Turkish Computational and Theoretical Chemistry

Turkish Comp Theo Chem (TC&TC)

Volume(Issue): 9(3) – Year: 2025 – Pages: 25-37

e-ISSN: 2602-3237

https://doi.org/10.33435/tcandtc.1055649

Received: 13.08.2024 **Accepted**: 19.09.2024 *Research Article*

CRISPR-Cas Single-Guide RNA Construction for Tail Fiber Gene Gp0047 of Salmonella Phage SSE-121

Erlia Narulita[a,b,c,1], Annisyah Nurmitha Oktarina[c], Rina B. Opulencia[d], Agung Haris Widianto[e], Riska Ayu Febrianti[a,b,f], Hardian Susilo Addy[a,b]

[a]Doctoral Program of Biotechnology, Postgraduate Program, University of Jember, Jember, Indonesia
[b]Center for Development of Advance Science and Technology, University of Jember, Jember, Indonesia
[c]Department of Biology Education, Faculty of Education, University of Jember, Jember, Indonesia
[d]Microbiology Division, Institute of Biological Sciences, University of the Philippines Los Baños, Los Baños, Philippines
[e]Master of Genome Analytics, Faculty of Science, Monash University, Melbourne, Australia
[f]Department of Agricultural Product Technology, Faculty of Agriculture Technology, University of Jember, Jember, Indonesia

Abstract: In this study, we conducted an in-silico study of single guide RNA (sgRNA) construction as the initial stage of CRISPR-Cas9-based genome editing on *Salmonella* bacteriophage SSE-121 to broaden its host range. The results of the study discovered 188 sgRNA candidates using CHOPCHOP prediction. All selected candidates were docked using free website-based docking tools HNADOCK, therefore the top seven candidates of sgRNAs were docked using HDOCK with the Cas9 protein. The molecular dynamics simulation of the most optimum sgRNA-Cas9 protein (docking score -387.43) and sgRNA-Cas9-DNA (docking score -431.58) were calculated using AMBER14 forcefield in YASARA Dynamics. Based on the in-silico evaluation, the gRNA6 was obtained as the optimum single guide RNA. The docking score with DNA targets was -592.23 and Cas9 protein was -387.43. The comparison between the Cas9 and Cas9-sgRNA complex showed that the binding effect of sgRNA could maintain better conformational stability than in the single form.

Keywords: CRISPR, genome editing, molecular docking, molecular dynamics

1. Introduction

Foodborne disease cases are still highly found in daily life. World data shows that one out of ten peoples are infected by foodborne diseases and up to 420,000 die every year. In 2019, Indonesia reached 20 million cases of food poisoning, including 7,244 people suffering from foodborne diseases, among those 3,281 felt unwell and five people passed away [1,2]. The common disease causing-agent can be classified into pathogenic bacteria, viruses, and parasites. *Salmonella enterica* species is the most species found in foodborne disease and its presence is independent of the location, season, and level of primary health of a country [3].

Commonly, therapy and treatment to overcome *Salmonella* infection is using antibiotics. But it will lead to a new problem, called bacterial antibiotic resistance. Several strains of *Salmonella* are resistant to fluoroquinolone, ceftriaxone, and three or more groups of antibiotics (multidrug resistance) [3-5]. According to Patra et al. [4] *Salmonella enterica Typhimurium* in South Asia countries (Malaysia, Thailand, Vietnam, Indonesia, Cambodia, Singapore, and the Philippines) are resistant to three or more groups of antibiotics

[1] Corresponding Authors
e-mail: erlia.fkip@unej.ac.id

(multi-drug resistance). Therefore, the application of antibiotics is not adequate to solve this problem. Another treatment for bacterial infection is using bacteriophage. Bacteriophage treatment for *Salmonella* using mice in the laboratory and chicken in the poultry sector was safe, efficient, and a great potential for solving bacterial infection. They used two kinds of bacteriophages that infected the Salmonella typhimurium. The bacteriophages were delaying and reducing the *Salmonella typhimurium* infection in the mice [6-8]. The host specificity directs the bacteriophage to minimize the impact on normal flora bacteria in the human body. The treatment for *Salmonella* using bacteriophage might have advantages also limitations due to its various serovars. An alternative way to solve the limitation of the host range is using a jumbo phage. According to Kwon et al. [9] *Salmonella* jumbo phage, pSal-SNUABM-04, has been isolated and can infect a broader host range. Accordance to the large genome size indicates the jumbo phage has more functional genes. Therefore, it is more effective in infecting bacteria and has a wide host range. However, jumbo phage isolation is difficult due to its large size so it is difficult to separate by filtration and diffusion in media to form plaques on the bacterial lawns [10,11].

Genome editing provides another alternative way to solve the bacteriophage host range limitation. There is three kind of genome editing methods: Zinc-Finger-Nuclease (ZFN), Transcription Activator–like Effector Nuclease (TALEN), and Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) and CRISPR-associated (Cas) protein [12]. The CRISPR and Cas system is the most widely used in molecular biology laboratories around the world because of their simple design, low cost, high efficiency, excellent reproducibility, and short-cycle advantages [13,14]. CRISPR-Cas systems are composed of two main components, single-guide RNA (sgRNA) and Cas proteins. Single guide RNA is very important for guiding Cas9 protein to the target gene sequence for cutting it down.

During the genome editing process, sgRNA and Cas9 protein complexes can be mobilized to specific locations in the genome and create double-strand breaks (DSBs) at specific sites. Two repair mechanisms can further repair these DSBs; the error-prone non-homologous end joint (NHEJ) pathway and the homology-directed repair (HDR) pathway [15]. According to Duong et al. [16] *Escherichia coli* phages were bioengineered using the CRISPR-Cas system. After successful CRISPR-Cas genome editing, the K11, M13, PPO1, ϕX174, T3, and T7 phages have a broader host range.

The potential target gene for modification is the phage tail fiber protein gene. Tail fiber protein is used for phage recognition to the host during the adsorption phase.17 Hopefully, after this gene was edited, the phage might be infected with more serovars of *Salmonella*. In this research, single guide RNA (sgRNA) will be designed for *Salmonella* phage as the first step of CRISPR-Cas genome editing in Salmonella phage.

2. Computational Method

2.1. Optimum single guide RNA Design

The whole genome of *Salmonella* phage SSE-121 was obtained from the NCBI (Accession ID: NC_027351.1). The target gene encodes the *Salmonella* phage tail fiber protein of SSE121 (Accession ID: YP_009148843.1). The Cas9 protein was retrieved from the RCSB PDB protein bank (https://www.rcsb.org/) with the entry code of 4ZT0.18 The sgRNA is constructed by using website-based sgRNA design tools CHOPCHOP (http://chopchop.cbu.uib.no/) [19]. RNA composer website (https://rnacomposer.cs.put.poznan.pl/) is used for the 3D structure visualization of the sgRNA candidate that has been selected.20

2.2. Single guide RNA molecular docking

This research used a molecular docking approach. First, the docking of candidate sgRNA with target CRISPR RNA (crRNA) using HNADOCK (http://huanglab.phys.hust.edu.cn/hnadock/) [21]. Second, the docking of selected sgRNA with Cas9 protein using HDOCK (http://hdock.phys.hust.edu.cn/) [22,23]. Specific residues used as binding sites are residue numbers 1-20 (sgRNA) and 775-908 (Cas9 protein). Third, the docking of complex Cas9-sgRNA with target DNA using the HDOCK server (http://hdock.phys.hust.edu.cn/). Specific residues used as binding sites are residue numbers 1-10 (sgRNA) and residue numbers 11-20 (DNA). The

docking results represent the binding affinity between the ligand and receptor. The interaction between docked compounds and proteins is visualized using BioVia Discovery Studio 2019 software [24].

