in silico Screening of Potential Spike Glycoprotein Inhibitors of SARS-CoV-2 with Drug Repurposing Strategy

Tianzi Wei
Department of Biology, Southern University of Science and Technology, Shenzhen, China
Hao Wang
Department of Obstetrics and Gynecology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Xueqing Wu
Department of Obstetrics and Gynecology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Yi Lu
School of Medicine, Southern University of Science and Technology, Shenzhen, China
Shenghui Guan
Department of Biology, Southern University of Science and Technology, Shenzhen, China
Fengquan Dong
Department of Cardiology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Chenle Dong
Department of Obstetrics and Gynecology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Guli Zhu
Department of Obstetrics and Gynecology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Yuzhou Bao
Department of Cardiology, Shenzhen University General Hospital, Shenzhen, Guangdong, China
Jian Zhang (zhangjian@sustech.edu.cn)
School of Medicine, Southern University of Science and Technology, Shenzhen, China
Guanyu Wang (wanggy@sustech.edu.cn)
Department of Biology, Southern University of Science and Technology, Shenzhen, China
Haiying Li (lihaiying111@hotmail.com)
Department of Cardiology, Shenzhen University General Hospital, Shenzhen, Guangdong, China

Research Article

Keywords: COVID19, SARS-CoV-2, Drug repurposing, Virtual screening, Spike glycoprotein, ACE2

Posted Date: March 19th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-17720/v1

Version of Record: A version of this preprint was published at Chinese Journal of Integrative Medicine on August 1st, 2020. See the published version at https://doi.org/10.1007/s11655-020-3427-6.

Abstract

COVID-19 has globally spread and has become a new pandemic, but there is still no effective drugs or vaccines to treat or prevent this disease. SARS-Cov-2 invades human cells through its spike proteins interacting with human ACE2 receptors. One strategy to prevent the virus from entering cells is the interruption of the viral spike protein interacting with ACE2. In such an emergent situation, drug repurposing is a promising method for rapid drug development. Here, we selected around 15000 molecular candidates including FDA-approved drugs from DrugBank and natural compounds from TCMSP to perform virtual screening for potential molecules that can target viral spike protein based on its crystal structure. In this article, we present the top 20 molecules with high binding affinity with spike protein, of which, digitoxin, a cardiac glycoside in DrugBank and bisindigotin in TCMSP, extracted from *indigo naturalis* and *polygoni tinctorii foliu*, have the highest docking scores. In addition, we also found that raltegravir, an HIV integrase inhibitor, has a relatively high binding score. Those molecules with high binding capacity to spike glycoprotein might be used by other researchers for further anti-COVID-19 drug development.

1 Introduction

Since the middle of December 2019, a novel coronavirus disease (COVID–19) outbroke in Wuhan, the capital city of Hubei Province of China, and rapidly spread throughout the country with the population movement of the Chinese New Year. Early studies utilized real-time PCR (RT-PCR) to confirm the samples collected from one patient with pneumonia were positive for pan-Betacoronavirus1. The virus isolated from human cell lines were observed containing typical crown-like shapes under the transmission electron microscope (TEM) with negative staining1.

Based on the metagenomic sequencing, the whole genome sequence of the virus was determined1,2. Bioinformatics analysis indicated the novel virus causing severe pneumonia is a new type of beta-coronavirus. The virus shares various typical genomic compositions with other beta-coronavirus family members and has the highest sequence homology (96% sequence similarity) to the SARS-like RaTG13 in bat1.

To date, COVID–19 cases were reported in more than 67 countries. As of 29 February 2020, a cumulative total of 83652 COVID–19 cases were reported around the world. Although the virus nucleic acid detection kits have been rapidly developed, effective drugs and vaccines are still ongoing studied with urgent needs. Facing such urgent situations, although necessary, developing new drugs or vaccines will cost a very long time from null, the drug reposition is more rapid and possible to discover anti-SARS- CoV–2 drugs from approved drugs3. Besides, the natural compounds from traditional Chinese medicine (TCM) are potential valuable pools for antivirus drug screening. Thus, rapid drug screening from these sources is an approach to anti-COVID–19 drugs.

Receptor recognition by coronavirus is the first and essential step for entering human cells. The homotrimeric spike glycoprotein (S protein), located on the envelope of the SARS-CoV–2, is responsible for receptor recognition. The S1 subunit of S protein, containing receptor-binding domain (RBD), directly interacts with the receptor on the

human cell membrane while the S2 subunit facilitates virus-cell fusion and entry4. Angiotensin-converting enzyme 2 (ACE2) was shown to be the receptor mediating SARS-CoV–2 invading human cells1. Thus, interrupting the interaction between S protein and ACE2 is a strategy to inhibit virus entry5.

Recently, the three-dimensional structures of the homotrimeric spike glycoprotein (PDB 6vsb))6 and ACE2-B0AT1 complex were solved by cryo-EM respectively. Besides, the complex structures of the receptor-binding domain (RBD) of S protein bound with human receptor ACE2 were reported by three Chinese research teams independently. All related coordinates have been published by these researchers in the first place. Avoiding inaccurate protein structure modeling, these accurate coordinates of protein structures make it possible for rapid drug screening and vaccine design.

