โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การค้นหาสารธรรมชาติและอนุพันธ์ที่ออกฤทธิ์ยับยั้งเอนไซม์เอ็นเอสทูบี/เอ็นเอสทรีเซอรีนโปทีเอสของเชื้อไวรัสซิกาด้วยระเบียบวิธีทางเคมีคอมพิวเตอร์
Discovery of Natural Compounds and Derivatives with Inhibitory Activity Against NS2B/NS3 Serine Protease of Zika Virus by Computational Approaches

ชื่อนิสิต นายวชิรพล มานิมนต์ เลขประจำตัว 6033083223

ภาควิชา เคมี

ปีการศึกษา 2563

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การค้นหาสารธรรมชาติและอนุพันธ์ที่ออกฤทธิ์ยับยั้งเอนไซม์เอ็นเอสทูบี/เอ็นเอสทรี เซอรีนโปทีเอสของเชื้อไวรัสซิกาด้วยระเบียบวิธีทางเคมีคอมพิวเตอร์

Discovery of Natural Compounds and Derivatives with Inhibitory Activity Against NS2B/NS3 Serine Protease of Zika Virus by Computational Approaches

โดย

นายวชิรพล มานิมนต์

รายงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต

ภาควิชาเคมี คณะวิทยาศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2563

โครงการ การค้นหาสารธรรมชาติและอนุพันธ์ที่ออกฤทธิ์ยับยั้งเอนไซม์เอ็นเอสทูบี/เอ็นเอสทรีเซอรีนโปทีเอสของเชื้อไวรัสซิกาด้วยระเบียบวิธีทางคอมพิวเตอร์

โดย นายวชิรพล มานิมนต์

ได้รับอนุมัติให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาเคมี
ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

คณะกรรมการสอบโครงการ

1. ศาสตราจารย์ ดร.พรเทพ สมพรพิสุทธิ์ ประธานกรรมการ
2. ผู้ช่วยศาสตราจารย์ ดร.ธนธรณ์ ขอทวีวัฒนา กรรมการ
3. ศาสตราจารย์ ดร.สุพจน์ หารหนองบัว อาจารย์ที่ปรึกษา
4. ผู้ช่วยศาสตราจารย์ ดร.ธัญญดา รุ่งโรจน์มงคล อาจารย์ที่ปรึกษาร่วม

รายงานฉบับนี้ได้รับความเห็นชอบและอนุมัติโดยหัวหน้าภาควิชาเคมี

.....
(ศาสตราจารย์ ดร.สุพจน์ หารหนองบัว)
อาจารย์ที่ปรึกษา

.....
(รองศาสตราจารย์ ดร.วรวีร์ โฮเว่น)
หัวหน้าภาควิชาเคมี

.....
(ผู้ช่วยศาสตราจารย์ ดร.ธัญญดา รุ่งโรจน์มงคล)
อาจารย์ที่ปรึกษาร่วม

วันที่ 27 เดือน พฤษภาคม พ.ศ. 2564

ชื่อโครงการ การค้นหาสารธรรมชาติและอนุพันธ์ที่ออกฤทธิ์ยับยั้งเอนไซม์เอ็นเอสทูบี/เอ็นเอสทรีเซอรีนโปทีเอสของเชื้อไวรัสซิกาด้วยระเบียบวิธีทางคอมพิวเตอร์
ชื่อนิสิตในโครงการ นายวชิรพล มานิมนต์ เลขประจำตัว 6033083223
ชื่ออาจารย์ที่ปรึกษา ศาสตราจารย์ ดร.สุพจน์ หารหนองบัว
ชื่ออาจารย์ที่ปรึกษาร่วม ผู้ช่วยศาสตราจารย์ ดร.ธัญญดา รุ่งโรจน์มงคล
ภาควิชาเคมี คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2563

บทคัดย่อ

โรคไข้ซิกาเกิดจากการติดเชื้อไวรัสซิกา (Zika virus) ซึ่งอยู่ในตระกูลฟลาวิไวรัส (flavivirus) โดยมียุงลายเป็นพาหะนำโรค มีการรายงานถึงความสัมพันธ์ของการเกิดความผิดปกติทางสมองของเด็กที่เกิดจากสตรีที่ติดเชื้อระหว่างตั้งครรภ์ ในปัจจุบันยังไม่มียาหรือวัคซีนที่ใช้รักษาโรคติดเชื้อไวรัสซิกาได้โดยตรง ดังนั้นจึงจำเป็นต้องพัฒนายาที่มีประสิทธิภาพในการรักษา โดยโปรตีนเป้าหมายที่สำคัญของไวรัส คือเอ็นเอสทูบี/เอ็นเอสทรีเซอรีนโปทีเอส (NS2B/NS3 serine protease) ซึ่งเป็นเอนไซม์ที่สำคัญในการจำลองแบบของไวรัส ในงานวิจัยนี้ผู้วิจัยได้ใช้ระเบียบวิธีทางเคมีคอมพิวเตอร์ ได้แก่ โมเลกุลาร์ด็อกกิ้ง (molecular docking) และการจำลองพลวัตเชิงโมเลกุล (molecular dynamics simulations) เพื่อใช้ค้นหาและศึกษาการจับกันระหว่างโปรตีนเป้าหมายกับกับสารประกอบธรรมชาติและอนุพันธ์ทั้งสิ้น 7 ชนิด ได้แก่ chalcone, flavonoid, iodovinyl sulfones, sulfonylated indeno quinolines, vinyl sulfones, quinolinone และ nitroolefin ที่บริเวณอัลโลสเตอริก (allosteric site) นอกจากนี้ยังคำนวณค่าพลังงานการยึดจับอิสระ (binding free energy) ของสารแต่ละตัวด้วยวิธี MM/GBSA ผลการศึกษาพบว่าสาร DN071_f ซึ่งเป็นสารอนุพันธ์ในกลุ่ม flavonoid สามารถจับกับเอนไซม์ได้ดีที่สุด โดยมีค่าพลังงานยึดจับอิสระเท่ากับ −17.17 ± 1.90 kcal/mol ดังนั้นผลที่ได้จากงานวิจัยนี้คาดว่าสาร DN071_f สามารถนำไปใช้ประโยชน์ในการออกแบบยาที่ออกฤทธิ์ยับยั้งเอนไซม์เอ็นเอสทูบี/เอ็นเอสทรีเซอรีนโปทีเอสของเชื้อไวรัสซิกาได้อย่างมีประสิทธิภาพในอนาคต

คำสำคัญ: โมเลกุลาร์ด็อกกิ้ง, การจำลองพลวัตเชิงโมเลกุล, เอ็นเอสทูบี/เอ็นเอสทรีโปรตีเอส, ไวรัสซิกา

Project Title Discovery of Natural Compounds and Derivatives with Inhibitory Activity Against NS2B/NS3 Serine Protease of Zika Virus by Computational Approaches
Student Name Wachirapol Manimont Student ID 6033083223
Advisor Name Professor Supot Hannongbua, Ph.D.
Co-advisor Name Assistant Professor Thanyada Rungrotmongkol, Ph.D.
Department of Chemistry, Faculty of Science, Chulalongkorn University, Academic Year 2020

Abstract

Zika virus (ZIKV) belongs to the *Flavivirus* genus primarily transmitted by *Aedes* mosquito. There have been reports of an association of child disorders among infected women during pregnancy. To date, there is no drug or vaccine that can directly combat ZIKV infection. Therefore, the design and development of effective drugs are extremely needed. Specifically, the targeted protein of this virus is NS2B/NS3 serine protease, which is an enzyme important for viral replication. In this study, we have used computational chemistry approaches including molecular docking and molecular dynamic (MD) simulation to screen and study binding interactions between protein and 7 groups of natural compounds at allosteric site. The groups of natural compounds include chalcone, flavonoid, iodovinyl sulfones, sulfonylated indeno quinolines, vinyl sulfones, quinolinone and nitroolefin. Moreover, binding free energy calculation based on the MM/GBSA method showed that DN071_f belonging to flavonoid category exhibited the highest binding affinity to ZIKV protease (–17.17 ± 1.90 kcal/mol). Therefore, the results presented here suggested that DN071_f can be used for further design of drug that effectively inhibits the NS2B/NS3 serine protease of ZIKV.

Keywords: Molecular docking, Molecular dynamics, NS2B/NS3 serine protease, Zika virus

ACKNOWLEDGEMENTS

First and foremost, thanks to myself, for my effort to complete the research successfully.

I would like to express my deep and sincere gratitude to Professor Supot Hannongbua, Ph.D., Assistant Professor Thanyada Rungrotmongkol, Ph.D., Bodee Nutho. Ph.D., and Bio Sim laboratory team, for giving me the opportunity to do research and providing invaluable guidance throughout this research.

My completion of this project could not have been accomplished without the support of Nhong, Mike, Yujin, Numnim, Pream, View, and all of my friends.

Finally, to my loving and supportive parents. The countless times you cooked a delicious meal. It was a great comfort and relief to know that you were willing to provide management of household activities while I completed my work. My heartfelt thanks.

TABLES OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Research rationality

The two most common flaviviruses, Dengue virus (DENV) and Zika virus (ZIKV) are mosquito-borne flaviviruses transmitted by *Aedes mosquito* that related to West Nile (WNV), and yellow fever virus (YFV) containing the positive-sense single-stranded RNA genome [1]. DENV is one of the most important infectious diseases, a disease caused by four viruses, defined as serotypes 1-4 in tropical and subtropical regions around the world, and there is no specific drug available in the market to combat its serious causes at the present [2]. Infection with each of the DENV serotypes may be asymptomatic in most cases; however, some patients may develop dengue fever that has a high fever and flu-like symptoms. Most seriously, secondary infections by a different DENV serotype culminate in the highest risk of development of severe dengue fever, characterized by bleeding order, increased vascular fragility, and permeability which are thought to associate with antibody-dependent enhancement (ADE) [3, 4]. Whereas patients infected with ZIKV generally shows mild symptoms and can resolve on their own within a few days. Although ZIKV infection does not affect the patient to death, there have been several reports of a relationship of fetal neurological disorders and infected women during pregnancy. It was found that children who were infected while in the mother's womb may develop brain developmental deficits and present a smaller head than normal, the so-called microcephaly. In addition, the main symptoms of ZIKV infection in adults are linked to Guillain-Barré syndrome, including muscle weakness and numbness in the arms and legs. This can cause serious complications if the muscles of the airways are affected.

Indeed, the genome of DENV and ZIKV is closely related to each other. The polyprotein of both viruses is translated into the three structural proteins (C, prM and E) and the seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). Specifically, NS3 is a serine protease that contains the catalytic triad H51, D75, and S135 in its active site and play a pivotal role in cleavage of the scissile junction within the polyprotein (**Figure 1.1**). To activate the proteolytic activity of NS3 protease, a small hydrophilic part of NS2B as a cofactor domain has been required to form as a complex with NS3 [5, 6]. Moreover, NS2B/NS3 serine protease is an enzyme important for viral replication and infectivity [7]. To this reason, NS2B/NS3 protease can be considered as an attractive therapeutic target for the development of anti-flaviviral agents [6, 8].

