A Computational Alanine Scanning of Heptapeptide RRRSAGM Targeting Dengue NS2B/NS3 Protease

Ahmad Ghazali Ismail1, Nadeem Abdalsatar Abdalrazaq1,2, Mohd Faizul Ridzal Mohd Rawi3, Suriani Mohamad1, Habibah A. Wahab1 and Ezatul Ezleen Kamarulzaman1,*

1School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
2College of Pharmacy, University of Uruk, Baghdad 10075, Iraq
3Malaysian Institute Pharmaceutical and Nutraceutical Malaysia, National Institutes of Biotechnology Malaysia, Halaman Bukit Gambir, 11700 Gelugor, Penang, Malaysia

(*Corresponding author’s e-mail: ezatulezleen@usm.my, ezatulezleen@gmail.com)

Received: 28 March 2021, Revised: 14 July 2021, Accepted: 14 August 2021

Abstract

The pathogenic dengue virus (DV) transmitted mainly by the Aedes aegypti mosquito has been the most rapidly spreading mosquito-borne viral disease and became a serious threat to global health issue in tropical and sub-tropical countries. The clinical symptoms ranging from mild to a fatal dengue hemorrhagic fever and dengue shock syndrome, characterized by failure of the circulation system that may lead to death. There are still no drugs or vaccines available in the market to prevent or treat dengue infections. Hence, the development of successful drugs that are safe and offer a long-lasting protection against dengue viruses is needed. Recently, the discovery of peptides with high specificity, selectivity, and efficacy; and relatively safer than small-molecule drugs have become a promising agent for a new drug. Previously, a molecular dynamic simulation reported that peptide NRRRRSAGMI from the capsid’s cleavage region had the most hydrogen bonds with NS2B/NS3 protease, thus it leads into a starting linear heptapeptide R(P3)-R(P2)-R(P1)-S(P1’)-A(P2’)-G(P3’)-M(P4’). This study aimed to investigate the importance and specificity of individual amino acid residue of the heptapeptide substrate on the binding affinity and interaction to NS2B/NS3 dengue protease through computational alanine scanning. It showed that the changing of P1 and P2 arginine to alanine resulted to higher values of binding affinity. Thus, the arginine is an important residue at the P1 and P2 positions for the binding of the substrate peptide to DENV NS2B/NS3 protease, hence the amino acid residues which are essential for the enzyme-substrate interactions can be predicted through computational alanine scanning.

Keywords: Computational alanine scanning, Molecular docking, Dengue virus, NS2B/NS3 protease, Peptide-based inhibitor

Introduction

Dengue is recognised as one of the most prevalent arthropod-borne viral diseases in human. In these few decades, the pathogenic dengue virus (DENV) has become a serious threat to global health due to the increased incidence of human infections worldwide and has rapidly spread in more than 100 countries. World Health Organisation (WHO) reported that over 3 billion of the world population live in dengue-endemic countries and more than 390 million dengue cases occur worldwide annually [1]. One of the main factors for dengue expansion to these countries includes increased international travel in the dengue endemic areas [2]. Although dengue is an age-old disease, the researchers are still looking for the drugs or vaccines to prevent or treat dengue infections. The only treatment for dengue disease is symptomatic treatment or intensive supportive therapy such as blood transfusion and plasma volume replacement, which can often save the lives of DHF patients [3].

Dengue virus belongs to the Flaviviridae family, in the genus flavivirus which also related to other human disease-causing viruses such as WNV (West Nile virus). There are 4 viral serotypes of dengue virus which are DV-1, DV-2, DV-3, and DV-4. The 4 serotypes are hardly distinguished based on clinical and pathological symptoms that they cause; however, they can be identified by monoclonal antibodies,
polymerase chain reaction (PCR), neutralization and hematological tests [4,5]. DENV virion is a small icosahedral enveloped virus with a diameter of ~50 nm and surrounded by a lipid envelope [6].

