Furin, a potential therapeutic target for COVID-19

Canrong WU,[a,1]Yueying YANG,[b,1]Yang LIU,[b]Peng ZHANG,[b]YaliWANG,[b]Qiqi WANG, [b] Yang XU,[b]Mingxue LI,[b]Mengzhu ZHENG,[a,*] Lixia CHEN,[b,*] &Hua LI[a,b,*]

[a]Hubei Key Laboratory of Natural Medicinal Chemistry and Resource Evaluation, School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

[b]Wuya College of Innovation, Key Laboratory of Structure-Based Drug Design & Discovery, Ministry of Education, Shenyang Pharmaceutical University, Shenyang 110016, China

[1] These authors contributed equally to this work.

*Corresponding author: Hua Li (E-mail: li_hua@hust.edu.cn).

Lixia Chen (syzyclx@163.com).

Mengzhu Zheng (mengzhu_zheng@hust.edu.cn).

Abstract

A novel coronavirus (SARS-CoV-2) infectious disease has broken out in Wuhan, Hubei Province since December 2019, and spread rapidly from Wuhan to other areas, which has been listed as an international concerning public health emergency. We compared the Spike proteins from four sources, SARS-CoV-2, SARS-CoV, MERS-CoV and Bat-CoVRaTG13, and found that the SARS-CoV-2 virus sequence had redundant PRRA sequences. Through a series of analyses, we propose the reason why SARS-CoV-2is more infectious than other coronaviruses. And through structure based virtual ligand screening, we foundpotentialfurin inhibitors, which might be used in the treatment of new coronary pneumonia.

1. Introduction

In December 2019, a series of acute respiratory diseases occurred in Wuhan, Hubei Province, China and then spread rapidly from Wuhan to other areas. As of February 17, 2020, a total of 71,444 patients have been diagnosed and 1,775 have died worldwide. This is caused by a novel coronavirus, which was named as "2019-nCoV" by the World Health Organization, and diseases caused by 2019-nCoV was COVID-19. 2019-nCoV, as a close relative of SARS-CoV, was classified as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses (ICTV) on February 11, 2020.

Coronaviruses (CoVs) are mainly composed of four structural proteins, including Spike (S), membrane (M), envelope (E) and nucleocapsid (N) [1]. Spike, a trimeric glycoprotein of CoVs, determines diversity of CoVs and host tropism, and mediates CoVs binding to host cells surface-specific receptors and virus-cell membrane fusion [2]. Current research found that SARS-CoV-2 belongs to the beta coronavirus genus, and speculated that it may interact with angiotensin-converting enzyme 2 (ACE2) on the surface of human cells through Spike protein, thereby infecting human respiratory epithelium cell [3]. Letko M and Munster Vthen identified the receptor for SARS-CoV-2 entry into human cells to be ACE2 [4].

Coronavirus Spike protein plays a key role in the early stages of viral infection, with the S1 domain responsible for receptor binding and the S2 domain mediating membrane fusion [5]. The process of SARS-CoV infecting the host involves two indispensable cleaving processes which affect the infectious capacity of SARS-CoV. First, Spike was cleaved into receptor-bound N-terminal S1 subunit and membrane-fusion C-terminal S2 subunit by host proteases at S1/S2 cleavage site (such as type II transmembrane serine protease (TMPRSS2), cathepsins B and L) [6,7]. Second, after CoVs are endocytosed by the host, the lysosomal protease mediates cleavage of S2 subunit (S2' cleavage site) and releases the hydrophobic fusion peptide to fuse with the host cell membrane [8].

Furin, a kind of proprotein convertases (PCs), is located in the trans-Golgi network (TGN) and activated by acid pH [9]. Furin can cleave precursor proteins with specific motifs to produce mature proteins with biological activity. The first (P1) and fourth (P4) amino acids at the N-terminus of the substrate cleavage site must be arginine "Arg-X-X-Arg ↓" (R-X-X-R，X: any amino acid, ↓:cleavage site). If the P2 position is basic lysine or arginine, the cleavage efficiency

can be improved by about 10 times [10]. Kibler KV *et al*. demonstrated that the Spike protein S1/S2 and S2′ cleavage sites of the infectious bronchitis virus (IBVs) Beaudette strain can be recognized by fruin, which is a distinctive feature of IBV-Beaudette with other IBVs and has stronger infection ability [11,12]. Based on the characteristics of furin's recognition substrate sequence, some short peptide inhibitors have been developed, such as Decanoyl-Arg-Val-Lys-Arg-chloromethylketone (Dec-RVKR-CMK) and modified α1-antitrypsin Portland (α1-PDX). However, the non-specific and irreversible inhibitory effects on all members of the PC family limit their application [10, 13]. No small molecule inhibitor of furin with good effect and high specificity has been found so far.

The epidemiological observations showed the infectious capacity of SARS-CoV-2 is stronger than SARS-CoV, so there are likely to be other mechanisms to make the infection of SARS-CoV-2 easier. We suppose the main possibilities as follows, first, SARS-CoV-2 RBD combining with ACE2 may have other conformations; second, the SARS-CoV-2 Spike protein can also bind to other receptors besides ACE2; third, Spike is more easily cleaved by host enzymes and easily fuses with host cell membrane. We compared the Spike proteins from four sources, SARS-CoV-2, SARS-CoV, MERS-CoV and Bat-CoVRaTG13, and found that the SARS-CoV-2 virus sequence had redundant PRRA sequences. Through a series of analyses, this study propose that one of the important reasons for the high infectivity of SARS-CoV-2 is a redundant furin cut site in its Spike protein.And through structure based virtual ligand screening, we proposed possible furin inhibitors, which might be potentially used in the treatment of COVID-19.

2. Methodology

2.1 Homology Spike protein blast and sequence alignment.

The Spike protein of(GB:QHR63250.1) was downloaded from NCBI nucleotide database. The protein sequence were aligned with whole database using BLASTp to search for homology viral Spike protein (Alogorithm parameters, Max target sequences: 1000, Expect threshold: 10). Multiple-sequence alignment was conducted in BLASTp online and analysis with DNAMAN and Jalview. The evolutionary history was inferred using the Neighbor-Joining method in MEGA 7 software package. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test wasdetermined by 500 replicates. The Spike protein sequence analyses were

conducted in snapgene view.

2.2 Furin cleavage site prediction

The prediction of furin cleavage sites were carried out in ProP 1.0 Server (http://www.cbs.dtu.dk/services/ProP/).

2.3 Compounds database

Approved drug database was from the subset of ZINC database, ZDD (ZINC drug database) containing 2924 compounds [14]. Natural products database was constructed by ourselves, containing 1066 chemicals separated from traditional Chinese herbals in own lab and natural-occurring potential antiviral components and derivatives. Antiviral compounds library contains 78 known antiviral drugs and reported antiviral compounds through literature search.

2.4 Homology modeling and molecular docking

Corresponding homology models predicted by Fold and Function Assignment System server for each target protein were downloaded from Protein Data Bank (www.rcsb.org). Alignment of two protein sequences and subsequent homology modeling were performed by bioinformatics module of ICM 3.7.3 modeling software on an Intel i7 4960 processor (MolSoft LLC, San Diego, CA). For the structure-based virtual screening, ligands were continuously resiliently made to dock with the targetthat was represented in potential energy maps by ICM 3.7.3 software, to identify possible drug candidates. 3D compounds of each database were scored according to the internal coordinate mechanics (Internal Coordinate Mechanics, ICM)[15]. Based on Monte Carlo method, stochastic global optimization procedure and pseudo-Brownian positional/torsional steps, the position of intrinsic molecular was optimized. By visually inspecting, compounds outside the active site, as well as those weakly fitting to the active site were eliminated. Compounds with Scores less than -30 or mfScores less than -100 (generally represents strong interactions) have priority to be selected. Protein-protein docking procedure was performed according to the ICM-Pro manual.

3. Results

3.1 Bioinformatics analysis reveals furin cut site in Spike protein of SARS-CoV-2

By sequence alignment of Spike protein sequence of SARS-CoV-2 with its highly homologous sequences, it was found that cleavage site Spike of SARS-CoV-2 had 4 redundant

amino acids-PRRA, and these were not found in those of high homology coronavirus, which formed a furin-like restriction site as RRAR(Figure S1). Through prediction in ProP 1.0 Server, it was found the sequence was indeed easily digested by furin(Figure S2). In order to explore the evolution of this sequence, we used the BLASTp method to find 1,000 homologous Spike sequences with homology from 100% to 31%, which all from beta CoVs. Multiple sequence alignments were performed on these thousands of Spike sequences. One sequence was selected from each highly homologous class (homology greater than 98.5%) for further sequence alignment, and about 155 sequences were finally selected. A homologous multiple sequence alignment was performed on these 155 sequences, and then a phylogenetic tree was constructed(Figure 1). It is found from the phylogenetic tree that the Spike of SARS-CoV-2 exhibited the closest linkage to those of Bat-SL-CoV and SARS-CoV, and far from those of MERS-CoV, HCoV-HKU1, HCoV-OC43. In general, most of the Spike protein in α-CoV does not have a furin cleavage site, most of that in gama-CoV has a furin cleavage site, and that in beta-CoV with or without furin cleavage site are common[16].

We performed furin digestion site prediction on the sequence of each type of coronavirus Spikethrough online software. It was found that all Spike with a SARS-CoV-2 Spike sequence homology greater than 40% did not have a furin cleavage site (Figure 1, Table 1), including Bat-CoV RaTG13 and SARS-CoV (with sequence identity as 97.4% and 78.6%, respectively). The furin cleavage site "RRAR" in SARS-CoV-2 is unique in its family, rendering by its unique insert of "PRRA". The furin cleavage site of SARS-CoV-2 is unlikely to have evolved from MERS, HCoV-HKU1, and so on. From the currently available sequences in databases, it is difficult for us to find the source. Perhaps there are still many evolutionary intermediate sequences waiting to be discovered.

By analysis of the SARS -CoV-2 Spike protein sequence, it was found that most features are similar to SARS-CoV. It has an N-terminal signal peptide and is divided into two parts, S1 and S2. Among them, S1 contains N-terminal domain and receptor binding region. And S2 is mainly responsible for membrane fusion. The C-terminal region of S2 is S2', containing a fusion peptide, Hetad repeat1, Hetad repeat 2, and a transmembrane domain(Figure 2). There are two cleavage sites between S1 and S2 ', named CS1 and CS2. However, there are some differences in this two cleavage sites.

