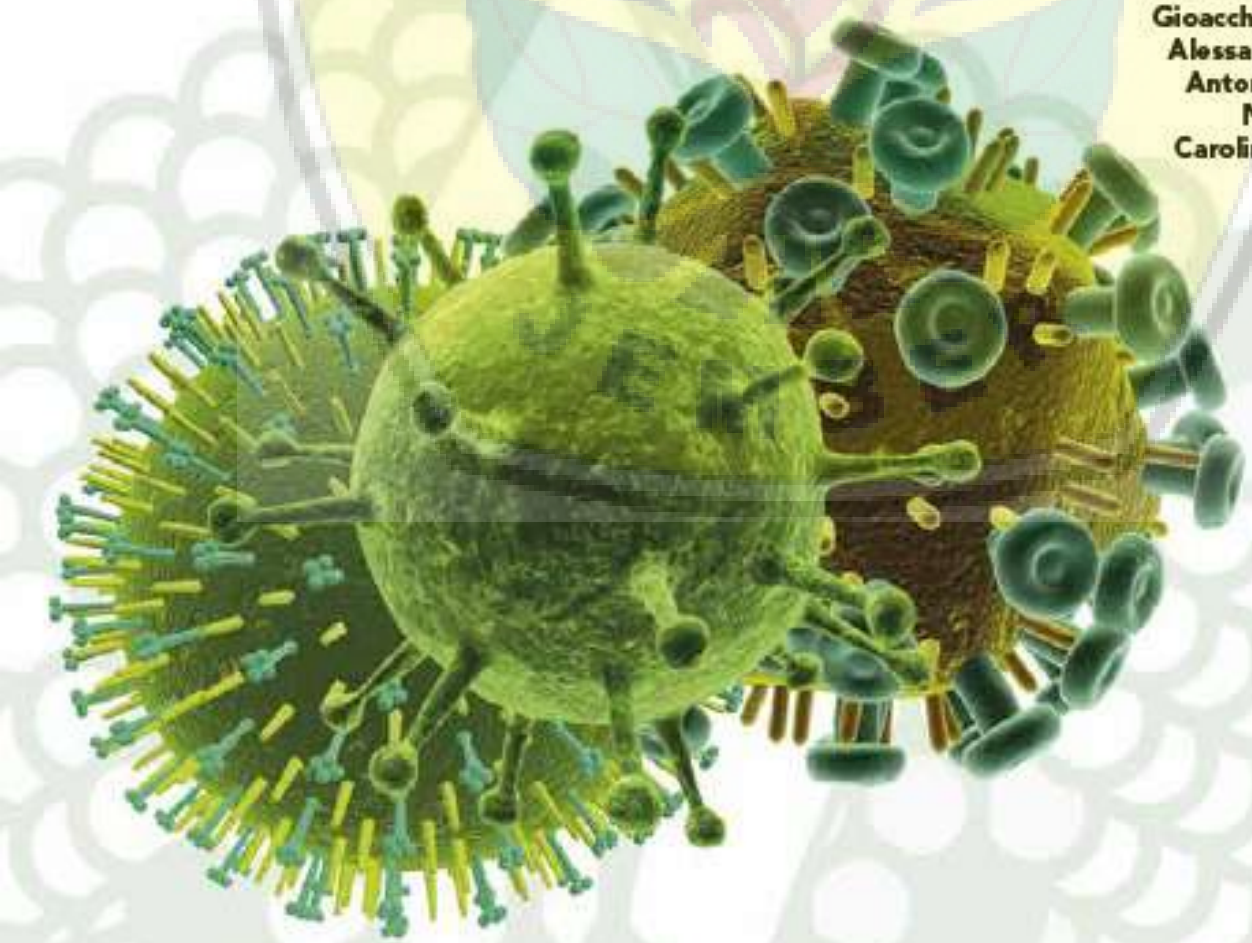
CODEN: MIBLDR 45 (3) 149-227 (2022)
ISSN 1121 - 7138

New microbiologica

quarterly journal
of basic and clinical
microbiological sciences

www.newmicrobiologica.org
vol. 45, no. 3, July 2022



EDIZIONI INTERNAZIONALI DI
EDIMES
Edizioni Mediche Scientifiche - Pavia

Journal Manager
Maria Carla RE

Publisher
Edizioni Internazionali srl
Div. EdiMeS Edizioni Medico-Scientifiche – Pavia
Via Riviera, 39 – 27100 Pavia -Italy
Phone: (+39) 0382-526253 r.a.
Fax: (+39) 0382-423120
e-mail: edint.edimes@tin.it

TABLE OF CONTENTS

FULL PAPERS

New Microbiologica, 45, 3, 173-180, 2022, ISSN 1121-7138

FULL PAPER

Construction of SHERLOCK-based sgRNA for SARS-CoV-2 Diagnostics from Indonesia

Erlia Narulita[1,2], Agung Haris Widianto[2], Syubbanul Wathon[3]

[1]*Center for Development of Advanced Science and Technology, University of Jember, Jember 68121, Indonesia;*
[2]*Department of Biology Education, University of Jember, Jember 68121, Indonesia;*
[3]*Department of Biology, University of Jember, Jember 68121, Indonesia*

SUMMARY

Handling a pandemic requires high sensitivity, high specificity, simple, fast, and flexible tests. However, conventional test methods (RT-PCR and Rapid Antigen) have weaknesses in test efficiency. Specific High sensitivity Enzymatic Reporter un-LOCKing (SHERLOCK), is a new technology that can detect nucleic acids even with limited sample preparation, but with high sensitivity, high specificity, rapidly, and flexibly. The key to the specificity of the SHERLOCK diagnostic method is the single guide RNA (sgRNA). The purpose of this study was to analyze the design of the SHERLOCK sgRNA, which has optimum potential to be used as a Cas13a marker to recognize the spike protein gene of the Receptor Binding Domain of the SARS-CoV-2 strain from Indonesia. The method used was an in-silico approach using genomic and proteomic data and molecular docking. This study used a sample of 37 genomic data representing 86 types of SARS-CoV-2 spike protein mutations in Indonesia. Based on the docking candidate results, sgRNA8 has the lowest energy to bind to the viral protospacer target SARS-CoV-2 and a high melting point value at 70.3°C, indicating that the sgRNA8 chain is the optimal candidate for sgRNA.