In general, flavivirus protease inhibitors can be classified into two groups: peptidomimetics and small molecules without substrate character [9]. However, the active site of the viral protease prefers basic amino acids (i.e., K or R residues) in peptide substrates. So, these positively charged

molecules cause limited oral bioavailability, leading to the difficulty in design of peptidomimetics against flavivirus NS2B/NS3 protease that are effective inhibitors *in vivo*. Alternatively, the search for novel allosteric inhibitors that do not bind to the enzyme active site with improved drug-likeness has attracted recent attention.

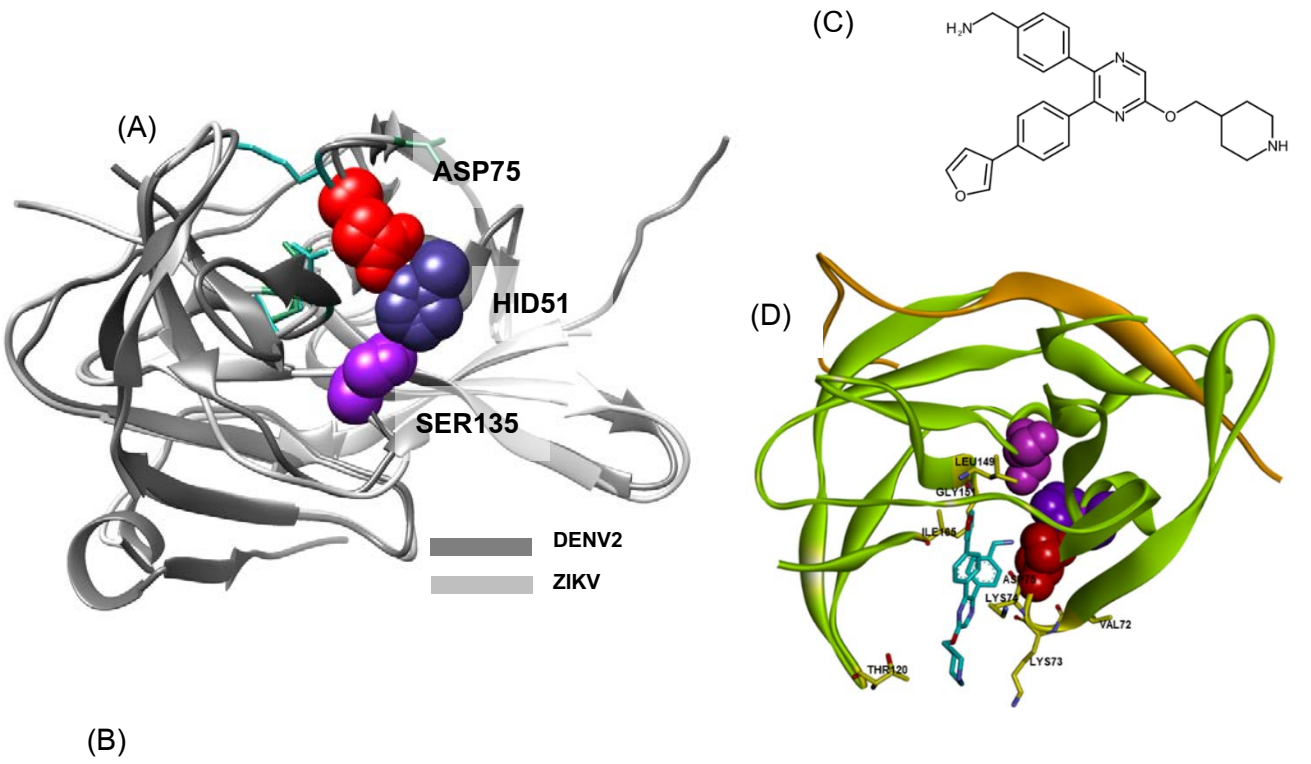


Figure 1.1 (**A**) Overlay structure between DENV2 and ZIKV NS2B/NS3 serine protease. (**B**) sequence alignment of allosteric site between where blue, red and pink colors highlights the catalytic residues His51, Asp75, and Ser135, respectively. (**C**) 2D chemical structure of SYC-1307. (**D**) 3D structure of a reference compound (SYC-1307) binding to the allosteric site of the DENV2 protease (PDB ID: 6MO0).

To date, there are no specific drugs available to prevent and treat DENV and ZIKV infection [10]. Likewise, a vaccine for ZIKV is unavailable, but the only dengue vaccine, Dengvaxia® (CYD-TDV), developed by Sanofi Pasteur that claims to be effective for protection against all four DENV serotypes has been licensed. Nevertheless, the vaccine has limited efficacy against acute DENV2 only 35–42%. Furthermore, getting vaccination may increase the chances of being hospitalized for people without

prior DENV infection as well as elevate the risk of severe dengue fever. Based on these reasons, the World Health Organization (WHO) has suggested to only use the DENV vaccine in person who has serological evidence of prior dengue infection [4].

Several studies report that drug-like small molecules could potentially inhibit DENV2 NS2B/NS3 as allosteric inhibitors (non-competitive inhibitors) and exert the disruption to the interaction between NS2B and NS3 [5, 6, 11-13]. Interestingly, SYC-1307 [1-(4-{3-[4-(furan-3-yl)phenyl]-5-[(piperidin-4-yl)methoxy]pyrazin-2-yl}phenyl)methanamine] (**Figure 1.1C**) was reported to possess potent cellular antiviral activity against ZIKV and DENV NS2B/NS3 protease, which could allosterically bind at a back pocket of the enzyme with the half maximal inhibitory concentration (IC_{50}) values of 200 and 590 nM, respectively [12].

In this project, we are interested in focusing on the use of natural compounds and their derivative from our in-house library for exploring the inhibitory activity toward NS2B/NS3 protease of DENV and ZIKV by computational techniques. In fact, the enzyme has a relatively small three-dimensional (3D) structure comprising approximately 230 amino acids [14]. This makes it easy to investigate the binding interactions of these known natural substances and derivatives to bind to NS2B/NS3 protease [15] and thereby stopping [16] or slowing down the life cycle of DENV and ZIKV [17]. Note that our in-house library consists of 7 groups of compounds including chalcone, flavonoid, iodovinyl sulfones, sufonylated indeno quinolines, vinyl sulfones, quinolinone and nitroolefin (548 in total). The core structure of each compound category is depicted in **Figure 1.2**.

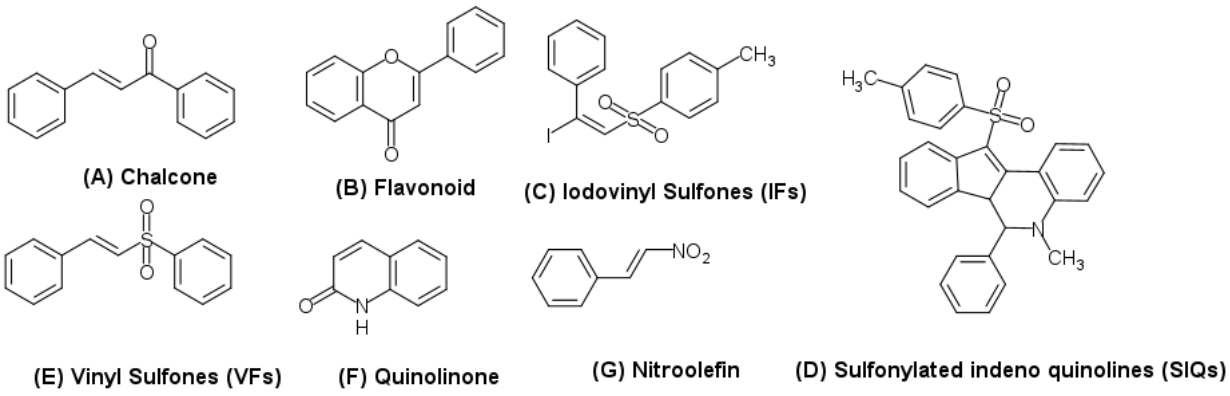


Figure 1.2 Chemical structure of (**A**) Chalcone, (**B**) Flavonoid, (**C**) Iodovinyl Sulfones (IFs), (**D**) Sulfonylated indeno quinolines (SIQs), (**E**) Vinyl Sulfones (VFs), (**F**) Quinolinone and (**G**) Nitroolefin.

1.2 Literature review

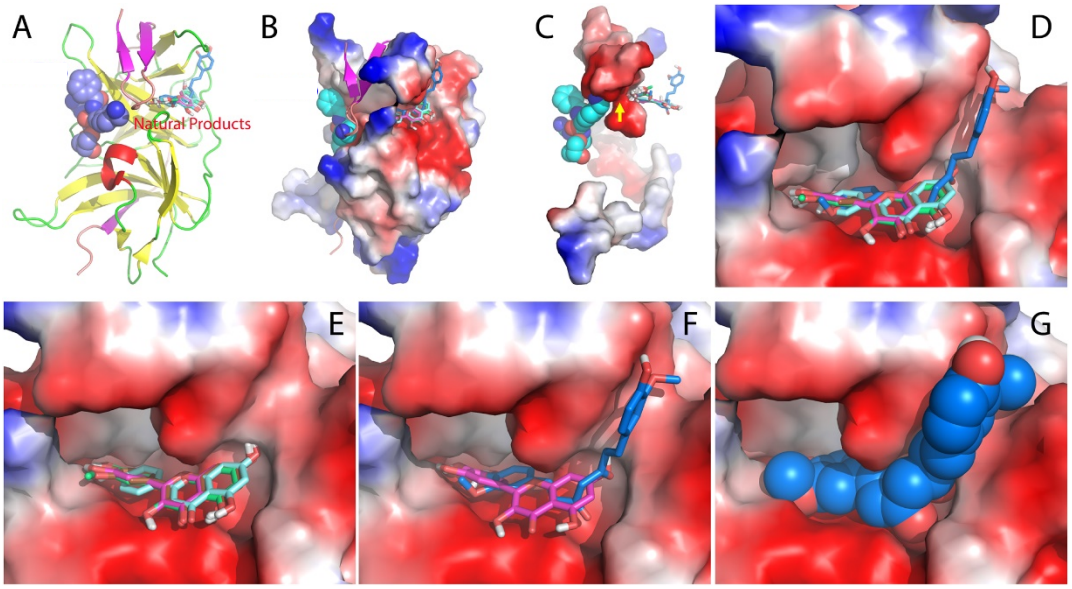
Several studies report that drug-like small molecules could inhibit NS2B/NS3 protease of flavivirus as allosteric inhibitors (non-competitive inhibitors), thereby exerting the disruption interaction between NS2B and NS3 by computational approaches. Then studies inhibit activity by cell-based assay. An example from the previous research that studies about interaction and inhibition between compound and NS2B/NS3 protease of flavivirus at the allosteric site as following details:

1.2.1 Solution conformations of ZIKV NS2B/NS3 protease and its inhibition by natural products from edible plants

This study showed that several natural products could bind to the allosteric site of NS2B/NS3 protease. The results has shown that there were six compounds which had significant inhibitory effects toward ZIKV NS2B/NS3 protease, including myricetin, quercetin, isorhamnetin, luteolin, apigenin, curcumin, daidzein, catechin and resveratrol [16], as shown in **Figure 1.3.** The study also investigated mode of inhibition and binding interactions of these bioactive products against ZIKV NS2B/NS3 protease. The parameter IC_{50} and inhibitory constant K_i were used to determine the inhibitory effect. Myricetin exhibited the highest inhibitory activity (IC_{50} of 1.26 μM and K_i of 0.77 μM), while apigenin exerted the weakest effect (IC_{50} of 56.32 μM and K_i of 34.02 μM), as shown in **Figure 1.3**. However, daidzein, catechin and resveratrol were not detected the inhibitory activity.