Unlike SARS-CoV, SARS-CoV-2 contains polybasic amino acids (RRAR) at the CS1 digestion site, and trypsin digestion efficiency will be significantly improved here[5]. More importantly, as mentioned above, this site can be recognized and cleaved by the furin enzyme. The cleavage of Spike protein promotes structural rearrangements of RBD for the adaptation to receptor, thus increasing the affinity[17]. More importantly, the digestion of Spike is an indispensable for membrane fusion of S2 part[18]. In this case, the efficiency of the SARS-CoV-2Spike protein cleavage is significantly higher than that of SARS-CoV, and the SARS-CoV-2Spike protein could be cut during the process of virus maturation (Figure 3). The receptor affinity and membrane fusion efficiency of SARS-CoV-2 would be significantly enhanced compared to that of SARS-CoV. The membrane fusion of SARS-CoV-2Spike protein is more likely to occur during endocytosis process. This may explains the current strong infectious capacity of SARS-CoV-2. So, the development of furin inhibitors may be a promising approach to block its transmissibility.
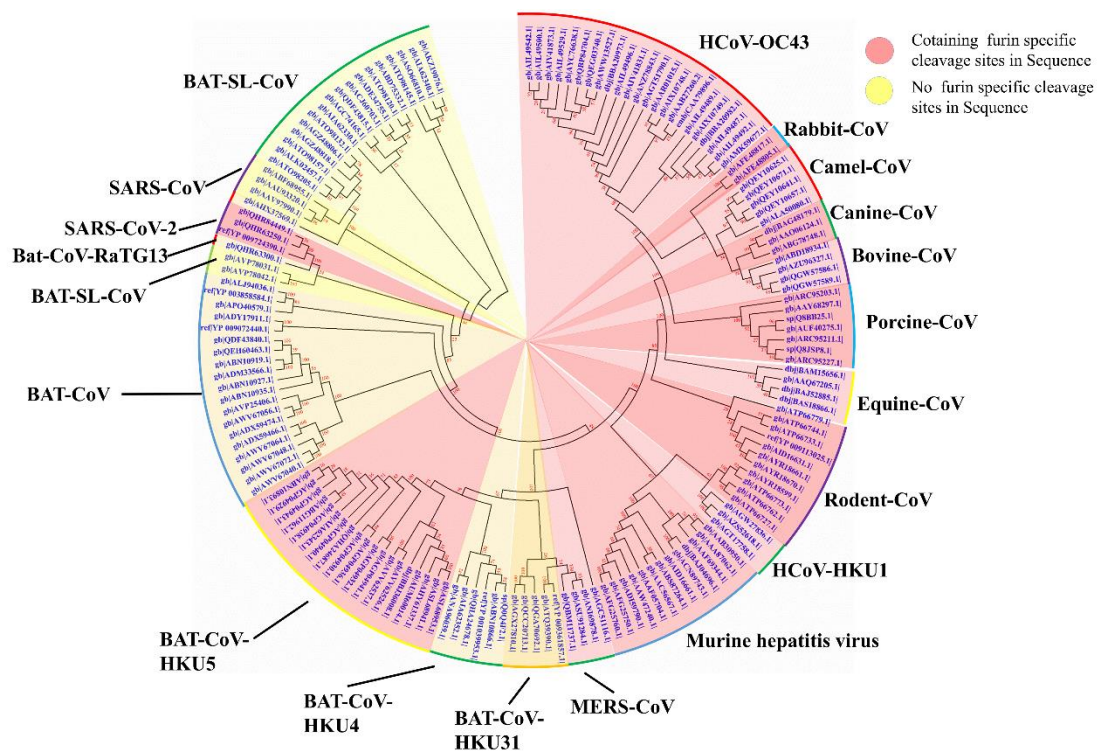
Figure1.Evolutionary relationships of taxa. The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were

computed using the Poisson correction method and in the units of the number of amino acid substitutions per site. The analysis involved 155 amino acid sequences. All positions containing gaps and missing data were eliminated. There are a total of 711 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. Those painted in red mean containing cleavage site in sequences and those painted in yellow mean no cleavage site in sequences.
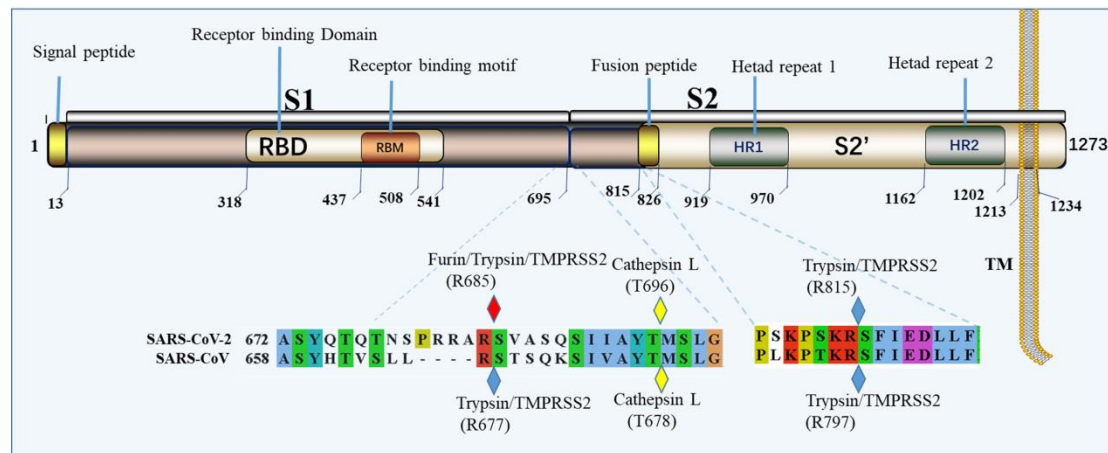

Figure 2.Sequence analysis of Spike protein in SARS-CoV-2. It contains an N-terminal signal peptide, S1 and S2. S1 contains N-terminal domain and receptor binding region. And S2 is mainly responsible for membrane fusion. The C-terminal region of S2 is S2', it contains a fusion peptide, HR1, HR2, and a transmembrane domain, the amino acid sequence numbers of every domain are annotated below them. Cleavage sites contained in SARS-CoV and SARS-CoV-2 are marked by rhombus.
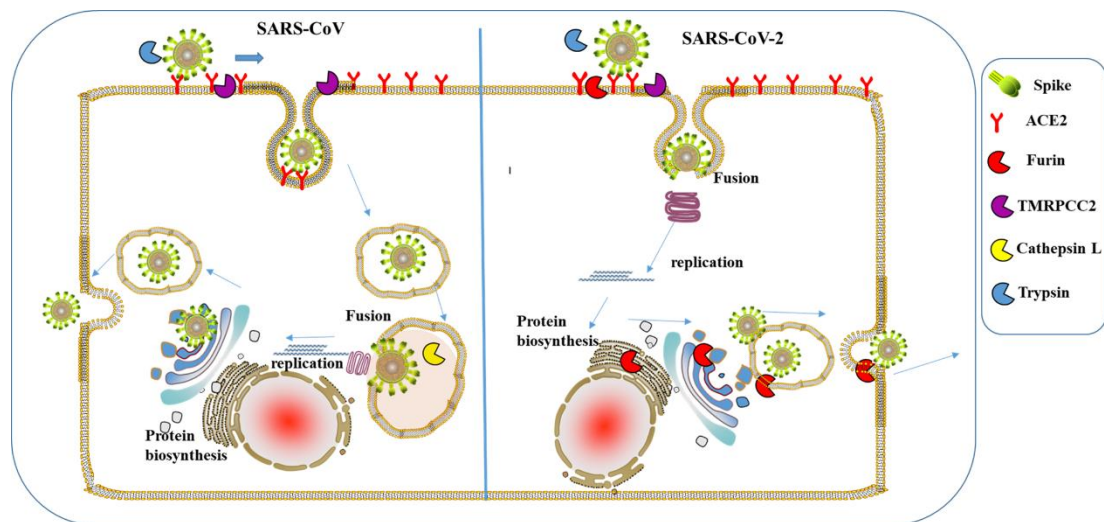
Figure 3.A schematic diagram of the process of SARS-CoV and SARS-CoV-2 infecting host cells.Those protease are presented by sector in different colors. Furin can cleaveSpike in the process of viral maturation.

Table 1.Furin cleavage probability of Spike sequence homology

Description	Accession no.	CS1 sequence	Furinscore[a]	Identity[b]
SARS-CoV-2	QHR63250.1	NSPRRAR/SV	0.620	100%
Bat-CoV-RaTG13	QHR63300.1	QTQTNSR/SV	0.151	97.4%
Bat-SL-CoV	AVP78042.1	HTASILR/ST	0.170	80.3%
SARS-CoV	ABF68955.1	QLTPAWR/IY	0.117	76.0%
Bat-CoV HKU5	AGP04941.1	PSARLAR/SD	0.697	37.1%
MERS-CoV	QBM11737.1	LTPRSVR/SV	0.563	35.0%
Rat-CoV	AFG25760.1	TAHRARR/SV	0.879	36.3%
MHV	ABS87264.1	TSHRARR/SI	0.861	36.9%
HCoV-HKU1	AGT17758.1	SSRRKRR/GI	0.744	36.8%
Rodent-CoV	ATP66727.1	TARRKRR/AL	0.795	37.3%
Beta-CoVsp	AYR18670.1	ATRRAKR/DL	0.753	35.9%
Equine-CoV	BAS18866.1	TARRQRR/SP	0.815	37.1%
Porcine-CoV	ARC95227.1	TSLRSRR/SL	0.758	36.1%
Bovine-CoV	QGW57589.1	TKRRSRR/AI	0.780	37.5%
Canine-CoV	ABG78748.1	TQRRSRR/SI	0.832	37.1%
Camel-CoV HKU23	ALA50080.1	IDRRARR/FT	0.718	36.5%
Rabbit-CoV HKU14	AFE48805.1	TLQPSRR/AI	0.629	37.7%
Human-CoV OC43	AMK59677.1	KTRRSRR/AI	0.720	36.8%

[a]Scores are predicted by ProP 1.0 Server. Scores above 0.5 mean furin cleavable.

[b]Identities compared with SARS-CoV-2 Spike protein.

3.2 Homology modeling and protein-protein docking calculation

In our previous studies (accepted by *ActaPharmaceuticaSinica B*), both SARS and SARS-CoV-2 spike RBD structures have been docked with human ACE2 to calculate their binding free energy. In that time, the complex structure of SARS-CoV-2 RBD with ACE2 was not available.Its energy was calculated based on the homology model generated from

SARS_RBD-ACE2 complex. The binding energy between the SARS-CoV-2 spike RBD and human ACE2 was -33.72 kJ mol[-1], and that between SARS-CoV spike RBD and ACE2 was -49.22 KJ mol[-1].This means the binding affinity between SARS-CoV-2 spike and ACE2 is weaker than that of SARS spike. During this manuscript was prepared, the structure of SARS-CoV-2 spike RBD-ACE2 complex was disclosed[19]. Based on this new real structure of SARS-CoV-2 spike RBD-ACE2 complex, we re-did the calculation and found that the binding free energy between SARS-CoV-2 spike RBD and ACE2 was -50.13 KJ mol[-1] (Figure S3). This means the binding affinity between SARS-CoV-2 spike and ACE2 is slightly stronger than that of SARS spike.By inspecting the crystal structure of SARS-CoV-2 RBD-ACE2 complex and SARS RBD-ACE2 complex, one can find that one key loop of SARS-CoV-2 RBD in the complex interface had very different conformation compared to that of SARS RBD and previous modeled SARS-CoV-2 RBD (Figure S4).

