2nd Bioinformatics and Biodiversity Conference
Volume 2022
http://dx.doi.org/10.11594/nstp.2022.2107

NST PROCEEDINGS

OPEN ACCESS

Conference Paper

In Silico Design and Validation of CRISPR-Cas13a System as a Potential Antiviral for SARS-CoV-2 in Indonesia

Alfero Putra Iryanto[1], Christy[1], Muhammad Farrel Ewaldo[2], Anggia Prasetyoputri[3,4], Ratih Asmana Ningrum[3,4], Riza Arief Putranto[5,6]*, Akhirta Atikana[3,4]

[1]Department of Biotechnology, Esa Unggul University
[2]Department of Biology, Diponegoro University
[3]Research Center for Biotechnology
[4]Research Organization of Life Sciences, National Agency for Research and Innovation (BRIN)
[5]Indonesian Research Institute for Biotechnology and Bioindustry
[6]Department of Biotechnology, Esa Unggul University

Corresponding author:
E-mail: riza.arief@esaunggul.ac.id

ABSTRACT

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a worldwide pandemic of coronavirus disease (COVID-19). Indonesia is one of the countries with large numbers of positive cases in Asia with certain dominant variants. Currently, there are no specific therapeutic agents against SARS-CoV-2. Therefore, the development of specific and effective therapeutic tools is urgently needed to overcome the pandemic. This study designed a CRISPR-Cas13a system strategy as a potential anti-SARS-CoV-2. We utilized comprehensive bioinformatics methods to identify a unique segment in the SARS-CoV-2 consensus sequence from Indonesia that is different from the related segment in the SARS-CoV. This unique segment was used as a specific target for SARS-CoV-2 Spike Protein to design a set of crRNA libraries. Off-target analysis and molecular docking simulation were performed to validate the specificity and to analyze interactions among the crRNA candidates, target RNA, and Cas13a. Our study identified a 17 amino acid unique segment on the Receptor Binding Domain (RBD) region. By using that unique segment, a total of 12 crRNA candidates were selected based on their GC content. Finally, based on the off-target and molecular docking validation, four crRNAs were selected as potential candidates for CRISPR-Cas13a-based antivirals. Although further validation with *in vitro* assays is important, the present study provides a comprehensive demonstration regarding the potential of CRISPR-Cas13a as a strategy for SARS-CoV-2 antiviral development. Considering the specific property of the CRISPR system, the present methodology can also be utilized to develop novel antiviral candidates for other RNA viruses.

Keywords: SARS-CoV-2, CRISPR-Cas13a, antiviral, in-silico, molecular docking

Introduction

Coronavirus disease 2019 (COVID-19) is a disease caused by SARS-CoV-2 that emerged in early December 2019, in Wuhan, China. The nature of the coronavirus is easy to infect humans and easily spreads almost all over the world. The virus is easily spread to every area, increasing the number of victims suffering from the COVID-19 and causing the outbreak of COVID-19 (pandemic).

Currently, no effective drugs are clinically approved. It is therefore pivotal to find strategies to prevent the virus infection, in particular the severity and fatality associated with it. The interaction that

How to cite:
Iryanto et al. (2022). In Silico design and validation of CRISPR-Cas13a system as a potential antiviral for SARS-CoV-2 in Indonesia. 2nd *Bioinformatics and Biodiversity Conference.* NST Proceedings. pages 44-51. doi: 10.11594/ nstp.2022.2107

occurs between the spike protein possessed by SARS-CoV-2 and its receptors on host cells in the human body is needed to initiate the infection mechanism of the COVID-19 disease. The SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2) as its cellular receptor, as do SARS-CoV and the coronavirus NL63 (Zhang et al., 2021).