In addition, molecular docking technique was utilized to obtain a better understanding of the experimental results and elucidate the structure-activity relationship (SAR) of the compounds. The docking results suggested that all six compounds bind to the pockets on the back of the active site of ZIKV protease, which were similar to the binding site of flavonoids binding to DENV2 NS2B/NS3 protease such as myricetin and quercetin [16].

(1)



(2)

	Inhibition at 500 μM	IC_{50} (μM)	K_i (μM)
Myricetin	Yes	1.3 ± 0.1	0.8 ± 0.1
Quercetin	Yes	2.4 ± 0.2	1.1 ± 0.1
Luteolin	Yes	2.7 ± 0.3	1.4 ± 0.1
Isorhamnetin	Yes	15.5 ± 0.7	6.2 ± 0.4
Apigenin	Yes	56.3 ± 0.9	34.0 ± 2.4
Curcumin	Yes	3.5 ± 0.2	2.6 ± 0.2

Figure 1.3 (**1**) (A) The crystal structure (PDB code of 5LC0) of Zika NS2B-NS3pro determined with an active site which six natural products (in sticks) were docked. B) Zika NS2B-NS3pro in complex with six natural products in which NS2B is displayed in ribbon and NS3pro domain in the electrostatic potential surface. (C) six natural products in complex with NS2B in the electrostatic potential surface. For clarity, only NS2B is displayed in the electrostatic potential surface. Expanded binding pockets of Zika NS2B-NS3pro in complex with all six compounds (D); with five flavonoids (E); with Myricetin and Curcumin (F); and only with Curcumin in spheres (G). (**2**) Inhibitory parameters of six natural products on ZIKV NS2B/NS3 protease [16].

1.2.2 Curcumin allosterically inhibits DENV NS2B/NS3 protease by disrupting its active conformation

solution conformations of ZIKV NS2B/NS3 protease and its inhibition by natural products from edible plants

This report investigated a key target for designing anti-flavivirus drugs by NMR and molecular docking [13]. This research has shown that curcumin allosterically inhibited the DENV protease by binding to a cavity with no overlap with the active site. Curcumin also showed a good inhibitory effect with an IC_{50} and K_i of 7.18 ± 0.62 μM and 4.35 ± 0.02 μM, respectively (**Figure 1.4B**). Moreover, MD simulations revealed that the binding of curcumin let to unfolding/displacing the characteristic β-hairpin of the C-terminal NS2B cofactor, thereby disrupting the closed (active) conformation of the protease, as shown in **Figure 1.4**.

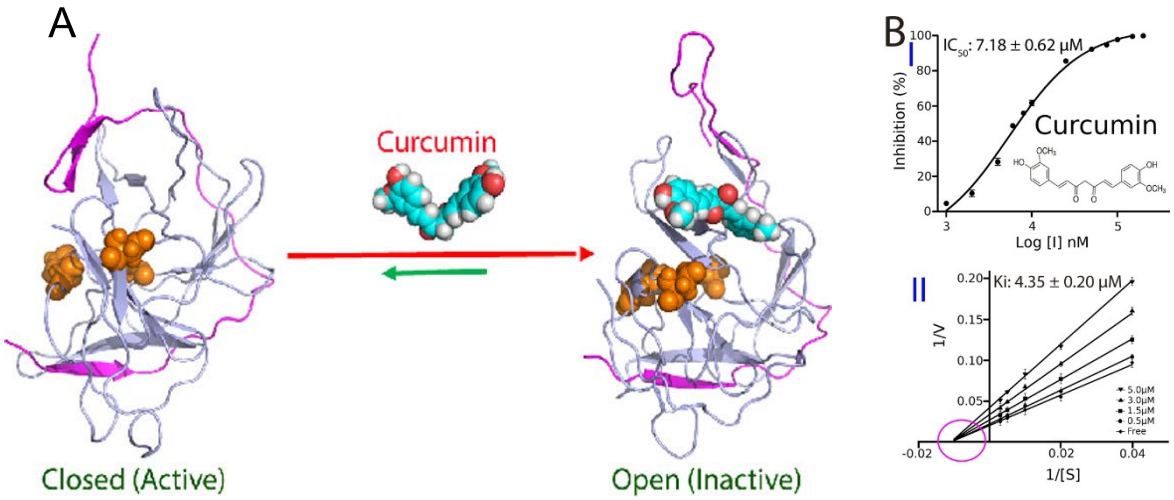


Figure 1.4 (**A**) Curcumin allosterically inhibits the Dengue NS2B/NS3 protease by disrupting its active conformation. (**B**) (I) the inhibitory data used for fitting IC_{50} value of curcumin, and (II) Lineweaver–Burk plot for determining inhibitory constant (K_i) of curcumin on the DENV NS2B/NS3 protease [13].

1.2.3 A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B/NS3 protease

This report studied about the conformational change of NS2B and characterized candidate allosteric inhibitors using split luciferase complementation (SLC) assay, as shown in **Figure 1.5** [11].

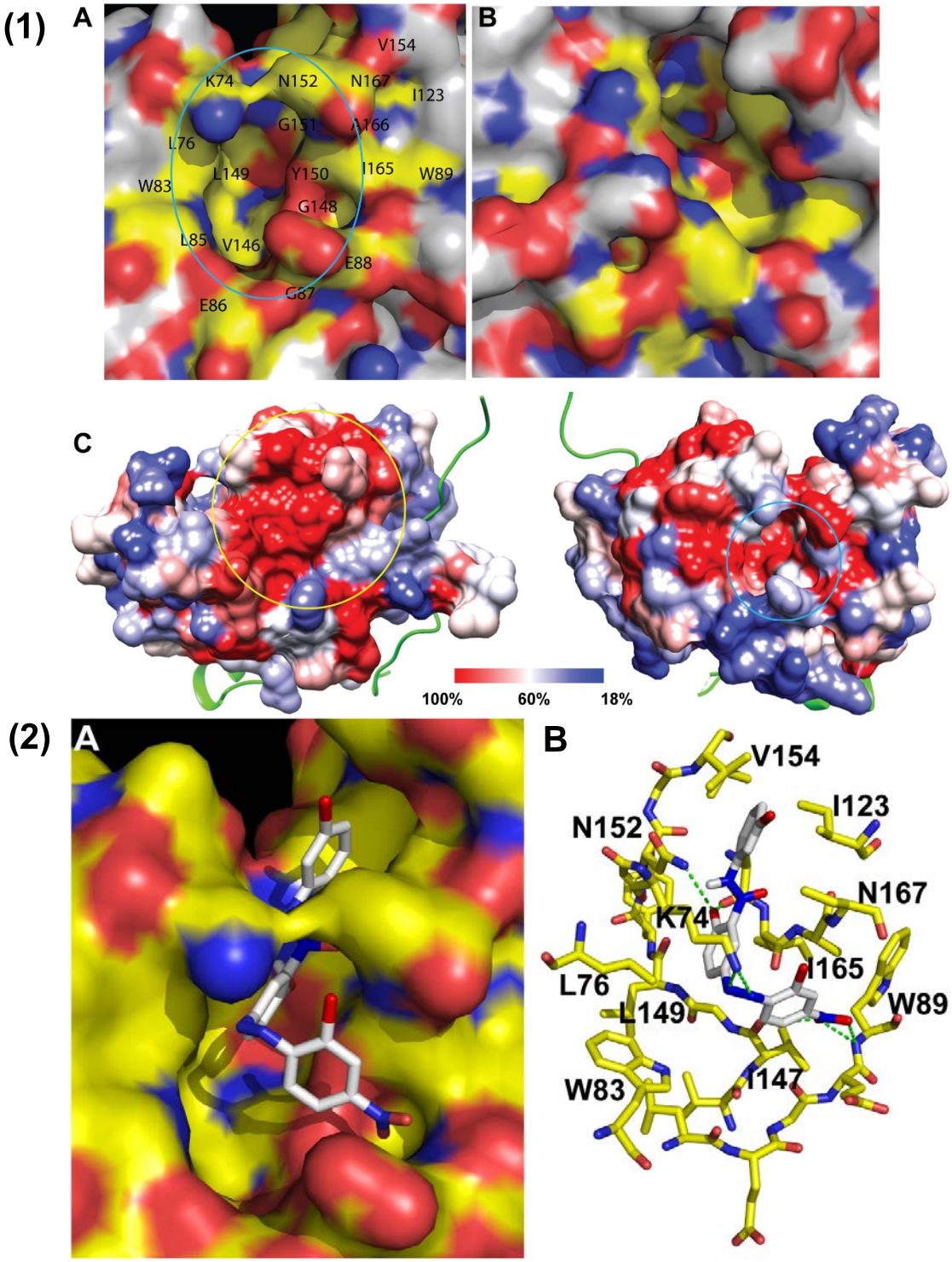


Figure 1.5 (1) A pocket unique to the inactive NS2B-NS3 structure. (A) Surface representation of the unique pocket of the inactive NS2B-NS3 structure. (B) Surface representation of the active NS2B-NS3 structure at the same area. Surface color was according to atomic color as follows unless otherwise stated: oxygen, red; nitrogen, blue, carbon of residues involved in pocket formation, yellow; other carbon, white. (C) Mapping of sequence conservation to surface of the DENV2 NS3 protease domain (PDB code: 2FOM); the structure is viewed from two angles, showing features in different regions.

Left: active site surface (yellow circle). Right: the unique pocket (cyan circle) of the inactive NS2B-NS3 structure. The surface is colored according to sequence conservation, resulting from the multiple sequence alignment. NS2B is displayed as a cartoon representation (green). (**2**) Predicted poses of compound NSC135618 in the allosteric pocket of the inactive NS2B-NS3 structure. (A): NS2B and NS3 are in surface representation with surface colors the same as in (**1**). [11].

A virtual screen of NCI library compounds was performed to identify allosteric inhibitors, followed by *in vitro* biochemical screening of the resultant candidates. Only three of these compounds including NSC135618, 260594, and 146771, significantly inhibited the protease of DENV2 *in vitro*, with IC_{50} values of 1.8 μM, 11.4 μM, and 4.8 μM, respectively. Only NSC135618 inhibits the conformational change of NS2B, which observed with suppressing a significant SLC enhancement, suggesting an allosteric inhibitor of flavivirus NS2B/NS3 protease.