DENV consists of a ~11 KB positive single-stranded RNA with short 5' and 3' non-coding regions at both ends. The viral genome is transcribed and translated by host cell ribosomes into a single polypeptide chain upon entering the host cell [7]. The RNA is covered by viral nucleocapsid, which is surrounded by a lipid bilayer coat [5]. Proteolytic cleavage of the polypeptide by host cell proteases, furin and signalase, and virus-encoded protease, NS2B/NS3 results in formation of 3 structural proteins (envelope (E), capsid (C) and pre-membrane/membrane ((prM/M) proteins) and 7 non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) [8,9].

The NS2B/NS3 protease is a 2-component protease and it has been reported that the optimal catalytic activity of NS3 protease is dependent on a cytosolic core sequence of the NS2B protein [10]. Furthermore, the structural stability of NS3pro depends upon the forty amino acid residues of the hydrophilic segment of NS2B cofactor [11]. The viral serine protease NS2B/NS3 complex plays an essential role in the cleavage of the viral precursor protein, which is indispensable for the maturation and pathogenesis of dengue virus [12,13]. Thus, the NS2B/NS3 complex is required for co- and post-translational processing into individual functional protein and therefore, this complex becomes a potential target for the design and development of anti-viral dengue.

Currently, research on substrates-based peptides for drug design and discovery appears as a promising starting point in the development of new drugs. Approximately, more than 140 peptide-based compounds were used as commercial drugs with more than 400 peptides are now in pre-clinical phase trial [14]. Despite the low stability, peptides are preferred for designing drug inhibitors because of their high activity and specificity [15-17]. The successful discovery and development of peptide inhibitors of hepatitis C virus protease which are now in clinical use has also led to the development of peptide-based compounds as clinical candidates against dengue virus protease [18].

A molecular dynamics (MD) simulation of NS2B/NS3 protease has been carried out with polypeptide substrates, Asn-Arg-Arg-Arg-Ser-Ala-Gly-Met-Ile (NRRRSAGMI) from the capsid’s cleavage region and it was reported that this peptide formed the highest number of hydrogen bonds with the protease [19]. Thus, this peptide sequence NRRRSAGMI leads to our interest of peptide-dengue protease interaction study, which then becomes a starting ligand into linear heptapeptide RRRSAGM (P3-P2-P1-P1’-P2’-P3’-P4’ positions) (Figure 1). In this study, we performed a computational alanine scanning based on the peptide RRRSAGM and the replacement of each amino acid residue at P3 to P4’ positions to alanine obtained another 6 different heptapeptides which are ARRSAGM, RARSAGM, RRASAGM, RRRAAGM, RRRSAAM and RRRSAGA. It has been reported that the amino acid at P3, P2 and P1 positions would give a good binding with NS2B/NS3 protease [20]. Thus, this work aims to study the binding interaction of peptide RRRSAGM and its substitution peptides with alanine against dengue NS2B/NS3 protease. It is hoped that the replacement of each amino acid with alanine can shed some light for us which amino acid is important in the binding with dengue protease. In addition, the interaction free energies in peptide-protease complexes will have usually been translated as indicating whether or not a particular side chain contributes to binding affinity.

**Figure 1** Linear heptapeptide RRRSAGM as at P³-P²-P¹-P¹’-P²’-P³’-P⁴’ positions.
Materials and methods

Preparation of the dengue protease

The dengue NS2B/NS3 protease homology model used in this study was downloaded from the available literature of Wichapong and co-workers [21]. For AutoDock, the dengue protease was prepared using AutoDock tools (ADT) version 1.5.6 [22]. All water, ions and unwanted molecules were withdrawn from the protease file. The protonation for the protease structure were assigned using ADT, all polar hydrogen atoms were added, default Kollman charges and solvation parameters were assigned to the protease atoms and further saved as pdbqt format file [23].