Here, we performed structure-based virtual screening to search the molecules from Drugbank and TCMSP databases to find the potential inhibitors targeting viral spike protein. The selected binding pocket neighbors the interface between viral spike protein and human receptor ACE2, which may further inhibit their interaction to prevent the virus from invading the human cells. Follow this computational work, more vigorous *in vitro* and *in vivo* experiments need to be performed. We hope this work will contribute to the discovery of anti-COVID–19 drugs.

2 Results

2.1 Pocket identified on Spike protein

On Feb. 20., 2019, Dr. Jianxun Qi's team published the crystal structure of the receptor- binding domain (RBD) of Spike protein complexed with ACE2 receptor on National Microbiology Datacenter (No. NMDCS0000001; PDB code:6LZG). According to this crystal structure, Discovery Studio 2016 was employed to detect binding pocket on the Spike protein. In this work, the second-ranked pocket was selected as its position near the protein-protein interface and may induce conformational change to intervene interaction (see Figure 1). The coordination of the pocket is: center x = –56.387, center y = 52.408, center z = 21.937; size x = 25, size y = 25, size z = 25.

2.2 Data source

Both 2191 compounds in DrugBank and 13026 compounds from TCMSP were prepared as candidates. Autodock vina was employed to perform *in silico* high- throughput screening. The compounds in DrugBank are all FDA-approved compounds. The compounds provided by TCMSP are natural compounds. Those are compounds in Traditional Chinese Medicine (TCM). All molecular structure files were optimized by force field MMFF94.

2.3 Docking results of molecules from DrugBank dataset.

To the DrugBank dataset, Table 1 lists the top 10 compounds with the highest binding energy. The specific screening results for each compound are given in supplementary.

The compound, digitoxin, is identified as having the strongest binding energy, –8.7kcal/mol at the binding site. Digitoxin is a cardiac glycoside. It has a known target, sodium/potassium-transporting ATPase subunit. Digitoxin can be used to treat patients with heart failure. Scientists have found that digitoxin also has capabilities to be a potential anticancer drug7,8. However, digitoxin can cause toxicity for the human body, like, nausea, anorexia, confusion and so on. Figure 2&3 shows the interaction of digitoxin with Spike protein.

Figure 4 & 5 represents the interaction between S protein and ACE2 and binding site of digitoxin on the complex.

2.4 Docking results of the TCMSP dataset.

The top 10 binding results of the TCMSP dataset are shown in Table 2. The complete binding information of TCMSP can be found in supplementary.

The compound having the highest binding energy in the TCMSP dataset with S protein is bisindigotin, which can be isolated from *Isatis indigotica* and *Polygoni Tinctorii Foliu.* Both herbs have effects of heat-clearing and detoxifying in the theories of Traditional Chinese Medicine. *Lsatis indigotica* is a folk medicine used to treat viral disease and inflammatory disease9. Besides, Wei et al. report bisindigotin can act as an antagonist to relieve the hepatoxicity caused by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, which is a carcinogen10. Figure 6&7 shows the interaction between Bisindigotin and RBD of S protein.

3 Discussion

The world is now confronted with a public health crisis caused by the SARS-CoV–2. The state-of-the-art techniques need to be applied to solve this pandemic. Normally, *de novo* drug development is a time-consuming project. Therefore, drug repurposing theories may be helpful to find the drug preventing coronavirus as soon as possible.

Here, we firstly chose spike protein rather than the ACE2 receptor for binding pocket search because ACE2 is expressed in various types of human cells and targeting ACE2 might cause more side effects. We observed the cavity neighboring the interface between viral spike protein and the human ACE2 receptor is the second pocket rank. Targeting this position may contribute to interrupt the interaction between S protein and the ACE2 receptor. Due to its high ranking and position advantages, we used this cavity for the next docking progress. Future work might focus on identifying a suitable cavity within the interface formed by viral spike protein and human receptor ACE2.

In this work, two databases, DrugBank containing FDA-approved drugs and TCMSP having TCM natural molecules, were used to explore potential candidates to prevent the SARS-CoV–2 from invading the human body. The docking results show that the molecules having the strongest binding energy from DrugBank and TCMSP are digitoxin and bisindigotin, respectively. Bisindigotin is a natural compound from TCM, *Isatis indigotica*. This is an herb used in folk to treat viral disease. It may have great potential to be served as a promising drug to treat COVID–19. More vigorous work remains to be done to test these molecules' effects.

Additionally, in the top 10 list of DrugBank docking results, it is worth to be noted that an HIV antiretroviral drug, raltegravir, is identified. This is a new class of HIV drugs, which can inhibit HIV integrase. On the other hand, in the top 10 list of TCMSP, in addition to bisindigotin, we observed that the majority of molecules exist in the herbs have effects on clearing heat-toxin, such as i*ndigo naturalis, isatidis radix, isatidis folium, stemonae radix, wikstroemiae indicae rasix, ginkgo semen, platycladi cacumen, forsythiae fructus, ranunculi ternati radix, canarii fructus, selaginella doederleinii hieron,* and *ginkgo folium.* According to theories of TCM, those herbs can be used to relieve symptoms of pneumonia.