Furthermore, the results from virus titer reduction assays demonstrated that NSC135618 is a broad spectrum flavivirus protease inhibitor, and can significantly reduce titers of DENV2, ZIKV, WNV and YFV infected on A549 cells that concentration of a compound gave half-maximal response (EC_{50}) in a low micromolar range. Nevertheless, the cytotoxicity of NSC135618 was only found in the moderate magnitude with the 50% cytotoxic concentration (CC_{50}) of 48.8 μM on A549 cells (**Figure 1.6**). Interestingly, NSC135618 also inhibited ZIKV in human placental and neural progenitor cells, which are essential in the pathogenesis of ZIKV infection.

Compound	Viruses	EC_{50} (μM)	CC_{50} (μM)	TI[a]
135618	DENV2	0.81	48.8	60
	ZIKV	1.0		49
	WNV	1.27		38
	YFV	0.28		174
260594	DENV2	13.5	>500	37
146771	DENV2	NE[b]	82	N/A

[a]TI: Therapeutic index defined as CC_{50}/EC_{50}.
[b]NE: Non-effective

https://doi.org/10.1371/journal.ppat.1006411.t002

Figure 1.6 Cell viabilities and antiviral activities of selected compounds on A549 cells [11].

Example report has shown that DENV protease inhibitors can inhibit the enzyme of the ZIKV virus, which belongs to the flavivirus virus genus. In addition, the mechanism of inhibition of the enzyme activity of flavivirus has been reported. The researchers used these information to discover

natural compounds and derivatives with inhibitory activity against ns2b/ns3 serine protease of zika virus by computational approaches for the development of drugs in the future.

1.3 Objective

To search for natural compounds and derivatives with inhibitory activity against NS2B/NS3 serine protease of ZIKV by computational approaches.

1.4 Expectation

Obtain natural compounds and derivatives with inhibitory activity against NS2B/NS3 serine protease of Zika and may have a high potential for further testing *in vitro* and *in vivo*.

1.5 Operation method

A. Study plan

1. Collecting related documents and information.
2. Preparing 3D structure of 548 natural compounds and their derivatives, as well as a structure of NS2B/NS3 serine protease of ZIKV**.**
3. Investigating binding affinity between enzyme NS2B/NS3 serine protease of ZIKV and each compound by molecular docking technique.
4. Screening the compounds that show a good binding affinity to the enzyme for further studying by means of molecular dynamic simulation in order to determine structural, dynamics and energetic properties.
5. Analyzing and comparing the results and summarizing the project.

B. Research period

Study plan	**Period of work (Starting April 2020)**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1 . Collecting related documents and information								
2 . Preparing 3D structure of 548 natural compounds and their derivatives, as well as a structure of NS2B/NS3 serine protease of ZIKV								
3. Investigating binding affinity between enzyme NS2B/NS3 serine protease of ZIKV and each compound by molecular docking technique								

4. Screening the compounds that show a good binding affinity to the enzyme for further studying by means of molecular dynamic simulation in order to determine structural, dynamics and energetic properties								
5. Analyzing and comparing the results and summarizing the project								

1.6 THEORIES

1.6.1 Molecular docking

Molecular docking technique is commonly used to investigate geometrical and energetic aspects of protein–ligand interactions. In this work, a molecular docking method was applied to predict the interaction energy and binding mode of compound bound to the binding site of the targeted enzyme. These predictions relate to the relative concentrations of the free enzyme, free ligand and complex that will exist in an equilibrium system. Binding affinity of protein–ligand complex in terms of binding free energy (ΔG_{bind}) is considered by the basic concept, as shown in Eq. (1).

$$[Enz]_{aq} + [Lig]_{aq} \overset{\Delta G_{bind}}{\longleftrightarrow} [EL]_{aq} \quad (1)$$

where $[EL]_{aq}$ is the concentration of enzyme–ligand complex, $[Enz]_{aq}$ is the concentration of enzyme and $[Lig]_{aq}$ is the concentration of ligand.

In addition, molecular docking method can be divided into two components: scoring function and search algorithm, as depicted in **Figure 1.7**.

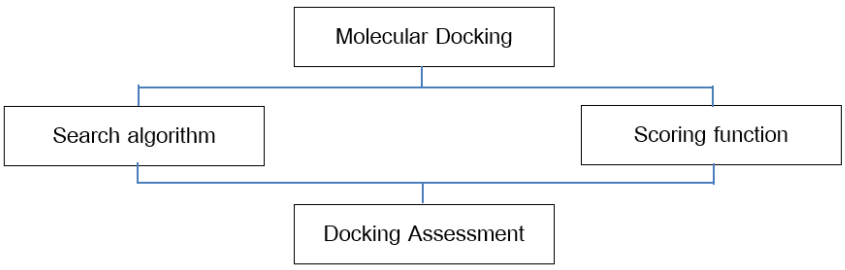


Figure 1.7 Scheme of the docking methodology

1.6.1.1 Scoring functions

The scoring function is used for evaluation of the binding affinities by observing the force field-based and/or knowledge-based. The force-field-based scoring function is referred to functional form and parameter set that is used in molecular modeling field, especially the energy function or interatomic potential. Meanwhile, the knowledge-based scoring function is a type of statistical observation that can be found from interaction of atoms or functional group. Besides, the random distribution is extracted from the basic knowledge-based which can be compared with the experimental data.

1.6.1.2 Search algorithm

The typical two types of search algorithm are stochastic and systematic search. The stochastic search method is involved in random changes to the state variables to a user - defined termination criteria is met for finding the local or global point. The systematic search method is involved in the finding of space at predefined intervals and deterministic. Many methods tend to find near or local minimum energy point [18].

1.6.2 Molecular mechanics

Molecular mechanics (MM) is a classical mechanics-based model, in which molecular structure is described by a series of atoms (point charges) connected to bonds (springs). The atomic size and the bond length can identify the types of atoms and bonds. MM framework is based on force field or empirical approximations, which are used for calculating interactions and evaluating the potential energy of the systems as a function of atomic coordination. The two potential energy functions of common force fields can be divided into two components: bonded and non-bonded interactions. The former interactions include covalent bond stretching (E_{bond}), angle bending (E_{angle}), and torsional rotation ($E_{dihedral}$), which are generally fixed throughout a simulation. Meanwhile, the latter interactions involve van der Waals (E_{VDW}) and electrostatic (E_{ele}) interactions. Therefore, the total potential energy (E_{total}) of a molecule [19] is a summation of each interaction term, as expressed in equations (2) - (7) :

$$E_{total} = E_{bond} + E_{angle} + E_{dihedral} + E_{VDW} + E_{ele} \quad (2)$$

Bond stretching energy: $E_{bond} = \sum_{bond} \frac{1}{2} k_b (r - r_0)^2$ (3)

Bond angle energy: $E_{angle} = \sum_{angle} \frac{1}{2} k_\theta (\theta - \theta_0)^2$ (4)

Torsional angle energy: $E_{dihedral} = \sum_{dihedral} \frac{V_n}{2}(1 + \cos(n\omega - \delta))$ (5)

Van der Waal interaction: $E_{vdw} = \sum_{i<j}^{atoms} 4\varepsilon \left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} \right)$ (6)

Electrostatic energy: $E_{ele} = \sum_{i<j}^{atoms} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$ (7)

1.6.3 Molecular dynamics

Molecular dynamics (MD) is a computer simulation technique which is commonly used to describe equilibrium and dynamics behaviors of biological system. In this technique, atoms and molecules are allowed to interact to each other for a given period of time, providing a view of the particle's movement. Moreover, MD can be an important tool for scientists to study the motion of individual atoms in a case which is not possible to perform in laboratory experiments [20].

MD simulation has been considered as "statistical mechanics by numbers" and "Laplace's vision of Newtonian mechanics" for prediction of future by animating nature's forces. This provides detailed insight into molecular motion at atomic level. MD simulation represents an interface between theoretical information and laboratory experiments, known as a virtual experiment. It combines the relationship between molecular structure, motion and function. All classical MD methods depend on force field for calculating interactions and evaluating the potential energy of the systems. A force field consists of both sets of equations used to calculate the potential energy and forces from particle coordinates, including a collection of parameters used in the equations mentioned in Molecular mechanics section [21] MD simulation is a useful technique in modern molecular modeling and allows us to follow and understand structure and dynamics of biological molecules with extreme information. This method is based on the calculation of the time dependent behavior of a molecular system. MD simulations can provide theoretical information in the context of fluctuations and conformational changes of biological systems (e.g., proteins, carbohydrate, fatty acids, and nucleic acids). Moreover, this method is commonly applied to determine the structures, dynamical behavior, and thermodynamics properties of biological molecules and their complexes taken from X-ray crystallography and from NMR experiments.

Classical MD simulation is an important method used to investigate the microscopic behaviors by integrating the motion of particles or molecules based on the Newton's second law equation of

motion. Fundamentally, the acceleration of particles can be calculated by the first-order derivative of velocity (v_i) or second-order derivative of the atomic position (r_i) with respect to time t (Eq. (8))

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m_i \frac{d^2 r_i}{dt^2} \tag{8}$$

where F_i is the total force exerted on the particle *i*, m_i and a_i are mass and acceleration of the particle *i* at given time t, respectively. In addition, the external force acting on the particle i can be obtained from the negative gradient of potential energy (*U*), as shown in Eq. (9)

$$F_i = -\nabla_i U \tag{9}$$

As mentioned earlier, the total potential energy is the summation between the bonded and nonbonded interactions (see Eq. (1)). In bonded term, the mechanical molecular model considers atoms as spheres and bonds as springs. The mathematics of spring deformation can be used to describe the ability of bonds to stretch, bend, and twist. In non-bonded term, intermolecular interactions consist of van der Waals attraction, steric repulsion and electrostatic attraction/repulsion.

In molecular system, the potential energy is the summation of MM energies (i.e., bonded and non-bonded interactions), as previously described. Again, bonded interactions compose of covalent bond-stretching, angle-bending, and dihedral angle potentials as described by harmonic oscillator function. Meanwhile, non-bonded interactions consist of electrostatic and van der Waals interactions, which are described by coulomb potential and Lennard-Jones potential, respectively. The sum of potential energy interactions [19] can be expressed as follows (Eqs. (10)-(11)):

$$U = (E_{bonds} + E_{angles} + E_{dihedrals})_{bonded} + (E_{ele} + E_{vdw})_{nonbonded} \tag{10}$$

$$U = \sum_{bond} \frac{1}{2} k_b (r - r_{eq})^2 + \sum_{angle} \frac{1}{2} k_\theta (\theta - \theta_{eq})^2 + \sum_{dihedral} \frac{V_n}{2}(1 + \cos(n\emptyset - \gamma)) + \sum_{i<j}^{atoms} \frac{q_i q_j}{D r_{ij}} + \sum_{i<j}^{atoms} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) \tag{11}$$

Currently, MD simulation can be achieved by several free or commercial software packages (e.g. AMBER, GROMACS, NAMD and CHARMM).