An effector ribonucleic acid RNA-guided clustered regularly interspaced short palindromic repeats (CRISPR) that target and cleave single-stranded RNA (ssRNA), named LwaCas13a, is gaining attention as a therapeutic and diagnostic tool in viral diseases and cancer (Fareh et al., 2021). This system consists of two components: RNA-guided RNase Cas13a and CRISPR RNA (crRNA) as the single targeting guide RNA. Unlike the others, Cas13a exerts a collateral cleavage effect, meaning that after recognizing its target RNA, activated Cas13a cleaves not only target RNA, but also surrounding non-target RNA. The application of Cas13a has been reported in RNA knockdown, RNA detection, and RNA capture (Aman et al., 2018). Based on the collateral cleavage mechanism, CRISPR/Cas-13a has also been used in SARS-CoV-2 diagnostics (Ganbaatar & Liu, 2021). Besides that, the CRISPR-Cas13a system has also been conducted to predict SARS-CoV-2 antiviral and it showed that the system has worked well in vitro (Wang et al., 2021). However, the study designed the crRNA system based on wild type sequence, which probably caused a mismatch in current dominant SARS-CoV-2 variants.

The present study utilized comprehensive bioinformatics methods to identify a unique region in the SARS-CoV-2 consensus sequence, specifically from Indonesia. This region is different from the related region in the SARS-CoV genome and can be used as a specific target for SARS-CoV-2 Spike Protein. In addition to that, a set of crRNA libraries was also designed based on this system. Off-target analysis and molecular docking simulation were performed to validate the specificity and to analyze the interactions among the crRNA candidates, target RNA, and Cas13a. Considering the specific property of the CRISPR system, the methodology proposed in this study can also be utilized to develop novel antiviral candidates for other types of RNA viruses.

Material and Methods

Selection of the target gene

Sequences from the high coverage viral genome database were obtained from the GISAID (https://www.gisaid.org/) in July – August 2021. A total of 3200 genomes originating from Indonesia were used in this study. A multiple sequence alignment was performed on these sequences using the MAFFT version 7 (Katoh et al., 2019) to obtain a consensus sequence with a threshold above 95%. The S gene sequence (Spike) was obtained from the consensus sequence and translated into protein. A pairwise alignment of that protein was performed with the Receptor Binding Domain (RBD) sequence to look for RBD in the obtained sequence. Alignment with SARS-CoV-1 was also performed to find the differences between SARS-CoV-1 and SARS-CoV-2 sequences. Finally, the superimposition of this system

Target gene validation

We used molecular docking to identify residues that play an important role in the interaction of RBD with ACE-2 in humans. Homology modeling was carried out between protein sequences obtained from the consensus compared to the databases using the Swiss Model (Waterhouse et al., 2018). Molecular Docking was performed using Cluspro 2.0 server (Kozakov et al., 2017) and the protein-protein interaction was checked using Ligplot version 4.5.3 (Wallace et al., 1996).

crRNA Design

The nucleotide sequences of protein sequences obtained in the present study were confirmed using SnapGene version 5.3.1 (from Insightful Science; available at snapgene.com) and translated into RNA sequence. This sequence was uploaded to the CRISPR-RT server to design the crRNA candidates (Zhu et al., 2017). The parameters for this design are Organism (SARS-CoV-2), PFS (default), and Off-

target (select specific default settings). The candidates with a GC content above 0.5 and the position in the front (5') region of the target gene were selected for further validation. Up to 12 candidates were selected in the present study. These candidates were validated using BLAST, to confirm their specificity.

In Silico Validation

The design of crRNA was analyzed using molecular docking, in the Hdock server (Yan et al., 2020). A docking score below the median value from all candidates was recorded to see which residues play an important role in the interaction of the CRISPR complex with the target RNA (Wang et al., 2021). In addition to that, the off-target effect was analyzed by performing BLAST for the selected crRNA candidates with the human genome and other coronaviruses genomes.

