. Thus, it can be speculated that the viral spike could be enzymatically processed in a similar pattern in all described hosts. Currently, there is no crystal structure for TMPRSS2; we applied computational algorithms to develop models for selected mammalian host. The template serine protease Hepsin (PDB code 5CE1) showed 35.7 % identity with the human TMPRSS2 sequence. TMPRSS2 serine protease domain from different hosts (mice, cat, pig, ferret, and dog) shares~83 % identity with the human protein. Interestingly, few structural differences were evident between the TMPRSS2 for each host. Fig. 4 displays the 3D structure of the serine protease domain generated for each enzyme where the residues His296, Asp345 and Ser441 are present as a catalytic triad. Table S4 summarizes the main structural parameters evaluated for each structure. The RMSD for the whole enzyme was around 1 Å and 2 Å for the serine protease domain. Assessment of the Ramachandran plots indicates a good overall geometry for the models, with  $\sim$  83 % of the residues in the most favored regions. The energy of protein-folding, as determined with ProSa-web, showed typical features of native structures with a z-score value of -7.00. All these structural parameters are in agreement with the homology between the models.

Some reports described that TMPRSS2 inhibition could be related to an antiviral effect. Indeed, for SARS-CoV-2 this was demonstrated *in vitro* (Hoffman et al., 2020). Based on that, this study was established to evaluate through molecular docking whether or not the inhibitors of human TMPRSS2 may affect the enzyme from other hosts. A small library was built by using Camostat and Nafamostat pharmacophore

abl	e 2	2			

Compound	TMPRSS2					
	Human	Dog	Pig	Cat	Ferret	Mice
ZINC348	-6,6	-6,4	-6,7	-5,5	-6,3	-6,7
ZINC865	-6,7	-6,2	-6,4	-6,6	-6	-6,7
ZINC1003	-6,8	-6,9	-6	-5,1	-6,1	-5,9
ZINC2062	-6,5	-6,5	-6,6	-5,6	-6,3	-6,5
ZINC22315	-7	-6,8	-6,5	-5,7	-6,2	-6,6
ZINC38545	-6,4	-5,9	-5,9	-5	-5,9	-6
ZINC119905	-6,7	-6	-6,5	-5,4	-6,7	-6,6
ZINC519080	-6,6	-5,9	-5,8	-5	-5,8	-6,2
ZINC1567130	-7,2	-6,3	-7	-5,6	-6,2	-6,9
ZINC1665640	-7	-6,4	-6,4	-5,8	-6,5	-6,3
ZINC1692348	-6,2	-6,3	-6,1	-6,9	-6,1	-6,1
ZINC1850213	-7,4	-6,6	-6,7	-5,9	-6,8	-6,6
ZINC3871842	-7,6	-6,4	-6,4	-5,7	-6,5	-7,2
ZINC3874467**	-8,1	-7,1	-8	-7,2	-7,4	-8,5
ZINC5419379	-7,6	-6,5	-6,6	-6,5	-7,2	-7,1

Results are expressed in Kcal/mol; The best interacting compounds are highlighted in bold;

<sup>\*</sup> Best inhibitor in all the hosts.

structures. The structural similarity search was carried out on the ZINC database, choosing only compounds that have already been evaluated in humans; in total, fifteen compounds met our criteria. Table 2 reveals the docking results with the serine protease domain of the mammalians TMPRSS2. The best-docked compound in all the structures was Nafamostat (ZINC3874467), with a binding energy around -7.5 Kcal/mol. Furthermore, the binding pose of Nafamostat in each active site for the mammalian hosts TMPRSS2 is shown in Fig. 5 (see also Figure S1). All compounds could produce pi-pi interactions and other electrostatic interactions allowing their position close to the catalytic residues (His296 and Ser441 in human). Thus, in theory, Nafamostat could interact with the mammalian TMPRSS2 and inhibit the proteolytic processing of the SARS-CoV-2 spike protein.

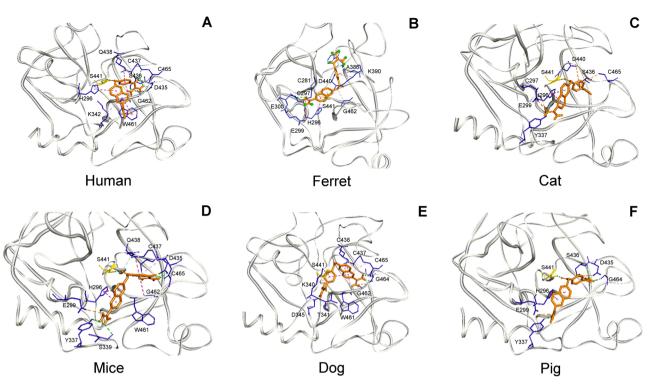


Fig. 5. Docking of Nafamostat on TMPRSS2 for the mammalians host. The best-docked compound with a binding energy around -7.5 Kcal/mol was Nafamostat. The active site for TMPRSS2 interacting with Nafamostat is shown for each enzyme. The main interaction and residues are shown for each complex. The residue (Ser) associated with the catalytic activity is shown in yellow (Ser441 in human).