1.6.4 Root mean square displacement (RMSD)

RMSD is the most used quantitative measure of the similarity between atomic coordinates of MD simulation at time with respect to those of initial structure. RMSD values are calculated by three coordinates (x, y, z), as expressed in Eq. (12).

$$RMSD(v,w) = \frac{1}{n}\sum_{i=1}^{n}\left((v_{ix} - w_{ix})^2 + \left(v_{iy} - w_{iy}\right)^2 + (v_{iz} - w_{iz})^2\right) \quad (12)$$

where *n* is atoms of protein, v and w are atomic coordinates at starting point (time = 0) and MD production (time = t), respectively.

1.6.5 Binding free energy

The molecular mechanics/generalized Born surface area (MM/GBSA) approach [22] was generally used to calculate the binding free energy (ΔG_{bind}) of the protein–ligand complex, as described previously [23]. Briefly, the free energy is computed as difference between the protein–ligand complex ($G_{complex}$) and its isolated forms ($G_{protein}$ and G_{ligand}), as shown in Eq. (13):

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \quad (13)$$

The ΔG_{bind} consists of the molecular mechanics (MM) energy in gas phase (ΔE_{MM}), solvation free energy (ΔG_{solv}), and entropic contribution term ($T\Delta S$), as given in Eq. (14):

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (14)$$

The ΔE_{MM} is achieved from the combination of electrostatic (ΔE_{ele}) and van der Waal (ΔE_{vdW}) energies, whereas the ΔG_{solv} is calculated using Eq. (15) below:

$$\Delta G_{solv} = G_{solv}^{ele} + G_{solv}^{nonpolar} \quad (15)$$

The G_{solv}^{ele} can be estimated using the GB model [24], whereas the $G_{solv}^{nonpolar}$ is calculated using solvent accessible surface area (SASA) [24], as shown in Eq. (16):

$$\Delta G_{solv}^{nonpolar} = \gamma SASA + b \quad (16)$$

Where γ and β are the experimental solvation parameters, which were set to 0.00542 kcal/mol·Å^2 and 0.92 kcal/mol, respectively [25]. Additionally, the contribution of each amino acid residue involved in ligand binding was evaluated using the per-residue decomposition free energy ($\Delta G_{bind}^{residue}$) analysis.

CHAPTER 2
RESEARCH AND METHODOLOGY

2.1 Structural preparation and molecular docking studies

The crystal structures of DENV NS2B/NS3 protease (PDB ID: 2FOM) and ZIKV NS2B/NS3 protease (PDB ID: 5GXJ) were obtained from Protein Data Bank (PDB). The one missing amino acid residue T168 in the DENV2 protease was built using Discovery Studio Visualizer 2020. The chemical structures were obtained from our in-house database. These compounds can be divided into 7 groups consisting of 226 chalcones, 58 flavonoids, 6 iodovinyl sulfones (IFs), 26 sulfonylated indeno quinolines (SIQs), 79 vinyl sulfones (VFs), 27 quinolinone, 15 nitroolefin and 110 structurally modified compounds (98 vinyl sulfones and 12 quinolinones that have been more binding by modify structure.). All the compounds were built using sketch molecules tools in Discovery Studio Visualizer 2020. The protonation state of each ligand was determined using Marvin Sketch, while all ionization amino acids of protein were predicted using PDB2PQR [26] in the APBS-PDB2PQR server at pH 7.4. Afterward, each compound was docked 5 times with 20 independent runs into the allosteric site of DENV and ZIKV protease using AutoDock VinaXB [27]. Note that the reference ligand (SYC-1307) located at the allosteric pocket of DENV protease was also defined as the binding site of ZIKV protease with setting the search space center (X, Y, Z) dimension to be -7.8 Å, -9.7 Å and 9 Å for DENV protease and -11 Å, -8.6 Å and 9 Å for ZIKV protease.

2.2 Molecular dynamics (MD) simulation

From molecular docking results, we selected the initial structures for subsequent MD simulations based on the top six lowest interaction energies. The chosen ligands were optimized at the HF/6–31(d) level of theory using Gaussian09 program [28]. The electrostatic potential (ESP) charges for the optimized ligands were calculated using the same method. Afterward, the restrained ESP charges and missing parameters of each compound were achieved by the antechamber and parmchk modules of AMBER16 [29], respectively. The ff14SB force field [30] was applied for protein, while the generalized AMBER force field (GAFF) [31] was applied to treat the ligand. The LEaP module of AMBER16 was used to add the missing hydrogens to the protein structure. The added hydrogen atoms were energy-minimized using the steepest descents (SD) and conjugated gradient (CG) methods. The complexes were solvated in the TIP3P water box [32] with a minimum distance of 13.0 Å away from the protein surface. Chloride or sodium ions were added to neutralize the simulated systems. Later, the water molecules were also minimized by the SD and CG algorithms.

MD simulations were conducted to gain an insight into the protein-ligand interactions and binding affinity of ligand to the protein at the molecular level. In this study, MD was simulated under periodic boundary condition with a time step of 2 fs. The particle mesh Ewald (PME) summation method was used to treat the electrostatic interaction [33], while the residue-based cutoff of 12 Å was employed for non-bonded interactions. The SHAKE algorithm [34] was applied to constrain chemical bonds involving hydrogen atoms. The Berendsen barostat was used to control the standard pressure of the system with a pressure-relaxation time of 1 ps [35]. The simulated systems were run for 100 ns. To analyze the system, root-mean-square deviation (RMSD), hydrogen bond interactions, key binding energy, number contact atoms and solvent-accessible surface area (SASA) were determined. Note that the CPPTRAJ module of AMBER16 was used to calculate the structural analyses. Whereas the molecular mechanics generalized Born surface area (MM/GBSA) was applied to calculate the binding free energy and the per-residue decomposition free energies of the complex structures using MMPBSA.py of AMBER16.

CHAPTER 3
RESULTS AND DISCUSSION

3.1 Molecular docking

SYC-1307 (see **Figure 1.1**) was an allosteric inhibitor of ZIKV and DENV NS2B/NS3 protease as introduced in the previous report [12]. It was thus chosen as a reference compound for comparing the binding strength with our compounds in NS2B/NS3 allosteric site. The assumption is that 'the lower docking energy, the higher binding affinity of compound'. Molecular docking results of 548 compounds were given in **Figure 3.1**. From the result, it was found that the interaction energy ($\Delta E_{\text{docking}}$) of reference compound (SYC-1307) is –8.1 kcal/mol. Interestingly, there were 19 compounds which had the $\Delta E_{\text{docking}}$ values less than –8.1 ± 1.0 kcal/mol, suggesting the possibly better allosteric inhibitor than SYC-1307. These compounds can be separated into 4 groups: (*i*) chalcones (11), (*ii*) flavonoids (8), (*iii*) quinolinone (1), and (*iv*) SIQ (1), as illustrated in **Figure 3.1**. The chalcones group included cpd26, cpd26_e, cpd28, cpd31, cpd44, cpd48, cpd48_e, cpd50, cpd50_e, cpd51 and cpd51_e. The flavonoids group comprised DN071_f, TP034, TP034_e, TP051, TTS09, TTS10, TTS11 and TTS12. The quinolinone group composed of compound 3v, and the SIQ group included compound SIQ20. Moreover, this study was compared to another report [36] regarding the discovery of compound with inhibitory activity against NS2B/NS3 protease of DENV, indicating that 17 compounds remained the same as our report. These 17 compounds were TP034_e, DN071_f, TP034, cpd48, cpd48_e, TTS10, cpd50_e, cpd51_e, cpd31, cpd44, TTS09, TTS12, cpd26_e, cpd28_e, cpd44_e, cpd49_e, and IP004 (–9.6, –9.3, –9.3, –8.9, –8.8, –8.7, –8.5, –8.5, –8.3, –8.2, –8.2, –8.2, –8.1, –8.0, –8.0, –8.0, and –8.0 kcal/mol, respectively), as shown in **Figure 3.2** (**A**). Among the overlapping compounds capable to bind both viral proteases, we selected the top ranked 6 compounds (**Figure 3.2** (**B**)) for further investigation by means of MD simulation. The 2D structure of compounds TP034_e, DN071_f, TP034, cpd48, cpd48_e and TTS10 are depicted in **Figure 3.3**. It should be mentioned that the docked configurations of these compounds within the allosteric site of DENV and ZIKV protease were in the similar manner (**Figure 3.4**).

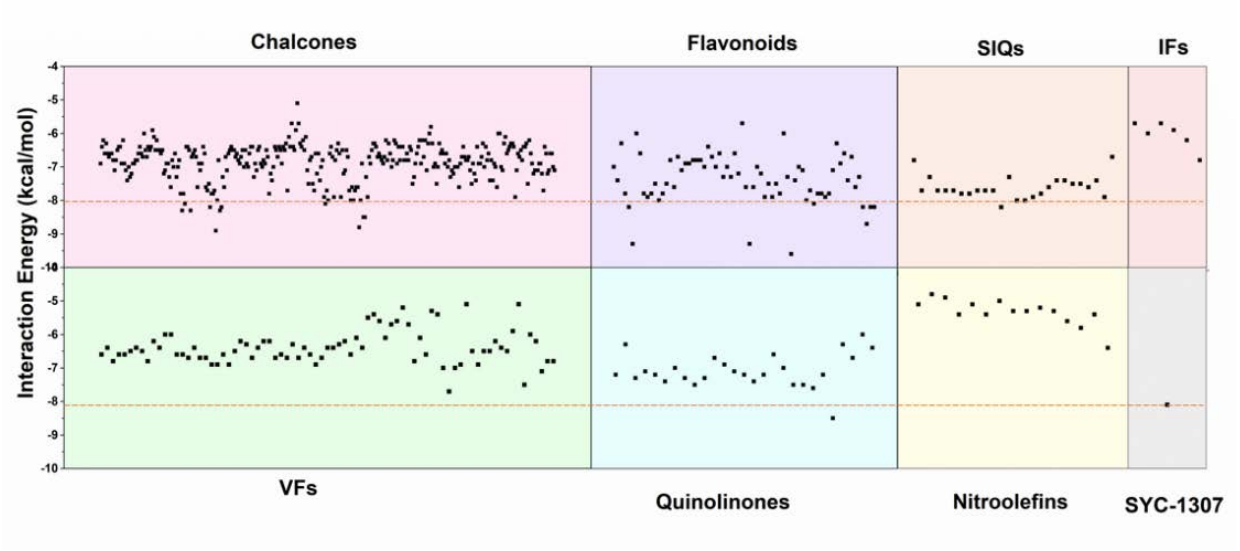


Figure 3.1 Molecular docking result of 548 compounds toward ZIKV NS2B/NS3 protease, where the points under the orange lines represent the compounds which are less than –8.1 kcal/mol.

