Molecular Dynamics Simulation of DENV RNA-Dependent RNA-Polymerase with Potential Inhibitor of Disulfide Cyclic Peptide

Usman Sumo Friend Tambunan, Raima Syahidah Noors,
Arli Aditya Parikesit, Elyana and Wahyu Ronggo
Department of Chemistry, Faculty of Mathematics and Natural Science,
University of Indonesia, Depok 16424, Indonesia

Abstract: Problem statement: Our researches have proposed two ligands of disulfide cyclic polypeptide, which are CDEEC and CDGSC as potential inhibitor of DENV RNA-dependent RNA-polymerase by molecular docking. **Approach:** Methodological approach was conducted to determine the best ligand to act as inhibitor. Molecular docking simulation was conducted without a solvent in which enzyme was made rigid and ligand was left free to find the most suitable conformation. In actual cellular system there is a solvent which makes the enzyme to have a dynamic movement. **Results:** Therefore in this study, Molecular Dynamic (MD) simulation was performed to estimate more reliable condition of enzyme-ligand complex. In this study, molecular dynamics simulation was performed during 5 ns with two different temperatures, 300 and 312 K. At the end of MD simulation at 300 K, CDEEC bound to two RdRp important residues, Arg-729 and Arg-737 while CDGSC didn't bind to any important residues. **Conclusion:** Simulation at 312 K also showed almost similar result. CDEEC was bound to two RdRP important residues, Arg-737 and Ser-710, whereas CDGSC didn't bind to any important residues. Based on the result of these two simulations, CDEEC is proposed as a better inhibitor of RdRp dengue virus and feasible to be developed as anti-dengue drug.

Key words: Anti-dengue, docking, molecular dynamics, polypeptide, RdRp

INTRODUCTION

The rapid developments in science have brought many changes in human life. As one example, advances in biological sciences and bioinformatics have brought a better understanding of the organism functions in cellular and molecular scale. As a result of this progress, most research in the pharmaceutical industry has started to identify suitable targets in the organism and to design drugs, which interact with the target (Franca *et al*., 2006). This type of drug designing is known as target oriented drug or rational drug design.

In a rational drug design, drug design process begins with knowing the structure of the target protein and then form a database that contains a collection of compounds that are expected to interact with the target protein. To determine which compounds that have the best interaction with protein target and become candidates for drug synthesis, a series of analyzing techniques is performed by using computer-assisted tool. Two of the most well known computational techniques in drug design process are docking and molecular dynamics simulations.

Docking techniques is designed to find the most suitable conformation of ligand and its receptor (Alonso *et al*., 2006). Molecular dynamics simulation is a computation approach in which atoms and molecules allowed to interact with each other during a certain time period so that system behavior can be observed (Nurbaiti *et al*., 2010). Fast and inexpensive docking protocols can be combined with accurate but more costly MD techniques to predict more reliable protein-ligand complexes. The strength of this combination lies in their complementary strengths and weaknesses (Alonso *et al*., 2006).

Dengue infection caused by dengue virus infection remains a public health problem in the world. Dengue virus (DENV) has infected 50-100 million people each year; with 500,000 patients suffer more severe disease manifestation, which is Dengue Haemorrhagic Fever (DHF) or Dengue Shock Syndrome (DSS). It resulted in approximately 20,000 deaths especially in children. Dengue virus is transmitted into the human body through the bite of female mosquito of the genus Aedes (Qi *et al*., 2008). DENV is a virus belonging to the family Flaviviridae, with the Flavivirus genus. This

Corresponding Author: Usman Sumo Friend Tambunan, Department of Chemistry, Faculty of Mathematics and Natural Science, University of Indonesia, Depok 16424, Indonesia

virus has four serotypes, known as DENV-1, DENV-2, DENV-3 and DENV-4. Infection by one serotype does not protect patients from infection of three other serotypes. Infection from one serotype will only worsen if the infection caused by three other serotype (Tomlinson *et al.*, 2009). Because of these four different serotypes, dengue vaccine development becomes difficult. Moreover vaccine development is also complicated by the lack of a suitable animal model for dengue. Therefore, antiviral is one of promising agent to cope with dengue virus.

DENV is an RNA virus genome RNA and RNA genome spans about 10.7 kb and contains a type I methyl guanosine cap structure at its 5' end but is devoid of a polyadenylated tail. The genomic RNA is translated into a single polyprotein, which is cleaved into three structural (C-prM-E) and seven nonstructural (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) proteins by both the viral and cellular proteases (Yap *et al.*, 2007). These proteins have been known to be the targets of antiviral inhibitors for prevention DENV.

Flavivirus NS5 of all types has at least three enzymes that are essential for viral propagation. Located on the N-terminal part of the NS5, there are approximately 320 residues, which are S-adenosylmethionine-dependent methyltransferase (MTase), which has a function as MTase and guaniltransferase Enzymes (Qi *et al.*, 2008). Section C-terminal residues 420-900 in the position of NS5 is the RNA-dependent RNA polymerase (RdRp) responsible for the synthesis of the intermediate RNA template for subsequent replication of positive strand RNA genome. Because human cells lack DNA or RNA-dependent RNA polymerase as in HIV-1 reverse transcriptase or Flavivirus RdRp, this enzyme is one of the most promising drug targets (Sampath and Padmanabhan, 2009).

An allosteric inhibitor for DENV RdRp, which was derived from n-sulfonylanthranilic acid, has been founded (Yin *et al.*, 2009). This compound can be identified to selectively inhibit viral polymerase dengue and computational studies based on molecular dynamics simulations of the RdRp complex with this compound showed that these compounds were bound to the allosteric between finger and thumb RdRp region.

Our research had worked on two distinct methods, first and second batches to generate two different ligand peptides. Molecular docking was conducted in a condition where enzyme was made rigid. Meanwhile in actual system, enzyme is not rigid because it is located in solute environment inside cells. Therefore, to obtain the conditions of enzyme-ligand complex that more resembles the real environment, molecular dynamics simulations were carried out.

MATERIALS AND METHODS

First batch method:
RdRp dengue virus enzyme crystal structure: Searching of RdRp Enzyme structure in PDB format was performed at Research Collaboratory for Structural Bioinformatics (RCSB) site (http://www.rcsb.org/pdb/). After the 3D structure was obtained, the analysis to determine the binding site was conducted. The binding site determination was performed using molecular modeling software.

Preparation of peptide ligands: Peptide ligands were drawn in 3D by using ACDLabs program. The peptide was modelled as cyclic peptide where cysteine residue was added at its end to form a disulfide bridge and it was composed of negatively charged amino acid residue, aspartic acid and glutamic acid. The 3D model Fig. 1-3 was saved in MDL Mol file format and optimized using VegaZZ Force field program, with TRIPOS and Gasteigger charges option. The conformation study was done by steepest descent and conjugate gradient method.

RdRp enzyme preparation: Water molecule, chlorine ion and tryethylene glycol was eliminated by using Pymol program. The force field CHARMM22_PROT optimization was conducted, with steepest descent and conjugate gradient methods, by using VegaZZ program.

Docking of peptide ligand and enzyme: The docking parameter was prepared by using AutoDock Tools. In the enzyme molecule, the polar hydrogen atom was added. In the ligand, the Gasteiger charge was added and every bond was rotated. The docking calculation was conducted in AutoDock 4.0 program, by using Lamarckian Genetic Algorithm (LGA). The utilized parameters are population sizes 150, energy evaluations $2,5.10^6$ and 50 times runs. The Grid box was prepared with 0,375 Å grid spacing and RMSD value of each cluster must not higher than 1 Å.