In order to further explore the possible mechanism how furin cleaves SARS-CoV-2 Spike, we perform protein-protein docking for furin and Spike. Although a Cryo-EM structure of SARS-CoV-2 Spike has been published in *bioRxiv* during this manuscript was prepared[20], the PDB coordinate was still not available so far. We already built a homology model of SARS-CoV-2 Spike in our previous paper submitted to another regular journal. SARS-CoV-2 Spike structure was built by using the SARS-CoVSpike structure as the temple (PDB code: 5X58)[21]. By superimposing the SARS-CoVSpike with the SARS-CoV-2 Spike, we can find that the major conformation differences between two structures are RBD domain, Arg685/677 loop region(furin/trypsin/TMPRSS2 cut site) and S2 loop region just after fusion peptide (Figure 4A).The trypsin/TMPRSS2 cut site of SARS-CoV was disordered and missing from the original Cryo-EM structure possibly due to its flexibility and without electro density. The "PRRA" inserting in SARS-CoV-2 in this region apparently generate the more flexible loop region and accessible cut site for protease. We performed protein-protein docking by setting SARS-CoV-2 Spikefurincleavage loop as the receptor, and furin active pocket as the ligand. The protein-protein docking results showed that furin acidic/negative active pocket can be well fitted onto the SARS-CoV-2 Spikebasic/positive S1/S2 protease cleavage loop with low energy (-18.43 Kcal/mol). This implies that the extra "PRRAR" loop of SARS-CoV-2 Spike renders it more fragile to the protease. And this may allow this site to be cut during the maturation, efficiently
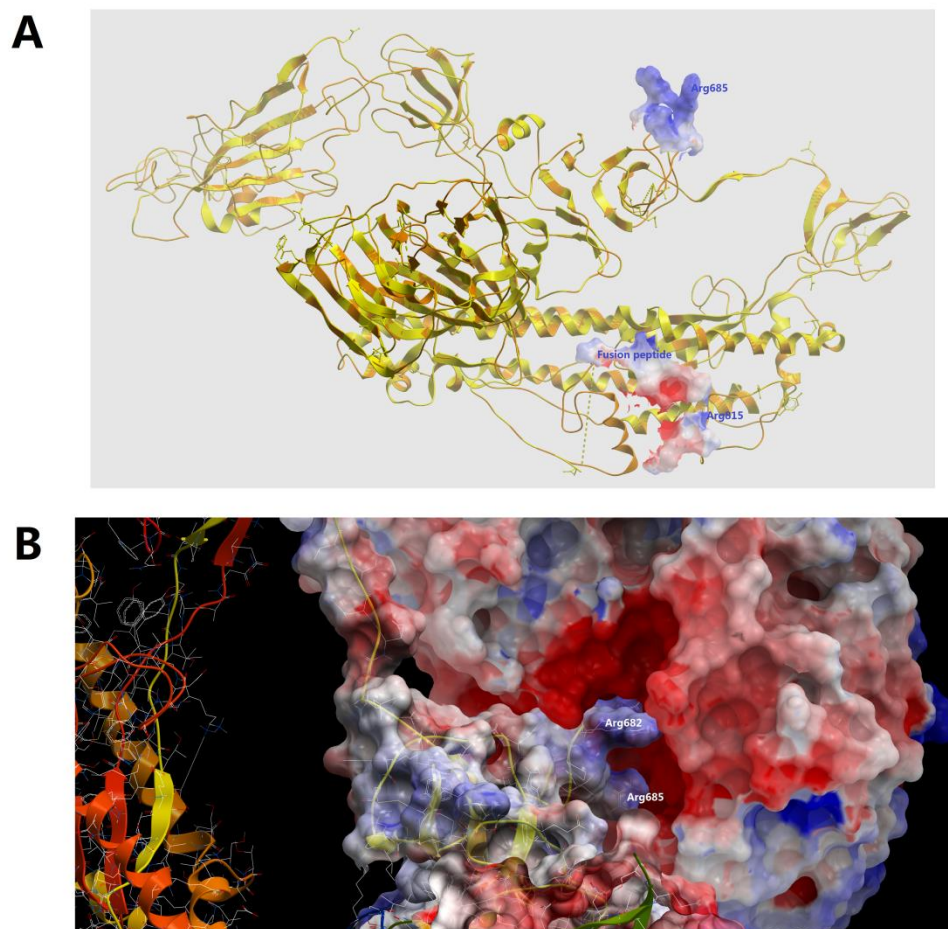
enhancing the infection efficiency.

Figure 4.Protein-protein docking model of SARS-CoV-2 Spike with furin. (A) Superimposition of SARS-CoVSpike and SARS-CoV-2 Spike. Two S1/S2 protease cleavage sites and fusion peptide were shown as electrostatic surface mode. (B) Furin was docked onto the putative furin cut site (Arg685) of SARS-CoV-2 Spike. Both domains are shown as electrostatic surface mode.
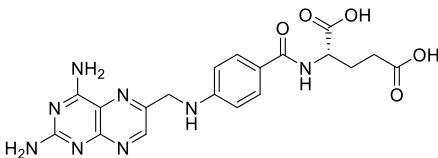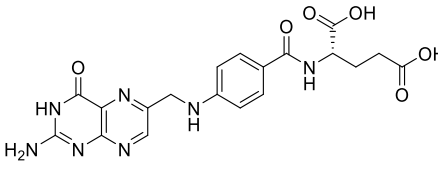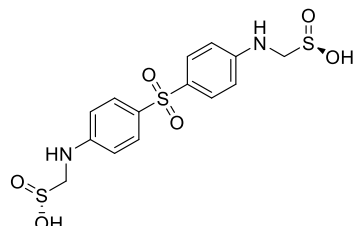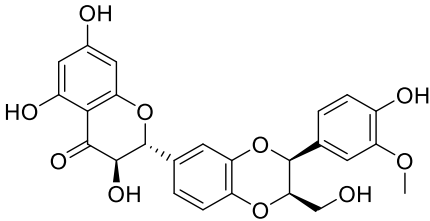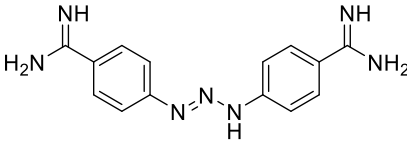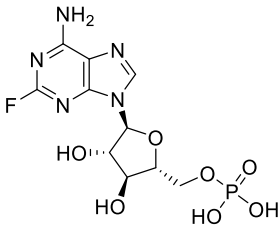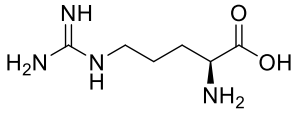
3.3. Virtual ligand screening of furin protein
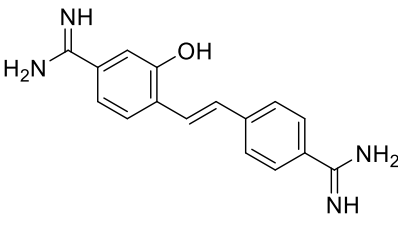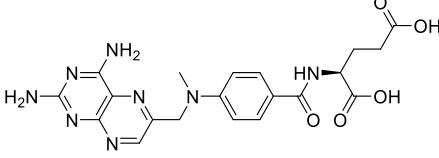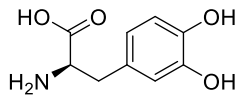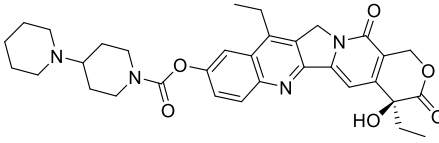
Structure-based virtual ligand screening method was used to screen potential furin protein inhibitors through ICM 3.7.3 modeling software (MolSoft LLC, San Diego, CA) from a ZINC Drug Database (2924 compounds), a small in-house database of natural products (including reported common antiviral components from traditional Chinese medicine) and derivatives (1066 compounds), and an antiviral compounds library contains 78 known antiviral drugs and reported antiviral compounds. Compounds with lower calculated binding energies (being expressed with scores and mfscores) are considered to have higher binding affinities with the target protein.
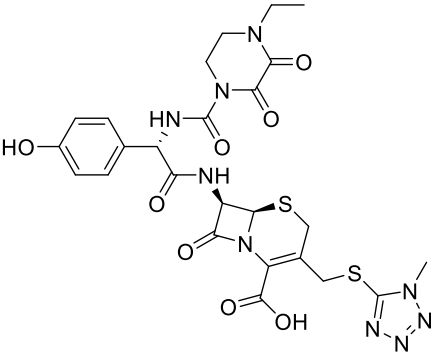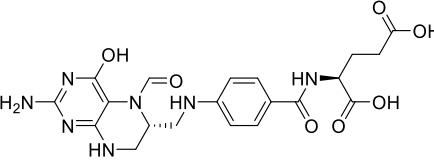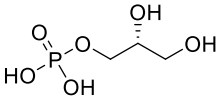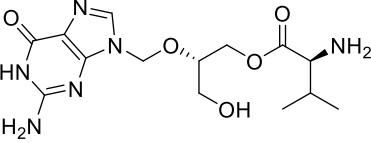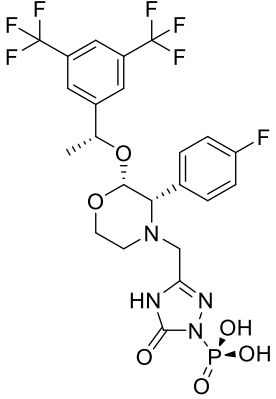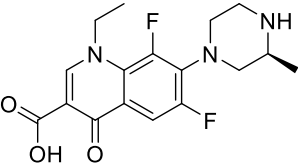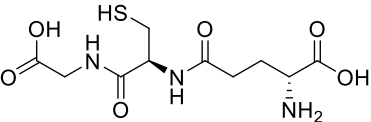
The screening results for the ZINC Drug Database (Table 2) showed that anti-tumor drugs Aminopterin, Fludarabine phosphate and Irinotecan, antibacterial drugs Sulfoxone,Lomefloxacinand Cefoperazone,antifungaldrug Hydroxystilbamidine, antivirus drugValganciclovir,hepatoprotective drugSilybin,folic acid supplementFolinic acid have higher binding affinity to furin with mfscores lower than -100 or Scores lower than -30.
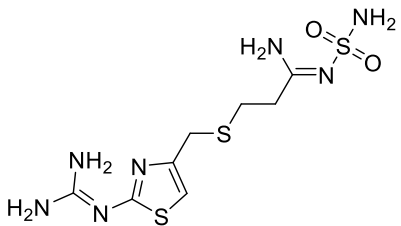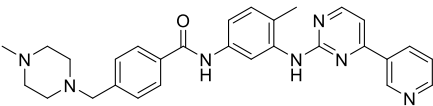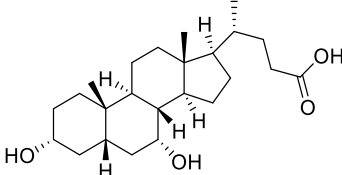
Here, we show one example of screen hits, Hydroxystilbamidine, which was predicted to bind in the active site of furin with low binding energy. In the generated docking model, Hydroxystilbamidine was well fitted into the binding pocket of the substrate and adopted similar conformation as substrate analogous inhibitor MI-52 in PDB model 5JXH[22],occupied two arms' position of MI-52 (Figure 5A). Asp159, Asp259 and Asp306 were predicted to form three hydrogen bonds with imine groups of compounds (Figure 5B). It looks like that Hydroxystilbamidine mimic at least two arginines. Weak hydrophobic interaction between His194, Leu227, the backbone of Trp254 and Asn295 with the compound may further stabilize its conformation.