A

compounds	Interaction energy(kcal/mol)
TP034_e	-9.6
DN071_f	-9.3
TP034	-9.3
cpd48	-8.9
cpd48_e	-8.8
TTS10	-8.7
cpd50_e	-8.5
cpd51_e	-8.5
cpd31	-8.3
cpd44	-8.2
TTS09	-8.2
TTS12	-8.2
cpd26_e	-8.1
cpd28_e	-8.0
cpd44_e	-8.0
cpd49_e	-8.0
IP004	-8.0

B

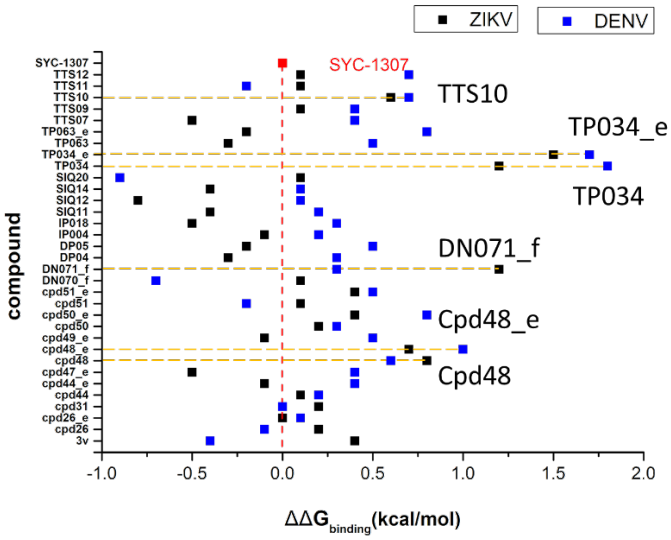


Figure 3.2 (**A**) Interaction energy of 17 compounds, (**B**) Top ranked 6 compounds chosen to perform MD simulation.

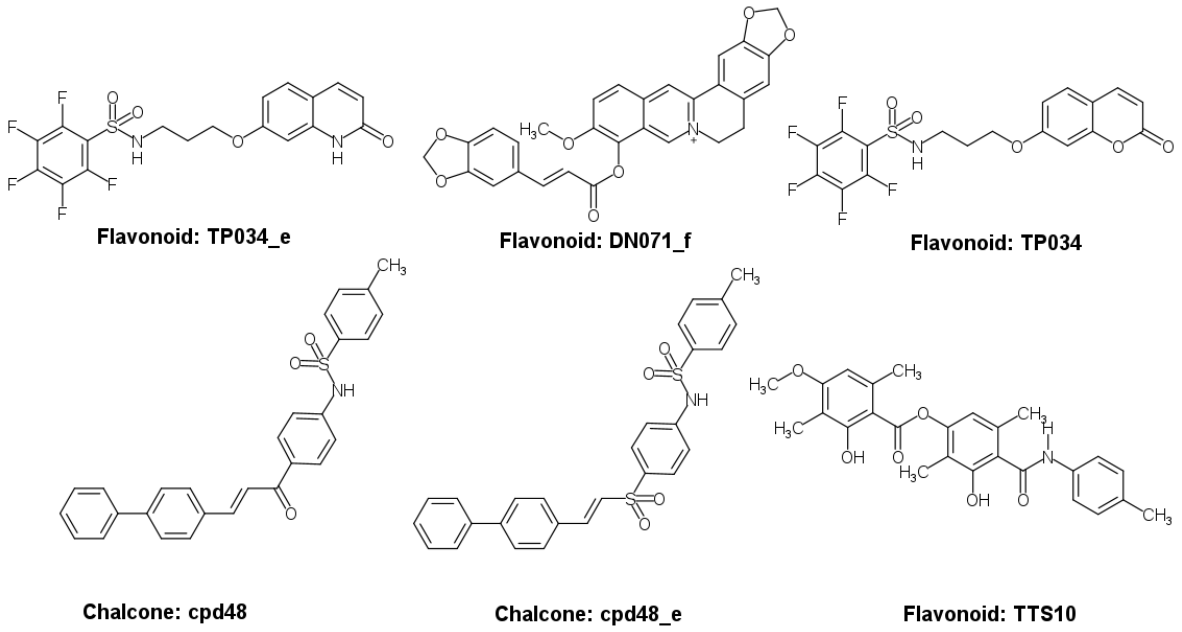


Figure 3.3 Chemical structure of TP034_e, DN071_f, TP034, cpd48, cpd48_e and TTS10

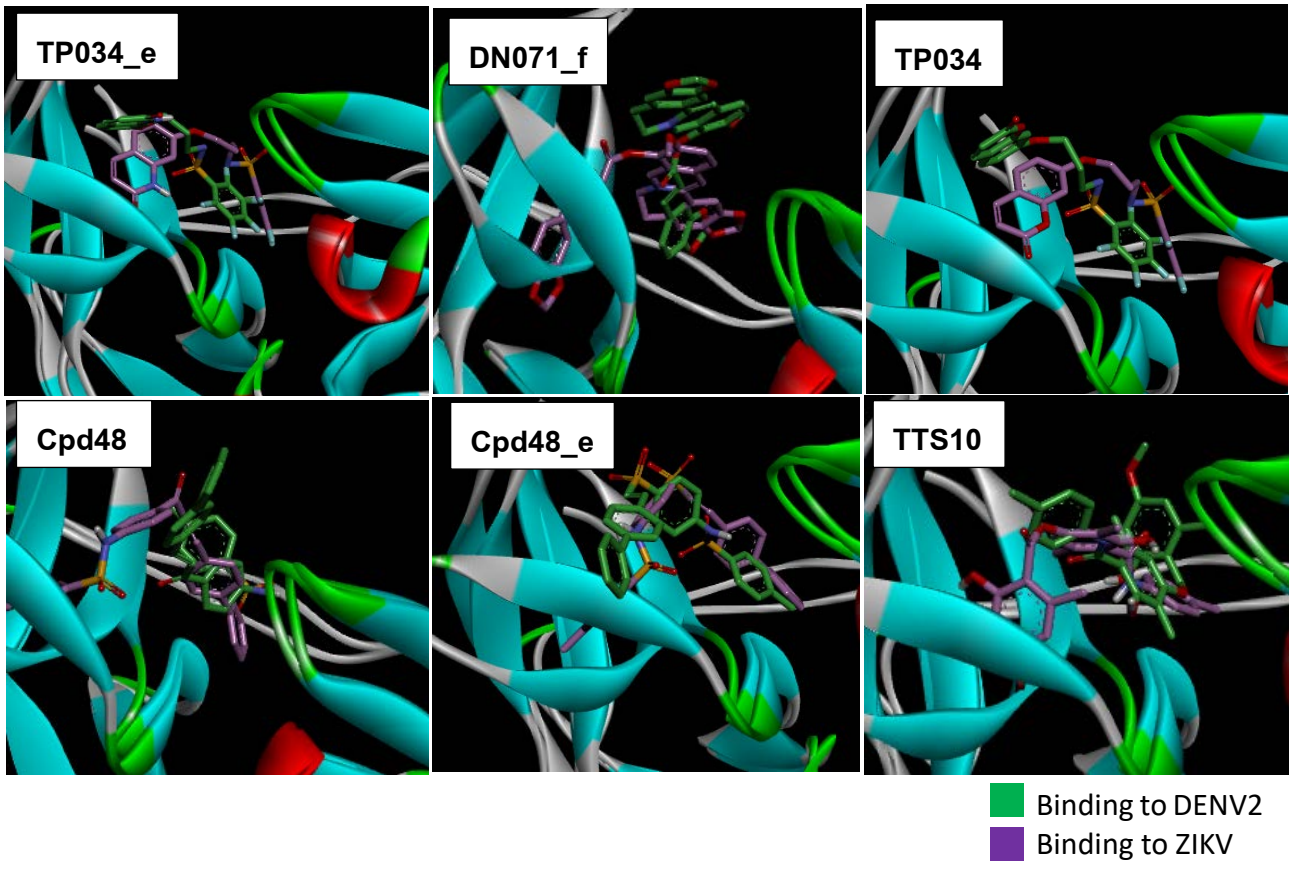


Figure 3.4 Orientation of 6 selected compounds binding to the allosteric site of DENV protease (green compound) and ZIKV protease (purple compound).

3.2 System stability and binding affinity of protein–ligand complex

In this study, we simulated NS2B/NS3–ligand complexes for 100 ns and investigated the system stability during MD simulation by calculation of root-mean-square deviation (RMSD) with respect to the initial structure, as shown in **Figure 3.5**. The RMSD results for heavy atoms of the protein (red line in **Figure 3.5**) showed that all the systems were likely to reach equilibrium at ~80 ns. Therefore, MD trajectories from the last 10 ns (90-100 ns) of the simulations were considered for subsequent binding free energy (ΔG_{bind}) calculation based on the MM/GBSA method (see below).

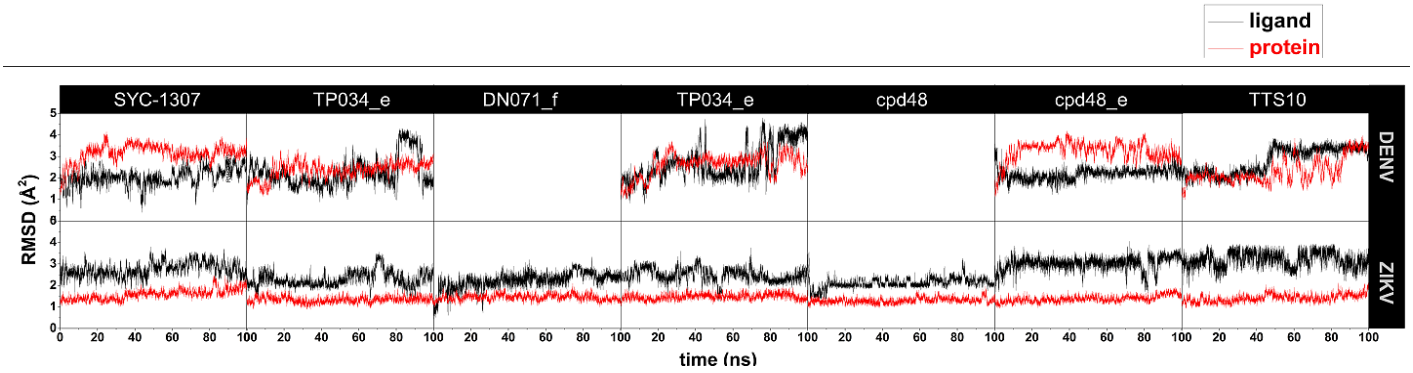


Figure 3.5 DENV and ZIKV RMSD result of SYC-1307, TP034_e, DN071_f, TP034, cpd48, cpd48_e, and TTS10.

By considering the molecular mechanics energy (ΔE_{MM}) in **Figure 3.6**, the result showed that van der Waals interaction ($\Delta \text{E}_{\text{vdW}}$) was important for molecular complexation between protein and ligand in all complexes rather than the electrostatic interaction ($\Delta \text{E}_{\text{ele}}$). Moreover, MM/GBSA-based binding free energy (ΔG_{bind}) was found in the order of DN071_f < TP034 < cpd48_e < cpd48 < SYC-1307 < TP034_e < TTS10 with the respective ΔG_{bind} of -17.17 ± 1.90, -15.87 ± 4.67, -14.16 ± 4.08, -12.92 ± 3.98, -12.39 ± 5.99, -7.41 ± 6.35, and 0.61 ± 1.57 kcal/mol. Obviously, the ΔG_{bind} values of DN071_f, TP034, cpd48 and cpd48_e were lower than that of SYC-1307. This result suggested that these four compounds showed the stronger binding affinity as compared to a reference compound SYC-1307, in which ZIKV NS2B/NS3–DN071_f complex exhibited the greatest binding efficiency. Note that the ΔG_{bind} values of TP034_e and TTS10 system were much higher than SYC-1307 system possibly due to an insufficient sampling of MD snapshots from the last 10-ns MD simulations, resulting in a somewhat high or even positive ΔG_{bind} value more than it should be.