Analysis of docking result toward peptide-RdRp enzyme complex: The docking analysis was conducted by examining the conformation which has the lowest energy value from the most populated cluster. Then, the $\Delta G_{binding}$ and Ki (inhibiton constant) values between peptide-enzyme was examined. This procedure was performed to describe the interaction, analyze the hydrogen bonding between peptide and enzyme and determine which enzyme residue that had certain contact with peptide ligand.
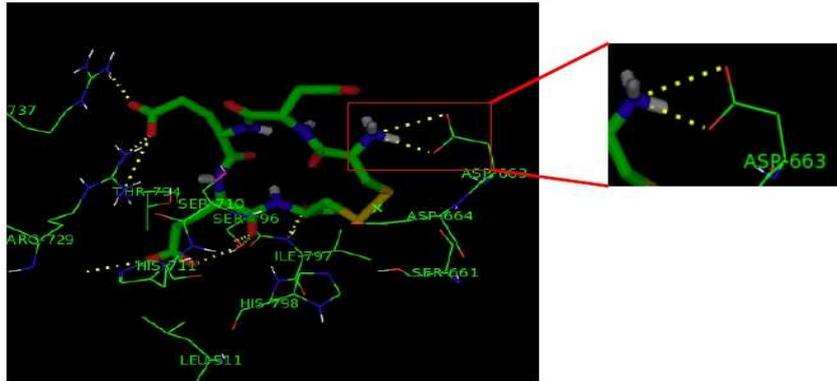
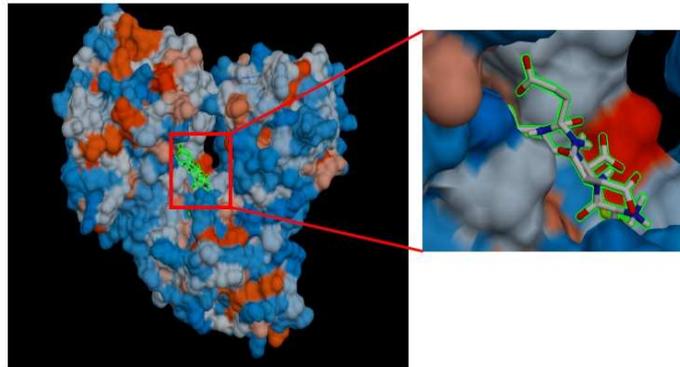Fig. 1: Salt Bridge Interaction between peptide ligand with Asp-663 residue

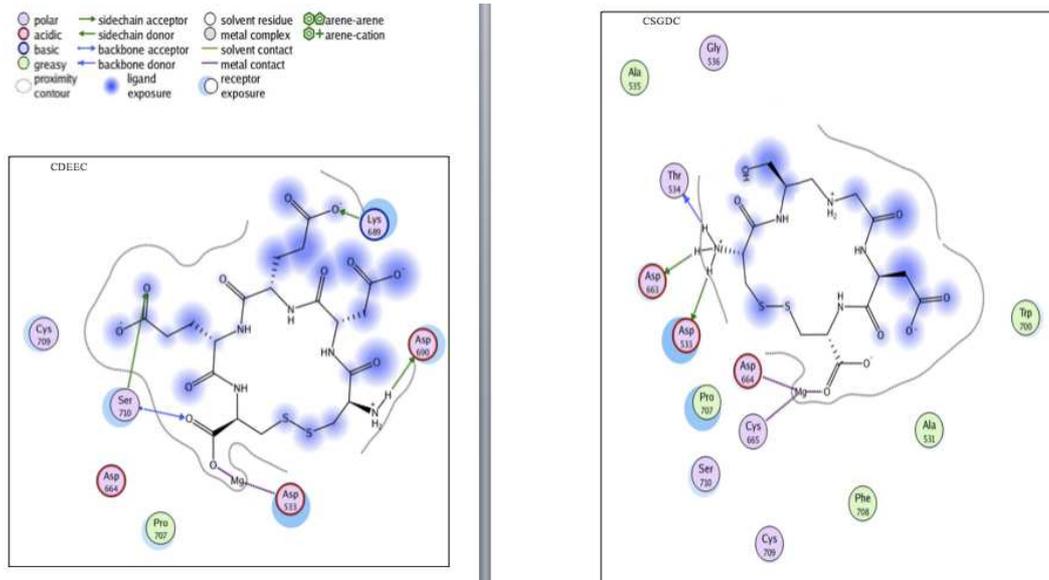Fig. 2: The position of peptide ligand CDEEC at the binding site

Fig. 3: CDEEC and CSGDC ligand interaction

Second batch method:
RdRp enzyme 3D structure: Dengue virus RdRp enzyme structure with ID 2J7U was downloaded from PDB database.

The parameter preparation of RdRp dengue virus enzyme: The preparation was conducted in accordance with the parameters from the first batch which were elimination of water molecule, chlorine ion and polyethylene glycol. These were performed to separate the enzyme from other irrelevant ions, which could hampering the catalytic process. Protonation was conducted to change the macromolecule ionization state with Protonate 3D option. The partial charges addition, hydrogen atom and gas phase solvation were utilized based upon the minimization energy of force field MMFF94x calculation. This process was conducted until the gradient RMS reached 0,05 kkal mol^{-1} Å. Other parameters were set to default values. This enzyme optimization process was performed by using MOE 2008.10 software.

The 3D structure design of cyclic peptide ligand as inhibitor: The tested 3D structure of peptide ligands were drawn by using ACDlabs. The peptide was cyclicized and every residue end was added with cysteine to form the disulfide bridge. The utilized amino acids are arginine, lysine, aspartic acid, glutamic acid, serine and glysine. All of them, except glysine, were chosen in order to form electrostatic and hydrogen interactions. Glysine was chosen in order to enhance the ligand flexibility. Then, the design was saved in MDL Molfile format. The ligand naming was based on the three residues between the cystein residues end. For example, CDEEC was only written as DEE.

Peptide ligand preparation as inhibitor: The ligand optimization was done by using MOE database viewer (dv). Every ligand was 'washed' in order to repair its 3D structure and charged by using MMFF94 force field calculation. The molecular energy structure minimization was done until the RMS gradient reached 0,001 kkal mol^{-1} Å. Other parameters were left at default value.

Peptide ligand docking with RdRp enzyme: The docking simulation was performed by using MOE-dock program. The ligand candidates database was arranged to interact with the chosen enzyme residues. They were Arg-737, Arg-729 and Ser-710. During this process, the enzyme was made rigid and the ligand was left free to rotate. The utilized placement method was triangle matcher, which is useful for generating ligand energy calculation for each $2,5.10^6$ iteration pose. The other parameters were in default values. The applied scoring function was London dG, which is for calculating the binding free energy. The result of this step would be manifested by arranging 10 retain. The next step, refinement, was using force field energy selection calculation with default configuration. The result of this last selection step was only displaying the most suitable molecule based on one retain. The docking result analysis was based on $\Delta G_{binding}$ (S) values. Ligands in population (~20% from the lowest $\Delta G_{binding}$ value) would be analyzed further in drug scan selection. Drug scan step was conducted by accessing http://service.bioinformatik.uni-saarland.de/edrugscan/. The result is a ligand which would be suitable as drugs and would be analyzed further.

Molecular dynamics simulation: The initialization step of molecular dynamics simulation was performed using MOE-dynamic. The utilized data were enzyme-ligand complex from the first batch, which would act as standard and the ligands which were suitable as drugs. The minimization energy calculation was conducted using MMFF94x force field by involving calculation toward solvation energy with Born implisit solvation. The enzyme-inhibitor complex charges optimation was conducted with partial charges option. The other parameter was set on default value, which was ensemble NVT and NPA algorithm for creating ensemble trajectory. The position, velocity and acceleration results were saved each 0,5 ps for further observation.

Ligand preparation: Both ligands that used in this study were derived from first and second batch, which were CDGSC and CDEEC. These two pentapeptide ligands were drawn in two dimensions using ACDLabs and cyclicized with cystein.

Docking phase: RdRp DENV structure needs to be optimized before docking process. This step was conducted in MOE 2008.10. The optimization was performed by changing the structure into its ionization state by protonate3D option, adding partial charge and minimizing the energy until RMS gradient 0,05 reached.

Meanwhile, ligands were also optimized by using MOE database viewer. Ligands were prepared with wash option to get the most favorable structure, next optimization was done by choosing MMFF94x forcefield to control molecular surface potential. Differ from DENV RdRp, minimization of ligands was carried out until RMS.

Docking simulation was performed by MOE-dock. The two ligands were arranged to interact with the selected enzyme residues, which were Arg-737, Arg-729 and Ser-710. These three residues are important residues of DENV RdRp. By choosing gas solvation state, the enzyme was made to be rigid and the ligand was free to rotate to gain the most suitable position. Placement method that was used is triangle matcher with number of return pose $2,5.10^6$ and other parameter was set to default.

Docking analysis: Result of docking simulation was saved in MOE database. This database was then analyzed to study the docking process. Analysis was carried out by comparing the binding energy between ligand and protein from the two ligands.