Received December 07, 2021 *Accepted* May 25, 2022

INTRODUCTION

The emergence of SARS-CoV-2 in Wuhan, China, in late 2019 resulted in the Coronavirus Disease 19 (COVID-19) pandemic, which led to Severe Acute Respiratory Syndrome (SARS) and is responsible for the deaths of millions of people worldwide (WHO, 2020; *Shereen et al*., 2020; Wu *et al*., 2020; Wu *et al*., 2020). Viruses predicted to originate from bats and infecting mammalian class interspecies can transmit disease through droplets, direct contact, aerosols in closed spaces, urine, and mother to fetus transmission (Ren *et al*., 2020; Yang *et al*., 2020). The SARS-CoV-2 infection causes symptoms in the form of fever, cough, sore throat, headache, fatigue, breathing difficulties, diarrhea, and pneumonia (Hui *et al*., 2004; Zhang *et al*., 2020). Pneumonia is an inflammation of the respiratory tract and lungs, causing the alveoli to be filled by excess mucus secretion (Thompson, 2016). Pneumonia in COVID-19 patients aged over 65 years risks causing death (Du *et al*., 2020). The mortality rate due to COVID-19 has reached 6.1% worldwide (Abdullahi *et al*., 2020). Indonesia is one of the countries with a higher mortality rate (8.5%) due to COVID-19 than the world average mortality rate; this figure is considered very high (Sipahutar and Eryando, 2020). SARS-CoV-2 had 99% genome similarity between the specimens and 85% with the SARS-CoV virus (Khan et al., 2020). SARS-CoV-2 mutates rapidly, and in October 2020, a reported 155,000 SARS-CoV-2 sequences were detected worldwide (GISAID, 2020). These mutations affect the development of vaccines, antivirals, and diagnostic tools (Abdullahi *et al*., 2020; Kumar *et al*., 2020).

Health protocols and social restrictions minimize the spread of the virus and reduce mortality (Chu *et al*., 2020). As a result, clinical tests were developed to increase the tests' efficiency and accuracy (Ward *et al*., 2020). The clinical test most widely used worldwide is testing specimens from a respiratory swab using reverse-transcriptase polymerase chain reaction (RT-PCR) (Zitek, 2020). However, a faster and more effective test method was developed over time using the rapid test, which only takes about 15 minutes; this method works by detecting the antigen response (Hoffman, 2020).

The rapid spread of the coronavirus requires accurate and rapid clinical tests, but further research has shown that both conventional clinical testing meth-

Key words:
Genomics, Nucleic Acids, Pandemics, SARS-CoV-2.

Corresponding author:
Erlia Narulita, Ph.D
E-mail: erlia.fkip@unej.ac.id

ods have drawbacks. Clinical tests using the RT-PCR or Swab Test method are accurate, with a sensitivity of 70%-98% and specificity of 95%-99.7% for detecting COVID-19 infection. The rapid-test checks for the presence of SARS-CoV-2 antigen or antibody in the blood. The rapid-test is fast and needs about 30 minutes to show the result, but its low sensitivity (60.5%-72.1%) is a limitation. The diagnostic results of the rapid test require further RT-PCR testing to confirm the results completely obtained (Brihn *et al*., 2021; Du *et al*., 2021; Mak *et al*., 2020). In practice, it requires special equipment and trained officers in the molecular field, along with higher costs and time, taking about 2.5 hours to get the test results (Asai, 2020; Kurniawati, *et al*., 2019; Mahony *et al*., 2004; Scohy *et al*., 2020).

Due to the weakness of the two conventional methods, it is necessary to develop new medical test methods that are more efficient in enriching diagnostic methods and increasing the accuracy of diagnosing SARS-CoV-2 infection. According to WHO, detecting a disease must be fast, specific, sensitive, instrument-free, and cost effective (Mustafa and Makhawi, 2020). Scientists recently discovered a specific genetic engineering method called CRISPR (Clustered Regularly Interspaced Palindromic Repeats) associated with the Cas protein. The CRISPR system is a self-defense mechanism in bacteria from foreign genes such as viral genes (Narulita, 2019; Vyas and Bernstein, 2019). CRISPR technology has also developed a nucleic acid detection technology called SHERLOCK (Specific High sensitivity Enzymatic Reporter un-LOCKing) based on CRISPR to detect nucleic acids from urine and saliva, serum, plasma, and blood with even limited sample preparation, with high sensitivity >95%, specificity >99%, and simpler, faster, and more flexible (Gao *et al*., 2021; Gronowski, 2018; Kellner *et al*., 2019). The SHERLOCK method answers this challenge as a clinical test method that is highly specific, sensitive, fast, cost-efficient, and easy to use for detecting pathogens such as SARS-CoV-2, Zika, Ebola, HPV, and M. tuberculosis (Wang *et al*., 2020). In general, the SHERLOCK diagnostic method has two stages: amplifying viral RNA in samples using an isothermal amplification reaction and detecting viral RNA in samples that have been amplified using CRISPR collateral reporter un-locking media (Joung *et al*., 2020).