2.3. Single guide RNA molecular dynamic simulation

The interaction and stability of Cas9-sgRNA and Cas9-sgRNA-DNA were calculated and visualized using the YASARA molecular dynamics application developed by Biosciences GmbH [25]. Three samples are used in this step, Cas9 as a single target, Cas9-sgRNA, and Cas9-sgRNA-DNA as a complex. The first step is to input each sample into the program using the Options menu, then the Macro & Movie is selected, and finally Set Target. Furthermore, macro input is carried out to do the molecular dynamics simulations that have been prepared previously in the temperature (310K) and physiological pH (7.4) [26]. In the next step, set the running time of 10,000 ps (10 ns) in macro md_run. This simulation uses the AMBER14 forcefield (because it has been optimized for DNA and RNA samples). The snapshot save is done every 25 ps. The results of potential energy analysis, number of hydrogen bonds in the solute, number of hydrogen bonds between solute and solvent, Root Mean Square Deviation (RMSD), and radius of gyration are obtained by running macro md_analyze. Furthermore, the Root Mean Square Fluctuation (RMSF) analysis was carried out using the macro md_analyzeres, meanwhile for the Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) analysis was used the macro md_analyze bind energy.

3. Results and discussion

3.1. Single guide RNA Design

The purpose of this study was to understand the construction process of single guide RNA in *Salmonella* phage, especially Salmonella phage SSE-121. This Salmonella phage has been granted a patent on March 9, 2010 with patent number US 7,674,467B2 by Alexander Sulakvelidze Towson, MD (US); Shanmuga Sozhamamnnan, Timmonium, MD (US); and Gary R. Pasternack, Baltimore, MD (US).27 Designing of single guide RNAs of Salmonella phage SSE 121 tail fiber gene using CHOPCHOP suggesting about 14 single guide RNAs based on efficiency on target, GC content, and self-complementarity (Table 1).

Molecular docking of single guide RNA candidate

The 3D structure of all sgRNAs was docked with the target crRNA using the DNA-RNA hybrid algorithm-free docking HNADOCK. The docking score used in HNADOCK describes the binding mode and binding affinity in the relative ranking of the complex model [21]. The docking results showed that the docking score among the fourteen sgRNAs with the target sequence is quite low ranging from -443.39 to -607.19 (Figure 1). These results indicated that the interaction between RNA and the target sequences occurs due to sequence complementarity. The median docking score of the 14 single guide RNA candidates was at a score of -522.15. Among fourteen suggested sgRNAs, only six predicted sgRNAs showed lower than average energy ranging from -543.61 to -607.19.

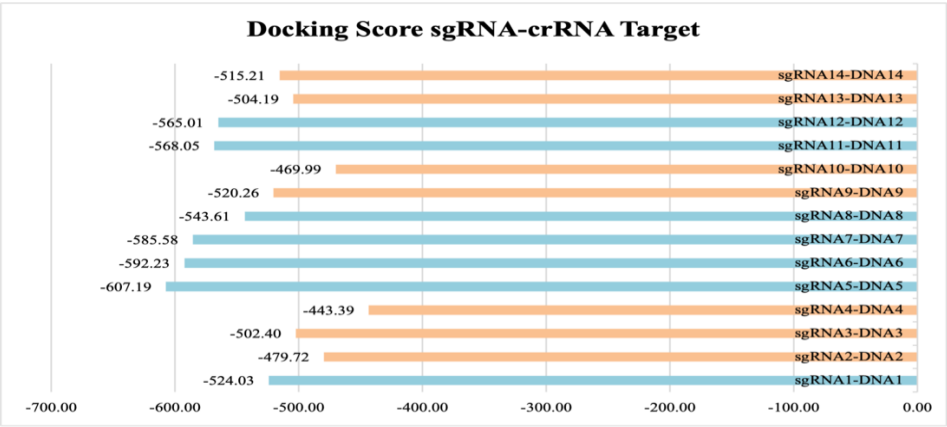

Figure 1. sgRNA – crRNA target docking score. Blue colour indicates the selected sgRNA based on average value for the next docking steps

Table 1. The selected final sgRNA candidates

Code	PAM	sgRNA sequences	Target DNA	Efficiency on Target	% GC content	Self-complementary
sgRNA1	CGG	GAGAUUGUACAACGACGCAG	CTCTAACATGTTGCTGCGTC	74,39	50	0
sgRNA2	TGG	CAUCACUGGUCCAUUCACAG	GTAGTGACCAGGTAAGTGTC	74,05	50	0
sgRNA3	TGG	GGGUUAACCAGUCUCAAUGG	CCCAATTGGTCAGAGTTACC	67,33	50	0
sgRNA4	TGG	ACAGCGGCAUGGUCCCCAAG	TGTCGCCGTACCAGGGGTTC	65,6	65	0
sgRNA5	CGG	CCAAACCCACAGCAUCCAGA	GGTTTGGGTGTCGTAGGTCT	65,34	55	0
sgRNA6	GGG	CUGGUCCAUUCACAGUGGGU	GACCAGGTAAGTGTCACCCA	63,98	55	0
sgRNA7	TGG	CUUGGAAACGCAACAAUCGG	GAACCTTTGCGTTGTTAGCC	63,32	50	0
sgRNA8	TGG	GCACAAAUGGGGUUCAAUGG	CGTGTTTACCCCAAGTTACC	60,54	50	0
sgRNA9	TGG	GGUGGGUUAACCAGUCUCAA	CCACCCAATTGGTCAGAGTT	60,19	50	0
sgRNA10	GGG	ACUGUUGCUGUUAGCCGCAG	TGACAACGAGAATCGGCGTC	59,77	55	0
sgRNA11	TGG	GUACAACGACGCAGCGGCGA	CATGTTGCTGCGTCGCCGCT	58,29	65	0
sgRNA12	AGG	GUACCCGCGAUCCACGGCAU	CATGGGCGCTAGGTGCCGTA	58,17	65	0
sgRNA13	CGG	GUCUGAGUACCCGCGAUCCA	CAGACTCATGGGCGCTAGGT	53,16	60	0
sgRNA14	AGG	ACUUGCCGAUGUUACAGUCC	TGAACGGCTACAATGTCAGG	54,56	50	0