Preparation of the peptide ligands

The peptide benzoyl-Norleucine-lysine-arginine-aldehyde (Bz-Nle-Lys-Arg-H) was used as a control ligand in this study because this peptide has been predicted to show activity and anti-dengue strength [24]. The control ligand, Bz-Nle-Lys-Arg-H was obtained from the PubChem database. The 2-dimensional structure of the backbone of studied peptide, RRRSAGM together with other 6 heptapeptides substrate (ARRSAGM, RARSAGM, RASAGM, RRRAAGM, RRSAAM and RRRSAGA (Figure 2) were built and converted into 3D structures using PerkinElmer ChemDraw Ultra 16.0 software. All peptide structures were then minimized (MM2 force field minimization) using the same program. The polar hydrogen atoms were added and the protonation states of an amino acid side chains were assigned to all peptide ligands using ADT depend on its environment in the protein. AutoTors in ADT was used to assign the flexibility of the peptide ligands. Since AutoDock can only assist 32 torsions in the docking run, the set up was adapted from previously published protocol whereby only the side chains of the peptide were assigned flexible while the amide bonds were set as non-rotatable part. The Gasteiger charges [25] were then added to the studied peptides and saved as a pdbqt file format for further docking.
The standard ligand Bz-Nle-Lys-Arg-Arg-H used for control docking was first removed and redocked to the DENV NS2B/NS3 protease homology model. The protocols were previously described in Nadeem and Ezatul, 2020 [26]. The alanine scanning docking of the proposed peptides were performed using the same parameters set up of the standard ligand Bz-Nle-Lys-Arg-Arg-H docking for AutoDock 4.2. The 3-dimensional grip maps along x, y and z axes were set at 60, 60 and 60 in order to cover the entire enzyme active site and to allow the peptidic inhibitor to move freely at 3D space. The center of ligand’s mass was set at x = 23.516, y = 43.373 and z = −1.358 coordinates, the energy range was set to 7, the number of modes was set to 20 and the other parameters were kept as default values. The calculation of docking was performed using Lamarckian Genetic Algorithm (LGA) with default parameters for 100 search runs. Upon completion, the analysis based on the clustering histogram was performed and the best conformation with the lowest estimated free energy and most populated cluster were selected for further analysis data. The 2D visualization of the selected best peptides with the dengue NS2B/NS3 protease from was prepared using Ligplot v.2.1 program [27], while the 3D visualization was prepared using a discovery studio visualizer v.17.2 2 [28].

Results and discussion

Validation of the control peptide

Computer-aided design has 2 classifications which are structure-based drug design, and ligand based drug design [29]. Computational docking is usually validated by ‘redocking’ process, where a complex of ligand-protein is separated and redocked again to the protein structure. This process is carried out to ensure that the docking parameters used can reproduce the binding mode of its original structure [30]. Thus, the validation of the docking procedure in this study was carried out by separating and redocking of peptide Bz-Nle-Lys-Arg-Arg-H into the dengue NS2B/NS3 protease. We found that the redocked conformation scored estimated free energy of binding (ΔGbind) of −5.96 kcal/mol. We also further verified the redocking process by visual inspection and we observed that the redocked Bz-Nle-Lys-Arg-Arg-H bound at the same binding location as that of the original conformation (Figure 3(A)) in a region close to the catalytic triad residues, (His51, Asp 75 and Ser 135) as shown by Wichapong et al. [21]. It shows that the residues of P1, P2, P3 and P1’ bind at desired S1, S2, S3 and S1’ substrate pockets of WNV NS2B/NS3 protease (Figure 3(B)). In addition, the docked conformation formed a network of hydrogen bond interactions with residues His51, Asp75, Gly82, Phe130, Ser135, Tyr161 and also made hydrophobic and van der Waals interactions with close contact residues particularly with residues Met84, Asp129, Tyr150, Gly151, Gly153 and Val155. The interaction with some residues for example Asp75, Ser135 and Gly153 were also reported in Nasution et al. [31]. This observation confirmed that the original conformation was acceptable reproduced by the docking procedures used and thus the same docking parameters were applied to run computational alanine scanning of the other heptapeptides.
**Figure 3** (A) An overview of the control ligand, Bz-Nle-Lys-Arg-Arg-H before docking (yellow) in comparison with its docked structure (blue) at the binding site of NS2B/NS3 dengue protease. The closed contact residues of the enzyme are shown as grey sticks and catalytic triads are highlighted with green colour. (B) The 3-dimensional structure of the docked tetrapeptide inhibitor bound at the S1, S2, S3 and S1’ at the substrate pockets of WNV NS2B/NS3 protease.