The key to the specificity of the SHERLOCK diagnostic method is the components in the form of CRISPR RNA (crRNA) and Trans-activator (tracrRNA) as a guide to specifically search for the target gene. Both (crRNA and tracrRNA) can be engineered into single cells. Single guide RNA (sgRNA) has two essential functions at once (crRNA and tracrRNA function). The nucleotide sequence at the 5' end of the sgRNA will bind to the target nucleotide by Watson-Crick base pairing, and the double-stranded structure at the 3' end of the sgRNA will bind to the Cas protein. The structure of this sgRNA chain can later be modified according to the desire to determine the target sequence to be addressed (Doudna and Charpentier, 2014; Yan *et al*., 2018).

This study targeted the gene Receptor Binding Domain SARS-CoV-2 samples from Indonesia and specimens from Indonesia because mutations in each country or region have unique characteristics (Lamptey, 2021). The RBD gene in the SARS-CoV-2 spike protein-coding chain was chosen as a target because it plays an important role in the infection process and also has a unique sequence of 76% (lowest) compared to the RBD spike protein (S) of other coronavirus members (Kaur *et al*., 2021; Kumar *et al*., 2021). The selected sequences are then used as a reference (query) in sgRNA construction (Karlapudi *et al*., 2018).

The construction of sgRNA was performed using the in-silico approach, a biological study process using a computer. The in-silico method has the flexibility, automation, and rapidity to evaluate many genes (Karlapudi *et al*., 2018; Palsson, 2000). This approach is also widespread in the medical field, targeting the study of biomolecules such as DNA, RNA, and protein (Chikhale, 2020).

MATERIALS AND METHODS

Genome retrieved and conserved RBD of SARS-CoV-2

The genomic sequences of SARS-CoV-2 specimens from Wuhan and genomes of the SARS-CoV-2 mutant specimens from Indonesia were obtained from GeneBank GISAID (https://www.gisaid.org) and NCBI (https://www.ncbi.nlm. nih.gov). The RBD gene in Indonesian specimens was identified based on data from the Wuhan specimen RBD gene sequences from the NCBI Genebank (https://www.ncbi.nlm.nih.gov/). The conserved region of the RBD gene in the SARS-CoV-2 specimen from Indonesia was obtained by the *MUSCLE Alignment* method by querying the SARS-CoV-2 RBD gene in the specimen from Wuhan.

Customization Guide RNA (sgRNA) targeting the SARS-CoV-2

A sequence from the conserved RBD gene was used as a target for constructing and customizing sgRNA candidate sequences using the CRISPR-RT web-based tool (http://bioinfolab.miamioh.edu/CRISPR-RT/interfaces/). First, the sgRNA candidate with the highest GC content (≥50%) was selected. The amount of GC content is used as a parameter for selecting sgRNA candidates because the GC content affects the bond strength and stability of the sgRNA chain; hence, it is crucial to sgRNA performance (Liu *et al*., 2020).

3D Structure Construction of sgRNA

Candidate sequences were constructed in a 3D structure along with the SARS-CoV-2 RBD gene target protospacer. This construction converts data in FASTA format into PDB format through RNAFitme website-based tools (http://rnafitme.cs.put.poznan.pl/) using the Cas13a design template from *Leptotrichia buccalis* (PDB ID: 5XWP). This 3D structure is then used in the docking test process to visualize the 3D structure of the constructed-sgRNA candidate using BIOVIA Discovery Studio software.

Molecular docking

The sgRNA candidates were tested for interaction of the bond between the sgRNA molecule and the Cas13a protein molecule with protein-RNA hybrid algorithm-free docking using HDOCK website-based tools (Wang *et al.*, 2021). The test was conducted to determine and analyze the bond energy formed by the Cas13-sgRNA interaction. Cas13a protein derived from the bacterium *Leptotrichia buccalis* C-1013-b (synthetic construct) was used (Liu *et al*., 2017). Therefore, the sgRNA candidates with a docking score lower than the median were selected (Wang *et al*., 2021).
The HDOCK web-based tool was used to determine and analyze the binding energy formed by the interaction of Cas13a-sgRNA with viral protospacer targets that form double-strand RNA (dsRNA) Cas13-sgRNA-target RNA (Wang *et al*., 2021). The Cas13-sgRNA-target RNA docking score indicates the sgRNA candidate with optimum ability to be used as a marker in SHERLOCK diagnostics. Both docking tests were carried out using the reference docking coordinates from Wang *et al*. (2021).

Validation and characterization of sgRNAs

Validation of optimum sgRNA candidate interactions and viral RNA RBD spike Protein SARS-CoV-2 was performed to ensure proper binding by docking Nucleic acid docking RNA-RNA complex structure modeling using HNADOCK website-based RNA to RNA docking tools (http://huanglab.phys.hust.edu.cn/hnadock/). Docking was done without the Cas13a protein to ensure that the sgRNA and target protospacer were complementary and could bind well.