SARS-CoV–2 is not stable as it is an RNA virus. Scientists have found that it may have mutations generating another subtype11. However, the critical proteins in the SARS- CoV–2 are still conservative. Currently, the possible mutations occurred will have a significant impact on our work. In the following research, we will investigate more molecules that rank below digitoxin and bisindigotin by binding energy. Moreover, *in vitro* tests will be carried out to verify the binding between S protein and the molecules identified by computational methods.

4 Methods And Materials

4.1 DrugBank

DrugBank is a drug data resource, which contained comprehensive drug information, including approved & experimental small-molecule data, biologics, and nutraceutical12. The latest version of DrugBank was released on Jan. 13 of 2020. It currently has 2,628 approved small drugs. Before performing virtual screening, those approved drugs with molecule weight larger than 500 like polypeptide were deleted, thus only 2,191 FDA- approved drugs were used to be screened.

4.2 TMCSP

TCMSP was developed by Northwest University in China, which is a platform providing information connecting natural molecules in TCM, targets, and diseases by a network. besides, this platform also provides related pharmacokinetic properties13. Based on this platform, in total 13026 small molecules from TCM herbs were used to be screened.

4.3 Software used

AutoDock vina 1.1.214 was used in this work to perform visual screening. PyMOL 2.3.3[15] and Discovery Studio 201616 were used to exhibit the ligand-receptor interactions.

Declarations

5. Acknowledgment

Great thanks to Dr. Jianxun Qi for working in solving the crystal structure of RBD- ACE2 complex. We have been approved by Dr. Jianxun Qi to use this crystal structure. In addition, we appreciate the support from Southern University of Science and Technology and Shenzhen University General Hospital

6. Ethics statement

This work was approved by the Institutional Review Board (IRB) of Southern University of Science and Technology.

7. Conflict declaration

All authors declare that there is no conflict of interest.

References

1. Zhou, *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, doi:10.1038/s41586-020-2012-7 (2020).
2. Wu, *et al.* A new coronavirus associated with human respiratory disease in China. *Nature*, doi:10.1038/s41586-020-2008-3 (2020).
3. Wang, *et al.* Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res*, doi:10.1038/s41422-020-0282-0 (2020).
4. Belouzard, , Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A* **106**, 5871- 5876, doi:10.1073/pnas.0809524106 (2009).
5. Zumla, A., Chan, J. , Azhar, E. I., Hui, D. S. & Yuen, K. Y. Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* **15**, 327-347, doi:10.1038/nrd.2015.37 (2016).
6. Wrapp, *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, doi:10.1126/science.abb2507 (2020).
7. Menger, *et al.* Trial watch: cardiac glycosides and cancer therapy. **2**, e23082 (2013).
8. Elbaz, H. A. *et al.* Digitoxin and its analogs as novel cancer **1**, 4 (2012).
9. Medicine, J. N. C. o. Dictionary of Traditional Chinese (1986).
10. Wei, Y. *et al.* Bisindigotin, a TCDD antagonist from the Chinese medicinal herb Isatis indigotica. **68**, 427-429 (2005).

11. Tang, *et al.* On the origin and continuing evolution of SARS-CoV-2. (2020).
12. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. **46**, D1074-D1082 (2018).
13. Ru, J. *et al.* TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. **6**, 13 (2014).
14. Trott, O. & Olson, A. J. J. J. o. c. c. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **31**, 455-461 (2010).
15. DeLano, L. J. C. N. o. p. c. Pymol: An open-source molecular graphics tool. **40**, 82-92 (2002).
16. Systèmes, D. J. D. S. B. S. D., CA, USA. BIOVIA, Discovery Studio Modeling Environment. (2016).

Tables

Due to technical limitations, Tables 1-2 are provided in the Supplementary Files section.

Figures

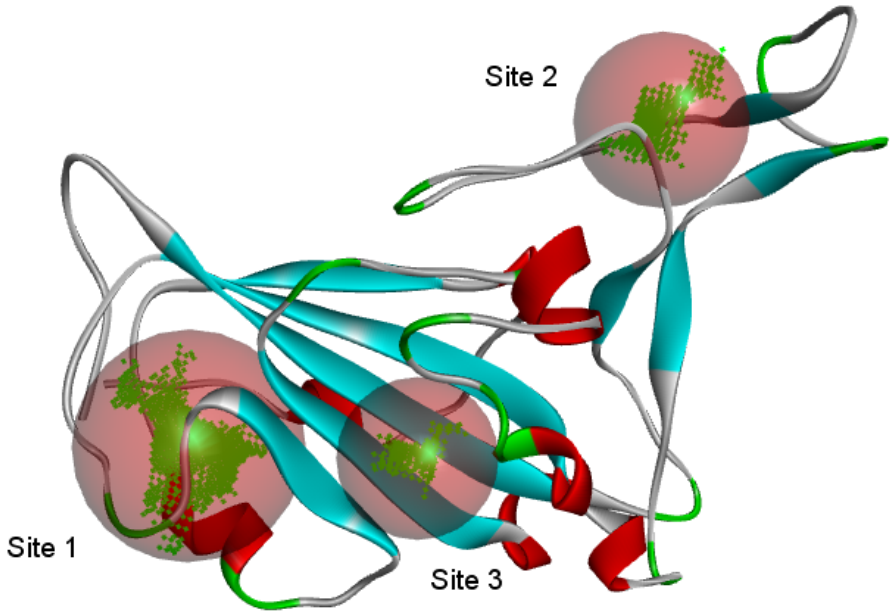

Figure 1

Top 3 pockets predicted by Discovery Studio. The pockets are labeled according to their rank.