Molecular dynamics simulation: Before doing MD simulation, enzyme-ligand complex was optimized with partial charge menu and energy minimization was performed until RMS gradient 0.05. To include water in simulation, solvation was set using Born solvation mode. MD simulation was done by choosing MMFF94x forcefield and NVT ensemble with 1 fs time step and sampling every 0.5 ps.

MD simulation was carried out at two different temperatures, 300 K and 312 K. Before main simulation was done, system was initialized by simulating for 30 ps at 300 K. Main simulation was set to 5 ns and the cooling phase takes 20 ps. For simulation at 312 K, heating phase was required and executed for 20 ps.

Molecular dynamics analysis: Analysis of MD result was performed by reviewing molecular dynamics database viewer. Ligands were marked by their residue contact with RdRp DENV and their total potential energy during simulation.

RESULTS AND DISCUSSION

First batch method:
Data mining of DENV-RdRp enzyme in PDB: The Protein 3D structure of RdRp enzyme of dengue virus was sought in Protein Data Bank. Two RdRp dengue virus enzyme structures were found, which are 2J7U and 2J7W. The physical difference between these two structures lies on the availability of ligands on each protein. Protein 2J7U didn't have any ligand, while 2J7W has 3'DGTP ligand. It is an inhibitor of RdRp enzyme on elongation process. The other difference lies in the crystalization resolution, 2J7U crystal has smaller resolution value than 2J7W crystal. The resolution value refers to the electron density. Based on it, crystal

with lowest resolution was chosen, because when the crystal is smaller, the coherence of separation level would eventually much better. Another feature of 2J7U is existence of Mg^{+2} ion, which has an important role in catalytic process, while 2J7W doesn't have it. Based on its amino acid arrangements, there is neither missing amino acids on the middle of its chain, nor breakage of the chain. The chosen enzyme 2J7U was superimposed with 2J7W by using PyMol and the obtained RMSD result is 0,365. RMSD is a likeness parameter between two structure, lower value means better structural likeness. According to Baxevanis (Baxevanis and Oullete, 2001), if the RMSD value is lower than 0,4, then both structures are essentially indistinguishable.

Peptide ligand preparation: The peptide was designed to have negative charge or acids, which involve aspartic acid (D) and glutamic acid (E). This is because negatively charged amino acid would help designed peptide to have strong interaction with the important residues in RdRp enzyme. These residues were Ser-710, Arg-729 and Arg-737, which are more positively charged. The chosen three amino acids on peptide ligands was based on principle that the amount of the amino acids in the peptide chain should be kept limited, in order to make the structure agile enough to pass through the paracelluar way. The combination between aspartic acid and glutamic acid on peptide design resulted 8 cyclic peptides with different amino acids composition. The resulted peptide cyclic ligands were CDDDC, CDDEC, CDEDC, CEDDC, CDEEC, CEDEC, CEEDC and CEEEC. The peptide ligand modeling was conducted by protonating the amino group and deprotonating the carboxyl group on it. This was needed because in physiological pH, the carboxyl group was entirely in R-COO⁻ form, while the amino group was in R-NH3⁺ form. The side chain of carboxyl group on aspartic acid residue and glutamic acid was deprotonated as well.

The docking result analysis: The docking process was conducted 50 times for each peptide ligand. The objective is to form 50 different conformations when peptide ligand binds to the enzyme. The AutoDock program will classify the same conformation in one cluster.

If the cluster has the most population, then it could be inferred that the cluster conformation was more favorable for ligand binding with its binding site. The most populated cluster should be the first rank. Cluster rank showed the ligand conformation group with the lowest $\Delta G_{binding}$ after several times of docking process. Low $\Delta G_{binding}$ values signify that the peptide ligand was

in the most stable conformation when bound with enzyme (The binding chance of 80-90%). When the most populated cluster was in the first cluster rank, the ligand-enzyme conformation is the most stable.

Based on existing data, more than half of the ligand fulfilled the most stable ligand conformation when they bound with the enzyme. There was some ligand, which ousted from the criteria, because they had dubious conformation. The ligands which fulfilled the criteria were CEEDC, CEDEC, CDEEC, CDEDC and CDDDC. After the ligands selection, the next process was to evaluate docking free energy value. If the rotatable binding value was smaller, the $\Delta G_{torsional}$ would be decreasing as well. When the rotatable bonds amount decreased by one point, $\Delta G_{torsional}$ would also decreased with constant value ~0.27 kkal mol^{-1}. This is in accordance with elimination of methylen group.

The $\Delta G_{intramolecular}$ values was affected by bond length, bond angle and dihedral angle of the ligand molecules. Based on data above, there is tendency that if the $\Delta G_{intramolecular}$ is increased (near positive value), the rotatable bond amount will be smaller. Then, we find a residue on enzyme which has ligand contact, by using Chimera program. Based on residual contact evaluation, it was perceived that those five ligands have contact with 3 important binding site residue. They were Ser-710, Arg-729 and Arg-737. The docking result showed that CDEEC ligand has the lowest $\Delta G_{binding}$ value among the others. It has the most residue contact, with total of 13 residues. Two of them were Asp-663 and Asp-664, which are catalytic site residues. It is expected that this ligand could deter the catalytic process. CDEEC peptide has hydrogen bond interaction with five other residues and they binded with Arg-729 and Arg-737. Besides of having hydrogen bond interaction with those residues, CDEEC peptide ligand was forming salt bridge with COO$^-$ group side chain with Asp-633. The salt bridge interaction is considered important for $\Delta G_{intermolecular}$ value, because its stabilisation value is stabilizing the hydrogen bond.

Based on docking result visualization, it is known that CDEEC peptide ligand was bound with RdRp enzyme inside the cavity. It is viral RNA entry when it want to begin initiation and elongation (NTP Tunnel).

It was inferred from the docking result, that the cyclic peptide ligand with CDEEC combination (Cys-Asp-Glu-Glu-Cys) could be applied as potential inhibitor to block the RdRp enzyme activity. The supporting conditions are as following:

- It has the lowest binding energy value among the ligands when bound with RdRp enzyme, which is -10,04 kkal mol^{-1}

- It has Ki value of nM scale (43,44 nM), indicates that stable peptide ligand-enzyme complex was formed
- It has the most contact with other residues and includes contact with catalytic site, Asp-663 and Asp-664, also Ser 710, Arg-729 and Arg-737, which have strong influence on RNA virus initialization
- It has hydrogen bond with some residues, for example with Arg-729 and Arg-737 and forming salt bridge interaction with Asp-663

Second batch method:
Searching 3D RdRp enzyme structure: Enzyme structure was downloaded from PDB. We found two RdRp enzyme structure, 2J7U and 2J7W. Downloaded RdRp enzyme structure was 2J7U, which is also used in the first batch as well.

Preparation of RdRp enzyme structure: The 2J7U crystal was prepared by eliminating water molecule, chlorine ion and polyethylene glycol by using MOE sequence editor. The water molecule was eliminated as a precondition for docking simulation. While chlorine ion and polyethylene glycol are additional molecules which was trapped when crystalization process. The docking simulation was executed after these molecules had been removed. The remaining molecules were the amino acids, ligands or enzyme cofactor (Zn^{2+} and Mg^{2+} ions). The next treatment was to convert the enzyme into the protonated state, by using protonate 3D. This protonate 3D application was utilized to change enzyme into ionized state level and exposing the position of hydrogen atom on the crystal. The existence of hydrogen atom was indispensable for the molecular mechanics process, molecular dynamics and electrostatic interaction calculation. However, most of the crystal structure didn't have hydrogen atom coordinate, because of the resolution limitation. The other reason was the existence of hydrogen atom and the ioinization level of certain group would affect the crystalization process. After protonation, the ions inside the enzyme could be seen clearly in accordance to its charges.

The next optimization proces was minimizing the enzyme energy calculation by appropriate force field with the system parameters, which is potential setup of MMFF94x. The minimization was done in order to remove bad contact, or high energy steric effect. The potential setup arrangement adjusted the hydrogen atom partial charge. It is useful to count potential electrostatic energy calculation. The applied salvation type for docking was gas phase. Energy minimization

process to remove bad contact was carried out in the absence of solvent. Then, MMFF94x force field were applied, until the RMS gradient had reached 0,05 kkal/mol Å, which was the most suitable value for protein. The other utilized parameters of MOE.2008.10 were set at default value.