RESULTS

Genome retrieval and conserved RBD gene SARS-CoV-2

The RBD gene of SARS-CoV-2 is located along a chain of genes coding for a structural spike protein (S) with a length of 223 amino acids at amino acid positions from the sequence 319 to 541. The RBD gene nucleotides sequence is located at sequences 22,517 to 23,186, with a length of 669 nitrogenous bases. Furthermore, the genome of SARS-CoV-2 mutant from Indonesia was also retrieved from GISAID. The number of samples used to represent a type of amino acid substitutions mutation was determined by the percentage of the number of its mutations found in the SARS-CoV-2 spike protein from the Indonesian sample; it is known that there were 86 types of mutations found from the Indonesian sample (Supplement 1). Moreover, as many as 37 samples from Indonesia presenting all types of amino acid mutations in the SARS-CoV-2 RBD gene were used (*Table 1*).
Alignment of 37 samples of genome from Indonesia and the RBD gene query sequence from Wuhan was performed by MUSCLE Alignment (100% conservation value parameter), which found seven conserved regions in the nucleotide sequence of the SARS-CoV-2 RBD with a length of 29bp or more (*Table 2*).

Single-Guide RNA (sgRNA) Customization

There were no mutations observed in the conserved regions of the SARS-CoV 2 nucleotide sequence (*Table 2*). These sequences are potentially excellent and efficient to be used as target RNA or protospacer sequences because mutation probabilities in this section are rare. These seven conserved sequences were used as queries in protospacer and crRNA construction with an in-silico approach using CRISPR-RT to obtain protospacer and crRNA candidates with

Table 1 - *Samples of SARS-CoV 2 from Indonesia.*

GISAID ID Number of SARS-CoV-2 Samples from Indonesia		
1. EPI_ISL_2233094	13. EPI_ISL_518819	25. EPI_ISL_2335090
2. EPI_ISL_1915533	14. EPI_ISL_1469250	26. EPI_ISL_2333654
3. EPI_ISL_2226648	15. EPI_ISL_2047531	27. EPI_ISL_2334474
4. EPI_ISL_1364467	16. EPI_ISL_2101164	28. EPI_ISL_2336149
5. EPI_ISL_2434991	17. EPI_ISL_1284303	29. EPI_ISL_2333810
6. EPI_ISL_2364675	18. EPI_ISL_2337254	30. EPI_ISL_2258218
7. EPI_ISL_2434168	19. EPI_ISL_2335456	31. EPI_ISL_2262280
8. EPI_ISL_1265168	20. EPI_ISL_2335146	32. EPI_ISL_2233109
9. EPI_ISL_2382412	21. EPI_ISL_2335872	33. EPI_ISL_1117452
10. EPI_ISL_1622404	22. EPI_ISL_2335021	34. EPI_ISL_2500494
11. EPI_ISL_1447354	23. EPI_ISL_2334094	35. EPI_ISL_2434157
12. EPI_ISL_2101042	24. EPI_ISL_2338534	36. EPI_ISL_2434061
		37. EPI_ISL_2429146

a high potential of specificity and stability. The key characteristics are that crRNA candidates must have GC content of ≥50% and only one on-target sequence (Wang *et al.*, 2021). The search results show that ten protospacer and crRNA candidates have GC content ≥50% and one on-target sequence (*Table 3*).

The value of GC content referenced in the selection of protospacer and crRNA is due to its effect on bond stability (G and C bases have three hydrogen bonds), which is stronger and more stable than that between A and U bases, which have only two hydrogen bonds (Liu *et al.*, 2020). In contrast, the on-target value is used as a reference to avoid the possibility of off-target binding.

Ten crRNA candidates (*Table 3*) were then designed for sgRNA sequences (tracrRNA + crRNA) by combining candidate crRNA sequences with tracrRNA sequences. TracrRNA sequences were obtained as crystal structures of Cas13a-complex (5XWP) retrieved from RSCB PDB. There were three chain sequences in the crystal structure of Cas13a-complex Leptotrichia buccalis C-1013-b synthetic construct (*Figure 1*). The first chain is an arrangement of amino acid chains that make up the Cas13a protein, the second chain is a nucleotide chain from guide RNA (sgRNA) that can be customized as needed, and the third chain is a nucleotide chain from the target RNA. The sgRNA sequences in the second chain were modified according to the crRNA and complementary to the SARS-CoV-2 RBD protospacer obtained in *Table 3, Figure 1.* CRISPR-complex structure. Furthermore, the sgRNA candidate sequences in *Table 4*, are designed from PDB structures using the sgRNA template from Cas13a *Leptotrichia buccalis* C-1013-b (*Figure 1* and *3a*).

Docking of candidate sgRNA with Cas13a

The docking test of the 3D structure of sgRNA (*Figure 3a*) and the Cas13a protein (5XWP) was conducted to

Table 2 - *Conserved regions of the RBD gene.*

Conserved Regions of the SARS-CoV-2 RBD Sequence from Indonesian Samples
>*Conserved*_RBD_*Region* 1 (107 *bases* | 22.518–22.623) AGAGTCCAACCAACAGAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCAC CAGATTTGCATCTGTTTATGCTTGGAA
>*Conserved*_RBD_*Region* 2 (85 *bases* | 22.626–22.710) AGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTGTTATGGA GTGT
>*Conserved*_RBD_*Region* 3 (102 *bases* | 22.712–22.813) TCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATC GCTCCAGGGCAAACTGGAAA
>*Conserved*_RBD_*Region* 4 (65 *bases* | 22.815–22879) ATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAA
>*Conserved*_RBD_*Region* 5 (47 *bases* | 22.919–22.965) GTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATA
>*Conserved*_RBD_*Region* 6 (52 *bases* | 23.076–23.126) TACCAACCATACAGAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACCAG
>*Conserved*_RBD_*Region* 7 (37 *bases* | 23.151–23.187) TCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCA

Table 3 - *Protospacer and crRNA candidate sequences.*

No.	*Protospacer + PFS (5'-3 ')*	*crRNA Complementary Seq. (5'-3')*	*GC*	*num target*	*numgenes*
1	GAUGAAGUCAGACAAAUCGCUCCAGGGCA	GCCCUGGAGCGAUUUGUCUGACUUCAUC	0:54	1	1
2	UCAGACAAAUCGCUCCAGGGCAAACUGGA	CCAGUUUGCCCUGGAGCGAUUUGUCUGA	0:54	1	1
3	CAGACAAAUCGCUCCAGGGCAAACUGGAA	UCCAGUUUGCCCUGGAGCGAUUUGUCUG	0:54	1	1
4	GAGGUGAUGAAGUCAGACAAAUCGCUCCA	GGAGCGAUUUGUCUGACUUCAUCACCUC	0:50	1	1
5	UGAUGAAGUCAGACAAAUCGCUCCAGGGC	CCCUGGAGCGAUUUGUCUGACUUCAUCA	0:50	1	1
6	AUGAAGUCAGACAAAUCGCUCCAGGGCAA	UGCCCUGGAGCGAUUUGUCUGACUUCAU	0:50	1	1
7	UGAAGUCAGACAAAUCGCUCCAGGGCAAA	UUGCCCUGGAGCGAUUUGUCUGACUUCA	0:50	1	1
8	GAAGUCAGACAAAUCGCUCCAGGGCAAAC	UUUGCCCUGGAGCGAUUUGUCUGACUUC	0:50	1	1
9	AAGUCAGACAAAUCGCUCCAGGGCAAACU	GUUUGCCCUGGAGCGAUUUGUCUGACUU	0:50	1	1
10	AGACAAAUCGCUCCAGGGCAAACUGGAAA	UUCCAGUUUGCCCUGGAGCGAUUUGUCU	0:50	1	1

Figure 1 - *CRISPR-complex structure.*

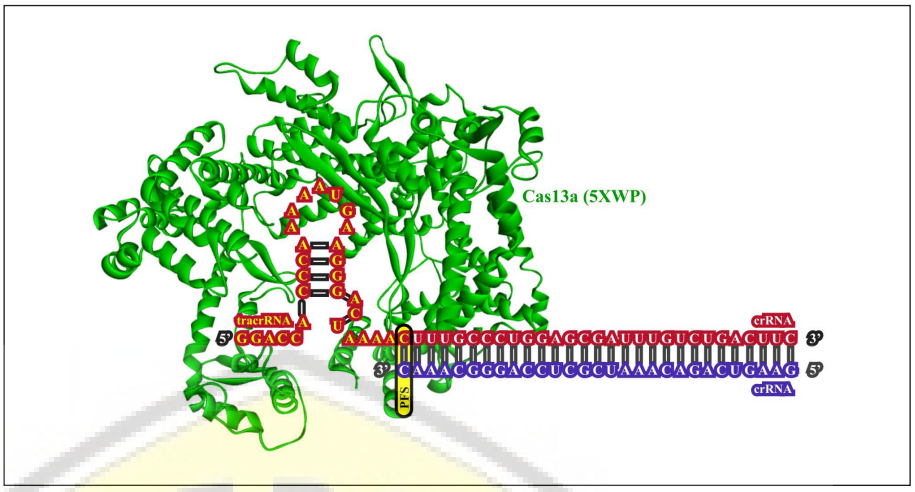


Table 4 - *sgRNA candidate sequences.*

No	*sgRNA candidate sequences*
1	Candidate 1
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACGCCCUGGAGCGAUUUGUCUGACUUCAUC
2	Candidate 2
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACCCAGUUUGCCCUGGAGCGAUUUGUCUGA
3	Candidate 3
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACUCCAGUUUGCCCUGGAGCGAUUUGUCUG
4	Candidate 4
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACGGAGCGAUUUGUCUGACUUCAUCACCUC
5	Candidate 5
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACCCCUGGAGCGAUUUGUCUGACUUCAUCA
6	Candidate 6
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACUGCCCUGGAGCGAUUUGUCUGACUUCAU
7	Candidate 7
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACUUGCCCUGGAGCGAUUUGUCUGACUUCA
8	Candidate 8
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACUUUGCCCUGGAGCGAUUUGUCUGACUUC
9	Candidate 9
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACGUUUGCCCUGGAGCGAUUUGUCUGACUU
10	Candidate 10
	>5XWP_2|Chains B[auth C], D[auth E]|RNA (59-MER)|synthetic construct (32630)
	GGACCACCCCAAAAAUGAAGGGGACUAAAACUUCCAGUUUGCCCUGGAGCGAUUUGUCU

examine the interaction between Cas13a as a receptor and the ten sgRNA candidates as ligands (*Figure 3b*). The results of docking Cas13a with ten sgRNA candidates showed various scores (*Figure 2a*). The lowest docking score was –1,548.32 (Cas13a-sgRNA 2) and the highest docking score was –1444.71 (Cas13a-sgRNA 4). The visualization results show that ten sgRNA candidates bind to the Cas13a protein in the correct position (*Figure 3b*). The median value of the docking score was –1540.46, after which several of the most efficient sgRNA candidates with a docking score lower than the median value were selected (sgRNA2, sgRNA6, sgRNA7, sgRNA8, and sgRNA10) (*Figure 2a*).