#### 3.3. GRP78

GRP78 is an ER chaperone that also has homeostatic functions by acting as a receptor in the cell membrane (Ge and Kao, 2019). MERS and HKU9, a bat coronavirus, use this protein as a co-receptor to improve the attachment to the cell. Recent reports indicated that GRP78 could serve as a secondary receptor for SARS-CoV-2. Therefore, we wanted to gain a deep understanding of GRP78 putative key role in host selectivity for SARS-CoV-2.

First, we did a sequence alignment between GRP78 derived from the different hosts evaluated in this work. GRP78 exhibit a very high degree of conservation: except for some differences in the N- and C- terminal regions of the protein, not involved in the interaction with SARS-CoV-2, the proteins from different mammalian hosts exhibit more than 99 % homology. In particular, the residues critical for the interaction with the spike were conserved in all the mammalian proteins (Fig. 6). Based on that, we employed the crystal structure for human GRP78 (PDB code 5E84) as a pattern to identify the binding site for the viral spike protein and its respective sequence and then to compare it to the other mammalian hosts. As shown in Fig. 7, GRP78 could be divided into three subunits: the alpha subdomain, the nucleotide-binding site, and the beta subdomain. Some reports established that the binding site for viral proteins is located in the beta subdomain. Our docking analysis confirmed that the SARS-CoV-2 viral spike protein could bind to the beta region with a binding energy around -14 Kcal/mol. The main residues and the binding pose are shown in Fig. 7 and Table S5. Moreover, this study performed a sequence analysis of the binding site for GRP78 to compare if there could occur any change affecting the viral spike protein binding. Our results pointed out that GRP78 displays a high homology sequence, over 95 %, in this region.

## 4. Discussion

SARS-CoV-2 gains access to the cell through interactions with the ACE2 receptor on the cell surface (Ortega et al., 2020; Zhang et al.,

2020). Some studies argue that interaction with ACE2 is the most important factor to determine host selectivity (Demogines et al., 2012; Li, 2013; Shang et al., 2020). As for SARS-CoV, mice cannot be infected with SARS-CoV-2. Hela cells transfected with mice ACE2 do not support SARS-CoV-2 replication (Zhou et al., 2020). However, if mice are genetically modified and express the human ACE2 receptor they become permissive to the infection (Bao et al., 2020; Dediego et al., 2008). Consequently, studying the interaction between the residues present in ACE2 and the viral spike protein may help us to understand better the factor associated with the host infection. Our sequence analysis of the ACE2 enzyme of several mammalian hosts showed substitution of specific amino acids residues of the contact region and particularly of those residues involved in the recognition of the SARS-CoV-2 RBD domain. The significance of the amino acid substitution of key residues Lys31, Glu35, Asp38, Met82, and Lys353 (from the human cell receptor) in the recognition of viral spike protein on ACE2 from several mammals, including bats, have been previously pointed out (Hou et al., 2010; Luan et al., 2020; Shang et al., 2020). This study evaluated the ACE2 protein for humans, mice, pigs, cats, and dogs. These proteins display a homology over 80 % and in the interacting domain around 70 % in comparison to the human counterpart. Our docking analysis showed that the viral spike protein could interact with all ACE2 proteins evaluated with a binding energy around -13Kcal/mol with the higher binding energy (lees stable complexes) for ACE 2 from mice and pig, in agreement with their no susceptibility to this infection. Nonetheless, a lower binding energy (more stable complex) was achieved with ACE2 from dogs. Meanwhile, the highest binding energy was detected with ACE2 from mice. Interestingly, neither mice nor pig could be infected by SARS-CoV-2 and dogs have low susceptibility. These results are in agreement with previous reports for the interaction of the SARS-CoV-2 viral spike protein and these hosts (Luan et al., 2020; Shen et al., 2020; Shi et al., 2020). The non-synonymous changes in mice and dogs ACE2 sequences reduce the hydrogen bonds between the  $\alpha 1$  and  $\alpha 2$ helixes and the viral spike. Thus, we proceeded to evaluate the permissive hosts, humans, cats, and ferrets. Some differences were found