Table 2. Potential furin inhibitors from ZINC drug database

No.	Drug Name	Structure	Pharmacological functions
1	Aminopterin		Anti-tumor
2	Folic acid		Vitamin B9, necessary material for the growth and reproduction of body cells
3	Sulfoxone		Antibacterial effect

4	Silybin		Hepatoprotective effect
5	Diminazene		Insecticidal effect
6	Fludarabine phosphate		Anti-tumor
7	L-Arginine		Nutritional supplement
8	Hydroxystilbamidine		Antifungal effect
9	Methotrexate		Antineoplastic, antirheumatic effects
10	L-dopa		Treatment of Parkinson's disease
11	Irinotecan		Anti-tumor

12	Cefoperazone		Antibacterial effect
13	Folinic acid		Folic acid supplement
14	Glycerol 3-phosphate		Intermediate for serine synthesis
15	Valganciclovir		Antivirus
16	Fosaprepitant		Treatment of nausea and vomiting induced by chemotherapy
17	Lomefloxacin		Antibacterial effect
18	Glutathione		Hepatoprotective effect

19	Famotidine		Treatment of gastrohelcosis
20	Imatinib		Anti-tumor
21	Chenodeoxycholic acid		Dissolving gallstones
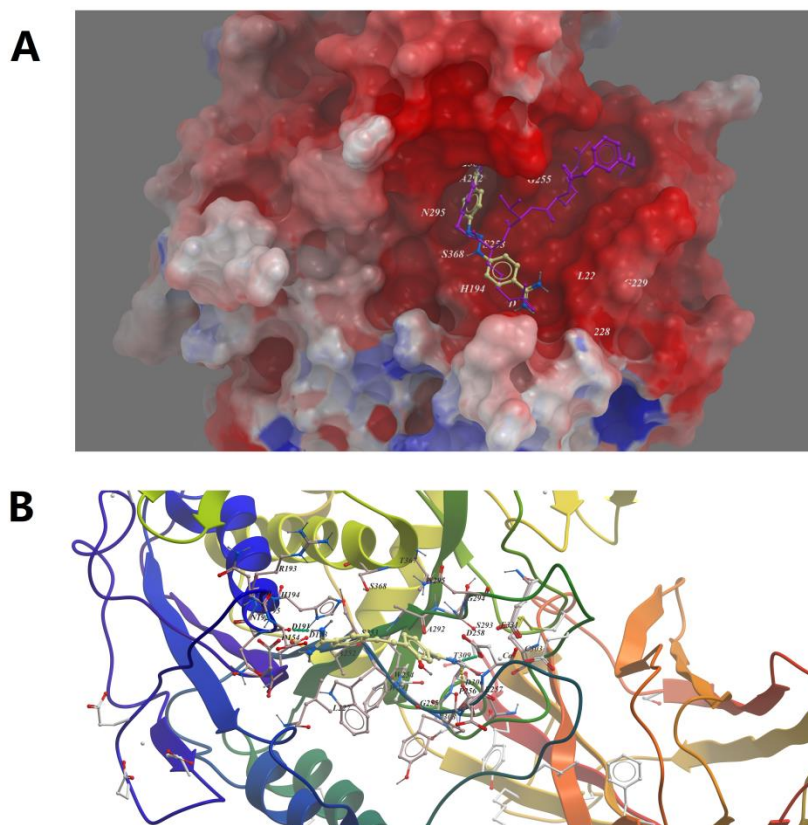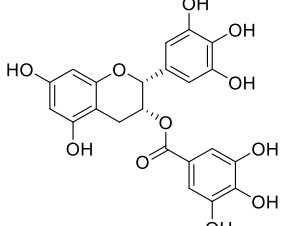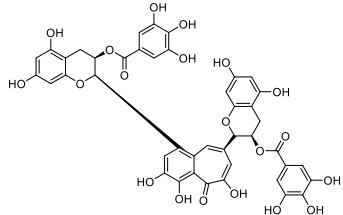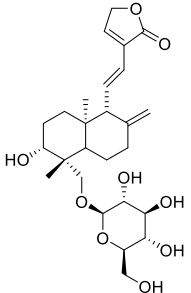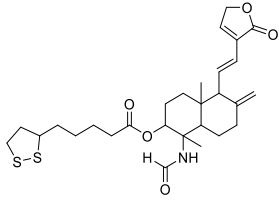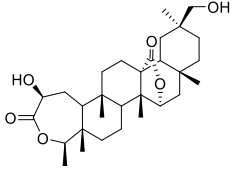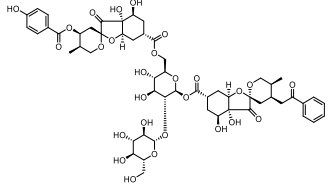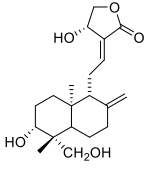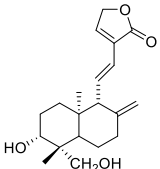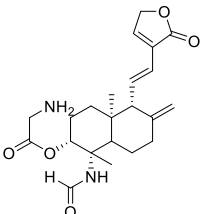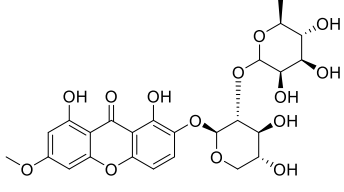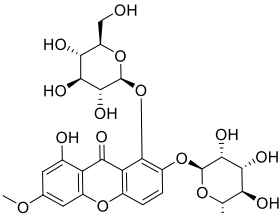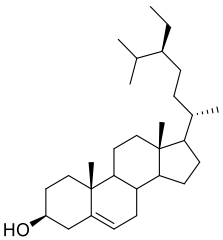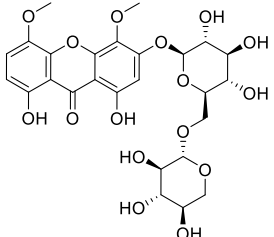
Figure 5.Low-energy binding conformations of Hydroxystilbamidine bound to furin generated by molecular docking. (A) Hydroxystilbamidinewas fitted well in the active pocket of human furin, and furin w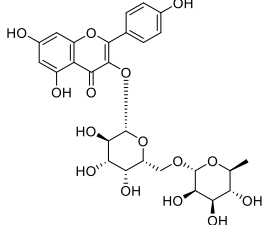as shown as electrostatic surface model. Hydroxystilbamidine (yellow) was overlapped with substrate analogue inhibitor MI-52 (purple).(B) Detailed view of Hydroxystilbamidinebinding in the activepocket of furin.

Another example was anticancer drug Imatinib. It was also predicted to bind in the active site of furin. In the generated docking model, Imatinib was fitted well in the binding pocket, and occupied the top two arms' position of MI-52 (Figure 6A). Two hydrogen bonds were predicted to form between the compound with Glu236 and Gly255. Weak hydrophobic interaction between Val231, Pro256, Trp254 and Gly294 and the compound was found (Figure 6B).

Figure 6. Low-energy binding conformations of Imatinib to furin generated by molecular docking. (A) Imatinib was fitted well in the active pocket of human furin, and furin was shown as electrostatic surface model. Imatinib (yellow) was overlapped with substrate analogue inhibitor MI-52 (purple).(B) Detailed view of Imatinib binding in the active pocket of furin.

Table 3.Potential furin inhibitors from in-house natural product database

No.	Drug Name	Structure	Pharmacological functions	Source
1	(-)-Epigallocatechin gallate		Antioxidation, anti-tumor, treatment of depression	*Camellia sinensis*
2	Theaflavin 3,3'-di-*O*-gallate		Antioxidant effect, anti-tumor, anti-virus	*Camellia sinensis*
3	Biorobin		Anti-virus	*Ficusbenjamina*
4	14-deoxy-11,12-didehydroandrographiside		Anti-virus, anti-inflammatory effect	*Andrographispanicu lata*
5	(1*S*,2*R*,4a*S*,5*R*,8a*S*)-1-formamido-1,4a-dimethyl-6-methylene-5-((*E*)-2-(2-oxo-2,5-dihydrofuran- 3-yl)ethenyl) decahydronaphthalen-2-yl 5-((*R*)-1,2-dithiolan-3-yl) pentanoate		Anti-virus, anti-inflammatory effect	Andrographolide derivatives

6	2β,30β-dihydroxy-3,4-seco-friedelolactone-27-lactone		Anti-virus	*Viola diffusa*
7	Phyllaemblicin G7		Anti-virus	*Phyllanthusemblica*
8	Andrographolide		Anti-virus, anti-inflammatory effect	*Andrographispaniculata*
9	14-deoxy-11,12-didehydroandrographolide		Anti-virus, anti-inflammatory effect	*Andrographispaniculata*
10	(1S,2R,4aS,5R,8aS)-1-formamido-1,4a-dimethyl-6-methylene-5-((E)-2-(2-oxo-2,5-dihydrofuran-3-yl)ethenyl) decahydronaphthalen-2-yl 2-aminoacetate		Anti-virus, anti-inflammatory effect	Andrographolide derivatives
11	2-[[2-*O*-(6-deoxy-*α*-L-mannopyranosyl)-*β*-D-xylopyranosyl]oxy]-1,8-dihydroxy-6-methoxy-9*H*-xanthen-9-one		Anti-virus, anti-inflammatory effect	*Swertiakouitchensis*
12	Kouitchenside J		Anti-virus, anti-inflammatory effect	*Swertiakouitchensis*

| 13 | Stigmast-5-en-3-ol | | Antioxidant effect | *Spatholobussuberect usdunn* |
| 14 | Kouitchenside F | | Anti-virus, anti-inflammatory effect | *Swertiakouitchensis* |

For the natural products (Table 3), a series of compounds with antivirus and anti-inflammation effects, such as (-)-Epigallocatechin gallateand Theaflavin 3,3'-di-*O*-gallatefrom*Camellia sinensis*,Biorobin from *Ficusbenjamina*, Andrographolide and 14-deoxy-11,12-didehydroandrographiside from *Andrographispaniculata*, one Andrographolide derivative (1*S*,2*R*,4a*S*,5*R*,8a*S*)-1-formamido-1,4a-dimethyl-6-methylene-5-((*E*)-2-(2-oxo-2,5-dihydrofuran-3-yl)ethenyl)decahydro naphthalen-2-yl 5-((*R*)-1,2-dithiolan-3-yl)pentanoate, 2*β*,30*β*-dihydroxy-3,4-seco-friedelolactone-27-lactone from *Viola diffusa*, Phyllaemblicin G7from*Phyllanthusemblica*, three xanthones2-[[2-*O*-(6-deoxy-*α*-L-mannopyranosyl)-*β*-D-xylopyranosyl]oxy]-1,8-dihydroxy-6-meth oxy-9*H*-xanthen-9-one, Kouitchenside J and Kouitchenside Ffrom *Swertiakouitchensis*exhibited high binding affinity to furin protein (mfscores< -100), suggesting the potential utility of these compounds in the treatment of SARS-CoV-2.