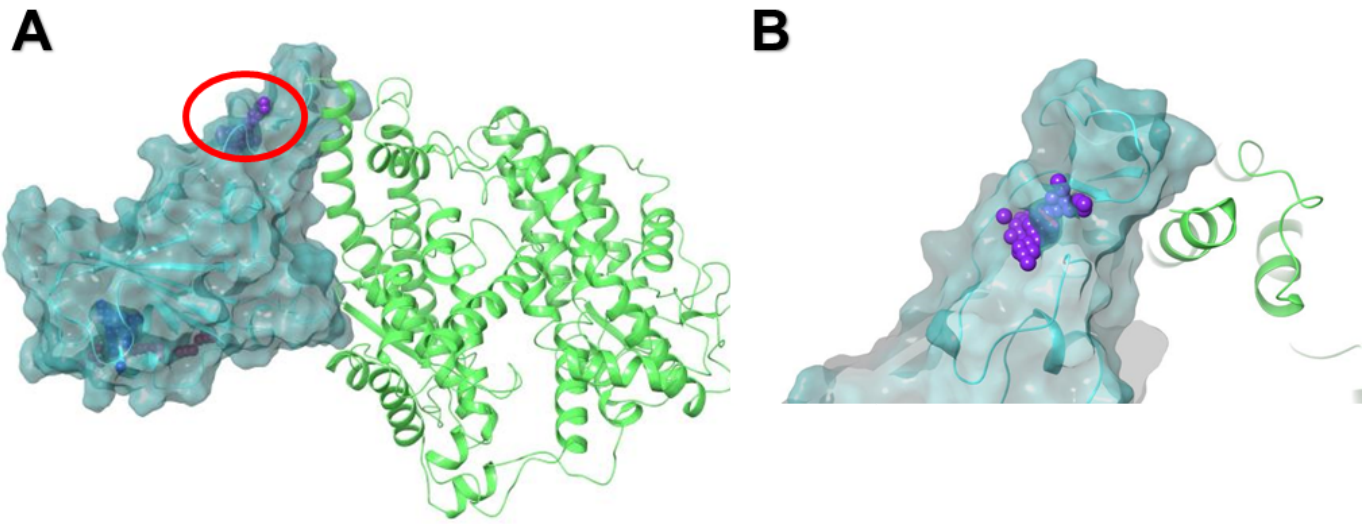

Figure 2

(A) the position circled by the red line is the pocket selected. The protein colored by green is ACE2; (B) The position consists of purple balls is the pocket selected.