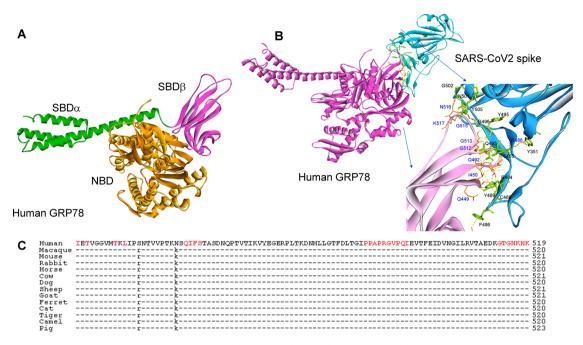
Human	MKLSLVAAMLLLLSAARAEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIANDQGNRITPSYVAFTPEGERLIGDAAKNQLTSN	NPENTVFDAKRLIGRTWNDPS	107
Macaque			107
Mouse	-mkft-v-aalg-vfffff		108
Rabbit Horse			107
Cow	v		108
Dog	vlgt		107 108 107
Sheep	vlqt		108 108
Goat	vlĝ		108
Ferret	vv		107
Cat			107
Tiger Camel	vg		107 107 107 107
Pig	-warws-stwapptrgkwgwwkg-ngawfperrgadcclgfsas		110
Human	VQQDIKFLPFKVVEKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAVVTVPAYFNDAQRQATKDAGTIAGLNVMF	RIINEPTAAAIAYGLDKREGE	217 217 218 217 217 218 217 218 217 218 217 218 218 218
Macaque			217
Mouse Rabbit			218
Horse			217
Cow			218
Dog			217
Sheep	v		218
Goat Ferret			218
Cat			217
Tiger			217
Camel			217
Camel Pig			217 217 217 217 220
Human	KNILVFDLGGGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQRVMEHFIKLYKKKTGKDVRKDNRAVQKLRREVEKAK.ALSSQHQARJ	IEIESFYEGEDFSETLTRAKF	326 327
Macaque Mouse		f	328
Rabbit			327
Horse	rr		328 327 327
Cow	<u>r</u>		328
Dog	rr		327
Sheep Goat	r		328 327 328 328 327 327 327 327 327
Ferret			320
Cat			327
Tiger Camel			327
Camel			327
Pig	rr		330
Human	PPT. MINDT. PD 900MV DUOVUT. PD 90T. VV 90T DPT UT. UCC900 T DV TOOT. UV PPPN/2V PD 90CT NDDP & UX VCA & UOA CUT. 9CD OD0CT. UT. U	LHUCDLULGTRUUGGUMURLT	436
Human Macague	${\tt BELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI$	LHVCPLTLG <mark>IET</mark> VGGVMTKLI	436
Macaque	EELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI	LHVCPLTLG <mark>IET</mark> VGGVM <b>TKL</b> I -d	436 437 438
Macaque Mouse Rabbit	EELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIFKIQQLVKEFFNGKEPSRGINFDEAVAYGAAVQAGVLSGDQDTGDLVLI g	LHVCPLTLGIETVGGVMTKLI -d	438 437
Macaque Mouse Rabbit Horse	EELNMDLFR.STMKPVQKVLED.SDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEP.SRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI g	-d -d	438 437 437
Macaque Mouse Rabbit Horse Cow	EELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI g	-d -d	438 437 437
Macaque Mouse Rabbit Horse Cow Dog	BELNMDLFR3TMKPVQKVLBD3DLKKSDIDEIVLVGGSTRIPKIQQLVKBFPNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVL1 g	d d d	438 437 437 438 437 438
Macaque Mouse Rabbit Horse Cow	EELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI g	d d d	438 437 437 438 437 438
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret	BELNMDLFR3TMKPVQKVLBD3DLKK3DIDEIVLVGGSTRIPKIQQLVKEFFNGKEP3RGINPDEAVAYGAAVQAGVLSGDQDTGDLVL1 g	d d d	438 437 437 438 437 438
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat	g	d d d	438 437 437 438 437 438
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger	EELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLI g		438 437 438 438 438 438 438 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel	g		438 437 438 438 438 438 438 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger	g		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Fig Human	g&		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Fig Human Macaque	g&		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Fig Human Macaque Mouse	g g g g g g g g g g g g g g		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Fig Human Macaque Mouse Rabbit	ga		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Pig Human Macaque Mouse Rabbit Horse Cow	g g g g g g g g g g g g g g		438 437 438 437 438 437 438 437 437 437 437 437 437 437
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Pig Human Macaque Mouse Rabbit Horse Cow	g4	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Camel Pig Human Macaque Mouse Rabbit Horse Cow Dog Sheep	g	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Tiger Pig Human Macaque Mouse Rabbit Horse Cow Dog Sheep Goat	ga	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret Cat Fig Human Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Ferret	g g snrvvvptknsqlfstasdnqetvtikvyegereltkonnligtpoltgippapegvpqievtfeidvngilevtaedkgtgnknkitt 	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Cat Tiger La Cat Camel Faret Human Mouse Rabbit Horse Cow Dog Sheep Sheep Sheep Cow Cat Cat Cat Cat Cat Cat Cat Cat Cat Cat	ga	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Macaque Rabbit Horse Cow Sheep Goat Ferret Cagmel Pig Masaque Masaque Masaque Nosse Cow Sheep Goat Cat Figer Camel Cat Figer Cat Cat Cat Cat Cat Cat Cat Cat Cat Cat	ga	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Cat Tiger La Cat Camel Faret Human Mouse Rabbit Horse Cow Dog Sheep Sheep Sheep Cow Cat Cat Cat Cat Cat Cat Cat Cat Cat Cat	g g s s s s s s s s s s s s s	d d d d d d d d d d d d d d d d d d d	44444444444444444444444444444444444444