(-)-Epigallocatechin gallate (EGCG) was predicted to bind in the active site of furin, as Imatinib, it occupied the top two arms' position of MI-52 (Figure 7A). Two hydrogen bonds were predicted formed between the compound with Asp258 and Ala292. Weak hydrophobic interactions between Pro256, Trp254 and Gly294 and the compound were predicted (Figure 7B).
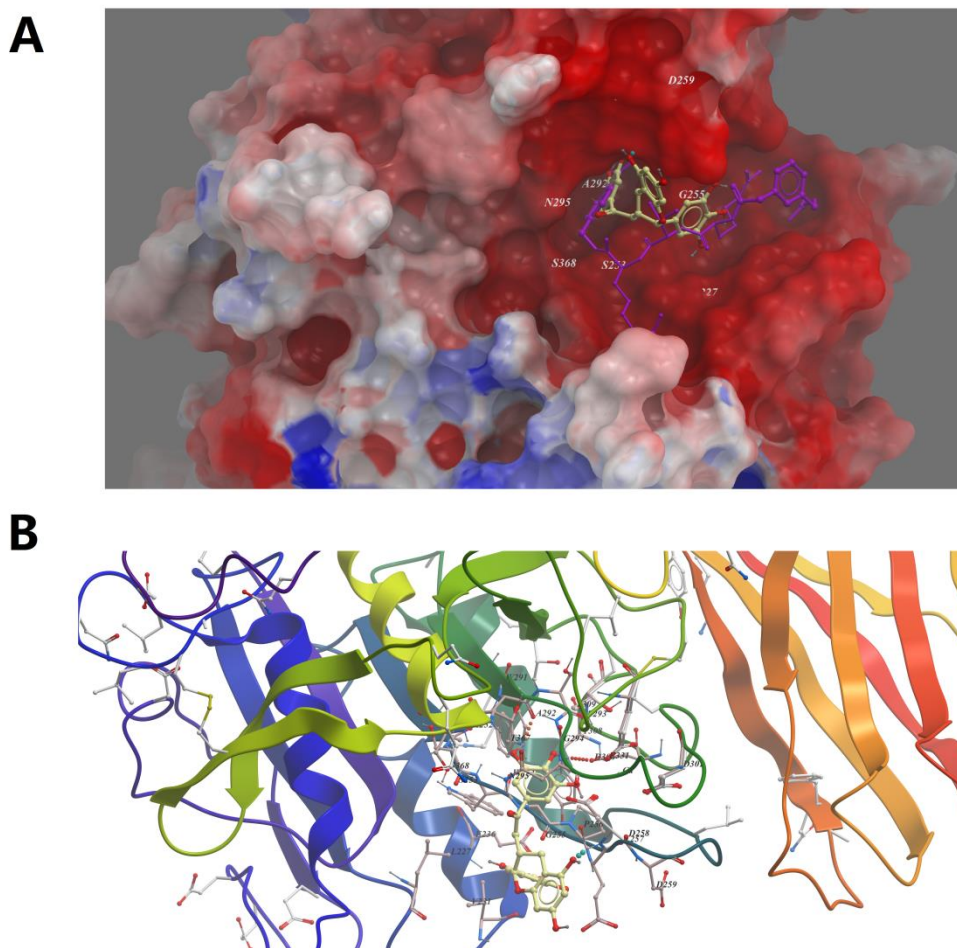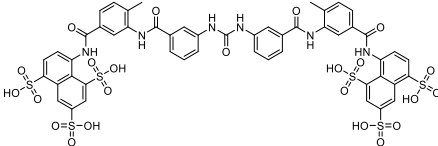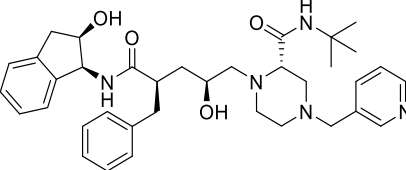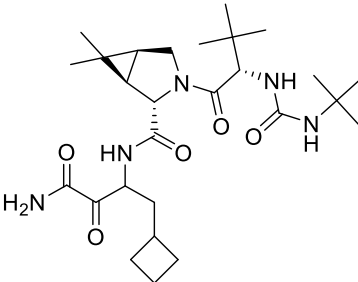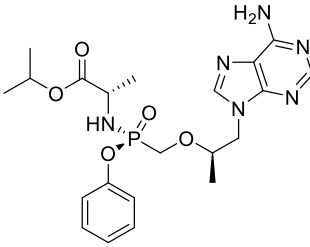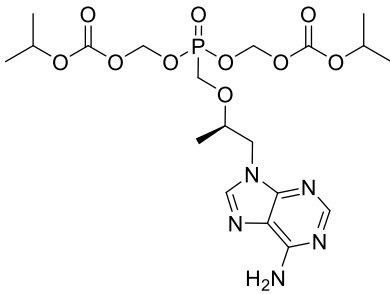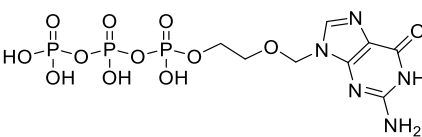
Figure 7. Low-energy binding conformations of ECCG to furin generated by molecular docking. (A) ECCGwas fitted well in the active pocket of human furin, and furin was shown as electrostatic surface model. ECCG (yellow) was overlapped with substrate analogue inhibitor MI-52 (purple).(B) Detailed view of ECCG binding in the activepocket of furin.
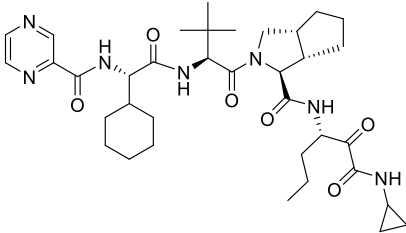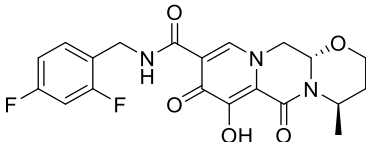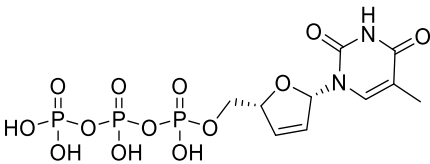
The database of 78 antiviral drugs including compounds already on the market and currently undergoing clinical trials to treat SARS-CoV-2 infections was further screened. The results were shown in Table 4. DNA topoisomerase II inhibitorSuramin treating hand-foot-and-mouth disease exhibited the highest affinity with furin (mfscore = 190.406). A series HIV-1 therapeutic drugs, such as Indinavir, Tenofoviralafenamide, TenofovirDisoproxil and Dolutegravir, and hepatitis C therapeutic drugs, Boceprevir and Telaprevir also have high binding affinity to furin.

Suramin was predicted to bind in the active site of furin with high binding mfScores. From generated docking model, Suramin occupied the top right arm and bottom arm positions of MI-52, it extended more to another adjacent pocket and covered almost all the surface areas for furin substrate binding (Figure 8A). Asp154, Asp228, Gly229, Ser253, Asp264, Glu271, Ile312, Lys449,

Arg490 and Asp530 were predicted to form 10 hydrogen bonds with the compound. Weak hydrophobic interactions may form between His194, Leu227, Tyr308, Trp531 and A532 with the compound (Figure 8B).

Table 4.Potential furin inhibitors from the common antiviral drugs database

No.	Drug Name	Structure	Pharmacological functions
1	Suramin		DNA topoisomerase II inhibitor
2	Indinavir		Human immunodeficiency virus Protease (HIV PR)
3	Boceprevir		Hepatitis C virus Serine protease NS3/4A (HCV NS3/4A) Modulator
4	Tenofoviralafenamide		HIV-1 nucleotide reverse transcriptase inhibitor
5	TenofovirDisoproxil		HIV, HBV nucleotide reverse transcriptase inhibitor
6	Acycloguanosine triphosphate		Thymidine kinase of herpesvirus

7	Telaprevir		Hepatitis C virus Serine protease NS3/4A (HCV NS3/4A) Modulator
8	Dolutegravir		Human immunodeficiency virus Integrase (HIV IN)
9	Maraviroc		1.C-C chemokine receptor type 5 (CCR5) 2.CCR5 messenger RNA(CCR5 mRNA)
10	Cobicistat		Inhibitor of cytochrome P450 3A (CYP3A) enzymes
11	Stavudine triphosphate		Nucleoside analogue reverse transcriptase inhibitor used in the treatment of HIV infection