The design of the cyclic peptide 3D structure as inhibitor: According to Yagi *et al.* (2007), there are six main amino acid as ligand candidate toward the target residue with total positive charge (arginine and lysine), they are: arginine, lysine, aspartic acid, glutamic acid, serine and glysin. The first five were chosen in order to form electrostatic and hydrogen interaction. Glysine was chosen in order to increase the ligand flexibility. The sequence of these six amino acids were combined in order to form cyclic pentapeptide.

Two amino acid at the end of the cyclic pentapeptide are cysteine. It was chosen to form cyclic disulfide bond. Shorter peptide chain (5 amino acids) and the cyclic bond were needed to improve its stability and delivery rate. The disulfide bond could stabilize protein until temperature above 100°C, by decreasing the entropy of protein randomness, or entropy effect (Nurbaiti *et al.*, 2010). Based on this explanation, it was expected that ligand candidate could reach the target with low rate of hydrolysis effect. Three other amino acids were the combination of six amino acids (Arg, Lys, Asp, Glu, Ser, Gly). Then, we obtained 216 cyclic peptapeptide ligands as inhibitor candidate. This candidate ligands would be named in accordance to its first character. These ligand candidates were drawn by ACDlabs in 2D zwitter ion. The result of the picture would be converted to 3D optimization (in 3D viewer), then it would be saved in MDL Molfile format.

Peptide ligand preparation as inhibitor: Ligand optimization by MOE wash parameter was performed, in order to repair ligand's structure and adding the explicit hydrogen atom. Wash function was applied in 2D structure to standarize length and bond angle. This was related to its potential energy, to achieve equilibrium state. The ligand was added with partial charge. Optimization was performed by applying energy calculation using MMFF94 force field. This is a calculation energy parameter for small organic molecules.

Enzyme and candidate ligand docking: The docking process of 216 ligand candidates with RdRp enzyme were carried out using MOE 2008.10. Enzyme was made rigid and ligand was left free to rotate (flexible docking). This was necessary to ensure that the docking

process was in accordance to lock and key mechanism. Docking was comprised of several steps. Couple of methods were available on each of it and new method could be easily added. Those steps were conformation analysis, placement, rescoring and refinement. The conformation analysis was done for observing the desired conformation on binding. MOE-Dock was conducting ligand conformation search by using every possible angle combination and the result was almost 5000 conformation. The placement method gave the ligand conformation pose. The utilized method was the default one, the triangle matcher. It was useful for producing the ligand energy calculation of each pose iteration. The maximum amount of ligand conformation evaluation pose were $2,5.10^6$ poses. This method was conducting random process on active side ligand pose in order to determine the optimal binding orientation.

The binding free energy calculation of binding orientation was utilizing London dG scoring function, with 10 times retain and no duplication. Retain was conducted in order to arrange the best ligand conformation amount. The ligand pose from placement step could be fixed in the refinement step. Refinement was the final evaluation step of free energy, by using force field Generalized Born Solvation Model (GB/VI). Force field refinement was much more accurate than GridMin, which utilized electrostatic calculation on minimization process. Henceforth, the process would eventually takes longer. The default arrangement of force field refinement was using pocket cut off 6 Å. It is the receptor distance which was applied in docking process. And the last retain was set to 1.

Analysis of binding free energy ($\Delta G_{binding}$): S is the final total calculation of the docking step, which represents $\Delta G_{binding}$ in kkal/mol. S value is equal with E_{refine}.

E_{refine} was represented in following equations:

$$E_{(x)} = E_{str} + E_{ang} + E_{stb} + E_{oop} + E_{tor} + E_{vdw} + E_{ele} + E_{sol} + E_{res}$$

Where:
E_{str} = Bond streching energy
E_{ang} = Angle energy
E_{stb} = Streching-bending energy
E_{oop} = Out-of-plane energy
E_{tor} = Torsional energy
E_{vdw} = Van der Waals energy
E_{ele} = Electrostatic energy
E_{sol} = Solvation energy
E_{res} = Restraint energy

Solvation energy calculation was removed, because the docking process was in gas phase. $\Delta G_{binding}$ (S) represents the strength between enzyme and ligand. When the value of $\Delta G_{binding}$ is decreasing, then the binding complex of Enzyme-ligand will be stronger.

Candidate ligands drugscan: Drugscan is the fastest measure for evaluating drug-like or lead-like compounds. This process was done in order to decrease the spending on screening process. The process could be hampered, when failure on ADMET (absorbtion, distribution, metabolism, excretion and toxicity) process was occured. The drugscan process was performed online, http://service.bioinformatik.uni-saarland.de/edrugscan/. The available 16 ligands were uploaded to that site for the screening process. The utilized screening was based on lipinski's rule of five (RO5).

RO5 is a screening parameter for predicting the drug oral consumption. It is safer than other two entry points (dermal and rectum), because it is meant to decrease the drugs side effect and the microorganism infections. It is suggested, that drug candidate must have acceptable ADMET properties in order to pass the first step clinical trial. The rules are molecular weight of maximum 500 g/mol, donor and acceptor hydrogen must not more than 5 and 10, C Log P (Calculated LogP) must be less than 5 (Lagorce *et al*., 2008).

The drug oral consumption must pass through the intestine's wall and transported into the blood stream. Then, it would penetrate the cell wall. Octanol is a model compound for the cell membrane, which would eventually helpful for determining the lipofilicity of drug molecule. Lipofilicity and partition coefficient (Log P) is a measurement to represent solutability and Log P value. The other precondition is drug molecular mass. If the molecular mass of drug compound is smaller, then it would aid the robustness of its difusion. 80% of drugs were having molecular mass less than 450 g mol^{-1}.

The drug molecules should be easily dissolved in the water, for smoothing its transportation in blood or cell's liquid. The solvability in water could be estimated from the amount of the donor hydrogen compared with the side alkyl chain. When we have more donor hydrogen, then it would be easier to dilute it. However, it would make the penetration toward cell membrane more difficult. The hydrogen bond was formed between three atoms, one hydrogen atom and two electronegative atom (Usually N or O atom). Hydrogen donor was hydrogen atom which had covalent bond with electronegative atoms. The screening result toward 16 ligands only resulted in one best ligand, CSGDC.

This ligand was in the seventh rank of G binding scores and having scores lower than CDEEC ligand.

Residual Contact analysis toward candidate ligands: We analyzed ligands with the lowest free binding energy, one that fulfilled drugscan precondition (CSGDC) and ligand resulted from the first batch (CDEEC). Based on the residual contact, both ligands had interaction with active site. It was close to the vicinity of Asp-533, Asp-663 and Asp-664 amino acids. The amount of CDEEC ligand residual contact was smaller than CSGDC. This indicates that binding of CSGDC to RdRp enzyme was preferable than CDEEC ligand.

Interaction analysis of candidate ligand toward RdRp enzyme: The ligand interaction toward RdRp enzyme was slightly different. CSGDC ligand was nearer to active site, while CDEEC was further from it. The active binding site was hydrophobic and CSGDC ligand could bind stronger because of its higher hydrophobicity than CDEEC ligand. There were 3 Hydrogen bonds with CSGDC, while CDEEC formed 4 Hydrogen bonds. These hydrogen bonds expressed the strength of the interaction complex. Although the amount of hydrogen bond between CSGDC and enzyme was fewer than the CDEEC complex binding, the CSGDC hydrogen bond occurred on two enzyme catalytic residues.

Two residues formed hydrogen bond with CSGDC ligand (asp-533 and asp-633) represent a general picture, that ligand was binded with the active site. Asp-533 was acting as general base and 3'hydroxil NTP group deprotonation and Asp-633 gives the best geometry for catalysis reaction.

Both ligands has interaction with Mg^{2+} ion. The function of this ion was still unknown, because it was found on non-catalytic position when the enzyme was activated. However, the calcium ion which has no role on WNV catalysis, were located at the same position. Those ions could have certain role on de novo initiation mechanism for facilitating the movement of nascent ds RNA, after the formation of first two nucleotides outside the active site. Based on both ligands interaction data, it could be seen that CSGDC ligand has larger potential to bind with enzyme active site. The existence of ionic bond between ligand and Mg^{2+} ion signifies that there was blocking interaction against the movement of nascent ds RNA.