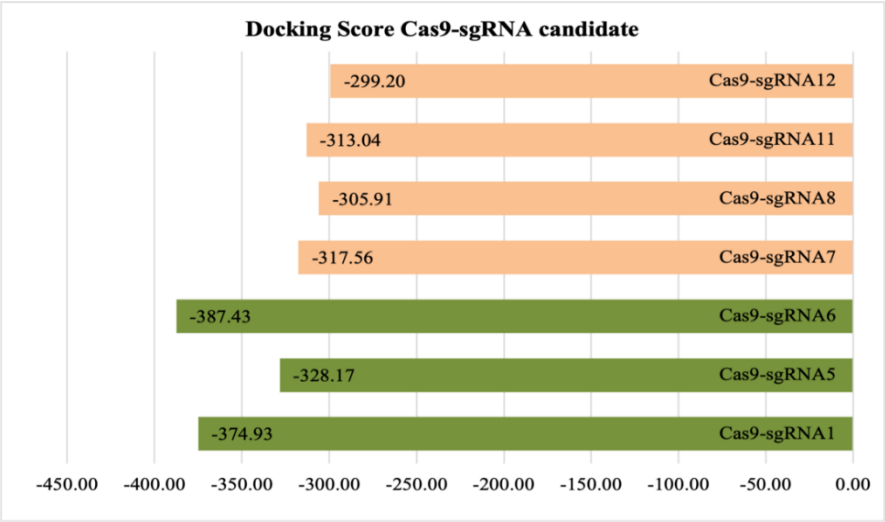

Figure 2. Cas9 protein–sgRNA docking scores. Green colour indicates the optimum candidate of single guide RNA based on average value of all candidates. The sgRNA6 is selected as the optimum single guide RNA

In addition, the docking of selected sgRNA with Cas9 protein using HDOCK (http://hdock.phys.hust.edu.cn/) showed that the binding energy between single guide RNA and Cas9 protein ranging between -299.20 to -387.43 with the highest and lowest value was showed by Cas9-sgRNA12 and Cas9-sgRNA6, respectively (Figure 2). Based on the best score, the Cas9-sgRNA6 was then chosen for further analysis.

According to the sgRNA predicted properties, the sgRNA6 had a GC content of 55% in crRNA and 38.8% in full sgRNA. Besides, the data also showed

that the melting point value of sgRNA6 was quite high, which is 68.8oC (Table 1). The protein-DNA/RNA docking summary of sgRNA with target DNA showed a specific binding site of the Cas9-sgRNAs complex protein interaction at the location on residue numbers 1-10 of sgRNA and residue numbers 11-20 of targeted-DNA. Docking summary (Table 2) and visualization (Figure 3) provided about 10 models of protein-DNA/RNA complex interaction. The model_1 has the most negative docking score (Table 2) and forms an interaction on an active site (bold font) with a distance ranging from 2.178 Å to 4.595 Å (Table 3). Hence, model_1 was used as a sample for the following molecular dynamic simulation.

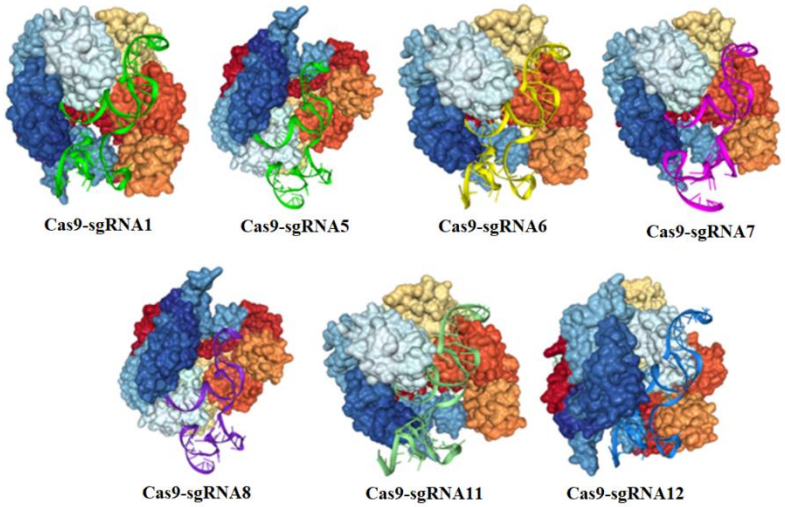

Figure 3. Cas9 protein – sgRNA docking visualization using HDOCK website

Table 2. Binding affinity between optimum sgRNA-Cas protein and DNA ligand

Rank	1	2	3	4	5	6	7	8	9	10
Docking Score	-431.58	-415.24	-411.91	-399.14	-396.54	-394.31	-384.38	-362.71	-352.8	-349.15
rmsd ligands (Å)	93.79	97.27	82.96	89.39	91.77	102.93	99.55	71.18	49.69	48.79
Interface residues	model_1	model_2	model_3	model_4	model_5	model_6	model_7	model_8	model_9	model_10

Table 3**.** Interaction residue amino acids produced on model_1 sgRNA-Cas9 protein with DNA ligand

Residue	Distance (Å)		Residue	Distance (Å)		Residue	Distance (Å)		Residue	Distance (Å)
235A - 8S	4.896		13B - 14S	4.354		32B - 19S	3.248		38B - 10S	3.93
235A - 9S	3.932		14B - 12S	4.22		32B - 20S	2.88		38B - 11S	3.044
252A - 9S	3.732		14B - 13S	1.058		33B - 12S	2,059		38B - 12S	2,631
252A - 10S	3,864		14B - 14S	3.21		33B - 13S	1.616		38B - 13S	3.77
254A - 10S	4.293		15B - 13S	2.813		33B - 14S	2.822		39B - 5S	4.127
255A - 9S	4.168		15B - 14S	2,069		33B - 15S	3,472		39B - 6S	3.89
255A - 10S	2,243		15B - 15S	3.203		34B - 12S	4.268		39B - 7S	3.271
255A - 11S	4.075		15B - 16S	4.678		35B - 11S	4.178		39B - 8S	3.074
258A - 11S	4.58		16B - 14S	4.86		35B - 12S	3.358		39B - 9S	4.166
262A - 9S	4.967		16B - 15S	2.986		36B - 9S	3.246		39B - 10S	4,532
264A - 9S	3.557		16B - 16S	2,594		36B - 10S	1.354		39B - 11S	2.56
9B - 10S	**3.95**		17B - 15S	3.852		36B - 11S	2.437		39B - 12S	4.253
9B - 11S	**2.939**		17B - 16S	1,792		36B - 12S	4.49		40B - 4S	4.651