Macaque Macaque Rabbit Horse Cow Sheep Goat Ferret Cagmel Pig Macaque Macaque Macaque Macaque Macaque Cow Sheep Sheep Cow Sheep Cow Sheep Cat Tiger Cat Cat Tiger Cat Cat Cat Cat Cat Cat Cat Cat Cat Cat	g g g g g g g g g g g g g g g g g g g	d d d d d d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Goat Fete Cape Human Macaque Macaque Human Macaque Dog Sheep Cow Dog Sheet Cow Dog Sheet Cow Dog Sheet Cow Dog Sheet Cow Human Horse Cow Human Horse Cow Human Horse Sheet Cow Com Cape Com Com Com Com Com Com Com Com Com Com	g 4 	d d d d d d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Bagoat Ferret Com Comel Forret Com Macaque Macaque Macaque Human Macaque Sheep Goat Ferret Cow Sheep Goat Ferret Com Sheep Goat Ferret Com Sheep Com Sheep Com Sheep Com Sheep Com Sheep Com Sheep Com Sheep Com Sheep Sheep Com Sheep Sheep Com Sheep She	g 4 	d d d d d d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Macuse Rabbit Horse Cow Sheep Goat Ferret Cat Fig Human Macaque Rabbit Horse Cow Dasep Goat Cat Fig Human Macaque Pig Human Macaque Pig Human Macaque Pig Human Macaque Natore Soat Cat Human Macaque Pig Human Macaque Natore Soat Cat Human Macaque Soat Cow Pig Sheep She	g 4 	d d d d d d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Macuse Rabbit Horse Cow Dog Sheep Goat Ferret Camel Pig Human Macaque Macaque Human Macaque Cow Sheep Goat Ferret Cabbit Horse Cabbit Horse Rabbit Horse	g 4 	d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Sheep Goat Ferret Car Ferret Camel Fig Mouse Rabbit Horse Cow Dog Goat Human Macaque Mouse Rabbit Horse Cow Macaque Mouse Rabbit Horse Cow	g 4 	d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Sheep Goat Ferret Car Ferret Camel Fig Mouse Rabbit Horse Cow Dog Goat Human Macaque Mouse Rabbit Horse Cow Macaque Mouse Rabbit Horse Cow	g 4 	d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Macuse Rabbit Horse Cow Sheep Goat Ferret Cat Tiger Camel Pig Macaque Rabbit Horse Cow Do Goat Ferret Camel Pig Human Macaque Pig Human Macaque Macaque Nacaque Cat Cat Ferret Cat Fig Human Macaque Nacaque Cow Do Sot Cat Cat Ferret Cow Do Sot Cow Cow Cow Cow Cow Cow Cow Cow Cow Cow	g 4 	d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Fratet Camel Fratet Camel Fig Rabbit Horse Cow Dog Scatt Fart Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Fart Camel Horse Cow Nacaque Human Macaque Horse Cow Dog Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Nacaque Horse Cow Nacaque Horse Cow Dog Scatt Camel Fart Camel Horse Cow Nacaque Horse Cow Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Fart Camel Cow Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Com Dog Scatt Camel Horse Com Camel Fart Horse Com Camel Fart Camel Horse Com Dog Scatt Camel Fart Camel Horse Com Camel Fart Camel Fart Camel Horse Com Camel Fart Camel Fart Camel Fart Camel Fart Com Com Scatt Camel Fart Com Com Scatt Camel Fart Com Com Scatt Camel Fart Com Com Scatt Com Com Com Com Com Com Com Com Com Com	g 4 	d d d d d d d d d d d d d d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Dog Sheep Fratet Camel Fratet Camel Fig Rabbit Horse Cow Dog Scatt Fart Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Fart Camel Horse Cow Nacaque Human Macaque Horse Cow Dog Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Nacaque Horse Cow Nacaque Horse Cow Dog Scatt Camel Fart Camel Horse Cow Nacaque Horse Cow Scatt Camel Fart Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Fart Camel Cow Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Cow Dog Scatt Camel Horse Com Dog Scatt Camel Horse Com Camel Fart Horse Com Camel Fart Camel Horse Com Dog Scatt Camel Fart Camel Horse Com Camel Fart Camel Fart Camel Horse Com Camel Fart Camel Fart Camel Fart Camel Fart Com Com Scatt Camel Fart Com Com Scatt Camel Fart Com Com Scatt Camel Fart Com Com Scatt Com Com Com Com Com Com Com Com Com Com	g 4 	d	44444444444444 333333333334 4444444444
Macaque Mouse Rabbit Horse Cow Dog Sheet Cow Postet Camel Piger Rabbit Horse Com Macaque Human Macaque Human Macaque Horse Com Postet Camel Piger Camel Post Com Postet Camel Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Post Post Com Com Post Com Com Com Com Com Com Com Com Com Com	g 4 	d	44444444444444 333333333334 4444444444
Macaque Macusque Rabbit Horse Cow Sheep Goat Cater Tiger Camel Ferret Camel Fag Rabbit Horse Cow Sheep Goat Cater Tiger Cater Tiger Cater Ferret Cater Ferret Cow Sheep Sheep Goat Cater Ferret Cater For Sheep Goat Cater For Sheep Cater Ferret Cater For Sheep Cater Sheep Cater Sheep Cater For Sheep Cater Sheep Cater Sheep Com Sheep Cater Sheep Cater Sheep Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Sheet Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet Cater Sheet She	g 4 	d d d d d d d d d d d d d d	44444444444444444444444444444444444444