Results and Discussion

CRISPR target region

The RBD in S gene is a common target to design neutralizing antibodies, inhibitors, and vaccines (Tai et al., 2020). However, drugs targeted ssRNA-viruses were reported to have constrained effects. Therefore, approaches that specifically target RNA viruses are important. In this study, we designed a pipeline that allows rapid identification of a specific site in the RNA virus target, thus can be used as a screening tool to select the best candidates for crRNA sequence.

The current study focuses on the design and development of a customized antiviral approach that targets the SARS-CoV-2 virus. The current work focuses on the S RBD section of the SARS-CoV-2 genome because of its importance in the viral invasion of human cells. This study described a highly specific region inside the S RBD of SARS-CoV-2, as well as sequence alignment and structural evaluation of the RBDs of SARS-CoV-2 and its close relatives, SARS-CoV. SARS-CoV-2 has a 51-bp region (17aa) that is unique to it (Figure. 1).

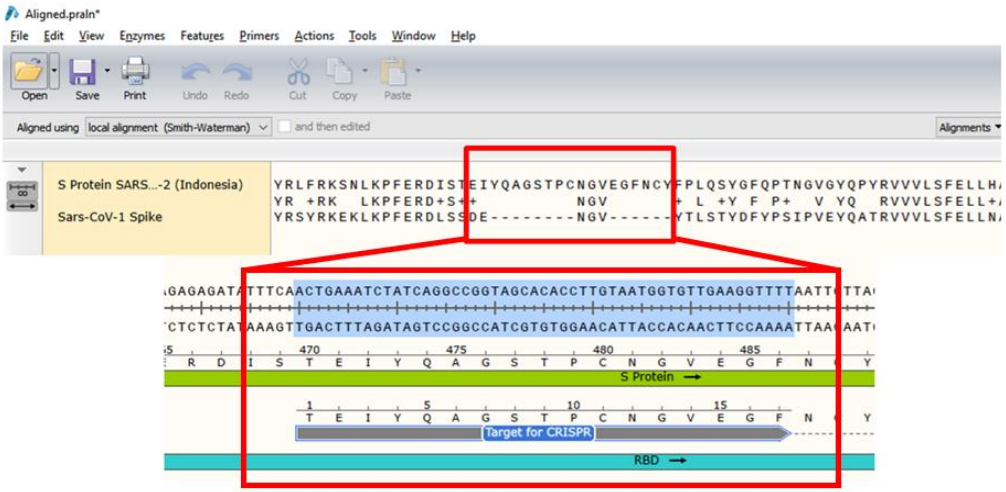

Figure 1. Unique region in the RBD from SARS-CoV-2 consensus sequence (red box), as the proposed target for the antiviral CRISPR system

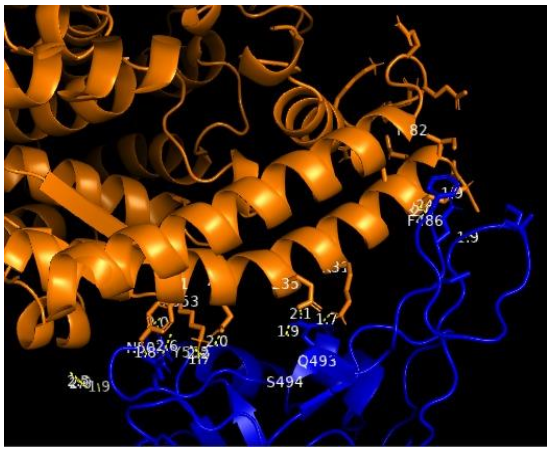

Figure 2. Protein-protein interaction between ACE2 (orange) and RBD (blue) from 3D modeling with the Indonesian consensus sequences of SARS-CoV-2. The label indicates a strong protein-protein interaction

The 3D structure model validation, using the RBD from protein modeling, showed the protein interaction with human ACE protein (Figure. 2). Several residues with strong bonding with each other were observed in this study, some of them are Serine494, Glutanime493, and Phenylalanine486. This result indicates that our proposed region has an important role in the viral infection process. Therefore, it could be a potential target to design an antiviral system.