9B - 12S	**4.595**		17B - 17S	2.296		36B - 13S	4.248		40B - 5S	2,193
10B - 11S	**2.178**		17B - 18S	3.049		36B - 14S	4.093		40B - 6S	4.964
10B - 12S	**2.488**		18B - 18S	3.643		37B - 8S	3.762		40B - 10S	3.978
11B - 11S	1975		18B - 19S	4.281		37B - 9S	1.65		40B - 11S	4.279
11B - 12S	1.706		19B - 19S	4.43		37B - 10S	0.913		41B - 2S	4.386
11B - 13S	3.414		19B - 20S	4.352		37B - 11S	4.704		41B - 3S	3.131
12B - 11S	2,068		28B - 20S	4.316		37B - 12S	3.466		41B - 4S	2,936
12B - 12S	2.846		29B - 20S	4.624		37B - 13S	2.718		53B - 3S	4.916
12B - 13S	2.515		30B - 20S	2,682		37B - 14S	4,559		65B - 1S	3.639
13B - 11S	4.351		31B - 19S	3.233		38B - 7S	3.79		66B - 1S	4.425
13B - 12S	2,585		31B - 20S	2.479		38B - 8S	2,776		66B - 2S	4.833
13B - 13S	2.408		32B - 18S	4048		38B - 9S	2,648			

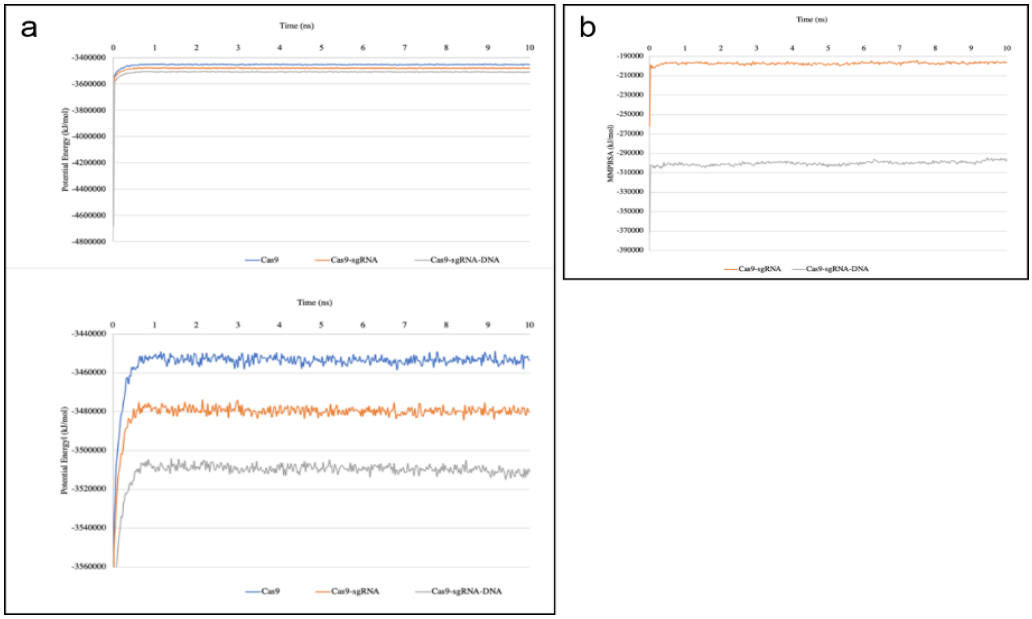

Figure 4. Graph of potential energy and free energy. a. Potential energy of single and complex proteins; b. Energy of free molecules in the Cas9-sgRNA and Cas9-sgRNA-DNA complexes based on MMPBSA value

3.2. Molecular Dynamic Simulation

3.2.1. Potential energy and free energy

The simulation of molecular dynamics generally showed that the average potential energy significantly increased the running time from 0 ns to 0.125 ns for Cas9 and 0.250 ns for Cas9-sgRNA and Cas9-sgRNA-DNA (Figure 4a). Furthermore, the Molecular Mechanics Poisson-Boltzmann Surface Area (MMPBSA) analysis showed that the free energy of implicit solvent at 10 ns of the production run was -197,467,244 kJ/mol for Cas9-sgRNA and -300,295,551 kJ/mol for Cas9-sgRNA-DNA, respectively (Figure 4b).

3.2.2. Total of hydrogen bond (H-bond)

The total of hydrogen bonds analysis using YASARA such as hydrogen bonds in solute (Fig. 5a) and hydrogen bonds between solute and solvent (Figure 5b) showed that about 1,164 hydrogen bonds on average formed in the Cas-sgRNA6 complex compared with Cas9 which was fewer about 77 hydrogen bonds. In addition, the number of hydrogen bonds was drastically increased while in the cell buffer condition about 3,399, 3,553, and 2,801 hydrogen bonds among Cas-sgRNA, Cas-sgRNA-DNA, and Cas9 with solvent, respectively. The result was also indicating that the hydrogen bonds between solute and solvent decreased over time.

3.2.3. Analysis of RMSD (Root Mean Square Deviation)

The trajectories of RMSD of the Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA display the fluctuated-increased values over the 3 Å about 0.525 ns for the Cas9, 0.752 ns for the Cas9-sgRNA, and 0.250 ns for the Cas9-sgRNA-DNA (Figure 6a).

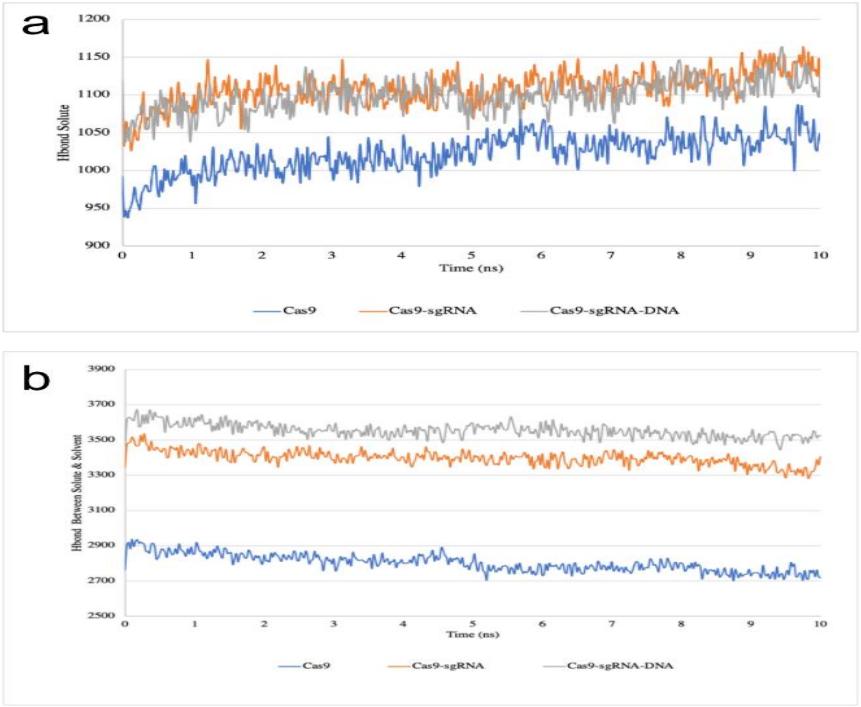

Figure 5. Graph of hydrogen bonds amount. a. Total of hydrogen bonds inside the Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA complex; b. Total of hydrogen bonds between Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA complex with solvent