**Computational alanine scanning**

Computational alanine scanning (CAS) was first introduced by Massova and Kollman (1991) which is normally provided a method to determine the changes in binding free energies of ligands to target macromolecules induced by replacement or mutations of each amino acid residues into alanine [32]. This strategy gives a key role in identifying of the hot interaction spots of small molecules with targets. The alanine scanning in this study has been carried out using computational docking of replacement or mutated each amino acid residue into alanine of the peptide RRRSAGM which then derived into 6 other heptapeptides substrate. This peptide sequence has been obtained from the polypeptide substrates, Asn-Arg-Arg-Arg-Ser-Ala-Gly-Met-Ile (NRARRSAGMI) from the capsid’s cleavage region in dengue NS2B/NS3 protease which have been shown the highest number of hydrogen bonds with the protease from a molecular dynamics simulation study [19].

In this study, we performed a computational alanine scanning based on the peptide RRRSAGM and replace each amino acid residue at P3-P2-P1-P1’-P2’-P3’-P4’ positions with alanine residue. By doing this, we obtained another 6 different heptapeptide substrates which were ARRSAGM, RARSAGM, RRASAGM, RRRAAGM, RRRSAAM and RRRSAGA. AutoDock has been chosen to calculate the binding energy of peptide-protease interaction in this study because AutoDock employs a semi-empirical free energy force field which was calculated using a large number of protein-inhibitor complexes for which both structure and inhibition constants, or Ki, are known. In addition, AutoDock provides rapid evaluation of the interaction energy using simpler force field and it could cover a wider area of conformational space [33].

Although AutoDock can only assist 32 torsions in the docking run, we set only the side chains of the peptide to rotate and freely move around the active site of the dengue protease while the amide bonds were set as non-rotatable part in order to reduce the number of torsions. We applied this strategy based on our previous study of using AutoDock to run several numbers of different peptide sequences targeting neuropilin-1 for anti-angiogenesis, which has proven useful when we confirmed with *in vitro* biological assay as compared to our docking results [34]. Thus, we hope that the same strategy would be applicable in this docking of alanine scanning of peptide RRRSAGM using AutoDock.

In the docking study of the heptapeptides substrate, we found that all the heptapeptides substrate bind at the same binding region, which occupied the S1, S2, S3 and S1’ substrate pockets of NS2B/NS3 protease as shown by original and docked conformations of standard peptide, Bz-Nle-Lys-Arg-Arg-H (Figure 4). It showed that the docking protocols used to dock the heptapeptides substrate are applicable for this docking study. Thus, we will further discuss the binding affinity and binding interaction of these heptapeptides with NS2B/NS3 protease.
Figure 4 Heptapeptide substrate sequences and their binding conformations at the active site pockets of NS2B/NS3 protease.

Table 1 shows the predicted Gibbs free binding energy ($\Delta G_{\text{bind}}$) values of the peptide RRRSAGM and its replacement of each residue into alanine with NS2B/NS3 protease. The $\Delta G_{\text{bind}}$ value is calculated based on the laws of thermodynamics, which involves the equilibrium constant (K) value. The lower or the more negative $\Delta G_{\text{bind}}$ value indicates the ligand-enzyme complex interaction is stronger and the complex is likely to be more stable. The stable ligand-enzyme complex is based on the stability and strength of non-covalent interactions (i.e hydrogen bonds) on the ligand-enzyme complex, which affects the molecular interaction of the complex and thus lowering the free binding energy value.