A recent study reported a rapid CRISPR-based diagnostic for COVID-19 as a novel diagnostic option for this new emerging virus. The CRISPR-based detection and diagnostic systems are not only limited to the COVID-19 but also applied for any types of pathogens, viruses, and fungi (Ganbataar & Liu 2021). The present study offers a more advanced use of the CRISPR-based system as an antiviral, specifically for COVID-19.

crRNA In Silico design and validation

In this investigation, a crRNA library was chosen as the target for developing SARS-CoV-2 antivirals. These crRNAs were created as a tiling library that covered the unique region's RNA sequence. In this work, the 12 crRNA sequences with the greatest GC contents (0.5) were chosen as candidates for further binding simulation (Table.1). This GC concentration is linked to the stability of double-stranded DNA or RNA. As a result, the approach provided in this work was proposed as a useful methodology that may be used without having to evaluate a huge number of probable crRNA sequences. However, gene annotation and mutation spot analysis revealed that these candidates cover mutations S477N, T478K, and E484K found in SARS-CoV-2 genome Delta variants. (Figure 3). This is probably due to the Delta variants being the most dominant variant worldwide, including in Indonesia.

Table 1. List of crRNA candidates

No	Protospacer+PFS	gRNA_start	gRNA_end	GC	num_targets	num_genes
1	ACUGAAAUCUAUCAGGCCGGUAGCACACC	1	29	0.5	1	1
2	CUGAAAUCUAUCAGGCCGGUAGCACACCU	2	30	0.54	1	1
3	UGAAAUCUAUCAGGCCGGUAGCACACCUU	3	31	0.5	1	1

To be continued….

4	AAAUCUAUCAGGCCGGUAG-CACACCUUGU	5	33	0.5	1	1
5	AAUCUAUCAGGCCGGUAG-CACACCUUGUA	6	34	0.5	1	1
6	AUCUAUCAGGCCGGUAG-CACACCUUGUAA	7	35	0.5	1	1
7	UCUAUCAGGCCGGUAG-CACACCUUGUAAU	8	36	0.5	1	1
8	AUCAGGCCGGUAGCACAC-CUUGUAAUGGU	11	39	0.54	1	1
9	CAGGCCGGUAGCACAC-CUUGUAAUGGUGU	13	41	0.57	1	1
10	AGGCCGGUAGCACACCUUGU-AAUGGUGUU	14	42	0.54	1	1
11	GCCGGUAGCACACCUUGU-AAUGGUGUUGA	16	44	0.54	1	1
12	CCGGUAGCACACCUUGU-AAUGGUGUUGAA	17	45	0.5	1	1