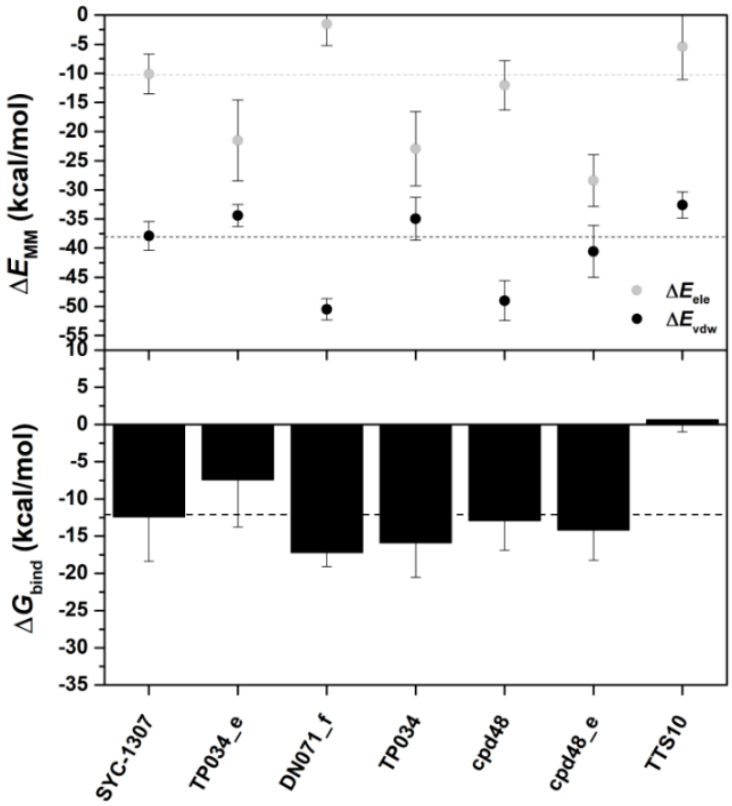


Figure 3.6 MM/GBSA-based binding free energy for each compound

3.3 Key binding residues involved in ligand binding

Apart from the total binding free energy calculation mentioned above, the per-residue decomposition energy ($\Delta G_{\mathbf{bind,residue}}$) based on the MM/GBSA method was also applied to determine the important residues involved in ligand binding. Note that the negative value and positive value represent the ligand stabilization and destabilization, respectively. The results showed that TRP69, GLN74, LEU76, TRP83, LEU85, ILE123, VAL147, ASN152, and VAL155 were considered as the key residues with the energy contribution lower than -1 kcal/mol (**Figure 3.7**). This is consistent with previous studies on identified allosteric inhibitors of DENV2 NS2B / NS3 protease through in silico molecular docking and found that residues GLN74 and ASN152 are essential for interactions with inhibitors [6]. These residues were all found in the NS3 domain. The lowest $\Delta G_{\mathbf{bind,residue}}$ value for SYC-1307, cpd48_e and TTS10 was observed at TRP83 (-2.94, -2.56, and -2.55 kcal/mol, respectively). Our results strongly agree with another report of DENV protease [36], suggesting that TRP83 is key residue for ligand binding of both ZIKV and DENV protease.

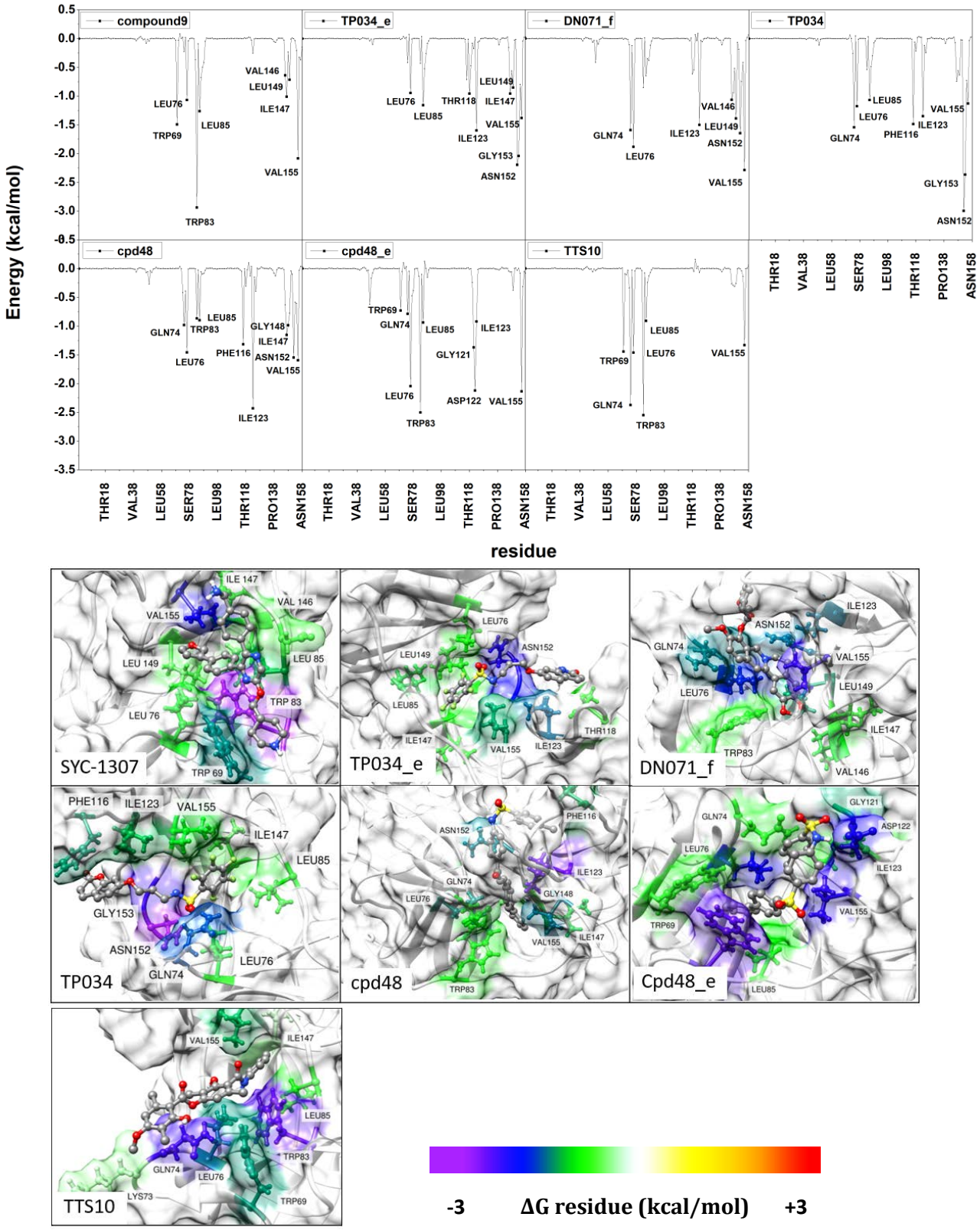


Figure 3.7 Per-residue decomposition $\Delta G_{\mathbf{bind,residue}}$ obtained from the MM/GBSA method for each system.

3.4 Hydrogen bond interaction

Hydrogen bond interaction is indeed important for molecular interaction between protein and ligand. The strength of hydrogen bonding formation can be determined by the percentage of hydrogen bonds occupation according to the two geometric criteria of (*i*) a proton donor (D) and acceptor (A) distance ≤ 3.5 Å and (*ii*) a D–H···A bond angle ≥ 120°. In addition, hydrogen bond occupations with the values of >80%, 50-79%, and <50% are defined as strong, medium, and weak hydrogen bond interactions, respectively. From the results, we found that the strong hydrogen bonds were formed by the residues ASN152 (86.8%) and GLY153 (98.3%) for TP034_e, ASN152 (90.3%) and GLY153 (91.5%) for TP034, and GLY121 (82.4%), ASP122 (99.2%) for cpd48_e, as shown in **Figure 3.8A**. Whereas weak hydrogen bonds were detected in the residues ASN158 (17.0%), TRP83 (2.6%), and ASP122 (7.7%) for SYC-1307, and residues GLY121 (7.7%) for TP034_e. Moreover, we found that weak hydrogen bonds were formed by residues THR118 with 12.4% for DN071_f, GLY121 with 15.7% for cpd48, and ASP71 with 17.9% for TTS10. Altogether, it can be assumed that hydrogen bonding played a minor role in stabilization of ZIKV NS2B/NS3-ligand complexes, which corresponded to the predominant van der Waal interactions as can be seen from the binding free energy calculation (see **section 4.2**).

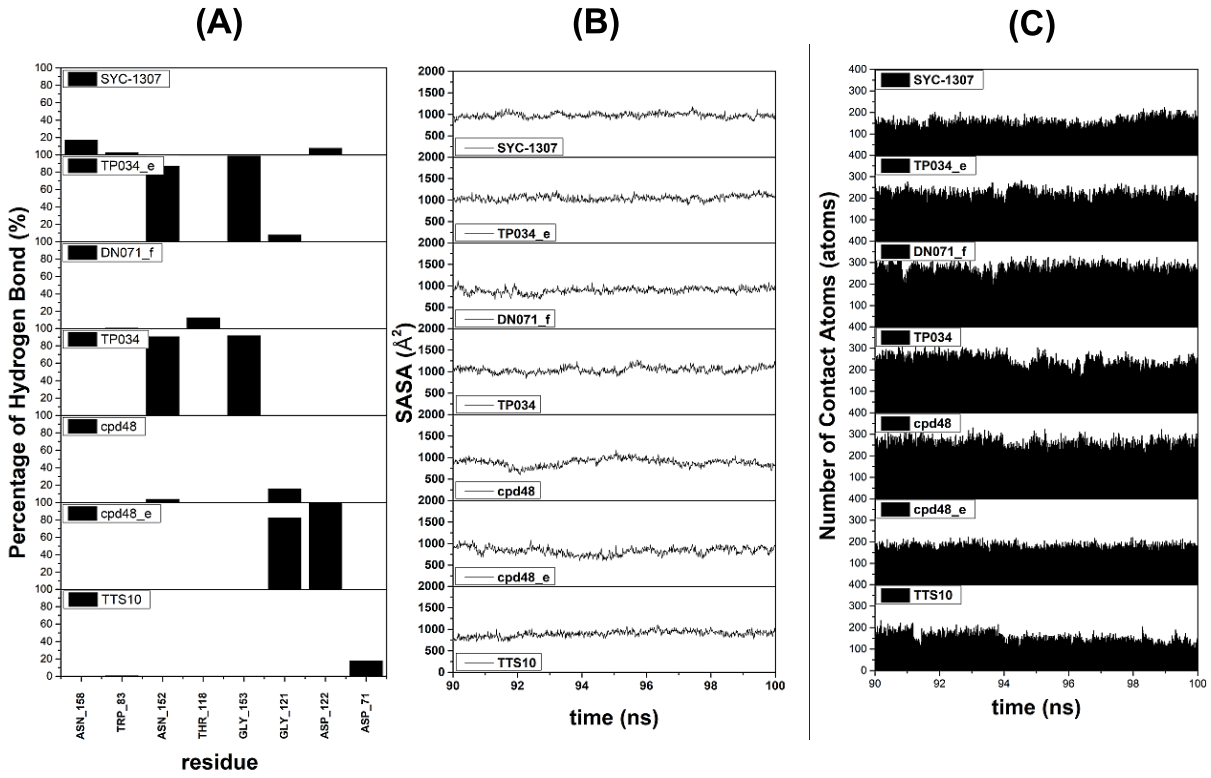


Figure 3.8 (**A**) Percentage of hydrogen bond. (**B**) SASA calculated for the amino acids of each ligand-binding pockets of NS2B/NS3 protease. (**C**) Number of contact atoms.