Figure 8.Low-energy binding conformations of Suraminto furin generated by molecular docking. (A) Suraminwas fitted well in the active pocket of human furin, and furin was shown as electrostatic surface model. Suramin (yellow) was overlapped with substrate analogue inhibitor MI-52 (purple).(B) Detailed view of Suraminbinding in the activepocket of furin.
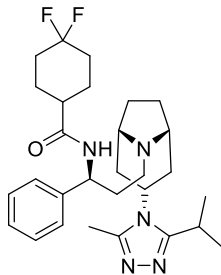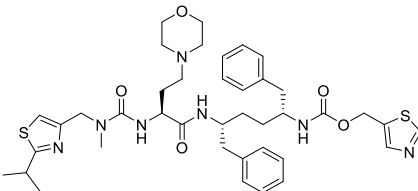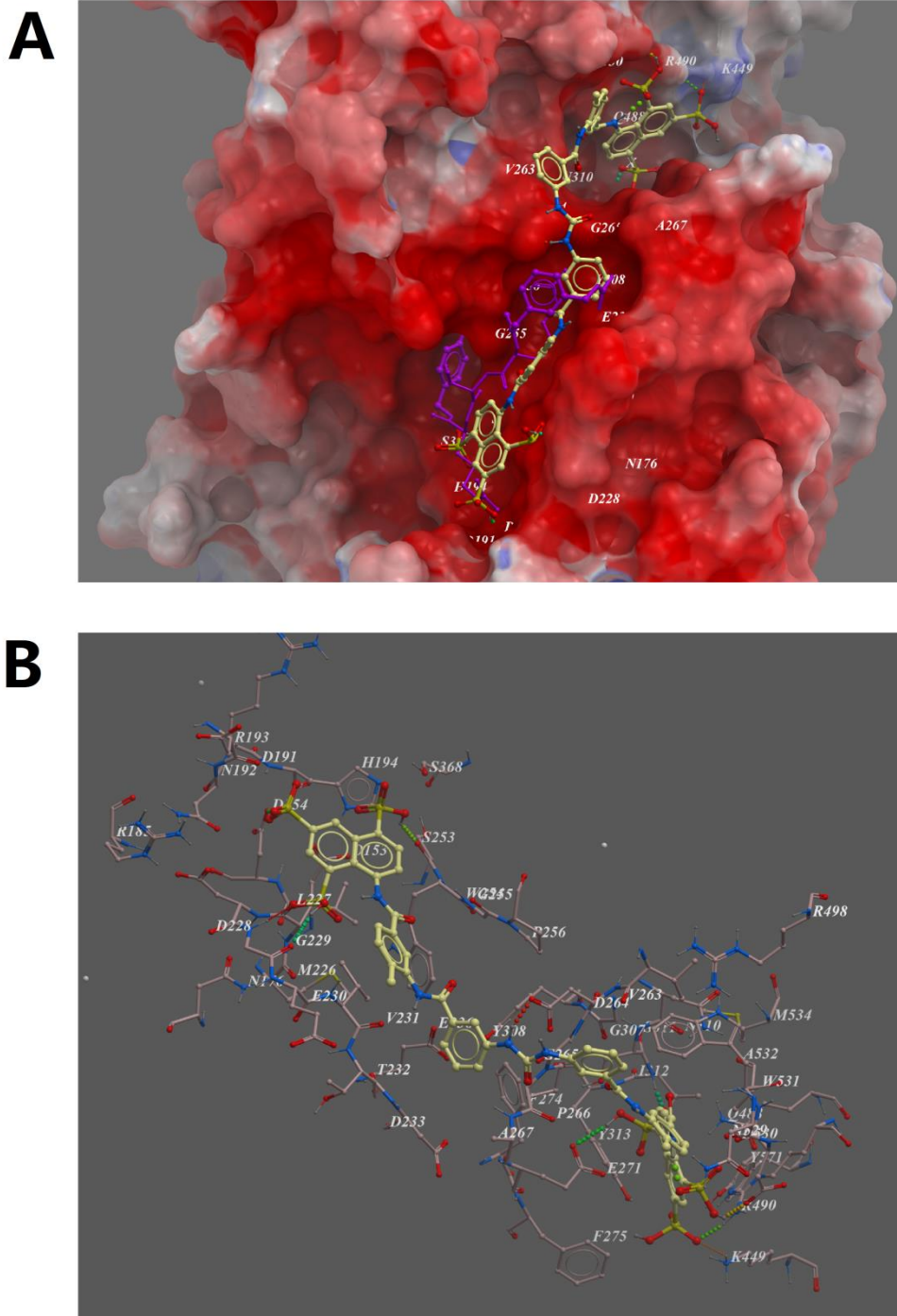
4. Discussion

Our previous study(accepted by *ActaPharmaceuticaSinica B*) analyzed the amino acid composition of the RBD domain of the ACE2 receptor of SARS-CoV-2 and Bat-CoVRaTG13. We found that several key amino acids determining binding were mutated in SARS-CoV-2, which are more similar tothat of SARS-CoV.The calculation results show that in the same conformation as the SARS-CoV protein, the binding energy of SARS-CoV-2 and ACE2 receptors was a litter higher, but this result cannot fully explain the epidemiologically high contagion, so we speculate (1)the RBD domain of SARS-CoV-2 may have other conformations; (2) there may be other receptors; and (3) there are other mechanisms that enhance infectivity. During this manuscript was prepared, the Cryo-EM structure of SARS-CoV-2 Spike was solved[20]. Comparing thestructure of SARS-CoV-2 with the Spike structure of SARS-CoV, combined with biophysical detection, they found that SARS-CoV-2 binds more strongly to cellular ACE2 receptors. Furthermore, the just disclosed crystal structure of SARS-CoV-2 RBD-ACE2 complex showed a distinct conformational change in the key loop of complex binding interface. And the binding free energy calculation indicated a slightly stronger binding for SARS-CoV-2 RBD compared to that of SARS RBD. These results confirm our guess that the conformational change of the RBD domain of SARS-CoV-2 leads to stronger binding. However, stronger receptor binding still can't fully explain the more infectious problem.

So we put forward these hypotheses: (1) SARS-CoV-2 can also bind to other receptors; (2) the lung may not be the earliest infection site; (3) SARS-CoV-2 is easier to cut and more easily fuse with cell membranes. Published in the Pubmed database, researchers performed RNA-seq analysis on tissue samples from 95 individuals' 27 different tissues. The results showed that ACE2 protein was highly expressed in the small intestine and duodenum, but the expression level in lung tissue is low (Figure S5). However, we analyzed the expression of furin and found that it is distributed in various organs with little difference in expression level. Combined with the possible infection mechanism of SARS-CoV-2, the widespread distribution of furin increases the SARS-CoV-2 infection of other organs. The possibility of other organ attack is consistent with the multiple symptoms observed in clinic of COVID-19.

Based on these three conjectures, we compared the Spike sequences from SARS-CoV-2, SARS-CoV, MERS-CoV and Bat-CoVRaTG13, and found that anextra "PRRA"insert near the

S1/S2 cleavage site. The "PRRA"insert and subsequent arginine (R) constitute a RRAR sequence that can be recognized and cleaved by furin-like proteases, which may be the reason why SARS-CoV-2 infection is stronger than SARS-CoV. What's more, we performed a homologous alignment and phylogenetic analysis of the SARS-CoV-2 sequence, and found that "PRRA"insert did not appear at any other close relatives of SARS-CoV-2, indicating that this insertwas completely novel in this genus virus. The existence of such a motif may allow Spikes to be cut into S1 and S2 by furin-like proteases before maturity, but not separated, which provides S1 with the flexibility to change the conformation to better fit the host receptor. According to Simmons G *et al.* studies, overexpression of furin can increase the activity of SARS-CoVSpike, but it will not cause Spike to be cleaved [23].This is consistent with our prediction.

Furthermore, Glowacka I*etal.*and Simmons G*et al.*studies have demonstrated that SARS-CoVSpikes can be activated by cleavage in two ways, including proteolytic activation by cathepsins B and L in host cells [24]. In addition, SARS-CoVSpike can be activated by TMPRSS2 cleavage on the host cell surface[6].What's more, MERA-CoV, S1/S1 and S2' cleavage sites cannot be cut by fruin[25].So we speculated that the activation of SARS-CoV-2 Spike can be through different protease cleavage pathways and these pathways can occur simultaneously in host cells. SARS-CoV-2 Spike can utilize host protease diversity to activate, which may explain the strong infectious capacity of SARS-CoV-2. As we can see in Figure 2, the Spike protein of SARS-CoV-2 can be cleaved at multiple stages, which greatly increases the efficiency of fusion. It is likely that the virus will fuse with the cell during endocytosis and release the genome. In addition, the binding ability of the cleaved Spike to the ACE2 receptor is also greatly enhanced [26].

According to our study, furin-like proteases may be potential drug targets for anti-SARS-CoV-2 treatment. At present, some peptide inhibitors have been developed and have good effects [27, 28]. To search potential inhibitors of furin-like proteases, we screened potential compounds from a ZINC drug database (2924 compounds), a small in-house database of natural products (1066 compounds), and existing antiviral drugs library (78 compounds) withfurinby virtual ligand screening. From the ZINC Drug Database, we found a series of anti-tumor, antibacterial, antivirus,hepatoprotective drugs, such as Aminopterin, Fludarabinephosphate,Sulfoxone,Irinotecan, Hydroxystilbamidine, Lomefloxacin, Cefoperazone, Valganciclovir,Imatinib, *etc*. might be used as furin inhibitors. For the natural products, some

flavonoids, diterpenoids, and steroids with antivirus and anti-inflammation effects, such as ECCG, Biorobin, Phyllaemblicin G7, Andrographolide and its derivatives, and xanthonesfrom the *Swertia*genus, *etc*.exhibited high binding affinity to furin protein. From the database of 78 antiviral drugs, a series of HIV-1, hepatitis C, and hand-foot-and-mouth disease therapeutic drugs, such as Indinavir, Tenofoviralafenamide, Tenofovir, Disoproxil, Dolutdegravir, Boceprevir, Telaprevir and Suraminalso showed high binding affinity to furin. These potentialfurin inhibitors and medicinal plants containing these compounds as major constituents might be useful for the treatment COVID-19. The further experiments to verify their efficiency *in viro* and i*n vivo* will be carried out in our future studies. What's more, combined administration of targeting different SARS-CoV-2 proteases with furin inhibitors may be an effective therapeutic strategy.

ACKNOWLEDGEMENTS

5. References

[1] Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. The coronavirus Spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* 2003; **77**: 8801-8811.

[2] Lu G, Wang Q, Gao GF. Bat-to-human: Spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends in Microbiology* 2015; **23**: 468-478.

[3] Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W, Hao P. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its Spike protein for risk of human transmission. *Sci China Life Sci* 2020.https://doi.org/10.1007/s11427-020-1637-5.

[4] Letko M, Munster V. Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV. *bioRxiv*, 2020.https://doi.org/10.1101/2020.01.22.915660.

[5] Belouzard S, Chu VC, Whittaker GR. Activation of the SARS coronavirus Spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A* 2009; **106**: 5871-5876.

[6] Glowacka I, Bertram S, Muller MA, Allen P, Soilleux E, Pfefferle S, Steffen I, Tsegaye TS, He Y, Gnirss K, Niemeyer D, Schneider H, Drosten C, Pohlmann S. Evidence that TMPRSS2 Activates the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Membrane Fusion and Reduces Viral Control by the Humoral Immune Response. *Journal of Virology* 2011; **85**: 4122-4134.

[7] Zhou Y, Vedantham P, Lu K, Agudelo J, Carrion R, Jr., Nunneley JW, Barnard D, Pohlmann S, McKerrow JH, Renslo AR, Simmons G. Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Res* 2015; **116**: 76-84.

[8] Kam YW, Okumura Y, Kido H, Ng LFP, Bruzzone R. Ralf AltmeyerCleavage of the SARS Coronavirus Spike Glycoprotein by Airway Proteases Enhances Virus Entry into Human Bronchial Epithelial Cells In Vitro. *PLoS ONE* 2009; **4**: e7870.