Fig. 6. Sequence analysis for GRP78 receptor from several mammalian species, including some SARS-CoV-2 putative host. Protein alignment of GPR78 regions showing in red the important residues involved in the interaction with SARS-CoV-2 RBD.

in terms of ionic interactions, between mice, dogs, and pigs (salt bridges, N–O pair, and long-range interactions) related to changes in some residues. In the human complex four ionic pairs are observed (a strong salt bridge and three long-range interactions); also, in cats and ferrets four ionic pairs are present. In non-permissive or less permissive species, only three of these ionic pairs are preserved. Altogether, the hydrogen bonds in the interface and the presence of ionic pairs between ACE2/SARS-CoV-2 RBD complexes appear to be an important stabilization factor.

The interaction with ACE2 seems essential for host selectivity. Nevertheless, the low-level expression of ACE2 in lungs compared to other tissues, such as the small intestine or colon, suggest that the virus uses other secondary receptors, increasing the probability to interact with ACE2 (Fagerberg et al., 2014). Recently, it was described that GRP78 could act as a secondary receptor for SARS-CoV-2 (Ibrahim et al., 2020). Also, this receptor had been reported as a secondary receptor for other coronaviruses (Chu et al., 2018). After the coronavirus spike protein interacts with ACE2 receptor occurs the formation of an endocytic vacuole. Subsequently, the viral spike protein is cleaved by TMPRSS2 and the fusion motif is exposed, allowing the fusion between

viral and lysosome membranes and the release of viral RNA into the cvtoplasm (Fehr and Perlman, 2015; Millet and Whittaker, 2015). These early stages in viral infection are essential to allow the virus infection in a specific host. However, less is known about whether TMPRSS2 and/or GRP78 may contribute to SARS-CoV-2 permissiveness in other hosts, modulating its viral infectivity. Some reports showed that TMPRSS2 is a host factor essential for pneumotropism and pathogenicity of H7N9 and H1N1 influenza virus in mice by cleaving of the hemagglutinin (Tarnow et al., 2014). Of note, SARS-Cov-2 can use both cathepsins and TMPRSS2 to prime the S protein for cell entry, depending on the target cell type, but only TMPRSS2 is required for SARS-Cov-2 infection of primary human lung cells (Hoffmann, 2020). Thus, as for H7N9 and H1N1, TMPRSS2 play main role in SARS-CoV2 activation in the lungs. In this study, was analyzed the TMPRSS2 protease in the different mammalian hosts through sequence analysis and structural modeling. Curiously, this protein appears to be conserved in these host species. The homology was around 70 % compared to humans. Benzoic acid derivatives as Camostat mesylate or Nafamostat can block the human protease in vitro and in vivo. Indeed, previous reports demonstrated that Camostat could block the entry of SARS-CoV-2 (Hoffman et al., 2020). Currently, clinical trials



**Fig. 7.** GRP78 and SARS-CoV-2 interaction region for some mammalian hosts. The crystal structure for the human GRP78 PDB code 5E84 is showed. A) The three main structural regions are denoted in the structure. The nucleotide-binding domain (NBD) and the substrate-binding subdomains betha (SBD $\beta$ ) and alpha (SBD $\alpha$ ). B) The best interaction pattern for GRP78 and the viral spike protein is showed. A closed-up of the interacting region and the residues involved are denoted. C) An alignment of the residues present in the binding site for GRP78 is shown.