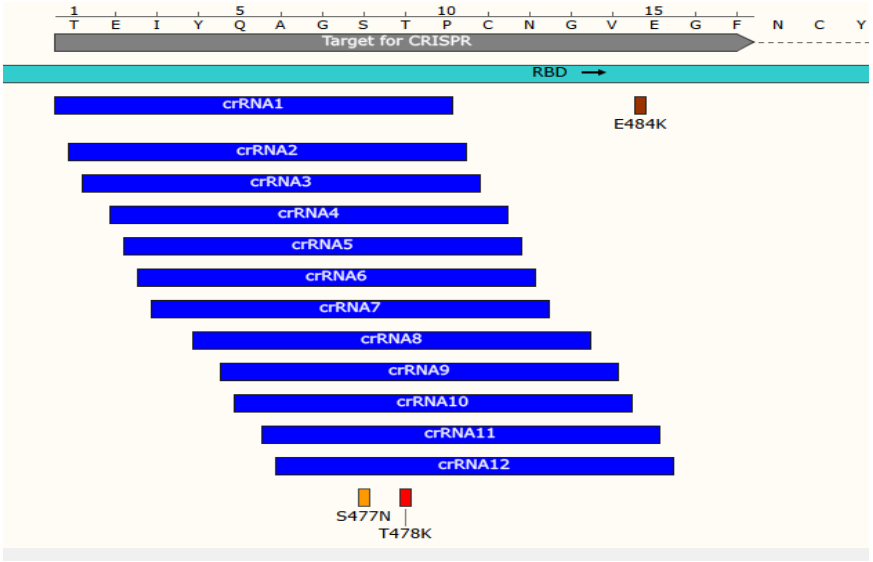

Figure 3. Position of crRNA candidates on RBD target region. The mutation of Delta variants is indicated as S477N, T478K, and E484K

The binding energy analysis compared the interactions of different crRNA sequences with the target RNA and Cas13a. Based on this analysis, 4 (four) candidates with the docking score below-median were selected as targets for antivirals, i.e., crRNA-5, crRNA-6, crRNA-9, and crRNA-12 (Figure 4). The specificity of these crRNA sequences showed a match between the crRNA sequences and SARS-CoV-2 S RNA, indicating crRNA candidates are specific for SARS-CoV-2. The off-target analysis against SARS-CoV-1 and human genome showed that the query coverage of these candidates is low, indicating that the crRNA will not be able to target SARS-CoV-1 and human genomes.

The present study observed that the crRNA-6 has the lowest docking score (Figure 4). This indicates that the crRNA-6 has a stronger bond with the Cas13a. In addition to that, the present study also observed the interaction between crRNA6, Cas13a, and target RNA virus (Figure 5). The docking

score of this system is negatively correlated with the interaction strength, which means interaction between the Cas13a and the target RNA requires less energy to interact, therefore, indicating good stability of this system.

This study identified a novel area in the Indonesian SARS-CoV-2 consensus sequence. Because the sequences provided in this work differed from those found in SARS-CoV-1, they can be exploited to develop a CRISPR-based antiviral that is specific for SARS-CoV-2. The collateral cleavage effect mechanism is extensively used in CRISPR-Cas13a applications. The creation of SHERLOCK, DETECTR, and SHINE as CRISPR-based methods to identify the SARS-CoV-2 infection is among the most popular (Aquino-Jarquin, 2021). The current study provided a method that has the potential to create COVID-19 medicines to support this CRISPR system. Instead of suppressing a specific viral activity, as existing antiviral medications do, the suggested method precisely targets the SARS-CoV-2 genome (Riva et al., 2020). The methodology proposed in this study not only contributes to genomic surveillance of the SARS-CoV-2 but can also be utilized to develop novel antiviral candidates for the other RNA viruses as a therapeutic system.

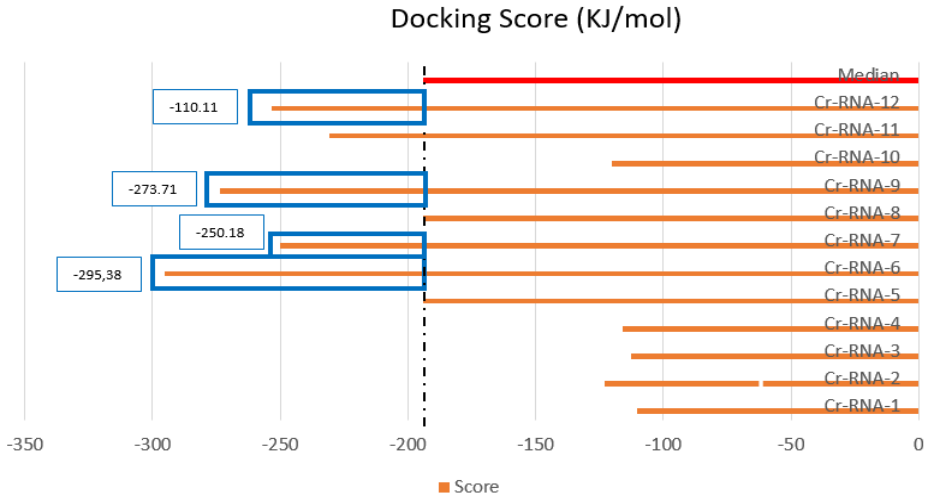

Figure 4. Docking score analysis of crRNA Candidates. The lowest score was observed in the crRNA6.