Docking of candidates with RBD

The five selected Cas13a-sgRNA 3D structures (*Figure 3b*) were redocked with target RNA using the protein-RNA docking process based on hybrid algorithm-free docking using HDOCK (Wang *et al.*, 2021). This second docking process was performed to test Cas13a-sgRNA in recognizing and binding to the target RNA protospacer (SARS-CoV-2 virus RBD gene). The docking scores and visualization of Cas13a-sgRNA receptors and target RNA ligand docking are presented in *Figure 2b* and *3c*, respectively. The docking scores with the lowest and highest variations were –1,343.68 (Cas13a- complex8) and –1,182.93 (Cas13a-complex2), respectively, while the median value was –1,276.04 (*Figure 2b*). Four sgRNAs were obtained (sgRNA2, sgRNA7, sgRNA8, and sgRNA10) as potential candidates. However, sgRNA candidate 8 had the highest potential as the guide RNA to bind with the target viral RNA as compared with other candidates (required lowest binding energy).

Validation and Characterization of sgRNA8

The optimum candidate sgRNA8 was then tested again for validation by nucleic acid RNA–RNA docking using HNADOCK. This docking phase tested the interaction between sgRNA and protospacer without Cas13a protein to ensure that sgRNA and protospacer can bind to each other and form a suitable structure. The docking results showed that candidate sgRNA8 can bind and recognize target RNA well, with a score of –1,220.89, and form a double-stranded structure that matches each other (*Figure 3d*). sgRNA8, which

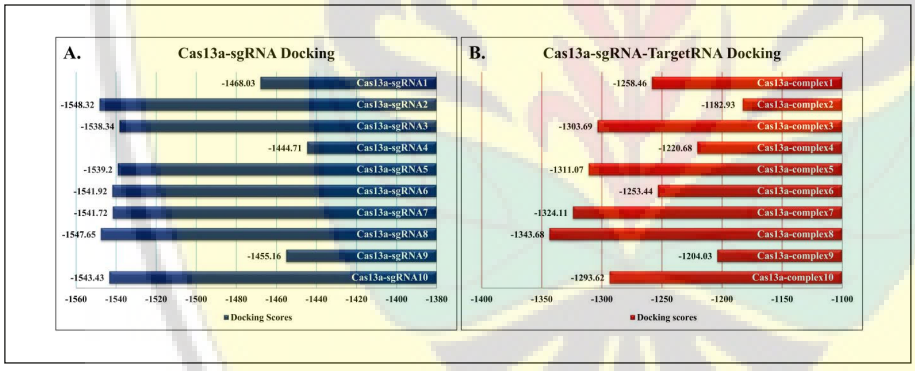


Figure 2 - *Docking scores (a) SgRNA-Cas13a docking. (b) SgRNA-Cas13a-targetRNA docking.*

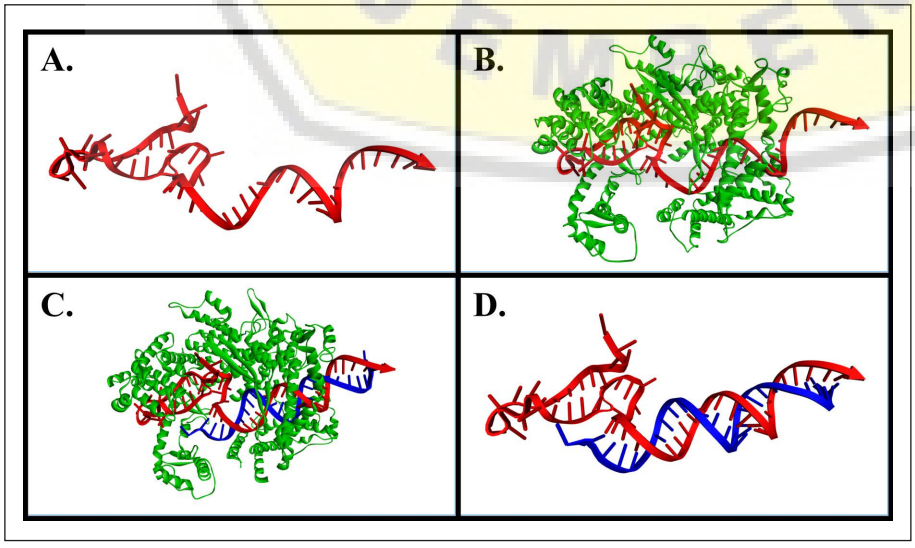


Figure 3 - *Molecule visualization (a) sgRNA candidate, (b) Cas13a-sgRNA docking, (c) Cas13a-sgRNA-target RNA docking, and (d) sgRNA validation.*

was the optimum guide RNA candidate, had a GC content of 49.2% (50% in crRNA) and showed a high melting point of 70.3°C.

DISCUSSION

The purpose of this study was to analyze the design of the SHERLOCK sgRNA, with optimum potential to be used as a Cas13a marker to recognize the Receptor Binding Domain spike protein (S) gene of the SARS-CoV-2 strain from Indonesia. The SHERLOCK diagnostic method uses sgRNA as a Cas13a nuclease pointer to find and cleave viral RNA followed by truncation of the signal RNA-probes on the active site of Cas13a as a fluorescent signal that a fluorometer can detect. This fact indicates that sgRNA plays a vital role in the specificity and diagnostic success of SHERLOCK to recognize the presence of viral RNA-protospacers (Mustafa dan Makhawi, 2020).

The docking results of sgRNA candidates with Cas13a show that the interaction energy was shallow, ranging from –1,548.32 to –1,444.71 (*Figures 2a* and *3b*). This is possible because sgRNA as the protein-binding base is very suitable and can form hairpin structures or loops that can attach to active site cavities of the Cas13a nuclease (Wang *et al*., 2021). The docking score variations were due to differences in the base arrangement structure of sgRNA, which affects the binding to Cas13a nuclease (Wang *et al*., 2021). The analysis of Cas13-sgRNA binding performance is based on the docking score that indicates the amount of energy required for the binding interaction between the receptor and ligand. The docking score correlates negatively with bond strength and stability; the lower the energy required, the more spontaneous and stable the bond that forms between the receptor and ligand (Wang *et al*., 2021). It is known that the median docking score of sgRNA and Cas13a is –1,540.46. Of the ten candidates, sgRNAs 2, 6, 7, 8, and 10 require lower energy than the median, indicating that candidate sgRNAs bind to Cas13a relatively more stably and have good potential for guiding Cas13a. Meanwhile, candidate sgRNAs 1, 3, 4, 5, and 9 were relatively less stable when bound to Cas13a.