Molecular dynamics simulation: Molecular Dynamics Simulation (MDS) of DEE and SGD ligands were performed only on initialization step.
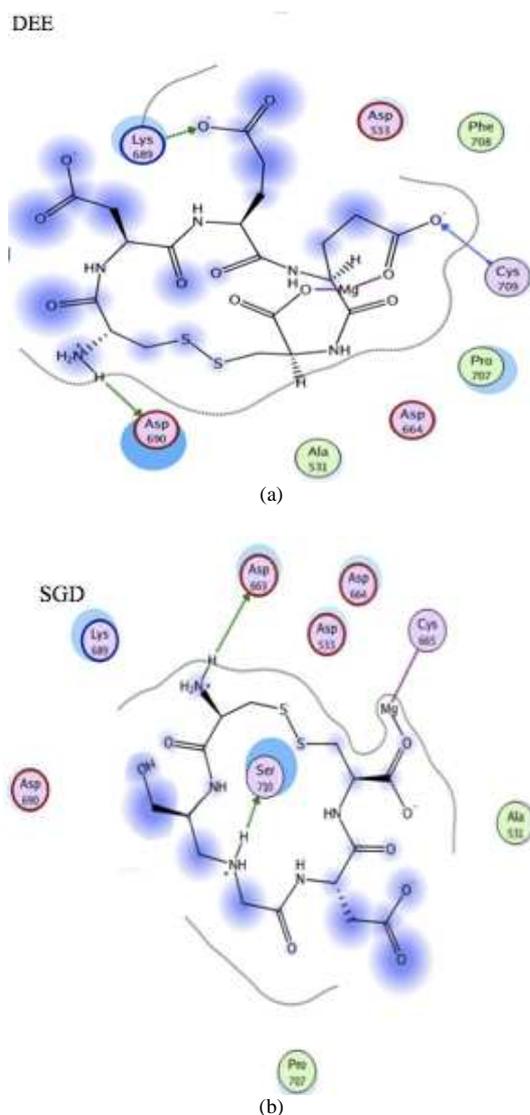
(a)

(b)

Fig. 4: Ligand interaction of (a) CDEEC and (b) CSGDC after 100 ps

This MDS preparation step was performed using enzyme-ligand complex which produced from docking simulation. The utilized enzyme-ligand complex was added with partial charges, optimized and minimized by using energy calculation (force field MMFF94x). The difference with the gas phase docking simulation, MDS was using born implicit solvation. It was necessary to appoint the solvent condition as medium and keep it distanced with simulation process. The born solvation was the only available solvation type on MOE.2008.10 which includes E_{sol} calculation on the system. It means that the simulation was conducted by using the solvent.

The utilized statistic specification for simulation would produce ensemble based conformation. The applied parameters were set to MOE defaut value, which were ensemble NVT (N, total atom; V, volume; T, temperature), constant temperature 300K and 101kPa pressure. This parameter was useful because in real experiment, it is much easier to adjust temperature. The employed NPA Algorithm was the most accurate and sensitive algorithm and it could set up the ensemble correctly.

The position, velocity and simulation acceleration was saved each 0,5 ps, until 100 ps. The simulation observation was done by examining the enzyme-ligand complex interaction between ligand atom with enzyme atom on the end of simulation (100 ps). The initialization of simulation time arrangement was based upon previous research (Balatsos *et al*., 2009). By the end of simulation, the system has not yet reached equilibrium state. The objective of MDS on this research is to conduct observation on ligand interaction.

In Fig. 4 (supplementary material), the observation of MDS toward both ligands showed that CSGDC ligand still had interaction with one active residue side, Asp-663. This was different in CDEEC ligand, which didn't has any interaction with the active site.

Based on drugscan and docking analysis, it could be inferred that cyclic peptide ligand CSGDC (Cis-Ser-Gli-Asp-Cis) could be appointed as potential competitive inhibitor toward RdRp enzyme active site, based upon these informations:

- Lower binding free energy compared with CDEEC ligand, which was -29.6122 kkal/mol
- Fulfill Lipinski rule of five
- More residue contacts than CDEEC ligand, which were Asp-533, Thr-534, Asp-663, Ala-531, Ala-535, Gly-536, Asp-664, Trp-700, Pro-707, Phe-708, Ser-710, Asp-664 and Cys-665
- Hydrogen bond with A motive (Asp-533) and C (Asp-663) on the enzyme catalytic site
- Ionic binding with Mg^{2+} ion, which would inhibit the movement of nascent ds RNA
- Based on MDS simulation, CSGDC ligand still had interaction with one active site residue, Asp-663

Ligand preparation of first and second batch: Both ligands from first and second Batch were drawn in two dimension using ACDlabs. Three dimensional optimization option was then performed to the ligands in order to get the most reliable ligands structure. Ligands were saved in MDL Molfile format and converted into MDL Mol using VegaZZ because MOE can only read ligands in MDL Mol format.

Table 1: Binding energy result

Ligand	S ($\Delta G_{binding}$) Kkal/mol
CDEEC	-24.2558
CDGSC	-35.7418

Docking phase: Docking between DENV RdRp with CDEEC and CDGSC was performed by choosing MOE-dock. The ligands was arranged to interact with three important residues of DENV RdRp, which are Ser-710, Arg-729 and Arg-737. These three residues play an important role in DENV life cycle by initiating replication with de novo mechanism. A high concentration of 3'dGTP is required for de novo initiation and within DENV RdRp, moiety of tP from 3'dGTP is coordinated by the three important residues mentioned before. We hoped that binding occurred between these three important residues with the proposed ligands, 3'dGTP can't hold onto DENV RdRp and as a result, initiation will not occur. As mentioned previously, at this phase enzyme was made rigid and ligands can move freely to get the most suitable conformation. With this setting, docking simulation consumes less time.

To get the best conformation, triangle matcher was chosen as placement method. Triangle matcher was used to place ligands at active sites based on charged group and spatial fit. Triangle matcher shows random ligand's movement at enzyme's active site to produce best bonding orientation (Cook *et al.*, 2009).

To analyze docking phase result, the binding energy and ligand interaction was reviewed.

Table 1 shows that both CDEEC and CDGSC had negative binding energy. These data shows that ligand's conformation obtained at enzyme-ligand complex was in the most favorable conformation. Lower binding energy that is shown by CDGSC indicates that interaction between CDGSC with DENV RdRp was more preferred than interaction between DENV RdRp with CDEEC.

In many recent papers, authors claimed that the scoring functions describing molecular docking experiments couldn't be quantitatively correlated to biological activities (Mazur *et al.*, 2010). Thus, in Fig. 5 we presented the two dimensional ligand-receptor diagrams, which allows one to qualitatively observe the binding between ligand and targeted residues.

From Fig. 5, it can be seen that both CDEEC and CDGSC had interaction that points to active sites located near Asp-663 and/or Asp-664. These two amino acid residues are member of GDD catalytic site. CDEEC bound with one of three important residues of DENV RdRp, which was Ser-710. While CDGSC interacted with other important residue, which was Arg-

729. CDGSC also made contact with Ser-710 although no binding occurred.

Figure 5 also shows that both ligands formed ionic bonding with Mg^{2+}. The function of this ion in DENV RdRP hasn't been known yet but the crucial role of Mg^{2+} ions in the catalysis of phosphodiester bond formation has long been known (Van Dijk *et al.*, 2004). Therefore, we proposed that binding with this ion strengthen inhibitory effects of ligands.

Molecular dynamics simulation: Besides getting the most reliable conformation of complex DENV RdRp with ligands, Molecular Dynamics Simulation (MDS) were also conducted as a refinement from docking result. MDS was not conducted only for complex of DENV RdRp with ligands, DENV RdRp was also simulated to visualize any differences occurred by adding ligands to enzyme system.

NVT ensembles were chosen and the experiment was done at constant temperature.

MDS for drug design was ideally conducted at temperature range 300-314 K. This temperature range was the possible range for human body temperature. In this study, MDS was done at 300 K and 312 K. 312 K is a temperature when human gets fever and as we all know, an early symptom of dengue is fever.