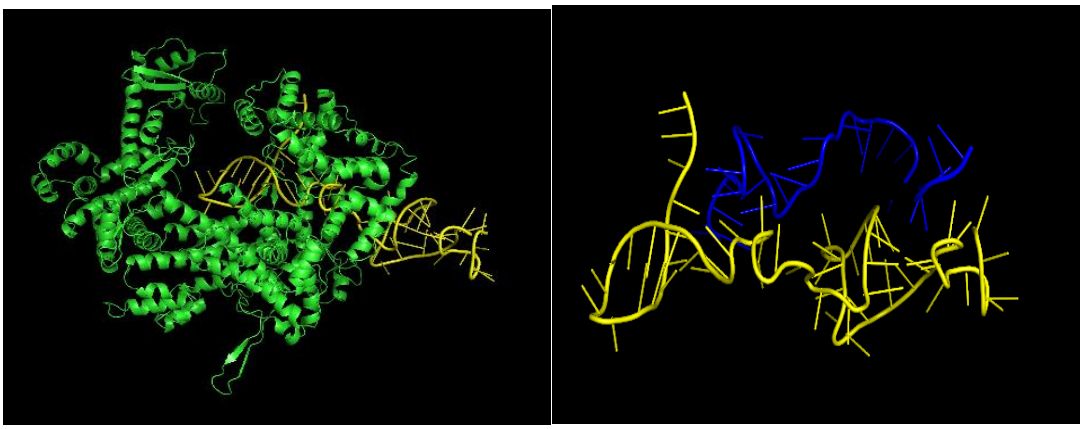

Figure 5. Protein-protein interaction between the crRNA-6 and the Cas13a (left) and the interaction between the crRNA-6 and the target RNA SARS-CoV-2 (right). The crRNA-6 was shown in yellow, the Cas13a was in green, and the target RNA virus was in blue.

Similar studies related to CRISPR-Cas13a based antivirals have also been conducted, such as prophylactic antiviral CRISPR in human cells (PAC-MAN), which mechanism is degrading viral RNA at the conserved region (Abbott et al., 2020). Another related CRISPR strategy also designed a specific segment targeting the SARS-CoV-2 genome (Wang et al., 2021). This study showed an in vitro test in which a candidate of crRNA interacted with the gene target. However, both studies were using a wild-type sequence so they might potentially have made a mismatch with another SARS-CoV-2 variant.

The present study proposed a crRNA candidate that is specific for SARS-CoV-2 based on the in-silico test. Therefore, it can be used as a candidate for SARS-CoV-2 antivirals. However, to validate this design needs further in vitro analysis, for example by applying this design to interact with a plasmid construct that contains the Spike genes. The in vitro analysis is important as a validation test of the CRISPR system as proposed in this study. Before this validation, it is also important to have specific primer pairs that cover the specific CRISPR region as proposed in this study. An ongoing study is conducted to design and develop primers for the in vitro validation of this proposed CRISPR system.

Conclusion

The present study highlights the pipeline development using an in-silico system to design SARS-CoV-2 antivirals based on the CRISPR-Cas13a system, targeting a specific region on RBD that is considered as a target for the antivirals. This study proposed a total of 12 candidates, in which 4 (four) were considered as potential candidates for target antivirals. Among all, the crRNA-6 has the lowest docking score, therefore selected as the most potent antiviral candidate targeting the SARS-CoV-2. The potential of crRNA-6 was confirmed by the RNA-RNA interaction analysis that showed good interaction with the target viral RNA. The approach proposed in this study can be further developed as the COVID-19 therapy. However, in vitro validation and further clinical trials are important to confirm the proposed CRISPR-based antiviral system.