3.5 Solvent accessibility at the binding pocket

Since the allosteric pocket of NS2B/NS3 protease was exposed to the solvent medium, water molecules could interfere the binding of ligand to the protease. To measure such effect, water accessibility in terms of solvent accessibility surface area (SASA) within 5.0-Å sphere of each compound was calculated along the last 10 ns (**Figure 3.8B**). Previously, it was reported that the high SASA value reflected the high number of molecules occupied in the protein binding pocket [37, 38]. From our results, we found that the SASA value of system was approximately 800 – 1000 Å. The average SASA value of SYC-1307, TP034_e, DN071_f, TP034, cpd48, cpd48_e and TTS10 complexes was 976.777, 1045.444, 894.548, 1042.397, 885.692, 829.645 and 895.344 $Å^2$, respectively. Moreover, the number of contact atoms around each ligand bound is calculated and shown in **Figure 3.8C**. The average of contact atoms for SYC-1307, TP034_e, DN071_f, TP034, cpd48, cpd48_e, and TTS10 complex was 156, 219, 274, 241, 257, 178 and 152 atoms, respectively. When we compared each system with SYC-1307 complex, it was found that TP034_e, DN071_f, TP034, and cpd48 was appeared to have more contact atoms than that of SYC-1307 system. So, this finding suggested that these four compounds could fit well within the allosteric site of the enzyme better than SYC-1307, thereby leading to the greater binding affinity toward the ZIKV protease. However, the cpd48_e and TTS10 systems exhibited the lower number of contact atoms than SYC-1307 system, corresponding to the weaker binding affinity against the protease.

CHAPTER 4
CONCLUSIONS

In the present study, the search for the novel natural compounds and derivatives with inhibitory activity against ZIKV NS2B/NS3 serine protease was studied by molecular docking and MD simulations. From molecular docking technique, there were six compounds which had the higher binding affinity than that of a reference compound SYC-1307. Thus, these six compounds were selected for performing MD simulations and calculating binding free energy. From the total binding free energy calculation based on the MM/GBSA method, it was found that DN071_f showed the strongest binding efficiency toward NS2B/NS3 protease, and the major driving force in forming molecular complexation between protein and ligand was van der Waals interactions. Whereas the hydrogen bonding formation played a minor role in stabilization of protein-ligand complex. The per-residue decomposition energy calculation displayed that the residues important for ligand binding were TRP69, LEU76, TRP83, LEU85, ILE123, VAL147, and VAL155. Moreover, the results from SASA and number of contact atoms calculations support that DN071_f is the best compound which strongly bind to the allosteric site of ZIKV protease. Therefore, the compound DN071_f could be severed as the promising novel compound toward ZIKV protease and further studied *in vitro* and *in vivo* testing in the future.

SUGGESTIONS AND FUTURE DIRECTION TO THE PROJECT

In the senior project of Biochemistry student, the analysis of the results regarding discover a compound with inhibitory activity against NS2B/NS3 protease of DENV still was not complete. Therefore, **Figure 3.5** is not complete. Because sometimes MD simulation server was down, so we can do MD simulation less than we need.

In the future, we will test the compound that highest efficiency inhibitor against NS3/NS2B protease of both ZIKV and DENV with enzyme-based assay for study toxicity.

REFERENCES

1. Shi, Y. and G.F. Gao, *Structural Biology of the Zika Virus.* Trends Biochem Sci, 2017. **42**(6): p. 443-456.
2. Mahajan, P., et al., *A multi-target approach for discovery of antiviral compounds against dengue virus from green tea.* Network Modeling Analysis in Health Informatics and Bioinformatics, 2020. **9**(1): p. 20.
3. Balasubramanian, A., et al., *High-throughput screening for the identification of small-molecule inhibitors of the flaviviral protease.* Antiviral Res, 2016. **134**: p. 6-16.
4. Hou, J., et al., *Sequential immunization induces strong and broad immunity against all four dengue virus serotypes.* NPJ Vaccines, 2020. **5**: p. 68.
5. Yildiz, M., et al., *Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus.* ACS Chemical Biology, 2013. **8**(12): p. 2744-2752.
6. Mukhametov, A., et al., *Allosteric pocket of the dengue virus (serotype 2) NS2B/NS3 protease: In silico ligand screening and molecular dynamics studies of inhibition.* Journal of Molecular Graphics and Modelling, 2014. **52**: p. 103-113.
7. Boldescu, V., et al., *Broad-spectrum agents for flaviviral infections: dengue, Zika and beyond.* Nature Reviews Drug Discovery, 2017. **16**(8): p. 565-586.
8. Nutho, B. and T. Rungrotmongkol, *Binding recognition of substrates in NS2B/NS3 serine protease of Zika virus revealed by molecular dynamics simulations.* Journal of Molecular Graphics and Modelling, 2019. **92**: p. 227-235.
9. Nitsche, C., *Strategies Towards Protease Inhibitors for Emerging Flaviviruses*, in *Dengue and Zika: Control and Antiviral Treatment Strategies*, R. Hilgenfeld and S.G. Vasudevan, Editors. 2018, Springer Singapore: Singapore. p. 175-186.
10. Cataneo, A.H.D., et al., *The citrus flavonoid naringenin impairs the in vitro infection of human cells by Zika virus.* Sci Rep, 2019. **9**(1): p. 16348.
11. Brecher, M., et al., *A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease.* PLoS Pathog., 2017. **13**: p. e1006411.
12. Yao, Y., et al., *Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease.* Journal of the American Chemical Society, 2019. **141**(17): p. 6832-6836.
13. Lim, L., et al., *Curcumin Allosterically Inhibits the Dengue NS2B-NS3 Protease by Disrupting Its Active Conformation.* ACS Omega, 2020. **5**(40): p. 25677-25686.

14. Yusof, R., et al., *Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro.* J Biol Chem, 2000. **275**(14): p. 9963-9.
15. Millies, B., et al., *Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases.* J Med Chem, 2019. **62**(24): p. 11359-11382.
16. Roy, A., et al., *Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants.* PLOS ONE, 2017. **12**(7): p. e0180632.
17. Nutho, B., A.J. Mulholland, and T. Rungrotmongkol, *The reaction mechanism of Zika virus NS2B/NS3 serine protease inhibition by dipeptidyl aldehyde: a QM/MM study.* Phys Chem Chem Phys, 2019. **21**(27): p. 14945-14956.
18. Morris, G.M. and M. Lim-Wilby, *Molecular docking.* Methods Mol Biol, 2008. **443**: p. 365-82.
19. Duan, Y., et al., *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.* J Comput Chem, 2003. **24**(16): p. 1999-2012.
20. Hospital, A., et al., *Molecular dynamics simulations: advances and applications.* Adv Appl Bioinform Chem, 2015. **8**: p. 37-47.
21. Ponder, J.W. and D.A. Case, *Force fields for protein simulations.* Adv Protein Chem, 2003. **66**: p. 27-85.
22. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities.* Expert Opinion on Drug Discovery, 2015. **10**(5): p. 449-461.
23. Mahalapbutr, P., V.S. Lee, and T. Rungrotmongkol, *Binding Hotspot and Activation Mechanism of Maltitol and Lactitol toward the Human Sweet Taste Receptor.* Journal of Agricultural and Food Chemistry, 2020. **68**(30): p. 7974-7983.
24. Kollman, P.A., et al., *Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models.* Accounts of Chemical Research, 2000. **33**(12): p. 889-897.
25. Sitkoff, D., K.A. Sharp, and B. Honig, *Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models.* The Journal of Physical Chemistry, 1994. **98**(7): p. 1978-1988.
26. Dolinsky, T.J., et al., *PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W665-7.
27. Koebel, M.R., et al., *AutoDock VinaXB: implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina.* J Cheminform, 2016. **8**: p. 27.

28. M.J. Frisch, G.W.T., H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, et al., *Gaussian 09*. 2009, Gaussian, Inc.: Wallingford CT.
29. Wang, J., et al., *Automatic atom type and bond type perception in molecular mechanical calculations.* J Mol Graph Model, 2006. **25**(2): p. 247-60.
30. Maier, J.A., et al., *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.* Journal of Chemical Theory and Computation, 2015. **11**(8): p. 3696-3713.
31. Wang, J., et al., *Development and testing of a general amber force field.* J Comput Chem, 2004. **25**(9): p. 1157-74.
32. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water.* The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
33. Tom Darden, D.Y., and Lee Pedersen, *Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems.* The Journal of Chemical Physics, 1993. **98**: p. 10089-10092.
34. Ryckaert, J.-P., G. Ciccotti, and H.J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes.* Journal of Computational Physics, 1977. **23**(3): p. 327-341.
35. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath.* The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
36. Vajarintarangoon, L., *In Silico Screening of Anti-Silico Screening of Anti Dengue Drug Candidates against NS2BNS3 Serine Protease of Dengue Virus* in *Department of Biochemistry, Faculty of Science*. 2021, Chulalongkorn University.
37. Panupong, M., et al., *Rosmarinic Acid as a Potent Influenza Neuraminidase Inhibitor: In Vitro and In Silico Study.* Current Topics in Medicinal Chemistry, 2020. **20**(23): p. 2046-2055.
38. Morris, K.F., et al., *Molecular dynamics simulation study of AG10 and tafamidis binding to the Val122Ile transthyretin variant.* Biochemistry and Biophysics Reports, 2020. **21**: p. 100721.

VITA

NAME	Mr. Wachirapol Manimont
DATE OF BIRTH	23 April 1999
PLACE OF BIRTH	Bangkok, Thailand
HOME ADDRESS	1/151, Saimai 23/1, Saimai road, Saimai, Bangkok 10220
EMAIL	wachirapol89@gmail.com
EDUCATION	2017 – present (B.Sc.) Department of chemistry, faculty of science, Chulalonkorn university, 2011 – 2016 Bodindecha (Sing Singhaseni)