After binding to the Cas13a nuclease, sgRNA binds to the target RNA based on a complementary sequence, thus docking between Cas13a-sgRNA and target RNA using HDOCK (Wang *et al*., 2021). The docking results showed that the interaction energy between sgRNA-Cas13a and viral RNA was deficient, ranging from –1,343.68 to –1,182.93 (*Figure 2b* and *3c*). The median docking score was –1,276.04. Of the five candidates, sgRNA 2, 7, 8, and 10 have lower energy than the median, so they were relatively more stable. These results are related to the study by Wang *et al*. (2021), which showed that candidates with the lowest docking scores have the potential to be good Cas13a guides to recognize and cleave target RNA with minimal off-target and side-cleavage effects.

However, candidate sgRNA8 had the highest potential to bind to viral RNA targets because it required the lowest energy to bind with target RNA (*Figure 2b and 3c*). The GC content in the sgRNA8 chain is 50% of its crRNA, which indicates that sgRNA8 met the minimum GC content value (Wang *et al*., 2021). The GC content in a nucleotide chain is crucial since it affects the bond strength, stability, and melting point of a nucleotide chain molecule. The melting point is the temperature used to accurately indicate the conversion between the solid and the liquid phase, which is crucial as a physical property to characterize a chemical compound (Tahir *et al*., 2002).

In summary, the sgRNA8 chain is the most appropriate candidate as compared to other candidates because it requires the lowest binding energy and has a high melting point of 70.3°C, which means that sgRNA8 is a stable molecule and has lower denaturation probability (Khandelwal and Bhyravabhotla, 2010). Furthermore, the CRISPR-Cas13a guide can be used as a highly efficient and specific RNA degradation tool for SARS-CoV-2 elimination and diagnosis in mammalian cells.

Acknowledgments

The authors express their gratitude toward Riska Ayu Febrianti, who assisted in using the molecular docking tools.

Compliance

The authors declare no conflicts of interest. This article does not contain any experiments involving animals or humans performed by any of the authors.

Author contributions

EN and AHW conceived the presented idea. EN, AHW, and SW performed the experiments, derived the models, and analyzed the data. SW assisted with the molecular docking measurements. All authors discussed the results and contributed to the final manuscript.

References

GISAID. Home page: SARS-CoV-2 data sharing via GISAID. https://www.gisaid.org/. [22 October 2020].

WHO. Coronavirus disease dashboard. https://covid19.who.int/. [Diakses 21 Oktober 2020].

Abdullahi IN, Emeribe AU, Ajayi OA, Oderinde BS, Amadu DO, Osuji AI. (2020). Implications of SARS-CoV-2 genetic diversity and mutations on pathogenicity of the COVID-19 and biomedical interventions. *Journal of Taibah University Medical Sciences*. **15** (4), 258.

Asai T. (2020). COVID 19: Accurate interpretation of diagnostic tests - A statistical. *Journal of Anesthesia*. 1.

Brihn A., Chang J., OYong K., Balter S., Terashita D., Rubin Z., et al. (2021). Diagnostic Performance of an Antigen Test with RT-PCR for the Detection of SARS-CoV-2 in a Hospital Setting - Los Angeles County, California, June–August 2020. *Morbidity and Mortality Weekly Report*. **70** (19), 702-705.

Chikhale H. (2020). Review on In-silico techniques: An approach to Drug discovery. *Current Trends in Pharmacy and Pharmaceutical Chemistry*. **2** (1), 24.

Chu D.K., Akl E.A., Duda S., Solo K., Yaacoub S., Schünemann H.J. COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *Lancet*. **395** (10242), 1973.

Doudna J., Charpentier E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*. **346** (6213), 1.

Du R.H., Liang L.R., Yang C.Q., Wang W., Cao T.Z., et al. (2020). Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: A prospective cohort study. *European Respiratory Journal*. **55** (5).

Du Z., Pandey A., Bai Y., Fitzpatrick M., Chinazzi M., et al. (2021). Comparative cost-effectiveness of SARS-CoV-2 testing strategies in the USA: A modelling study. *Lancet Public Health*. **6**, 185.

Gronowski A.M. (2018). Who or what is SHERLOCK? *EJIFCC*. **29** (3), 201.

Hoffman T, Nissen K, Krambrich J, Rönnberg B, Akaberia D, et al. (2020). Evaluation of a Covid-19 igm and igg rapid test; an efficient tool for assessment of past exposure to Sars-Cov-2. *Infection Ecology and Epidemiology*. **10** (1), 1.

Hui D., Chan M., Wu A., Ng P. (2004). SARS (SARS): Epidemiology and clinical features. *Postgrad Med Journal*. **80**, 373.

Joung J., Ladha A., Saito M., Segel M., Bruneau R., et al. (2020). Point-of-care testing for Covid-19 using SHERLOCK diagnostics. *MedRxiv*. 2.

Karlapudi A., **C VT,** Tammineedi J., Srirama K., Kanumuri L., Kodali V. (2018). *In silico* sgRNA tool design for CRISPR control of quorum sensing in Acinetobacter species. *Genes and Diseases*. **5**, 124.