Acknowledgment

This study was financially supported by the LPDP Research Mandatory COVID-19 (No. 8/FI/P.KCOVID-19.2B3/X/2020) coordinated by Akhirta Atikana. This study is also part of the research collaboration between the Research Center for Biotechnology, BRIN, and the Faculty of Health Sciences, Esa Unggul University. The support from the Deputy of Life Sciences, BRIN and Dekan Fakultas Ilmu-ilmu Kesehatan, Esa Unggul University, are greatly appreciated.

References

Abbott, T. R., Dhamdhere, G., Liu, Y., Lin, X., Goudy, L., Zeng, L., Chemparathy, A., Chmura, S., Heaton, N. S., Debs, R., & Pande, T. (2020). Development of CRISPR as an antiviral strategy to combat SARS-CoV-2 and influenza. *Cell*, *181*(4), 865-876. doi: 10.1016/j.cell.2020.04.020.

Aman, R., Ali, Z., Butt, H., Mahas, A., Aljedaani, F., Khan, M.Z., Ding, S., & Mahfouz, M. (2018). RNA virus interference via CRISPR/Cas13a system in plants. *Genome biology, 19*(1), 1-9.

Aquino-Jarquin, G. (2021) Recent progress on rapid SARS-CoV-2/COVID-19 detection by CRISPR-Cas13-based platforms. *Drug Discov Today, 26*(8), 2025-2035

Fareh, M., Zhao, W., Hu, W., Casan, J. M., Kumar, A., Symons, J., Zerbato, J. M., Fong, D., Voskoboinik, I., Ekert, P. G., & Rudraraju, R. (2021). Reprogrammed CRISPR-Cas13b suppresses SARS-CoV-2 replication and circumvents its mutational escape through mismatch tolerance. *Nature communications*, *12*(1), 1-16. doi: 10.1038/s41467-021-24577-9.

Ganbaatar, U., & Liu, C. (2021) CRISPR-Based COVID-19 testing: Toward next-generation point-of-care diagnostics. *Frontiers in cellular and infection microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.663949

Katoh, Rozewicki, & Yamada. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics, 20,* 1160-1166. doi: 10.1093/bib/bbx108.

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., & Vajda, S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols, 12*(2), 255-278. doi: 10.1038/nprot.2016.169.

Riva, L., Yuan, S., & Yin, X. (2019). Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature, 586*, 113–119.

Tai, W., Zhang, X., He, Y., Jiang, S., & Du, L. (2020) Identification of SARS-CoV RBD-targeting monoclonal antibodies with cross-reactive or neutralizing activity against SARS-CoV-2. *Antiviral Res., 179,* 104820

Wallace, A. C., Laskowski, R. A., & Thornton, J. M. (1996). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng., 8,* 127-134.

Wang, L., Zhou, J., Wang, Q., Wang, Y., & Kang, C. (2021). Rapid design and development of CRISPR-Cas13a targeting SARS-CoV-2 spike protein. *Theranostics, 11*(2), 649.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes*. Nucleic Acids Res., 46*(W1), W296-W303.

Yan, Y., Tao, H., He, J., & Huang, S-Y. (2020) The HDOCK server for integrated protein-protein docking. *Nature Protocols, 15*(5), 1829-1852. doi: https://doi.org/10.1038/s41596-020-0312-x.

Zhu, H., Richmond, E., & Liang, C. (2018) CRISPR-RT: a web application for designing CRISPR-C2c2 crRNA with improved target specificity. *Bioinformatics*, *34*(1), 117-119.

Zhang, Q., Xiang, R., Huo, S., Zhou, Y., Jiang, S., Wang, Q., & Yu, F., (2021) Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduction and Targeted Therapy, 6*(1), 1-19.