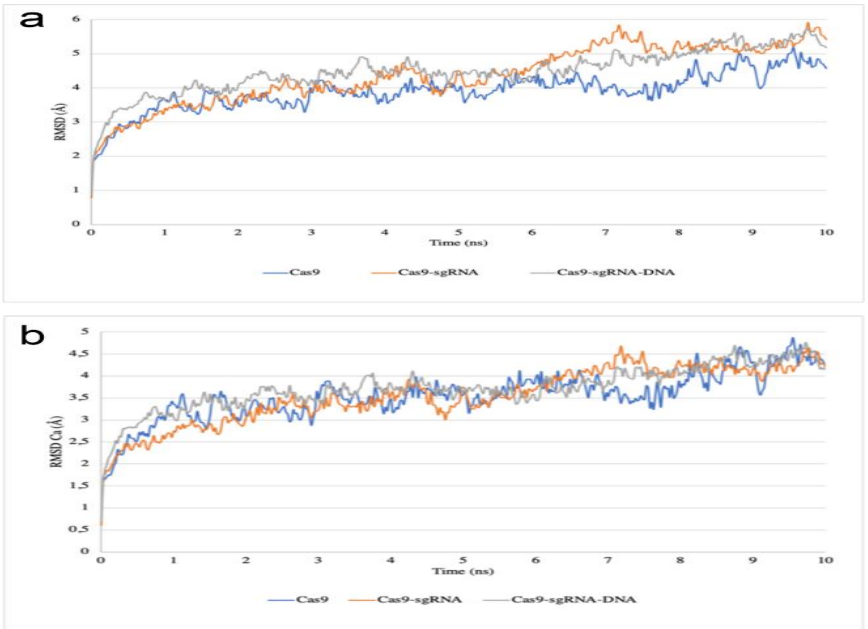

Figure 6. Graph of RMSD. a. Single and complex total RMSD; b. RMSD Cα of Cas9, Cas9-sgRNA and Cas9-sgRNA-DNA complexes

The data also showed that the RMSD value of the Cas9 was higher than the Cas9-sgRNA and the Cas9-sgRNA-DNA complexes, which was 3.905 Å compared with 4.369 Å and 4.482 Å, respectively.

In addition, the alpha-carbon RMSD value of all samples (the Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA complexes) showed similar trajectories pattern as well as RMSD with less production time and value (Figure 6b). The alpha-carbon RMSD values for each sample was 3.556 Å for the Cas9, 3.558 Å for the Cas9-sgRNA, and 3.724 Å for the Cas9-sgRNA-DNA.

3.2.4. Analysis of radius of gyration (RG)

The pattern of average RG values for the Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA complexes showed a contradictory pattern between a single sample (Cas9 alone) with complex samples (both Cas9-sgRNA or Cas9-sgRNA-DNA). The trajectory RG pattern of a single sample of Cas9 was fluctuate-increase while the complex samples were fluctuated-decrease (Figure 7).

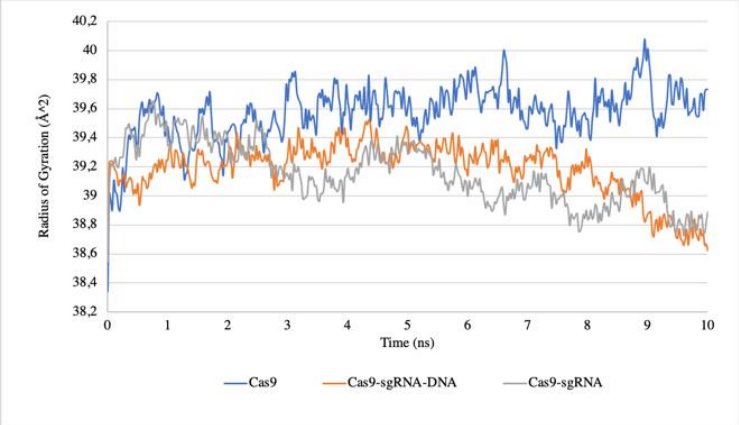

Figure 7. Graph of the comparison in RG patterns between single Cas9, Cas9-sgRNA, and Cas9-sgRNA-DNA complexes

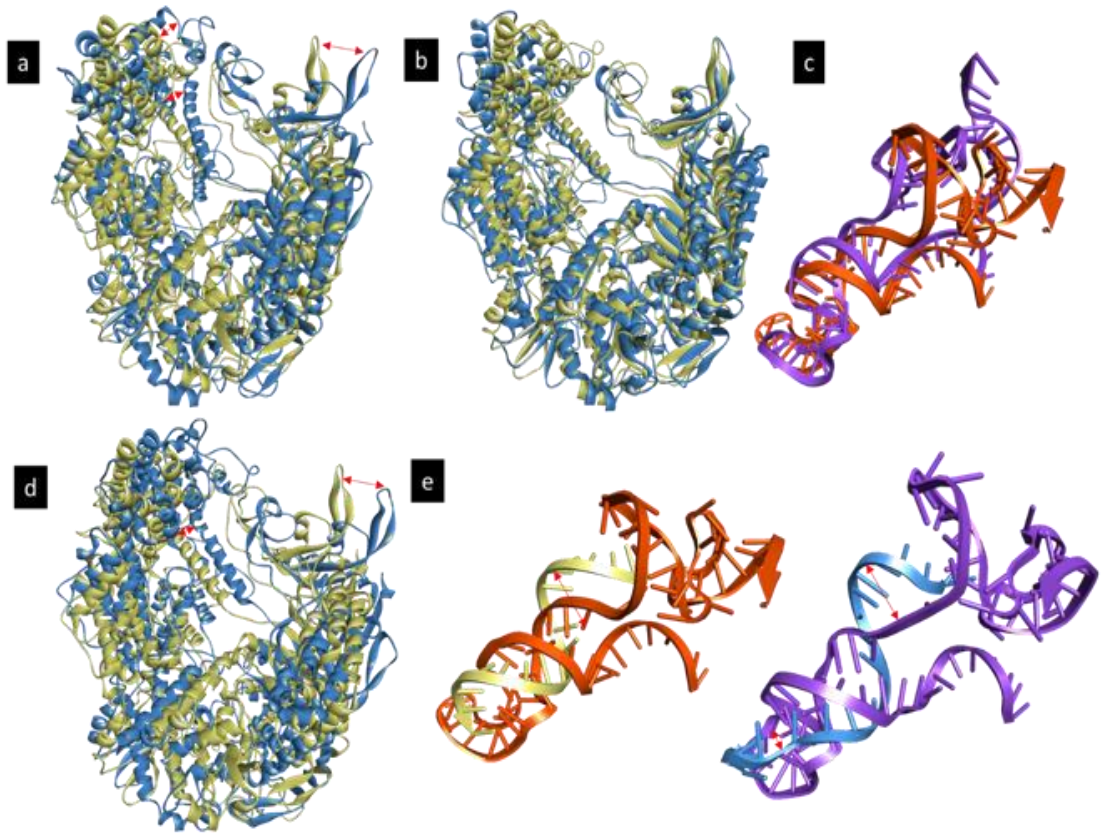

Figure 8. Visualization of superimpose position. a. Cas9; b. Cas9 on the Cas9-sgRNA complex; c. sgRNA on the Cas9-sgRNA complex; d. Cas9 on the Cas9-sgRNA-DNA complex; e. sgRNA on the Cas9- sgRNA-DNA before (yellow and orange) and after (blue and purple) simulation using molecular dynamics.