Molecular dynamics analysis: There are several ways to analyze MDS result. In this study we were concerned to review the total potential energy plot of complex conformation and ligand interaction to study the interaction between ligands and DENV RdRp. In proteins, conformation refers to a three dimensional arrangement from a group of atoms that can be changed without altering covalent bond (Bhagavan, 2002).

Total potential energy plot could be used to overview system conformation changes during simulation. Any damage to enzyme structure such as denaturation will affect total potential energy plot. From both of Fig. 6 and 7 (Supplementary material), all system gave similar plot during simulation. It means that complexion with proposed ligand did not damage enzyme's structure. Enzyme was stable with or without ligands complexes on it. If we compare 300 K and 312 K plots we can also see that the plots are similar. It means that at 312 K, enzyme still has activity and no structural damage occurred.

By reviewing only total potential energy plot, we couldn't compare our proposed ligands because both of them give similar plot. Therefore, we analyzed ligand interaction after simulation to find out which ligand is better. During simulation, thousands possible complex conformations were being examined and we could assume that the last one was the best conformation.
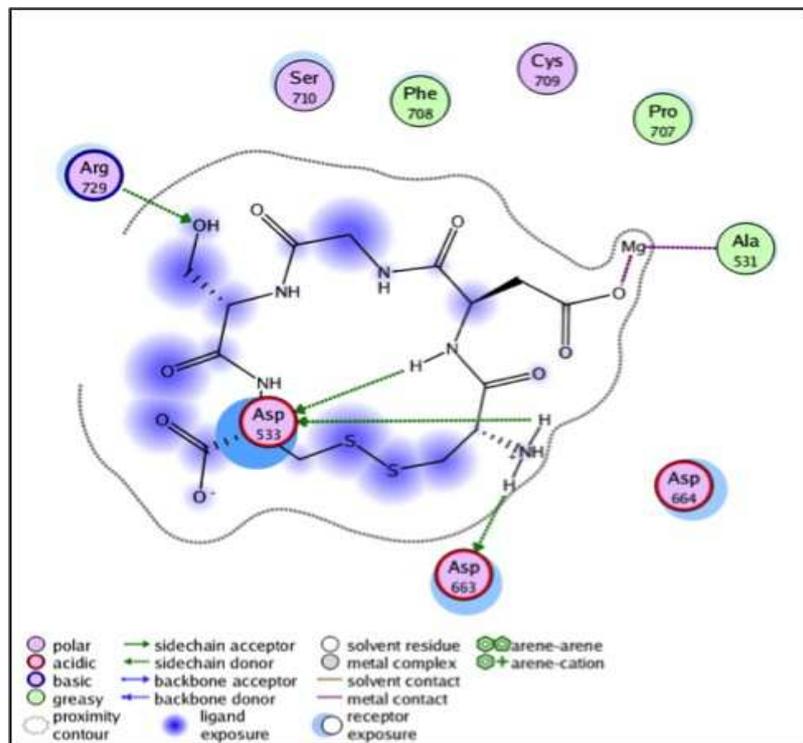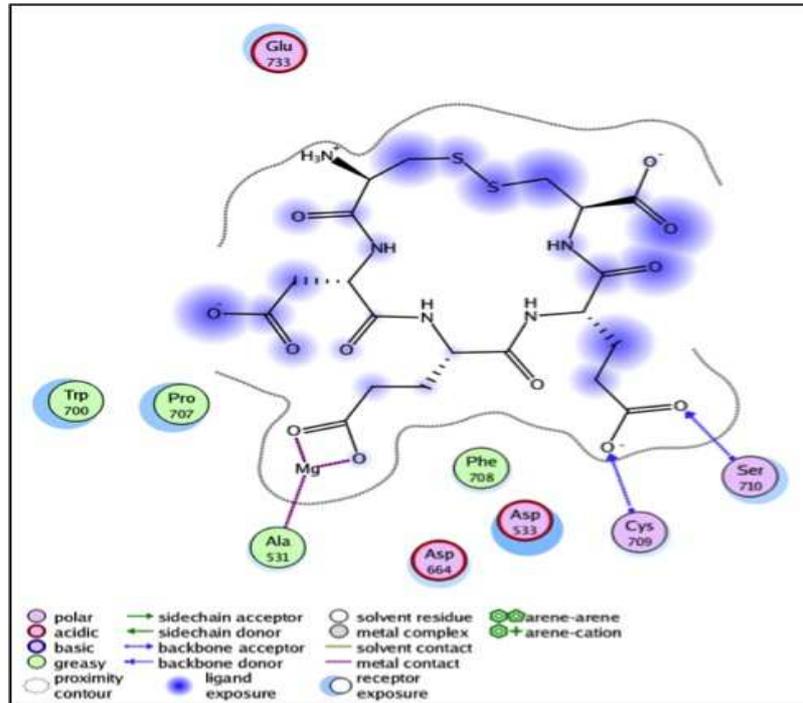
(a)

(b)

Fig. 5: Show interaction between DENV RdRp with: (a) CDEEC; (b) CDGSC
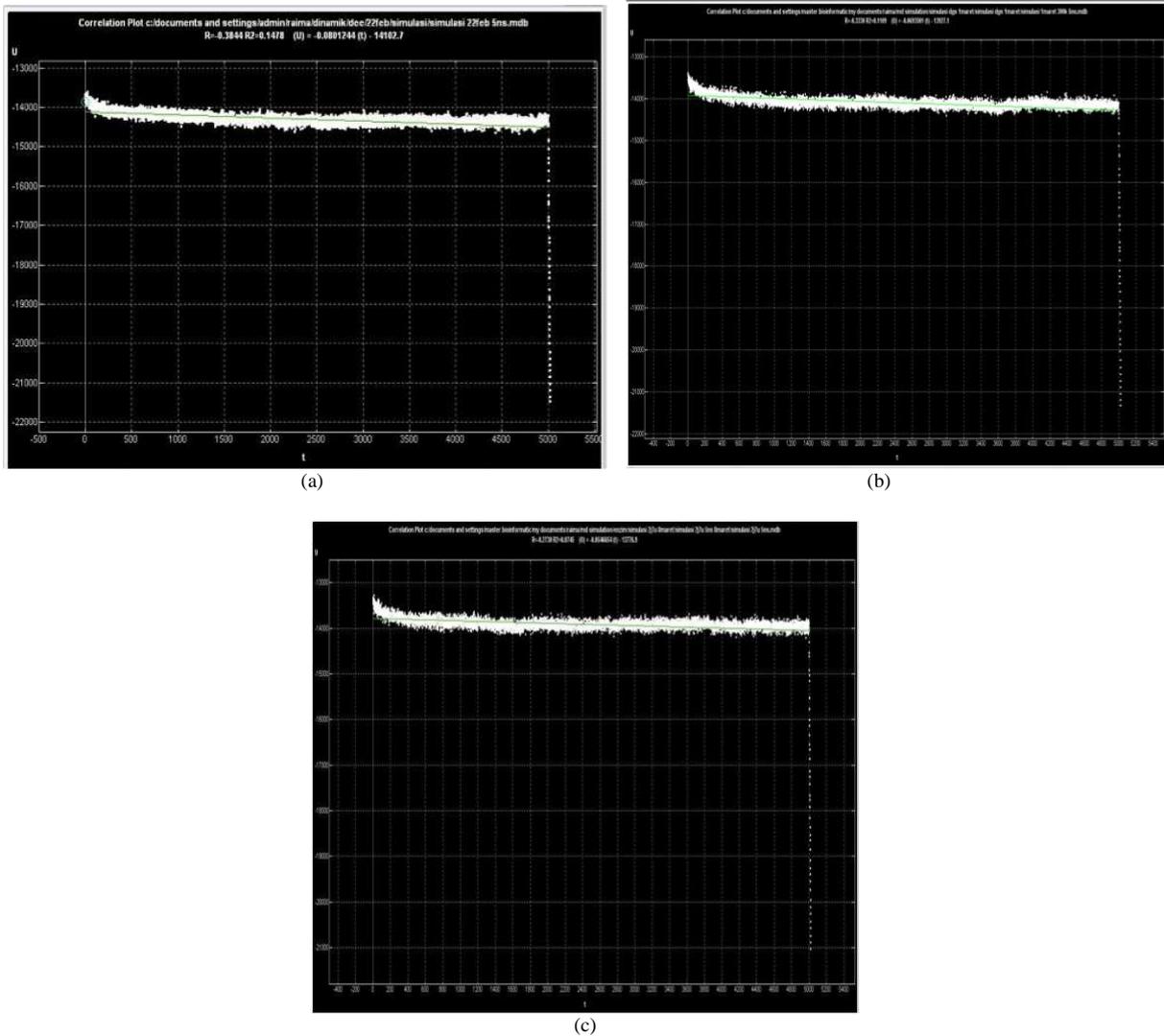
(a)

(b)

(c)

Fig. 6: Total potential energy plot during MDS at 300 K (a). DENV RdRp; (b) DENV RdRp complex with CDEEC; (c) DENV RdRp complex with CDGSC

Figure 8 shows interaction between both proposed ligands with DENV RdRp. We arranged ligands to contact with three important residues. We can see from the diagram that there were changes in ligand's orientation during simulation. These changes were probably caused by the effect of solvent in dynamic movement of enzyme.