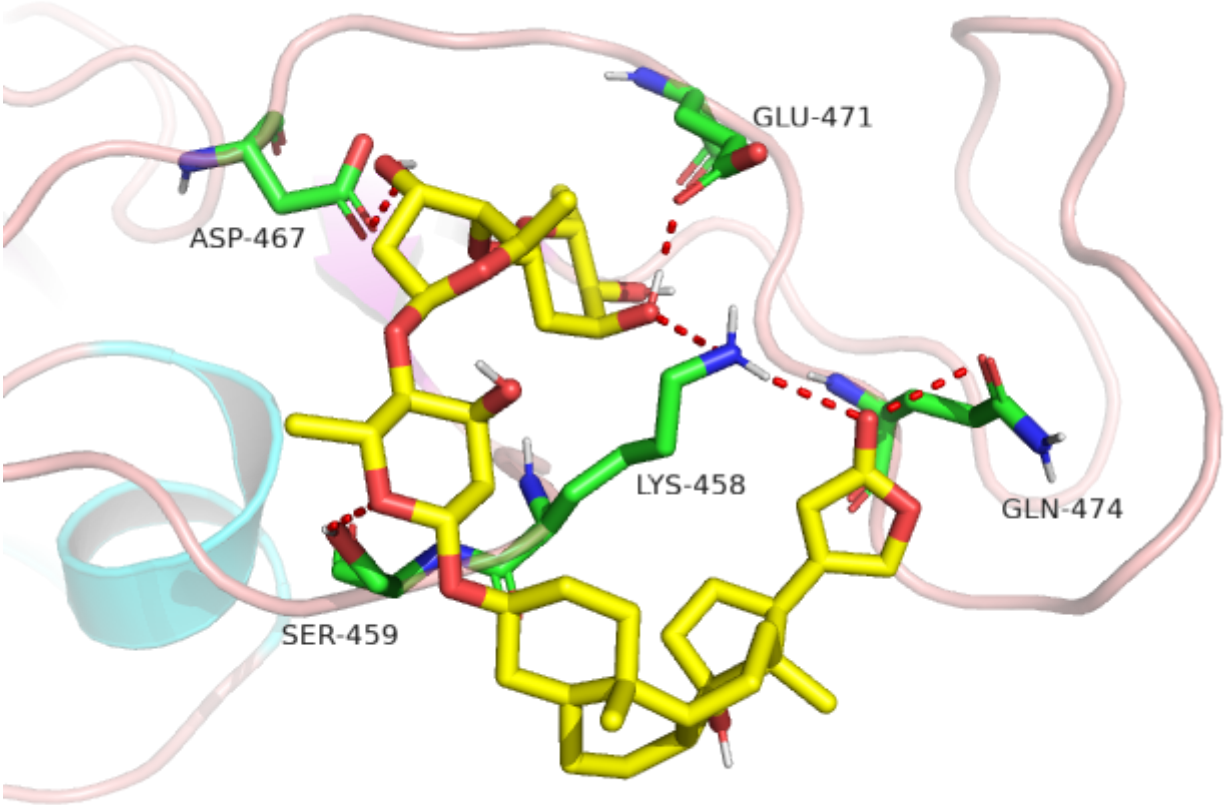

Figure 3

3D interaction between Digitoxin (yellow) and RBD of S protein (green sticks, pink loops, and blue helix). The red dotted line represents polar contacts.