[9] Feliciangeli SF, Thomas L, Scott GK, Subbian E, Hung CH, Molloy SS, Jean F, Shinde U, Thomas G. Identification of a pH sensor in the furin propeptide that regulates enzyme activation. *J Biol Chem* 2006; **281**: 16108-16116.

[10] Henrich S, Cameron A, Bourenkov GP, Kiefersauer R, Huber R, Lindberg I, Bode W, Than ME. The Crystal Structure of the Proprotein Processing Proteinase Furin Explains Its Stringent Specificity. *Nat Struct Biol* 2003; **10**: 520-526.

[11] Tay FP, Huang M, Wang L, Yamada Y, Liu DX. Characterization of cellular furin content as a potential factor determining the susceptibility of cultured human and animal cells to coronavirus infectious bronchitis virus

infection. *Virology* 2012; **433**: 421-430.

[12] Yamada Y, Liu DX. Proteolytic activation of the Spike protein at a novel RRRR/S motif is implicated in furin-dependent entry, syncytium formation, and infectivity of coronavirus infectious bronchitis virus in cultured cells. *J Virol* 2009; **83**: 8744-8758.

[13] Matsuyama S, Shirato K, Kawase M, Terada Y, Kawachi K, Fukushi S, Kamitani W. Middle East Respiratory Syndrome Coronavirus Spike Protein Is Not Activated Directly by Cellular Furin during Viral Entry into Target Cells. *J Virol* 2018; **92**.

[14] Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling* 2012; **52**: 1757-1768.

[15] Abagyan R, Totrov M, Kuznetsov D. ICM-A New Method for Protein Modeling and Design_ Applications to Docking and Structure Prediction from the Distorted Native Conformation. *J Comput Chem* 1994; **15**: 488-506

[16] Millet JK, Whittaker GR. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research* 2015; **202**: 120-134.

[17] Walls AC, Xiong X, Park Y-J, Tortorici MA, Snijder J, Quispe J, Cameroni E, Gopal R, Dai M, Lanzavecchia A, Zambon M, Rey FA, Corti D, Veesler D. Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* 2019; **176**: 1026-1039.

[18] Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Corbett KS, Graham BS, McLellan JS, Ward AB. Pre-fusion structure of a human coronavirus Spike protein. *Nature* 2016; **531**: 118-121.

[19]http://nmdc.cn/?from=groupmessage#/resource/detail?no=NMDCS0000001

[20]Daniel Wrapp, Nianshuang Wang, Kizzmekia S. Corbett, Jory A. Goldsmith, Ching-Lin Hsieh, Olubukola Abiona, Barney S. Graham, View ORCID ProfileJason S. McLellan. Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *bioRxiv*. 2020. https://doi.org/10.1101/2020.02.11.944462.

[21] Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X, Gao GF. Cryo-EM structures of MERS-CoV and SARS-CoV Spike glycoproteins reveal the dynamic receptor binding domains.*Nat Commun*. 2017; **10**:15092.

[22] Dahms SO, Arciniega M, Steinmetzer T, Huber R, Than ME. Structure of the unliganded form of the proprotein convertase furin suggests activation by a substrate-induced mechanism. *Proc Natl Acad Sci* 2016; **113**:11196-11201.

[23] Simmons G, Bertram S, Glowacka I, Steffen I, Chaipan C, Agudelo J, Lu K, Rennekamp AJ, Hofmann H, Bates P, Pohlmann S. Different host cell proteases activate the SARS-coronavirus Spike-protein for cell-cell and

virus-cell fusion. *Virology* 2011; **413**: 265-274.

[24] Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, Bates P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad Sci U S A* 2005; **102**: 11876-11881.

[25] Van Lam van T, Ivanova T, Hardes K, Heindl MR, Morty RE, Bottcher-Friebertshauser E, Lindberg I, Than ME, Dahms SO, Steinmetzer T. Design, Synthesis, and Characterization of Macrocyclic Inhibitors of the Proprotein Convertase Furin. *ChemMedChem* 2019; **14**: 673-685.

[26] Parka J-E, Lib K, Barlana A, Fehrc AR, Perlmanb S, McCray PB Jr, Gallagher T. Proteolytic processing of Middle East respiratory syndrome coronavirus Spikes expands virus tropism.*Proc Natl Acad Sci U S A* 2016; **113**: 12262–12267

[27] Dahms SO, Hardes K, Steinmetzer T, Than ME. X-ray Structures of the Proprotein Convertase Furin Bound with Substrate Analogue Inhibitors Reveal Substrate Specificity Determinants beyond the S4 Pocket. *Biochemistry* 2018; **57**: 925-934.

[28] Dahms SO, Jiao GS, Than ME. Structural Studies Revealed Active Site Distortions of Human Furin by a Small Molecule Inhibitor. *ACS Chem Biol* 2017; **12**: 1211-1216.

Furin, a potential therapeutic target for COVID-19
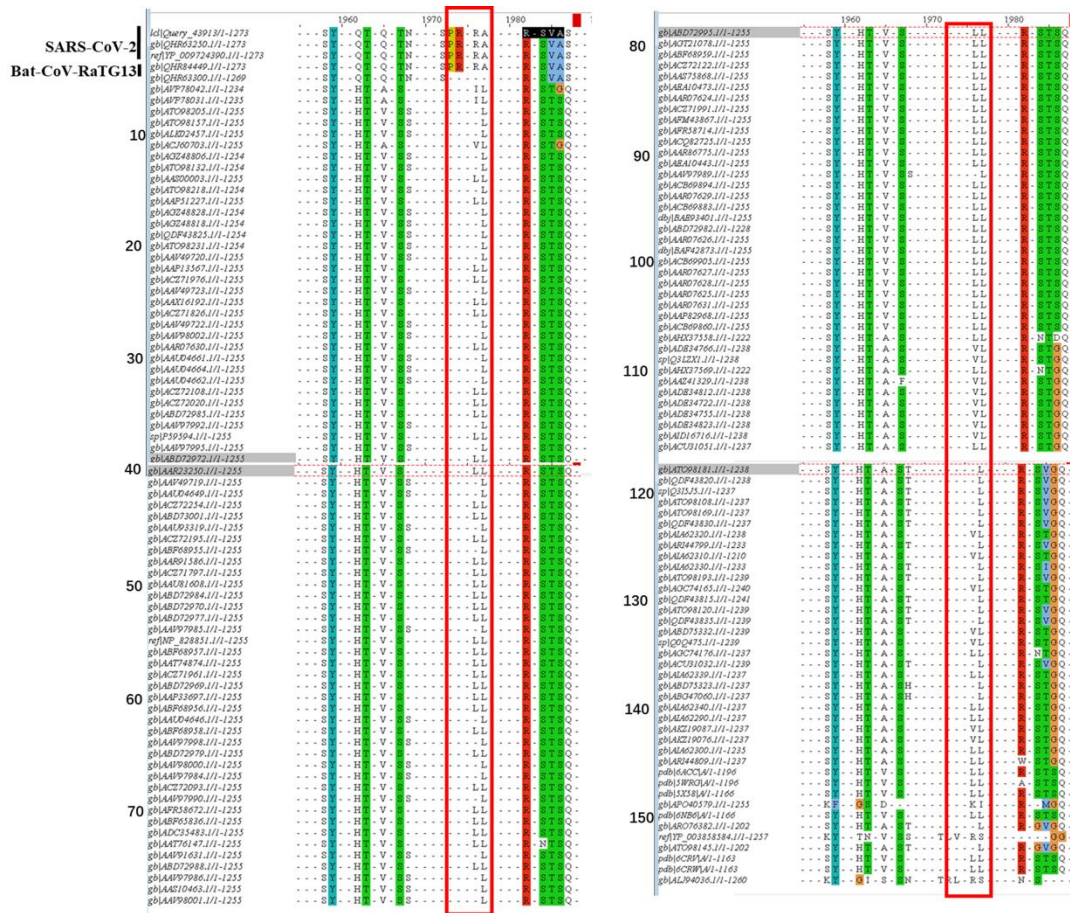
Supplementary information

Figure S1. Multiple sequence alignment of 1000 Spike proteins. These 156 proteins were ranked according to their homology with SARS-2.The sequence corresponding to PRRA in SARS-CoV-2 in each sequence is marked in the red box.