are evaluating the efficacy of these compounds against SARS-CoV-2 in (Clinicaltrials.gov code NCT04321096, NCT04353284, vivo NCT04353284). Since the proteases evaluated here exhibit a high homology in the active site, these compounds could block TMPRSS2 by a similar mechanism. Our molecular docking results showed that Nafamostat was the best-docked compound in the catalytic site of the mammalian TMPRSS2. SARS-CoV-2 can infect cats and, with low permissiveness, dogs thus these domestic animals could play an important role as reservoirs of the virus. The fact that TMPRSS2 is being evaluated as a possible alternative therapeutic against SARS-CoV-2 let us hypothesize that these proteases could be evaluated as targets in these intermediary hosts. These therapeutic options may have a significant impact on the pandemic control. Modulating the viral replication through cell targets would have several advantages in the development of new therapeutic agents (Ortega et al., 2013). Besides, understanding if other host factors produce an increase or decrease in the susceptibility/infectivity to SARS-CoV-2 could help us identify new therapeutic targets.

For its part, GRP78 was described recently as a possible secondary receptor for SARS-CoV-2. In theory, this receptor helps the virus to concentrate on the cell membrane and then interact with ACE2. This protein was also assessed in this work as a possible key factor in host selectivity. The sequence analysis demonstrated that this receptor is also highly conserved in mammalians. In fact, the region that was reported as a possible binding site for SARS-CoV-2 exhibit over 95 % homology between all mammal species examined. Therefore, in theory, SARS-CoV-2 could interact with GRP78 in all these hosts, enhancing its interaction with ACE2.

Collectively, the results present in this work help to clarify some important aspects related to SARS-CoV-2 host susceptibility. These findings contribute to explain the infection events that occurred recently in some pets and additionally to identify possible therapeutic targets. These targets would also help to control the pandemic distribution in intermediary hosts like cats and dogs.

Finally, one interesting outcome of this work was related to the dog infection permissiveness. Sequence analysis and docking results showed that this host could be infected. However, the available *in vivo* data

indicate that this host has low susceptibility (Shi et al., 2020). Thus, changes or variation in other secondary receptors (such as integrin Sigrist et al., 2020 and/or CD147, Wang et al., 2020,) or cellular restriction factors associated with the immune innate response could be related to this low susceptibility. Therefore, we need further studies to understand better the early stages of this viral infection and the role of the host restriction factors. Altogether, this novel knowledge about the virus biology will contribute to propose new therapeutic targets. Furthermore, our investigation enhances the search for other putative targets that may have a significant impact on the control of the current SARS-CoV-2 pandemic.

## **Funding source**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

#### **CRediT** authorship contribution statement

H.R. Rangel: Conceptualization, Data curation, Investigation, Methodology, Validation, Writing - review & editing. J.T. Ortega: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. S. Maksoud: Investigation, Methodology, Writing - review & editing. F.H. Pujol: Conceptualization, Supervision, Data curation, Visualization, Writing - review & editing. M.L. Serrano: Conceptualization, Data curation, Project administration, Supervision, Visualization, Writing original draft, Writing - review & editing.

### **Declaration of Competing Interest**

The authors declare no conflict of interests.

# Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2020.198154.

#### H.R. Rangel et al.

#### References

Ahn, D.G., Shin, H.J., Kim, M.H., et al., 2020. Current status of epidemiology, diagnosis, therapeutics, and vaccines for novel coronavirus disease 2019 (COVID-19). J. Microbiol. Biotechnol. 30 (3), 313–324. https://doi.org/10.4014/ jmb.2003.03011.

Amanat, F., Krammer, F., 2020. SARS-CoV-2 vaccines: status report. Immunity 52 (4), 583–589. https://doi.org/10.1016/j.immuni.2020.03.007.

Arnold, K., Bordoli, L., Kopp, J., et al., 2006. The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201. https://doi.org/10.1093/bioinformatics/bti770.

Bao, L., Deng, W., Huang, B., et al., 2020. The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. Nature. https://doi.org/10.1038/s41586-020-2312-y.

Belouzard, S., Chu, V.C., Whittaker, G.R., 2009. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc. Natl. Acad. Sci. U.S.A. 106 (14), 5871–5876. https://doi.org/10.1073/pnas.0809524106.

Chu, H., Chan, C.M., Zhang, X., et al., 2018. Middle East respiratory syndrome coronavirus and bat coronavirus HKU9 both can utilize GRP78 for attachment onto host cells. J. Biol. Chem. 293 (30), 11709–11726. https://doi.org/10.1074/jbc. RA118.001897.

Dediego, M.L., Pewe, L., Alvarez, E., et al., 2008. Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice. Virology 376 (2), 379–389. https://doi.org/10.1016/j.virol.2008.03.005.

Demogines, A., Farzan, M., Sawyer, S.L., 2012. Evidence for ACE2-utilizing coronaviruses (CoVs) related to severe acute respiratory syndrome CoV in bats. J. Virol. 86 (11), 6350–6353. https://doi.org/10.1128/JVI.00311-12.