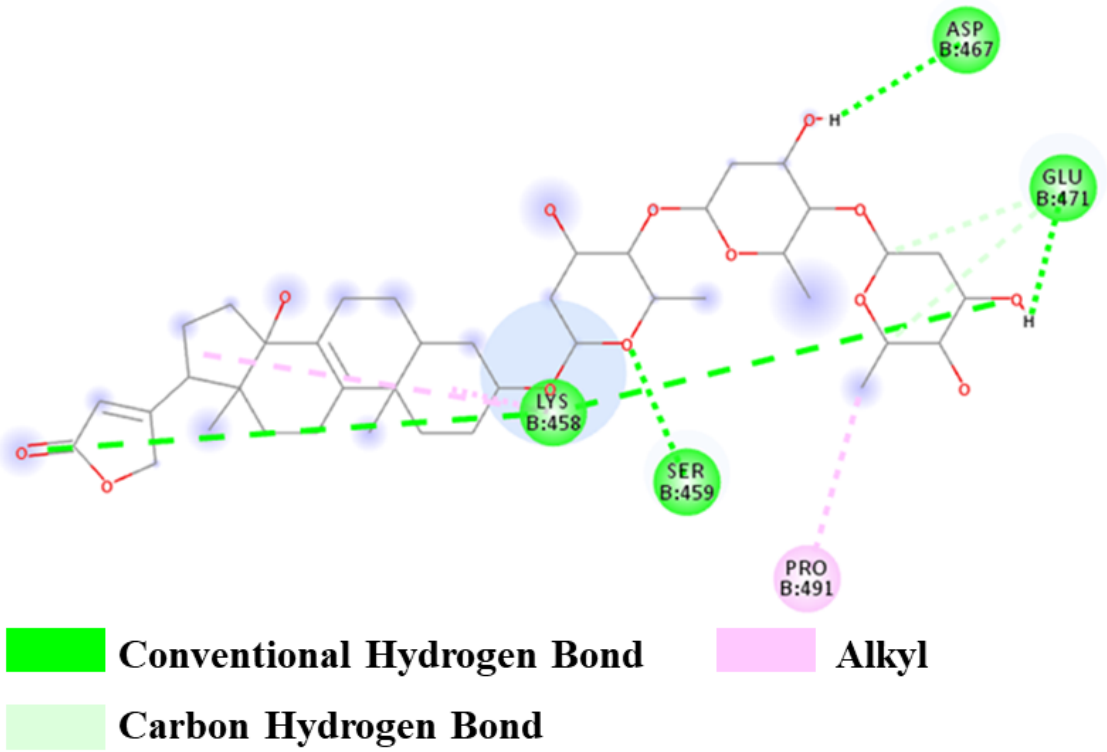

Figure 4

2D interaction between Digitoxin and RBD of S protein

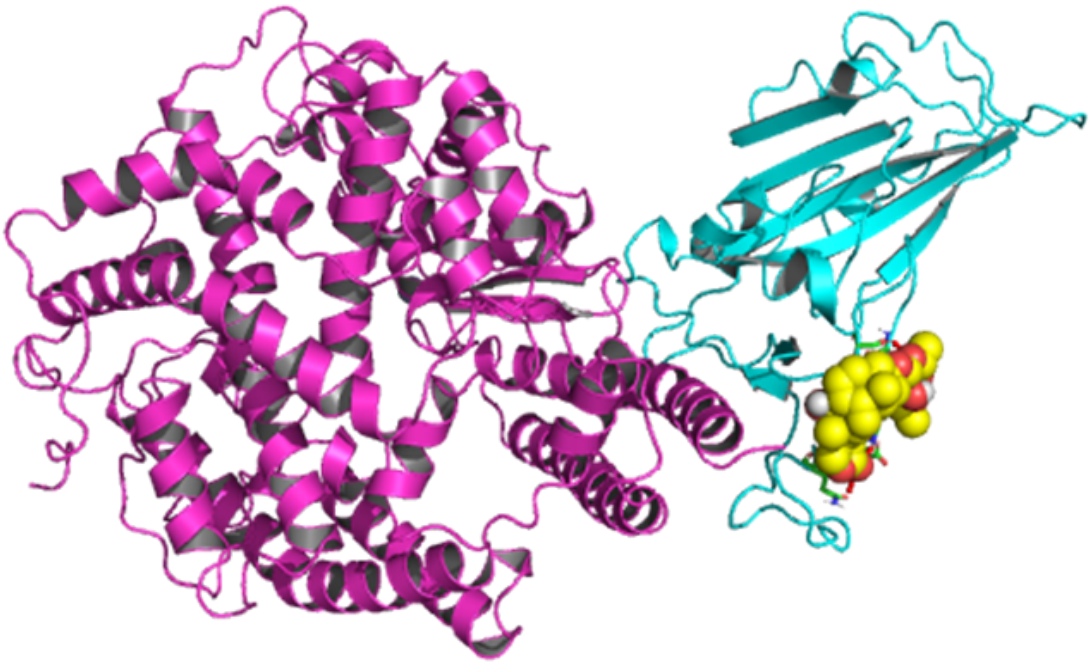

Figure 5

Digitoxin on the interface between RBD and ACE2. The protein colored by magenta is ACE2; The blue part is the S1 subunit of S protein. The balls colored by yellow, red and white is digitoxin.

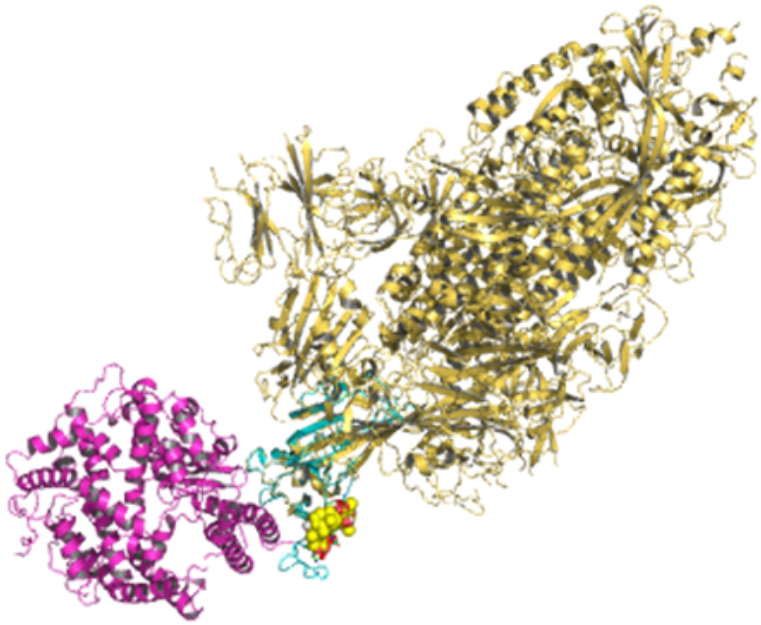

Figure 6

Digitoxin on the interface between S protein and ACE2. The protein colored by magenta is ACE2; The blue part is the S1 subunit of S protein. The balls colored by yellow, red and white is digitoxin. The trimer S protein (golden) is superposed to the RBD-ACE2 complex.