| Sequence | 457 | YLYRLFR|KS | 0.179 | . |
Sequence	458	LYRLFRK	SN	0.100	.
Sequence	462	FRKSNLK	PF	0.069	.
Sequence	466	NLKPFER	DI	0.154	.
Sequence	509	VCGYQPYR	VV	0.091	.
Sequence	528	ATVCGPK	KS	0.059	.
Sequence	529	TVCGPKK	ST	0.146	.
Sequence	535	KSTNLVK	NK	0.066	.
Sequence	537	TNLVKNK	CV	0.093	.
Sequence	557	VLTESNK	KF	0.057	.
Sequence	558	LTESNKK	FL	0.091	.
Sequence	567	PFQQFGR	DI	0.127	.
Sequence	577	DTTDAVR	DP	0.076	.
Sequence	634	QLTPTWR	VY	0.098	.
Sequence	646	SNVFQTR	AG	0.103	.
Sequence	682	TQTNSPR	RA	0.111	.
Sequence	683	QTNSPRR	AR	0.146	.
Sequence	685	NSPRRAR	SV	0.620	*ProP*
Sequence	733	LPVSMTK	TS	0.065	.
Sequence	765	FCTQLNR	AL	0.076	.
Sequence	776	IAVEQDK	NT	0.071	.
Sequence	786	EVFAQVK	QI	0.066	.
Sequence	790	QVKQIYK	TP	0.056	.
Sequence	795	YKTPPIK	DF	0.079	.
Sequence	811	ILPDPSK	PS	0.058	.
Sequence	814	DPSKPSK	RS	0.071	.
Sequence	815	PSKPSKR	SF	0.333	.
Sequence	825	EDLLFNK	VT	0.053	.
Sequence	835	ADAGFIK	QY	0.094	.
Sequence	847	LGDIAAR	DL	0.105	.
Sequence	854	DLICAQK	FN	0.064	.
Sequence	905	AMQMAYR	FN	0.169	.
Sequence	921	VLYENQK	LI	0.065	.
Sequence	933	FNSAIGK	IQ	0.067	.
Sequence	947	TASALGK	LQ	0.077	.
Sequence	964	ALNTLVK	QL	0.079	.
Sequence	983	LNDILSR	LD	0.068	.
Sequence	986	ILSRLDK	VE	0.129	.
Sequence	995	AEVQIDR	LI	0.081	.
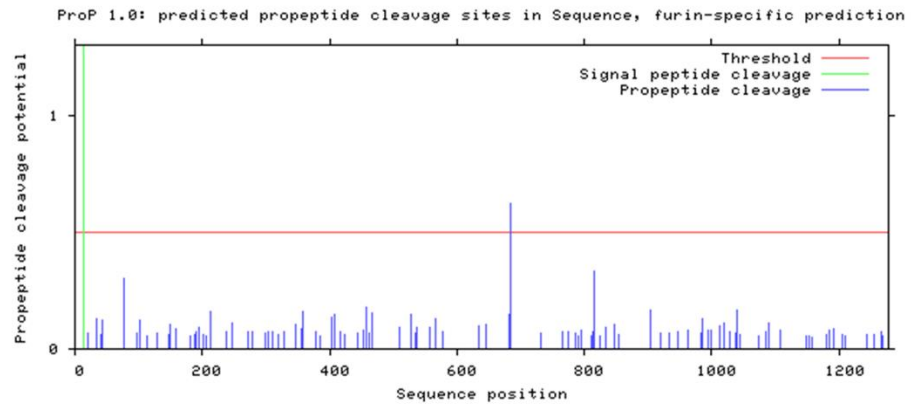
Figure S2. Result of furin cleavage site pridiction of Spike protein in SARS-CoV-2, which predicted by online method ProP 1.0
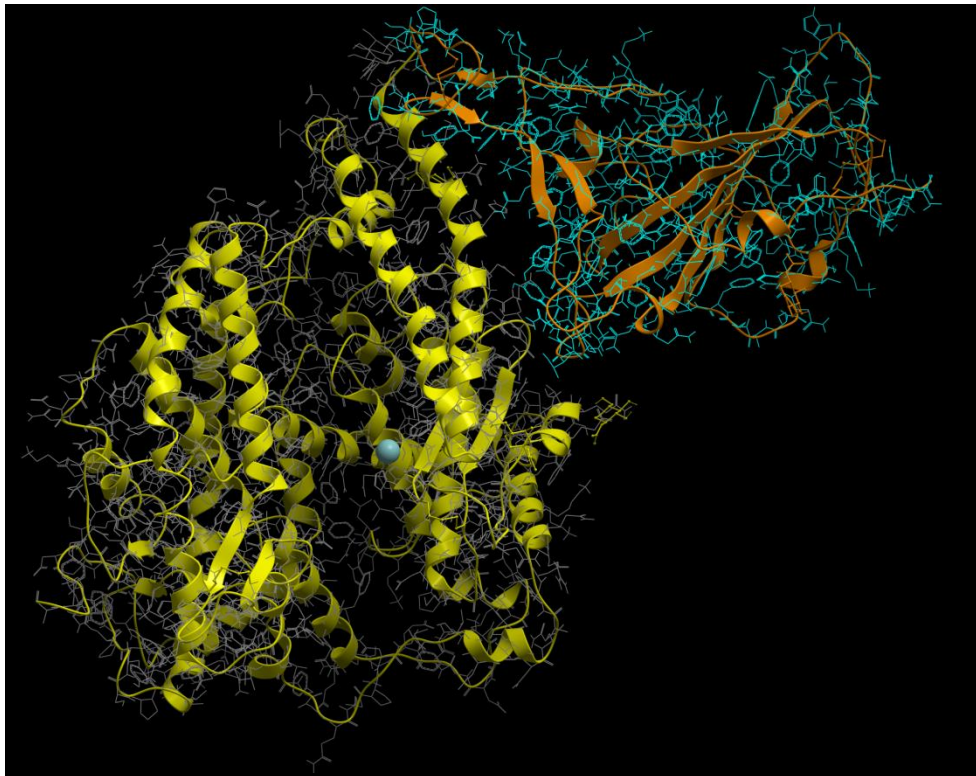
Server.

Figure S3. Protein-protein dockingcalculation model of SARS-CoV-2 spike RBD (light blue) with human ACE2 (yellow),

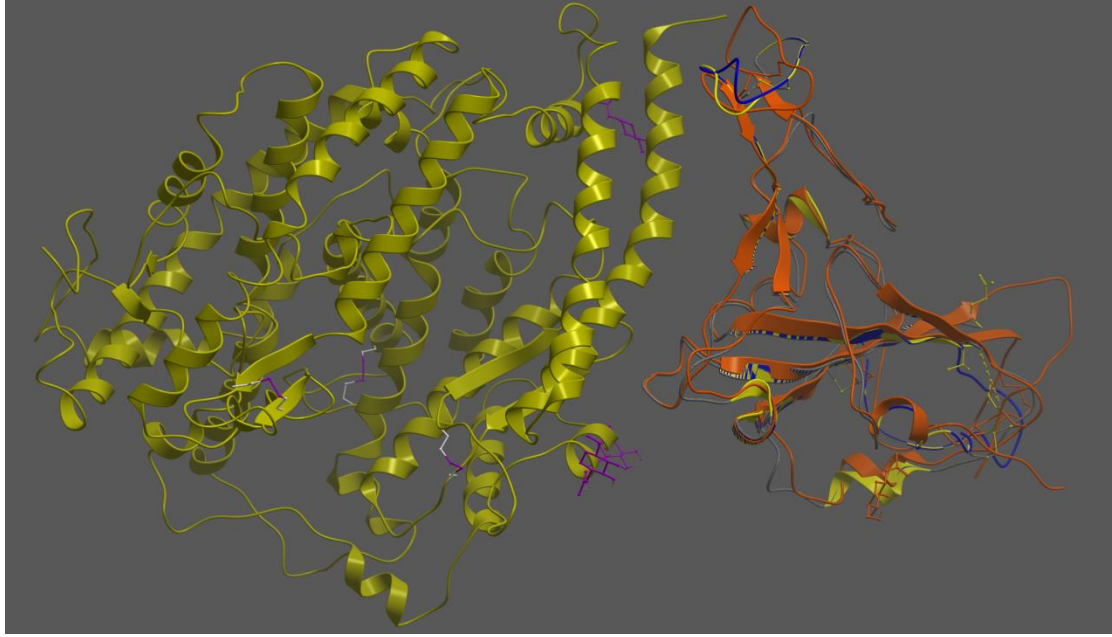original RBD conformation was shown in orange. The calculated free energy is -50.13 Kcal/mol.

Figure S4. Comparison of SARS-CoV-2 spike RBD (orange) and SARS spike RBD (yellow). The complex with ACE2 (left part, yellow) was shown. The homology model of SARS-CoV-2 spike RBDbuilt from SARS spike RBD was shown as blue.
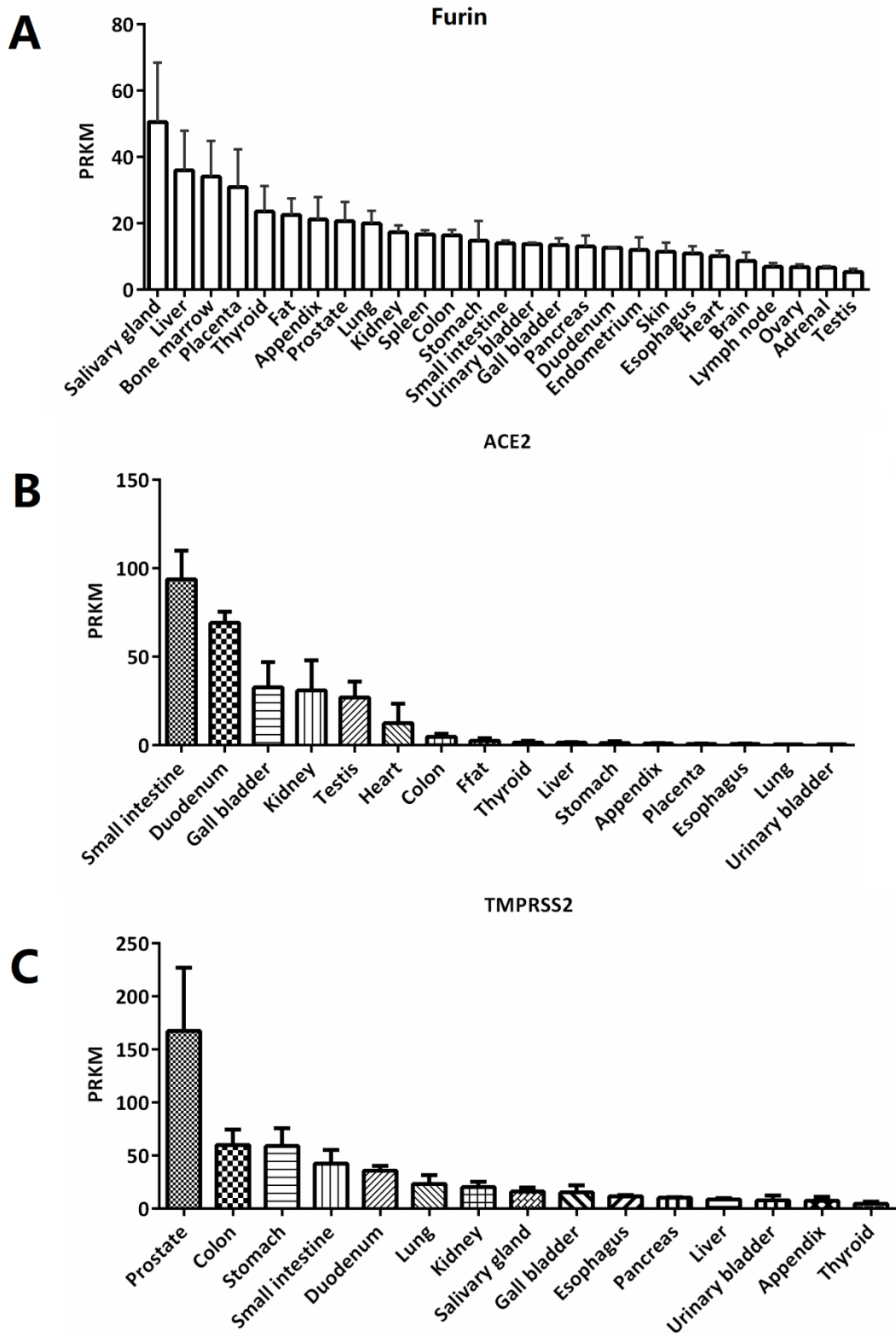
Figure S5. Expression levels of Furin, ACE2 and TMPRSS2 in various tissues. The data is from pubmed [1-3].

References

[1] Angiotensin I converting enzyme 2 (ACE2) expression level in human tissues using HPA RNA-seq method. [DB/OL].(2020-02-03) [2020-02-17]. https://www.ncbi.nlm.nih.gov/gene/59272/?report=expression.

[2] Furin, paired basic amino acid cleaving enzyme (Fruin) expression level in human tissues using HPA RNA-seq method. [DB/OL].(2019-12-21) [2020-02-17]. https://www.ncbi.nlm.nih.gov/gene/5045/?report=expression.

[3] Transmembrane serine protease 2 (TMPRSS2) expression level in human tissues using HPA RNA-seq method. [DB/OL].(2019-12-21) [2020-02-17]. https://www.ncbi.nlm.nih.gov/gene/7113?report=expression.