After simulation phase, CDEEC showed seven hydrogen bonds with DENV RdRp residues. CDEEC bind with two important residues of DENV RdRp, which were Arg-737 and Arg-729. In this simulation phase, CDEEC also maintained its ionic bond with Mg^{2+}. CDGSC showed a little bit different result. It has

ionic bond with Mg^{2+} but it doesn't have any contact with any important residues of DENV RdRp.

Ligand interaction after MDS at 312 K can be seen on Fig. 9. Like MDS 300 K, CDEEC showed hydrogen bond with Arg-737, but no longer bound with Arg-739. We can see from the diagram that although CDEEC didn't bind Arg-739, it bounds another important residue, Ser-710. Meanwhile CDGSC also showed the same result like previous simulation and didn't bind any important residues of DENV RdRp. At 312 K MDS, we can see that both CDEEC and CDGSC maintain their ionic bond with Mg^{2+}.
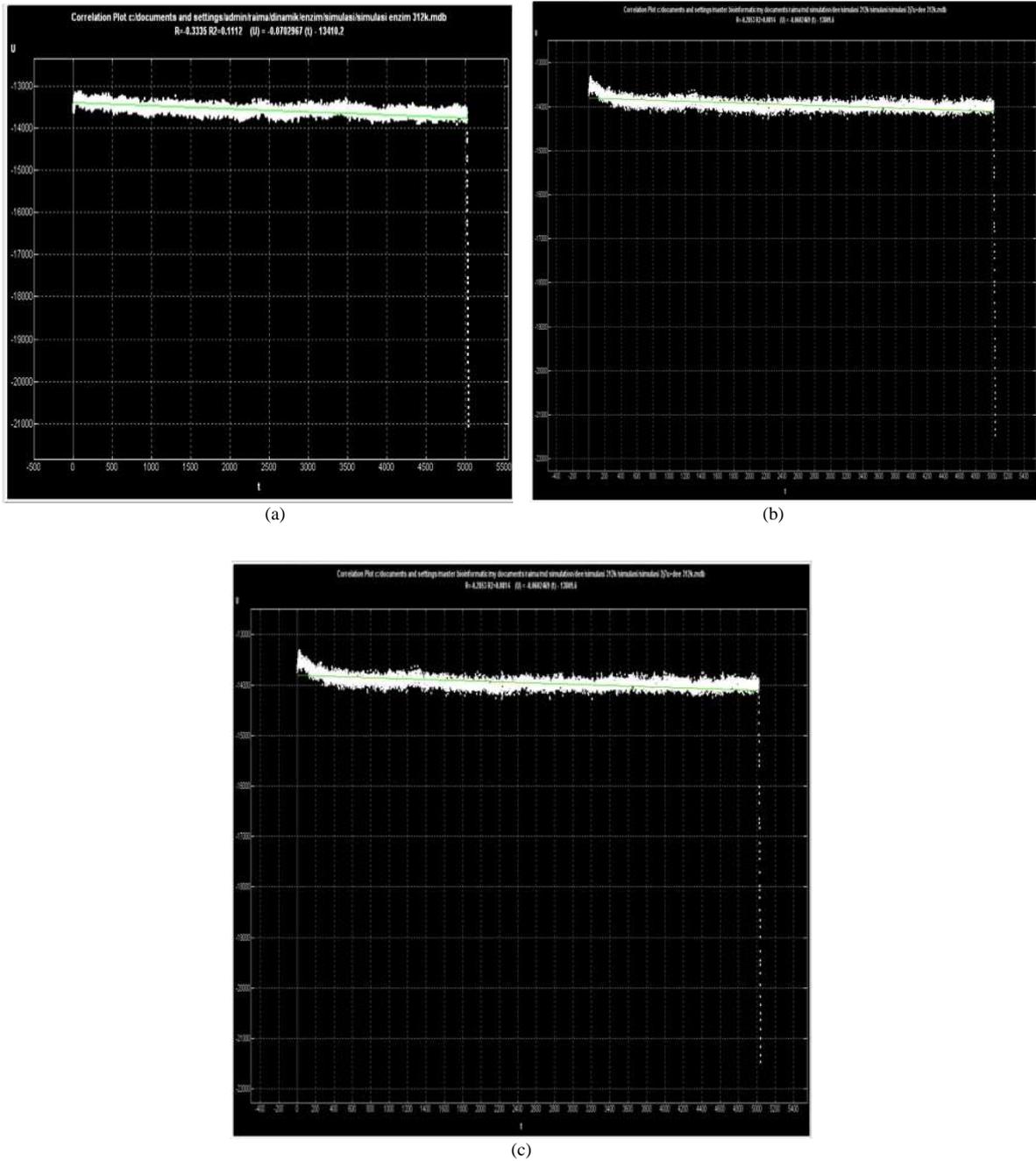
59

(a)

(b)

(c)

Fig. 7: Total potential energy plot during MDS at 312 K (a). DENV RdRp ; (b) DENV RdRp complex with CDEEC; (c) DENV RdRp complex with CDGSC

To make it easier to understand ligand interaction in various phase, we can see data at Table 2. Table 2 shows comparison of ligand interaction during different phase of simulation. According to Table 2, CDEEC always interacted with one or two of three important residues of DENV RdRP. While CDGSC only showed interaction with important residues of DENV RdRp at docking phase only.
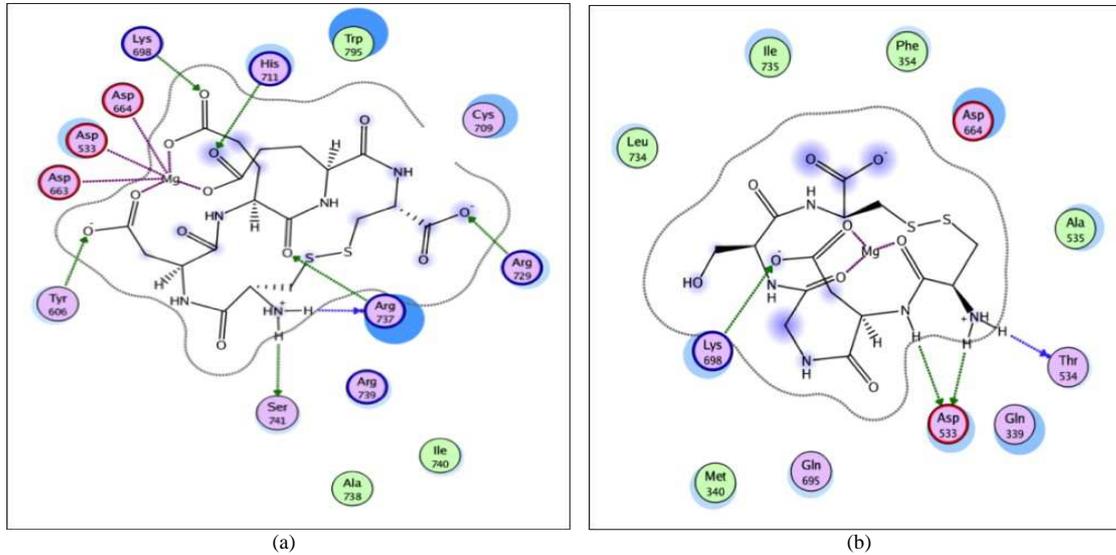
Fig. 8: Ligand Interactions after MDS at 300 K between DENV RdRp with; (a) CDEEC; (b) CDGSC
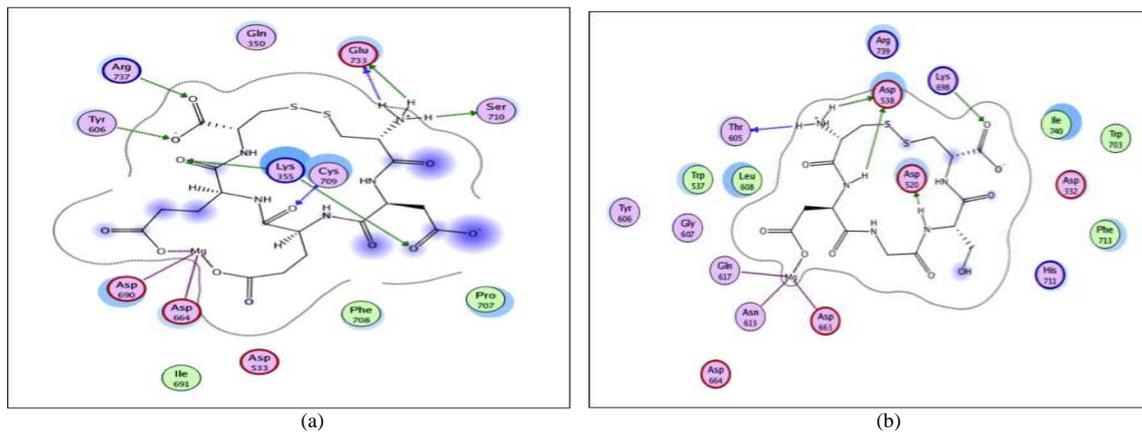