Kaur N., Singh R., Dar Z., Bijarnia R., Dhingra N., Kaur T. (2021). Genetic comparison among various coronavirus strains for the identification of potential vaccine targets of SARS-CoV2. *Infection, Genetics and Evolution*. **89**, 6.

Kellner M., Koob J., Gootenberg J., Abudayyeh O., Zhang F. (2019). SHERLOCK: Nucleic acid detection with CRISPR nucleases. *Nat Protoc*. **14** (10), 2.

Khan M.I., Khan Z., Baig M.H., Ahmad I., Farouk A.E., et al. (2020). Comparative genome analysis of novel coronavirus (SARS-CoV-2) from different geographical locations and the effect of mutations on major target proteins: An in silico insight. *PloSOne*. **15** (9), 1.

Khandelwal G., Bhyravabhotla J. (2010). A phenomenological model for predicting melting temperatures of DNA sequence. *PloSOne*. **5** (8), 1.

Kumar S., Nyodu R., Maurya V., Saxena S. (2020). Morphology, genome organization, replication, and pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Medical Virology: from Pathogenesis to Disease Control.* 25.

Kurniawati M., Sumaryam, Hayati N. (2019). Aplikasi polymerase chain reaction (PCR) konvensional dan real time-pcr untuk deteksi virus vnn (viral nervous necrosis) pada ikan kerapu macan (Epinephelus fuscoguttatus). *Jurnal Techno-Fish*. **3** (1), 19.

Lamptey J., Oyelami F., Owusu M., Nkrumah B., Idowu P., et al. (2021). Genomic and epidemiological characteristics of SARS-CoV-2 in Africa. *Plos Neglected Tropical Diseases*. **15** (4), 1.

Liu G., Zhang Y., Zhang T. (2020). Computational approaches for effective CRISPR guide RNA design and evaluation. *Computational and Structural Biotechnology Journal*. **18**, 37.

Liu L., Li X., Ma J., Li Z., You L., et al. (2017). The molecular architecture for RNA-guided RNA cleavage by Cas13a. *Cell*. **170**, 714-726.

Mahony J., Petrich A., Louie L., Song X., Chong S., et al. (2004). Performance and cost evaluation of one commercial and six in-house conventional and real-time reverse transcription-PCR assays for detection of severe acute respiratory syndrome coronavirus. *Journal of Clinical Microbiology.* **42** (4), 1471.

Mak G., Cheng P., Lau S., Wong K., Lau C., et al. (2020). Evaluation of rapid antigen test for detection of SARS-CoV-2 virus. *Journal of Clinical Virology.* **129**, 3.

Mustafa M., Makhawi A. (2020). Sherlock and Detectr: 1 CRISPR-Cas systems as potential rapid diagnostic tools for emerging infectious diseases. *Journal of Clinical Microbiology.* **2**.

Narulita E. (2019). *Bioteknologi untuk strata 1 berbasis riset*. Yogyakarta: LaksBang PRESSindo.

Palsson B. (2000). The challenges of in silico biology. *Nature Biotechnology*. **18**, 1147.

Ren L.L., Wang Y.M., Wu Z.Q., Xiang Z.C., Guo L., et al. (2020). Identification of a novel coronavirus causing severe pneumonia in human: A descriptive study. *Chinese Medical Journal*. **133** (9), 1015-1024.

Scohy A., Anantharajah A., Bodéus M., Kabamba-Mukadi B., Verroken A., et al. (2020). Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. *Journal of Clinical Virology.* **129**, 1.

Shereen M.A., Khan S., Kazmi A., Bashir N., Siddique R.. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research.* **24**, 91-92.

Sipahutar T., Eryando T. (2020). COVID-19 case fatality rate and detection ability in Indonesia. *Jurnal Kesehatan Masyarakat Nasional.* **1**, 15.

Tahir I., Wijaya K., Yahya M.U. (2002). Quantitative relationships between molecular structure and melting point of several organic compounds. *Indonesian Journal of Chemistry.* **2** (2).

Thompson A. (2016). Pneumonia. *JAMA*. **315** (6), 626.

Vyas V., Bernstein D. (2019). An introduction to CRISPR-mediated genome editing in fungi. *Journal of Microbiology dan Biology Education.* **20** (3), 1.

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), 654.

Wang X., Shang X., Huang X. (2020). Next-generation pathogen diagnosis with CRISPR/Cas-based detection methods. *Emerging Microbes and Infections*. **9**, 1687.

Ward S., Lindsley A., Courter J., Assa'ad A. (2020). Clinical testing for COVID-19. *Journal Allergy Clinical Immunology.* **146** (1), 23.

Wu C., Chen X., Cai Y., Xia J., Zhou X., et al. (2020). Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Internal Medicine.* **180** (7), 938.

Wu D., Wu T., Liu Q., Yang Z. (2020). The SARS-CoV-2 outbreak: What we know. *International Journal of Infectious Diseases*.

Yan J., Chuai G., Zhou C., Zhu C., Yang J., et al. (2018). Benchmarking CRISPR on-target sgRNA design. *Briefings in Bioinformatics.* **19** (4), 721-722.

Yang Y., Xiao Z., Ye K., He X., Sun B, et al. (2020). SARS-CoV-2: Characteristics and current advances in research. *Virology Journal.* **17** (117), 3.

Zhang X., Cai H., Hu J., Lian J., Gu J., et al. (2020). Epidemiological, clinical characteristics of cases of SARS-CoV-2 infection with abnormal imaging. *Jurnal of Infectious Diseases*. **94**, 83.

Zitek T. (2020). The appropriate use of testing for COVID-19. *Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health.* **1**.