Elfiky, A.A., 2020. Human papillomavirus E6: host cell receptor, GRP78, binding site prediction. J. Med. Virol. https://doi.org/10.1002/jmv.25737, 10.1002/jmv.25737.

Fagerberg, L., Hallström, B.M., Oksvold, P., et al., 2014. Analysis of the human tissuespecific expression by genome-wide integration of transcriptomics and antibodybased proteomics. Mol. Cell Proteomics 13 (2), 397–406. https://doi.org/10.1074/ mcp.M113.035600

Fehr, A.R., Perlman, S., 2015. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol. Biol. 1282, 1–23. https://doi.org/10.1007/978-1-4939-2438-7\_1.

Ge, R., Kao, C., 2019. Cell surface GRP78 as a death receptor and an anticancer drug target. Cancers (Basel) 11 (11), 1787. https://doi.org/10.3390/cancers11111787.

Glowacka, I., Bertram, S., Müller, M.A., et al., 2011. 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), 4122–4134. https://doi.org/10.1128/JVI.02232-10.

Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18 (15), 2714–2723. https://doi. org/10.1002/elps.1150181505.

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 (2), 271–280. https://doi.org/10.1016/j.cell.2020.02.052 e8.

Hou, Y., Peng, C., Yu, M., et al., 2010. Angiotensin-converting enzyme 2 (ACE2) proteins of different bat species confer variable susceptibility to SARS-CoV entry. Arch. Virol. 155 (10), 1563–1569. https://doi.org/10.1007/s00705-010-0729-6).

Ibrahim, I.M., Abdelmalek, D.H., Elshahat, M.E., et al., 2020. COVID-19 spike-host cell receptor GRP78 binding site prediction. J. Infect. 80 (5), 554–562. https://doi.org/ 10.1016/j.jinf.2020.02.026.

Jindadamrongwech, S., Thepparit, C., Smith, D.R., 2004. Identification of GRP 78 (BiP) as a liver cell expressed receptor element for dengue virus serotype 2. Arch. Virol. 149 (5), 915–927. https://doi.org/10.1007/s00705-003-0263-x.

Kleine-Weber, H., Elzayat, M.T., Hoffmann, M., et al., 2018. Functional analysis of potential cleavage sites in the MERS-coronavirus spike protein. Sci. Rep. 8 (1), 16597. https://doi.org/10.1038/s41598-018-34859-w.

Kumar, S., Nussinov, R., 2002. Relationship between ion pair geometries and electrostatic strengths in proteins. Biophys. J. 83 (3), 1595–1612. https://doi.org/ 10.1016/S0006-3495(02)73929-5.

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 581, 215–220. https://doi.org/ 10.1038/s41586-020-2180-5.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., et al., 1993. PROCHECK-a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291. https://doi.org/10.1002/jmr.2618.

Letko, M., Marzi, A., Munster, V., 2020. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nat. Microbiol. 5 (4), 562–569. https://doi.org/10.1038/s41564-020-0688-y.

Li, F., 2013. Receptor recognition and cross-species infections of SARS coronavirus. Antiviral Res. 100 (1), 246–254. https://doi.org/10.1016/j.antiviral.2013.08.014.

Liu, C., Zhou, Q., Li, Y., et al., 2020. Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. ACS Cent. Sci. 6 (3), 315–331. https://doi.org/10.1021/acscentsci.0c00272.

Luan, J., Lu, Y., Jin, X., et al., 2020. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. Biochem. Biophys. Res. Commun. 526 (1), 165–169. https://doi.org/10.1016/j. bbrc.2020.03.047. Millet, J.K., Whittaker, G.R., 2015. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. Virus Res. 202, 120–134. https://doi.org/ 10.1016/j.virusres.2014.11.021.

Nain, M., Mukherjee, S., Karmakar, S.P., et al., 2017. GRP78 is an important host factor for japanese encephalitis virus entry and replication in mammalian cells. J. Virol. 91 (6), e02274–16. https://doi.org/10.1128/JVI.02274-16.

Ortega, J.T., Rangel, H.R., Pujol, F.H., 2013. Cellular targets as an alternative mechanism in antiviral therapy. Interciencia 38 (12), 836–842.

Ortega, J.T., Serrano, M.L., Suárez, A.I., et al., 2019. Antiviral activity of flavonoids present in aerial parts of *Marcetia taxifolia* against Hepatitis B virus, Poliovirus, and Herpes Simplex Virus *in vitro*. EXCLI J. 18, 1037–1048. https://doi.org/10.17179/ excli2019-1837.

Ortega, J.T., Serrano, M.L., Pujol, F.H., et al., 2020. Role of changes in SARS-CoV-2 spike protein in the interaction with the human ACE2 receptor: an *in silico* analysis. EXCLI J. 19, 410–417. https://doi.org/10.17179/excli2020-1167.