Fig. 9: Ligand Interaction after MDS at 312 K between DENV RdRp with; (a) CDEEC; (b) CDGSC

Table 2: Ligand interaction comparison

Ligand	Docking	300 K MDS	312K MDS
CDEEC	Ser-710, Cys- 709, Ala-531, Mg, Pro-707, Trp- 700, Glu-733	Arg 729, Arg-737, Ser-741, Tyr-606, Lys-698, His-711, Mg (Asp-533, Asp-663, Asp-664) Trp-795, Cys-709, Arg-739, Ile-740, Ala-738	Glu-733, Ser-710, dgn Lys-355, Cys-709, Arg-737, Tyr-606, Mg (Asp-664, Asp-690) Gln-350, Pro-707, Phe-708, Asp-533, Ile-691
CDGSC	Arg-729, Asp-533, Ala-531, Mg, Asp-663, Asp-664, Ser-710, Phe-708, Cys-709, Pro-707	Thr-534, 2 Asp-533, Mg, Lys-698 Ile-735, Phe-354, Asp-664, Ala-535, Gln-339, Gln-695, Met-340, Leu-734	Thr-605, Asp-538, Lys-698, Asp-520, Mg (Gln-617, Asp-663, Asp-613) Arg-739, Ile-740, Trp-703, Asp-332, Phe-713, His-711, Asp-664, Gly-607, Tyr-606, Leu-608

CONCLUSION

DENV RdRp succeeded in maintaining its stable three-dimensional conformation during simulation and ligands did not affect DENV RdRp stability. The first and second batches method produced two different optimized ligands and methodological comparison was conducted to determine which one is the best. Docking result showed that CDGSC is better than CDEEC due to lower binding energy. This indicates that binding between DENV RdRp with CDGSC is preferred over CDEEC. However, MDS showed different result with

docking phase. After both simulation at 300 and 312 K, CDGSC didn't bind any important residues. Meanwhile CDEEC bind with two important residues. At 300 K, CDEEC binds with Arg-737 and Arg-729. At 312 K, CDGSC binds with Arg-737 and Ser-710. We can assume that the presence of solvent affect CDGSC interaction with DENV RdRp. From total potential energy plot, we can infer that both ligands inhibit DENV RdRp by interaction with its important residues without affecting the stability of the enzyme.

ACKNOWLEDGEMENT

REFERENCES

Alonso, H., A.A. Bliznyuk and J.E. Gready, 2006. Combining docking and molecular dynamic simulations in drug design. Med. Res. Rev., 26: 531-568. DOI: 10.1002/med.20067

Balatsos, N.A., D. Vlachakis, P. Maragozidis, S. Manta and D. Anastasakis *et al*., 2009. Competitive inhibition of human poly(A)-specific ribonuclease (PARN) by synthetic fluoro-pyranosyl nucleosides. Biochemistry, 48: 6044-6051. DOI: 10.1021/bi900236k PMID: 19472977

Baxevanis, A.D. and B.F.F. Oullete, 2001. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. 2nd Edn., John Wiley and Sons, New York, ISBN: 0471383910, pp: 470.

Bhagavan, N.V., 2002. Medical Biochemistry. 4th Edn., Harcourt/Academic Press, San Diego, ISBN: 0120954400, pp: 1016.

Cook, I.T., T.S. Leyh, S.A. Kadlubar and C.N. Falany, 2009. Structural rearrangement of SULT2A1: effects on dehydroepiandrosterone and raloxifene sulfation. Horm. Mol. Biol. Clin. Invest., 1: 81-87. DOI: 10.1515/HMBCI.2010.012

Franca, T.C.C., A. Wilter, T.C. Ramalho, P.G. Pascutti and J.D. Figueroa-Villar, 2006. Molecular dynamics of the interaction of *Plasmodium falciparum* and human serine hydroxymethyltransferase with 5-formyl-6-hydrofolic acid analogues: design of new potential antimalarials. J. Braz. Chem. Soc., 17: 1383-1392. DOI: 10.1590/S0103-50532006000700028

Lagorce, D., O. Sperandio, H. Galons, M.A. Miteva and B.O. Villoutreix, 2008. FAF-Drugs2: Free ADME/tox filtering tool to assist drug discovery and chemical biology projects. BMC Bioinformatics, 9: 396-396. DOI: 10.1186/1471-2105-9-396

Mazur, P., T. Magdziarz, A. Bak, Z. Chilmonczyk and T. Kasprzycka-Guttman *et al*., 2010. Does molecular docking reveal alternative chemopreventive mechanism of activation of oxidoreductase by sulforaphane isothiocyanates? J. Mol. Model., 16: 1205-1212. DOI: 10.1007/s00894-009-0628-5 PMID: 20024690

Nurbaiti, S., H. Nagao, H. Saito, R. Hertiadi and M.A. Martoprawiro *et al*., 2010. Domain motions of Klenow-like DNA polymerase I ITB-1 in the absence of substrate. Int. J. Integrative Biol., 9: 104-110.

Qi, R., Z. Ling and C. Chi, 2008. Biological characteristics of dengue virus and potential targets for drug design, Acta Biochim. Biophys. Sin., 40: 91-101. DOI: 10.1111/j.1745-7270.2008.00382.x

Sampath, A. and R. Padmanabhan, 2009. Molecular targets for flavivirus drug discovery. Antivir. Res., 81: 6-15. DOI: 10.1016/j.antiviral.2008.08.004

Tomlinson, S.M., R.D. Malmstrom and S.J. Watowich, 2009. New approaches to structure-based discovery of dengue protease inhibitors. Infect Disord. Drug Targets, 9: 327-343. PMID: 19519486

Van Dijk, A.A., E.V. Makeyev and D.H. Bamford, 2004. Initiation of viral RNA-dependent RNA polymerization. J. Gen. Virol., 85: 1077-1093. DOI: 10.1099/vir.0.19731-0

Yagi, Y., K. Terada, T. Noma, K. Ikebukuro and K. Sode, 2007. *In silico* panning for a non-competitive peptide inhibitor. BMC Bioinformatics, 8: 11-11. DOI: 10.1186/1471-2105-8-11 PMID: 17222344

Yap, T.L., T. Xu, Y.L. Chen, H. Malet and M.P. Egloff *et al*., 2007. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. J. Virol., 81: 4753-4765. DOI: 10.1128/JVI.02283-06

Yin, Z., Y.L. Chen, R.R. Kondreddi, W.L. Chan and G. Wang *et al*., 2009. N-sulfonylanthranilic Acid Derivatives as allosteric Inhibitors of Dengue Viral RNA-dependent RNA polymerase. J. Med. Chem., 53: 7934-7937. DOI: 10.1021/jm901044z