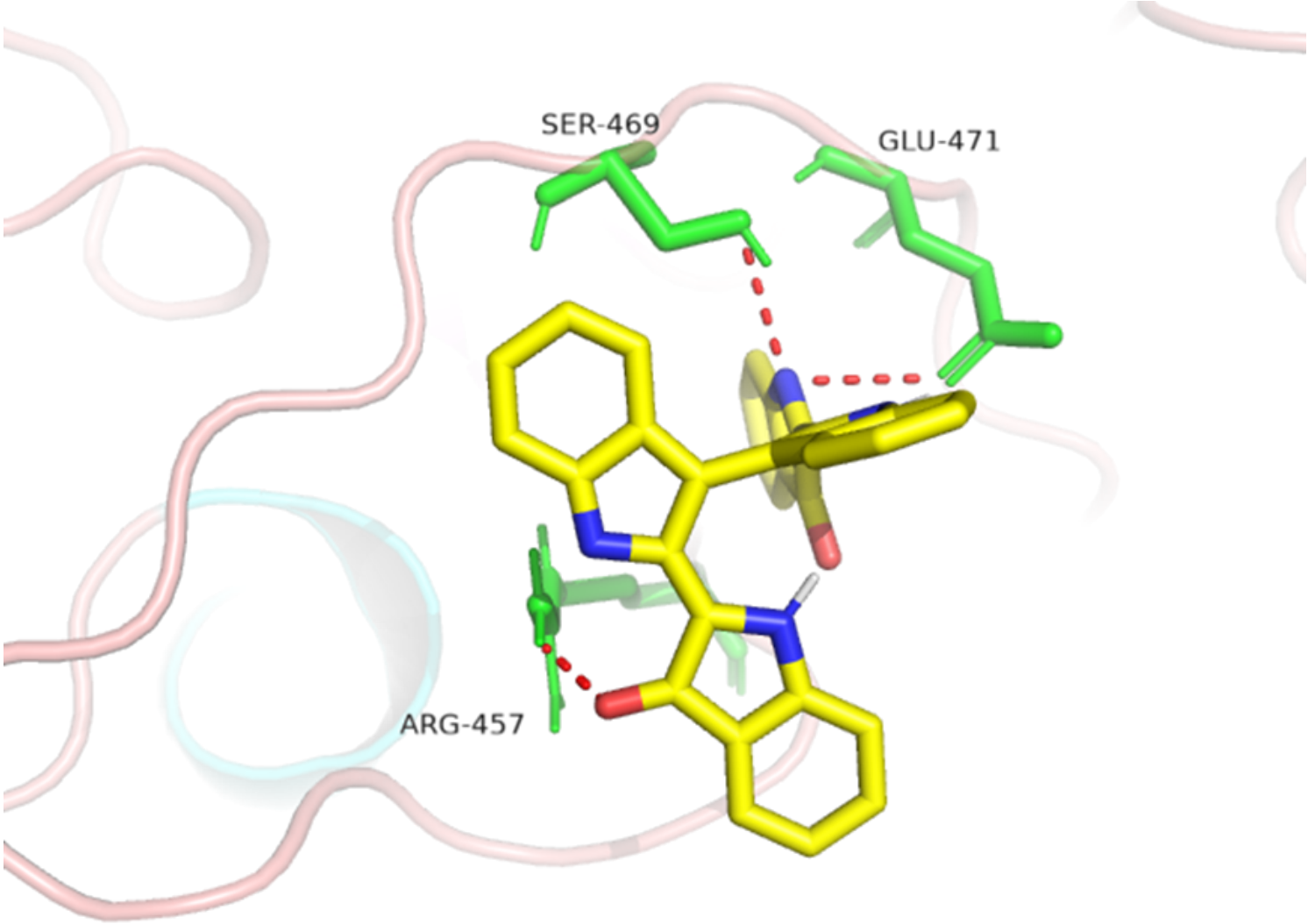

Figure 7

3D interaction between bisindigotin (yellow) and RBD of S protein (green sticks, pink loops, and blue helix). The red dotted line represents polar contacts.