3.2.5. Visualization of Cas9-sgRNA and Cas9-sgRNA-DNA Complex Superimposition

The results of the superimposition of all samples before and after the simulation showed the effect of sgRNA binding on the conformational stability of the Cas9 protein (Figure 8). It also visualized the changes in the configuration of sgRNA (Figure 8c) and DNA (Figure 8e) as a ligand and also showed that the bound of sgRNA with DNA changed the conformation of the Cas9 protein and predicted an initiation change of DNA configuration to begin the cleavage process.

All samples showed the RMSF values of the Cas9 structure amino acid residues below 3 Å, such as 2.089 Å for the Cas9, 2.183 Å for the Cas9-sgRNA, and 1.956 Å for the Cas9-sgRNA-DNA (Figure 9a). Whereas, the RMSF trajectory values of the sgRNA structural nucleotides in the Cas9-sgRNA and Cas9-sgRNA-DNA complexes increased at nucleotide number 40-75 with the trajectory of Cas9-sgRNA-DNA complexes relatively below the RMSF value of Cas9-sgRNA complex (Figure 9b). Moreover, the result also showed that the declining pattern of RMSF trajectory values occurred at the residues of 1 to 10 (Figure 9b). Meanwhile, among the active site of DNA located at the nucleotide residues number 11-20, the nucleotide residues number 16-20 showed RMSF values of more than 3 Å (Figure 9c).

3.3. Discussion

The process of constructing single guide RNA for Salmonella phage in silico is needed to select candidate single guide RNA to be used in the in vitro research. The construction starts with the step to determine the optimum single guide RNA (sgRNA) using online tools such as CHOPCHOP [28]. The design of single guide RNA (only for crRNA) using the CHOPCHOP online tools suggests candidates of single guide RNAs based on the off-target score, the efficiency on target score, GC content, and self-complementarity (Table 1). Zhu & Liang [29] and Konstantakos et al. [30] suggest the consideration of selecting the efficient sgRNA (20 nucleotides) is influenced by the amount of GC (varying from 40 to 80%), with more than 50% (G) located next to protospacer adjunction motif (PAM) of NGG will increase the on-target efficiency. In this screening, 14 selected candidates have been obtained for molecular docking (Table 1).

The molecular docking data of sgRNA–crRNA target (Fig 1) indicates the interaction between RNA and the target sequences occurs due to sequence complementarity. The two sequences interact via hydrogen bonds, adenine-thymine with 2 hydrogen bonds and guanine-cytosine with 3 hydrogen bonds [31]. In RNA, thymine is substituted with uracil, so that the interactions that occur are adenine-uracil and guanine-cytosine [32]. There are differences in binding affinity due to differences in the arrangement of nitrogenous bases and the number of bonds that occur in each sgRNA and target sequence because binding energy is very sensitive to differences in structure [33]. In addition, the docking also suggests that six sgRNA candidates (Figure 1) with the docking score under the average docking score are the stable sgRNA's candidates. According to the Du et al. [34] that lower binding energy indicated that the bond or interaction of complementary sequence occurs will be more stable.

Furthermore, the molecular docking using the protein-RNA hybrid algorithm-free docking HDOCK describes binding affinity, but not absolute binding affinity [22] In this result, sgRNA6 is selected as the optimum single guide RNA candidate because it has the lowest docking score (Figure 2). The lower binding energy indicated that the binding of sgRNA with Cas9 protein is the most stable compared to the binding of other sgRNA candidates. This is consistent with result of Wang et al. [35] and Narulita et al. [36] that the optimum single guide RNA obtained through the screening process for CRISPR-Cas is RNA with a relatively lower docking value than other sgRNA candidates, both sgRNA docking scores with Cas protein and target crRNA.

The content of GC will affect the bond strength, stability, and melting point of a molecule. The more GC content, the more hydrogen bonds formed between the two nucleotides, consequently the bond becomes more stable [37]. The melting point value of sgRNA6 is quite high indicated that sgRNA6 is a good candidate. The higher the melting point of the molecule, the more stable the molecular

structure and more resistant to degradation or mutation of the nucleotide chain arrangement [38].

The value of MMPBSA is depending on the amount of gas phase energy, free energy solvation, and the contribution of the entropy configuration of the solute. The closer to the positive, the more stable and stronger the binding energy [39]. The Cas9-sgRNA-DNA complexes has lower stability compared with Cas9-sgRNA suggesting that the initiation process of attaching sgRNA to the target DNA remining many nucleotides that are exposed to the cell and not formed a bond yet. The more the number of hydrogen bonds between a sample, the better its tendency to be hydrophilic. The increase in hydrogen bonds occurs due to the formation of new hydrogen bonds between sgRNA and target DNA. The addition of hydrogen bonds is expected to make the complex more stable.

RMSD is an analysis score that provides information on conformational changes in a macromolecule that acts as a receptor after the interaction process with a particular ligand.40 RMSD is data that is sufficient to represent the stability of the sample under simulation conditions. The dynamic stability in question is the absence of significant conformational changes, which is better known as the unfolding process. RMSD can also be used as a standard deviation of conformational change, with standard < 2 Å and > 2 Å which are generally applied to the results of docking. The RMSD standard for a simulated protein receptor is 3 Å, if the RMSD value of a protein ≥ 3 Å, it is indicating that the protein has undergone a conformational change that is much different from its native condition. The trend of total RMSD value of the complex is higher because of the change in ligand conformation, in the form of sgRNA and DNA. The similar pattern of RMSD is also reported for the apoCas9, Cas9–sgRNA, and Cas9–sgRNA–DNA complex with the target of REC domain, NUC lobe components, RuvC and HNH domains [41].

The radius of gyration (RG) is a parameter that describes the equilibrium conformation of the entire simulation system. The RG value which can also be explained as the radius of rotation of the dynamic movement of a complex of both proteins and protein compounds against the solvent, becomes one of the ways to predict the simulation of the solubility of the sample in a solution or liquid solvent [42]. The lowest value indicated the folded protein condition, while the highest value indicates the protein conformational condition when unfolded [43].

The results of the comparison between all samples showed that the binding of gRNA6 significantly changed the RG pattern of the Cas9 protein and the binding of DNA changed the RG of the Cas9-sgRNA complex (Figure 7). In addition, the complex pattern periodically shows a decrease in value predicted as folding conformation in Cas9 to protect the sgRNA, so that it is not exposed to solvent molecules. The folding conformation formation process supports the purpose of sgRNA in the delivery process to DNA because it will increase the stability and bioavailability of the complex [44]. A gradually decreasing profile was also formed in the Cas9-sgRNA-DNA complex. RMSF is a score that provides information on conformational changes in more detail because it is associated with fluctuations in the level of amino acid residues and nucleotide ligands [45].