Ortega, J.T., Serrano, M.L., Pujol, F.H., et al., 2020b. Unrevealing sequence and structural features of novel coronavirus using *in silico* approaches: the main protease as molecular target. EXCLI J. 19, 400–409. https://doi.org/10.17179/excli2020-1189.

Ortega, J.T., Serrano, M.L., Jastrzebska, B., 2020c. Class a g protein-coupled receptor Antagonist Famotidine as a therapeutic alternative against SARS-CoV2: an in Silico Analysis. Biomolecules 10 (6), E954. https://doi.org/10.3390/biom10060954.

Parvez, M.K., Parveen, S., 2017. Evolution and emergence of pathogenic viruses: past, present, and future. Intervirology 60 (1–2), 1–7. https://doi.org/10.1159/ 000478729

Pedretti, A., Villa, L., Vistoli, G., 2004. VEGA–an open platform to develop chemo-bioinformatics applications, using plug-in architecture and script programming. J. Comput. Aided Mol. Des. 18 (3), 167–173. https://doi.org/10.1023/b: jcam.0000035186.90683.f2.

Phillips, J.C., Braun, R., Wang, W., et al., 2005. Scalable molecular dynamics with NAMD. J. Comput. Chem. 26 (16), 1781–1802. https://doi.org/10.1002/jcc.20289.

Pierce, B.G., Wiehe, K., Hwang, H., et al., 2014. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics 30, 1771–1773. https://doi.org/10.1093/bioinformatics/btu097.

Ricke, D., Malone, R.W., 2020. Medical Countermeasures Analysis of 2019-nCoV and Vaccine Risks for Antibody-Dependent Enhancement (ADE) (2/27/2020). https:// doi.org/10.2139/ssrn.3546070. https://ssrn.com/abstract=3546070.

Sánchez-Linares, I., Pérez-Sánchez, H., Cecilia, J.M., et al., 2012. High-Throughput parallel blind virtual screening using BINDSURF. BMC Bioinformatics 13 (Suppl 14 (Suppl 14)), S13. https://doi.org/10.1186/1471-2105-13-S14-S13.

Shang, J., Ye, G., Shi, K., et al., 2020. Structural basis of receptor recognition by SARS-CoV-2. Nature 581 (7807), 221–224. https://doi.org/10.1038/s41586-020-2179-y.

Shen, M., Liu, C., Xu, R., et al., 2020. SARS-CoV-2 Infection of Cats and Dogs? Preprints, 2020040116.

Shi, J., Wen, Z., Zhong, G., et al., 2020. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. Science 368 (6494), 1016–1020. https://doi.org/10.1126/science.abb7015.

Sigrist, C.J., Bridge, A., Le Mercier, P., 2020. A potential role for integrins in host cell entry by SARS-CoV-2. Antiviral Res. 177, 104759 https://doi.org/10.1016/j. antiviral.2020.104759.

Tarnow, C., Engels, G., Arendt, A., et al., 2014. TMPRSS2 is a host factor that is essential for pneumotropism and pathogenicity of H7N9 influenza A virus in mice. J. Virol. 88 (9), 4744–4751. https://doi.org/10.1128/JVI.03799-13.

Tu, Y.F., Chien, C.S., Yarmishyn, A.A., et al., 2020. A review of SARS-CoV-2 and the ongoing clinical trials. Int. J. Mol. Sci. 21 (7), E2657. https://doi.org/10.3390/ ijms21072657.

Vanommeslaeghe, K., Hatcher, E., Acharya, C., et al., 2010. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J. Comput. Chem. 31, 671–690. https://doi.org/10.1002/ jcc.21367.

Wang, M., Wey, S., Zhang, Y., et al., 2009. Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders. Antioxid. Redox Signal. 11 (9), 2307–2316. https://doi.org/10.1089/ars.2009.2485.

Wang, Q., Zhang, Y., Wu, L., et al., 2020. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 181 (4), 894–904. https://doi.org/10.1016/j. cell.2020.03.045 e9.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. W407–W410. https://doi.org/10.1093/nar/gkm290, 35(Web Server issue).

Xue, L.C., Rodrigues, J.P., Kastritis, P.L., et al., 2016. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics 32 (23), 3676–3678. https://doi.org/10.1093/bioinformatics/btw514.

Yan, R., Zhang, Y., Li, Y., et al., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367 (6485), 1444–1448. https://doi.org/ 10.1126/science.abb2762.

Zhang, H., Penninger, J.M., Li, Y., et al., 2020. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med. 46 (2020), 586–590. https://doi.org/10.1007/s00134-020-05985-9.

Zhou, P., Yang, X.L., Wang, X.G., et al., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579 (7798), 270–273. https://doi. org/10.1038/s41586-020-2012-7.