From the quantitative results based on the $\Delta G_{\text{bind}}$ value, most of the heptapeptides which having arginine at P1 and P2 positions showed low $\Delta G_{\text{bind}}$ values except for 2 heptapeptides with the mutation or replacement of arginine into alanine at P1 and P2 positions which showed slightly higher $\Delta G_{\text{bind}}$ values. Yusof and co-workers (2000) demonstrated that the residues of Arg-Arg in P1 and P2 locations next the cleavage site are responsible for the high binding affinity with the protease enzyme [35]. Later, Li et al. (2005) suggested that a strong preference was observed in the residue of essential amino acids (Arg or Lys) at P1 position, and also amino acids of Arg > Thr > Gl/Asn/Lys were preferred for a P2 position [20]. In addition, previously reported docking and an alanine scanning study suggested that the P2 Arg is necessary to form important interactions with NS2B/NS3 protease, designating that Arg at P2 position plays a critical role in the binding with the protease. This is supported in another study, Yotmanee et al. (2015) found that substrates P1 and P2 play a significant role in binding the NS2B/NS3 pro-type 2 virus [19]. Thus, our computational alanine scanning demonstrated a consistent finding as previously reported data that arginine at P1 and P2 positions are important and necessary in the binding with NS2B/NS3 protease.

The summary of the binding interaction of the RRRSAGM and the other 6 heptapeptides with NS2B/NS3 protease is also shown in Table 1, while the visualization of the interaction is shown in Figure 5. The interaction of the heptapeptides and the NS2B-NS3 protease can occur not only in the hydrogen bond form, but also with other non-covalent interaction, such as hydrophobic and Van der Waals interactions, which can contribute to the effectiveness in the binding affinity between ligand and protein complex. The hydrogen bond is defined as the intermolecular force that occurs between high electronegative atoms with hydrogen atoms that covalently bonded to an electronegative atom [36]. The electronegative atom can be oxygen, nitrogen, or sulphur. Generally, the interactions that occurred in the ligand-enzyme complex are weak, compared to covalent bond.