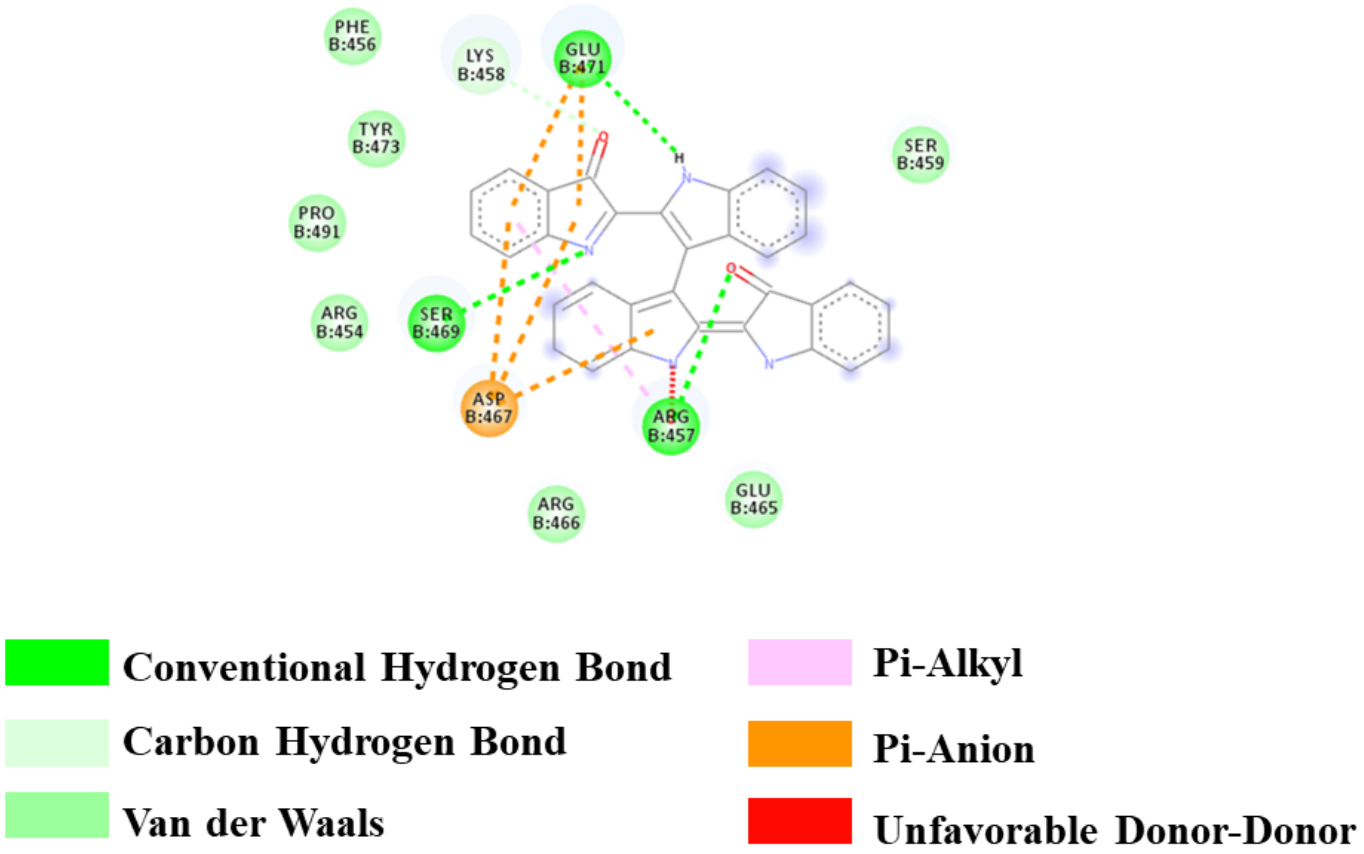

Figure 8

2D interaction between bisindigotin and RBD of S protein

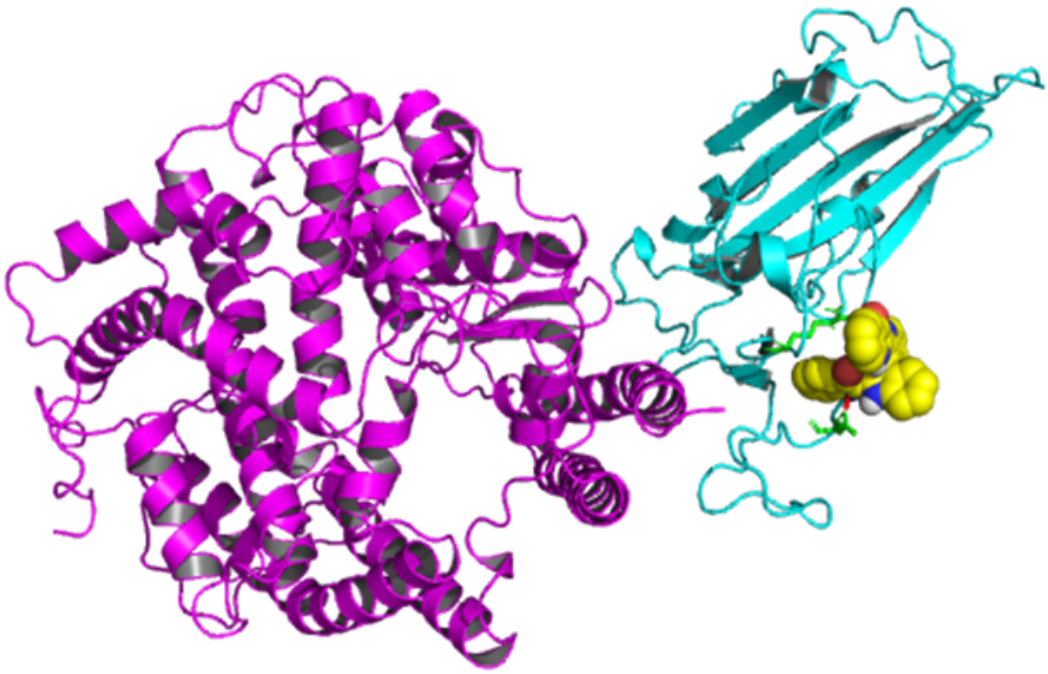

Figure 9

Bisindigotin on the interface between RBD of S protein and ACE2. The blue part is the S1 subunit of S protein. The balls colored by yellow, red and white are bisindigotin; The protein colored by magenta is ACE2.

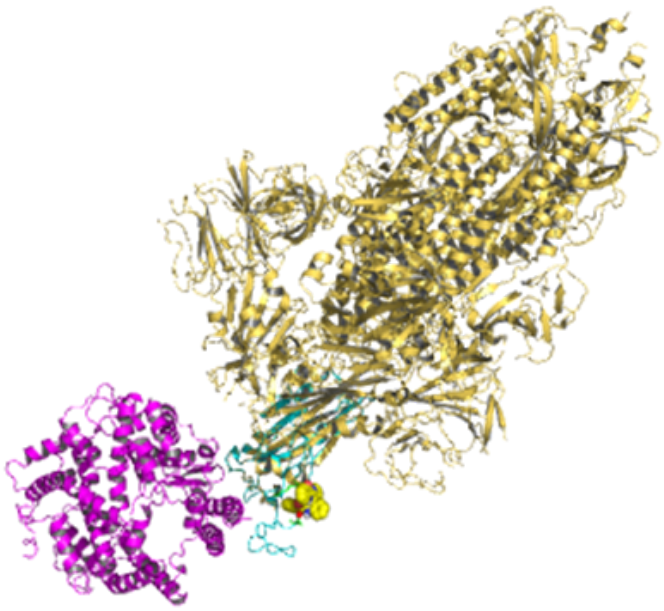

Figure 10

Bisindigotin on the interface between S protein and ACE2. The protein colored by magenta is ACE2; The blue part is the S1 subunit of S protein. The balls colored by yellow, red and white is digitoxin. The trimer S protein (golden) is superposed to the RBD-ACE2 complex.

Supplementary Files

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

- supplementarydata.rar
- table1.png
- table2.png
- supplementalmaterial.zip