4. Conclusions

In conclusion, genome editing was used to expand the host range of bacteriophage infection. This in-silico study was a preliminary step in CRISPR-Cas9-based genome editing of *Salmonella* bacteriophage SSE-121 to expand its host range. Based on the computational evaluation of the docking score, one optimum single guide RNA was obtained, the sgRNA6. These in silico studies of sgRNA construction can determine optimal sgRNAs for Salmonella phages and are expected to minimize errors for the next phase of trials. Furthermore, this optimum sgRNA can be used as a reference to do further in vivo and in vitro trials to clarify the efficiency of this CRISPR-Cas9-based genome editing for expanding bacterial host range. Finally, to overcome antibiotic resistance in bacterial infections, we could develop CRISPR-based modified bacteriophage as a strong potential tool against it.

ACKNOWLEDGEMENT

This work was part of Percepatan Guru Besar grant with contract number 4338/UN25.3.1/LT/2022 and Visi Universitas grant with contract number 3791/UN25.3.1/LT/2023 under the name of Erlia Narulita. The content is solely the responsibility of

the authors and does not necessarily represent the official views of the Institution.

References

[1] Badan Pengawas Obat dan Makanan, *Laporan Tahunan BPOM* [Annual Report of BPOM]., Jakarta, 2019.

[2] T.R.P. Lestari, Penyelenggaraan Keamanan Pangan sebagai Salah Satu Upaya Perlindungan Hak Masyarakat sebagai Konsumen [The implementation of food safety to protect the rights of the community as consumers], Jurnal Masalah-Masalah Sosial, 11 (2020) 57–72. https://doi.org/10.22212/aspirasi.v11i1.1523

[3] M.P. Herrera-Sánchez, R.E. Castro-Vargas, L.C. Fandiño-De-Rubio, R. Rodríguez-Hernández, I.S. Rondón-Barragán, Molecular identification of fluoroquinolone resistance in *Salmonella* spp. Isolated from broiler farms and human samples obtained from two regions in Colombia, Vet. World, 14 (2021) 1767–1773. https://doi.org/10.14202/vetworld.2021.1767-1773

[4] S.D. Patra, N.K. Mohakud, R.K. Panda, B.R. Sahu, M. Suar, Prevalence and multidrug resistance in *Salmonella enterica Typhimurium*: an overview in South East Asia, World J. Microbiol. Biotechnol., 37 (2021) 1–17. https://doi.org/10.1007/s11274-021-03146-8

[5] J. Wain, J.A. Simpson, L.T.D. Nga, T.S. Diep, P.T. Duy, S. Baker, N.P.J. Day, N.J. White, C.M. Parry, Bactericidal activities and post-antibiotic effects of ofloxacin and ceftriaxone against drug-resistant *Salmonella enterica serovar Typhi*, J. Antimicrob. Chemother., 76 (2021) 2606–2609. https://doi.org/10.1093/jac/dkab215

[6] J. Gomez-Garcia, A. Chavez-Carbajal, N. Segundo-Arizmendi, M.G. Baron-Pichardo, S.E. Mendoza-Elvira, E. Hernandez-Baltazar, A.P. Hynes, O. Torres-Angeles, Efficacy of *Salmonella* Bacteriophage S1 Delivered and Released by Alginate Beads in a Chicken Model of Infection, Viruses, 13 (2021) 1932. https://doi.org/10.3390/v13101932

[7] Q. Lamy-Besnier, L. Chaffringeon, M. Lourenço, R.B. Payne, J.T. Trinh, J.A. Schwartz, A. Sulakvelidze, L. Debarbieux, Prophylactic administration of a bacteriophage cocktail is safe and effective in reducing *Salmonella enterica serovar Typhimurium* burden in vivo, Microbiol. Spectr., 9 (2021) e0049721. https://doi.org/10.1128/spectrum.00497-21

[8] K. Wessels, D. Rip, P. Gouws, *Salmonella* in chicken meat: consumption, outbreaks, characteristics, current control methods and the potential of bacteriophage use, Foods, 10 (2021) 1–20. https://doi.org/10.3390/foods10081742

[9] J. Kwon, S.G. Kim, H. Kim, S.S. Giri, S.W. Kim, S.B. Lee, S.C. Park, Isolation and characterization of *Salmonella* jumbo-phage psal-snuabm-04, Viruses, 13 (2021). https://doi.org/10.3390/v13010027

[10] L.M. Iyer, V. Anantharaman, A. Krishnan, A.M. Burroughs, L. Aravind, Jumbo phages: A comparative genomic overview of core functions and adaptions for biological conflicts, Viruses, 13 (2021). 1–42. https://doi.org/10.3390/v13010063

[11] H.X. Zhang, Y. Zhang, H. Yin, Genome Editing with mRNA Encoding ZFN, TALEN, and Cas9, Mol. Ther., 27 (2019) 735–746. https://doi.org/10.1016/j.ymthe.2019.01.014

[12] A.R. Weinheimer, F.O. Aylward, Infection Strategy and Biogeography Distinguish Cosmopolitan Groups of Marine Jumbo Bacteriophages, ISME J., 16 (2022) 1657–1667. https://doi.org/10.1038/s41396-022-01214-x

[13] Y. Xu, Z. Li, CRISPR-Cas systems: Overview, innovations and applications in human disease research and gene therapy, Comput. Struct. Biotechnol. J., 18 (2020) 2401–2415. https://doi.org/10.1016/j.csbj.2020.08.031

[14] N. Roshanravan, H. Tutunchi, F. Najafipour, M. Dastouri, S. Ghaffari, A. Jebeli, A glance at the application of CRISPR/Cas9 gene-editing technology in cardiovascular diseases, J. Cardiovasc. Thorac. Res., 14 (2022) 77–83. https://doi.org/10.34172/jcvtr.2022.14

[15] R. Ceccaldi, B. Rondinelli, A.D. D'Andrea, Repair Pathway Choices and Consequences at the Double-Strand Break, Trends Cell Biol., 26 (2016) 52–64. https://doi.org/10.1016/j.tcb.2015.07.009

[16] M.M. Duong, C.M. Carmody, Q. Ma, J.E. Peters, S.R. Nugen, Optimization of T4 phage engineering via CRISPR/Cas9, Sci. Rep., 10 (2020) 1–9. https://doi.org/10.1038/s41598-020-75426-6

[17] A. Schmidt, W. Rabsch, N.K. Broeker, S. Barbirz, Bacteriophage tailspike protein-based assay to monitor phase variable glucosylations in *Salmonella* O-antigens, BMC Microbiol., 16 (2016), 1–11. https://doi.org/10.1186/s12866-016-0826-0

[18] F. Jiang, K. Zhou, L. Ma, S. Gressel, J.A. Doudna, A Cas9-guide RNA complex preorganized for target DNA recognition, Science, 348 (2015) 1477–1481. https://doi.org/10.1126/science.aab1452

[19] T.G. Montague, J.M. Cruz, J.A. Gagnon, G.M. Church, E. Valen, CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing, Nucleic Acids Res., 47 (2014) W171-W174. https://doi.org/10.1093/nar/gku410