Figure 5(A) shows that R3 of heptapeptide RRRSAGM formed a network of hydrogen bond with Asp129 and Gly159 of S1 and S3 pockets and no hydrogen bond or hydrophobic interaction were made with catalytic triad residues. For heptapeptide ARRSAGM, hydrogen bond was shown between R2 and Ser83 of S2 pocket, R3 with Asp129 and Phe130 of S1 pocket and the other hydrogen bond was formed at SAGM and catalytic triad His51 and Ser135, Arg54, Asn152 and Gly153 of S3 pocket (Figure 5(B)). This peptide occupied S1-S3 pockets of NS2B/NS3 protease through hydrogen bond and hydrophobic interaction. A substitution of alanine at P2 position had made heptapeptide RARSAGM lost the interaction with catalytic triad residues and also with S2 pocket. R1 made hydrogen bonds with Asn152 and Gly153
of S3 pocket, while R3 with Asp129 and Gly159 of S1 and S3 pockets (Figure 5(C)). R2 of heptapeptide RRASAGM formed 2 hydrogen bonds with Met84 of S2 pocket and one hydrogen bond with each Gly153 and Val155 of S3 pocket, while, SAGM made hydrogen bonds with residues His51, Gly151, Asn152 and Tyr161 (Figure 5(D)). One of hydrophobic interaction was observed with catalytic triad Ser135 and the other residues were Asp129, Phe130, Ser131, Pro132, Thr134, Tyr150 and Val154 of S1 and S3 pockets. A replacement of alanine at P1’ position made R1 heptapeptide RRRAAGM formed hydrogen bonds with Phe130, Ser131 and Ser135 of S1 pocket, R2 with Ser83 of S2 pocket, while its C-terminal made hydrogen bonds with Met84 and Gly153 of S2 and S3 pockets (Figure 5(E)). For heptapeptide RRSSAMA, the R1 formed a network of hydrogen bond with catalytic triad His51, Gly82 of S2 pocket and Asn152 of S3 pocket, while R2 made hydrogen bonds with Met84 of S2 pocket and the SAMA part made 1 hydrogen bond with Gly151 of S3 pocket (Figure 5(F)). This peptide lost hydrogen bond with S1 pocket however it showed hydrophobic interaction through Asp129, Pro132 and Ser135. The others hydrophobic interactions were observed at Gly151, Val154, Tyr161 of S3 pocket and Ile36 of S1’ pocket. Figure 5(G) shows R1 of heptapeptide RRSSAGA formed hydrogen bonds with Phe130, Ser131 and Tyr150 of S1 and S3 pockets. It is observed that R2 made hydrogen bonds with His51 of catalytic triad residue and Asp81 of S2 pocket, while the hydrogen bond was observed between R3 and Ser83 of S2 pocket. RRSSAGA shows that both hydrogen bond and hydrophobic interaction were occupied at S1-S3 pockets of NS2B/NS3 protease.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Free energy of binding (kcal/mol)</th>
<th>Type of interaction</th>
<th>Residues involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRRSAGM</td>
<td>−4.64</td>
<td>Hydrogen bond</td>
<td>Asp129, Asn152, Gly153, Gly159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Ile86, Met84, Ser83, Gly151, Val155, Val154, Tyr161</td>
</tr>
<tr>
<td>ARRASGM</td>
<td>−4.38</td>
<td>Hydrogen bond</td>
<td><strong>His51</strong>, Arg54, Asp75, Ser83, Asp129, Phe130, Ser135, Gly153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Gln35, Ile36, Met84, Ser131, Pro132, Val155, Gly151, Tyr161</td>
</tr>
<tr>
<td>RARSAGM</td>
<td>−2.19</td>
<td>Hydrogen bond</td>
<td>Asp129, Asn152, Gly153, Gly159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Arg54, Pro132, Gly151, Asn152, Gly159, Tyr163</td>
</tr>
<tr>
<td>RRASAGM</td>
<td>−2.82</td>
<td>Hydrogen bond</td>
<td><strong>His51</strong>, Met84, Gly151, Asn152, Gly153, Val155, Tyr161</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Asp129, Phe130, Ser131, Pro132, Ser135, Thr134, Val154, Tyr150</td>
</tr>
<tr>
<td>RRRAAGM</td>
<td>−5.00</td>
<td>Hydrogen bond</td>
<td>Ser83, Met84, Phe130, Ser131, Ser135, Gly153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Asp81, Gly83, Pro132, Gly151, Asn152, Val155, Tyr161</td>
</tr>
<tr>
<td>RRSSAAM</td>
<td>−4.62</td>
<td>Hydrogen bond</td>
<td><strong>His51</strong>, Met84, Gly151, Asn152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Ile36, Asp129, Pro132, Ser135, Gly151, Val154, Tyr161</td>
</tr>
<tr>
<td>RRSSAGA</td>
<td>−5.20</td>
<td>Hydrogen bond</td>
<td><strong>His51</strong>, Asp75, Asp81, Ser83, Phe130, Ser131, Gly151, Tyr161</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrophobic</td>
<td>Trp50, Gly83, Pro132, Ser135, Gly151, Asn152, Val154, Val155, Tyr161</td>
</tr>
</tbody>
</table>

**Note:** Bold residues indicate the amino acid of catalytic triad, His51, Asp75 and Ser135.
Figure 5 The 2-dimensional structure of the docked conformation of the heptapeptide substrates with the NS2B/NS3 protease binding site using LigPlot+ (A) RRRSAGM, (B) ARRSAGM, (C) RARSAGM, (D) RRASAGM, (E) RRRAAGM, (F) RRRSAAM and (G) RRRSAGA.

Conclusions

In this study, amino acid residues of the heptapeptide substrate RRRSAGM and its substituted amino acid into alanine were performed using molecular docking. The essential amino acid that contributed for the enzyme-substrate interactions can be predicted through computational alanine scanning approach. It was found that S1 and S2 pockets show strong preferences for basic amino acids, whereas the S3 pocket is less selective for basic residue. In addition, docking results show that the arginine at P1 and P2 positions of the heptapeptide substrate contribute significantly to the substrate binding affinity.

Acknowledgements

The authors thank the Ministry of Education (MOE), Malaysia for the financial support of this study through the Fundamental Research Grant Scheme FRGS/1/2015/SKK09/USM/02/3 (203/PFARMASI/6711520).

References