[20] M. Antczak, M. Popenda, T. Zok, J. Sarzynska, T. Ratajczak, K. Tomczyk, R.W. Adamiak, M. Szachniuk. New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure, Acta Biochim. Pol., 63 (2016) 737–744. https://doi.org/10.18388/abp.2016_1329

[21] J. He, J. Wang, H. Tao, Y. Xiao, S.Y. Huang, HNADOCK: a nucleic acid docking server for modeling RNA/DNA-RNA/DNA 3D complex structures, Nucleic Acids Res., 47 (2019) W35–W42. https://doi.org/10.1093/nar/gkz412

[22] Y. Yan, D. Zhang, P. Zhou, B. Li, S.Y. Huang, HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy, Nucleic Acids Res., 45 (2017) W365–W373. https://doi.org/10.1093/nar/gkx407

[23] Y. Yan, H. Tao, J. He, S.Y. Huang, The HDOCK server for integrated protein–protein docking, Nat. Protoc., 15 (2020) 1829–1852. https://doi.org/10.1038/s41596-020-0312-x

[24] Jejurikar, L. Bhagyashree, S.H. Rohane, Drug designing in discovery studio, Asian J. Res. Chem., 14 (2021) 135–138. https://doi.org/10.5958/0974-4150.2021.00025.0

[25] E. Krieger, G. Vriend, New ways to boost molecular dynamics simulations, J. Comput. Chem., 36 (2015) 996–1007. https://doi.org/10.1002/jcc.23899

[26] N. Shekhar, P. Sarma, M. Prajapat, P. Avti, H. Kaur, A. Raja, H. Singh, A. Bhattacharya, S. Sharma, S. Kumar, A. Prakash, B. Medhia, In silico structure-based repositioning of approved drugs for spike glycoprotein S2 domain fusion peptide of SARS-CoV-2: rationale from molecular dynamics and binding free energy calculations, mSystems, 5 (2020) 1–16. https://doi.org/10.1128/mSystems.00382-20

[27] A. Sulakvelidze, S. Sozhamamnnan, G.R. Pasternack, *Salmonella* bacteriophage and uses thereof, (2020).

[28] W.A. Hardiyani, A. Wafa, W.I.D. Fanata, H.S. Addy, Design and Construction of Single Guide RNA for CRISPR/Cas9 System Based on the xa13 Resistance Gene in Some Varieties of Rice (*Oryza sativa*), J. Trop. Plant Pests Dis., 23 (2023) 47-55. https://doi.org/10.23960/jhptt.12347-55

[29] H. Zhu, C. Liang, CRISPR-DT: designing gRNAsfor the CRISPR-Cpf1 system with improved target efficiency and specificity, Bioinformatics, 35 (2019) 2783–2789. https://doi.org/10.1093/bioinformatics/bty1061

[30] V. Konstantakos, A. Nentidis, A. Krithara, G. Paliouras, CRISPR–Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning, Nucleic Acids Res., 50 (2022) 3616–3637. https://doi.org/10.1093/nar/gkac192

[31] C. Fonseca-Guerra, F.M. Bickelhaupt, J.G. Snijders, E.J. Baerends, Hydrogen bonding in DNA base pairs: Reconciliation of theory and experiment, J. Am. Chem. Soc., 122 (2020) 4117–4128. https://doi.org/10.1021/ja993262d

[32] J.P. Cerón-Carrasco, A. Requena, E.A. Perpète, C. Michaux, D. Jacquemin, Double proton transfer mechanism in the adenine-uracil base pair and spontaneous mutation in RNA duplex, Chem. Phys. Lett., 484 (2009) 64–68. https://doi.org/10.1016/j.cplett.2009.11.004

[33] Ajay, M.A. Murcko, Computational Methods to Predict Binding Free Energy in Ligand-Receptor Complexes, J. Med. Chem., 38 (1995) 4953–4967. https://doi.org/10.1021/jm00026a001

[34] X. Du, Y. Li, Y.L. Xia, S.M. Ai, J. Liang, P. Sang, X.L. Ji, S.Q. Liu, Insights into protein–ligand interactions: Mechanisms, models, and methods, Int. J. Mol. Sci., 17 (2016) 1–34. https://doi.org/10.3390/ijms17020144

[35] L. Wang, J. Zhou, Q. Wang, Y. Wang, C. Kang, Rapid design and development of CRISPR-Cas13a targeting SARS-CoV-2 spike protein, Theranostics., 11 (2020) 649–664. https://doi.org/10.7150/thno.51479

[36] E. Narulita, A.H. Widianto, S. Wathon, Construction of SHERLOCK-based sgRNA for SARS-CoV-2 Diagnostics from Indonesia, New Microbiol., 45 (2022) 173–180.

[37] H. Chen, C.K. Skylaris, Analysis of DNA interactions and GC content with energy decomposition in large-scale quantum mechanical calculations, Phys. Chem. Chem. Phys., 23 (2021) 8891–8899. https://doi.org/10.1039/d0cp06630c

[38] F. Qi, D. Frishman, Melting temperature highlights functionally important RNA structure and sequence elements in yeast mRNA coding regions, Nucleic Acids Res., 45 (2017) 6109–6118. https://doi.org/10.1093/nar/gkx161

[39] M. Aldeghi, A. Heifetz, M.J. Bodkin, S. Knapp, P.C. Biggin, Predictions of ligand selectivity from absolute binding free energy calculations, J. Am. Chem. Soc., 139 (2017) 946–957. https://doi.org/10.1021/jacs.6b11467

[40] X. Cheng, I. Ivanov, Molecular Dynamics, Methods Mol. Biol., 929 (2012) 243–285. https://doi.org/10.1007/978-1-62703-050-2_11

[41] P.V. Zhdanova, A.A. Chernonosov, D.V. Prokhorova, G.A. Stepanov, L.Y. Kanazhevskaya, V.V. Koval, Probing the dynamics of *Streptococcus pyogenes* Cas9 endonuclease bound to the sgRNA complex using hydrogen-deuterium exchange mass spectrometry, Int. J. Mol. Sci., 23 (2022) 1129. https://doi.org/10.3390/ijms23031129

[42] M.Y. Lobanov, N.S. Bogatyreva, O.V. Galzitskaya, Radius of gyration as an indicator of protein structure compactness, Mol. Biol., 42 (2008) 623–628. https://doi.org/10.1134/S0026893308040195

[43] E. Yamamoto, T. Akimoto, A. Mitsutake, R. Metzler, Universal relation between instantaneous diffusivity and radius of gyration of proteins in aqueous solution, Phys. Rev. Lett., 126 (2021) 128101. https://doi.org/10.1103/PhysRevLett.126.128101

[44] X. Xu, D. Duan, S.J. Chen, CRISPR-Cas9 cleavage efficiency correlates strongly with target sgRNA folding stability: From physical mechanism to off-target assessment, Sci. Rep., 7 (2017) 143. https://doi.org/10.1038/s41598-017-00180-1

[45] L. Martínez. Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis, PLoS One, 10 (2015) e0119264. https://doi.org/10.1371/journal.pone.0119264