STRUCTURE BASED DESIGN AND SYNTHESIS OF POTENTIAL DENGUE VIRUS NS2B-NS3 PROTEASE INHIBITORS

SUMAN SINHA

UNIVERSITI SAINS MALAYSIA

2018

STRUCTURE BASED DESIGN AND SYNTHESIS OF POTENTIAL DENGUE VIRUS NS2B-NS3 PROTEASE INHIBITORS

by

SUMAN SINHA

Thesis submitted in fulfillment of the requirement

for the degree of

Doctor of Philosophy

April 2018

DEDICATION

This thesis is dedicated to my mother, who is my first teacher and my daughter, who is the source of my eternal pride and inspiration...

ACKNOWLEDGEMENT

Foremost I feel thankful to my research supervisor Professor Dr.Habibah A.Wahab for the being the beacon of light in the crux of my life. Under her guidance, I realize I have evolved as a confident and able professional as ever. I will never forget some of the advices I received from her over this time of PhD studentship. They remain as some of the key lessons of my life. I express my sincere gratitude to my mentor for being so patient with me, for motivating me and for listening to me whenever I wanted to discuss. I take this opportunity to also thank my co-supervisor Professor Dr. Hasnah Osman for her able guidance and all the motivation. Her guidance has helped me all throughout my studentship.

I thank Dr.Amirah Binti Mohd Gazzali for helping me with the manuscripts with her constructive criticism and open discussions.

I thank all laboratory technicians in Dept. of Pharmaceutical Technology (En. Shamsudin and En.Rosli), Dept. of Pharmaceutical Chemistry (En. Hamid, En. Fizal, En. Zaenudin and En. Annuar) and NMR centre, School of Chemical Sciences (Mr.Nizam) for helping me during experiments.

I thank my fellow lab mates who has graduated before me for all the discussions we had and all the experiences I shared with them. I thank my current lab mates, specially, Nadhira, Mira, Iza, Ema, Kesavan for all the discussions and the nice time we had together. I thank Arshad and Ragheed for being such a nice companion, although be it during the final phase of my stay in my lab.

I thank my family for being so patient with me all throughout my journey of doctoral studies. Thank you so much for being by my side.

TABLE OF CONTENTS

ACKNOWLEDGEMENT **ii**

TABLE OF CONTENTS **iii**

LIST OF TABLES **v**

LIST OF FIGURES **vi**

LIST OF SCHEMES **viii**

LIST OF ABBREVIATION& SYMBOLS **ix**

ABSTRAK **xi**

ABSTRACT **xiii**

CHAPTER 1: INTRODUCTION **1**

1.1 Statement of the Problem 1

1.2 Objectives 2

1.3 The Structure and Genome of Dengue Virus 2

1.3.1 Life cycle of dengue virus and role of important proteins 4

1.3.2 Dengue virus NS2B-NS3 protease 6

1.3.3 Proteolytic mechanism of NS2B-NS3 protease 8

1.4 NS2B-NS3 Protease Inhibitors 10

1.4.1 Peptide inhibitors 10

1.4.2 Peptide-conjugates 14

1.4.3 Small molecule inhibitors (natural and synthetic) 15

1.4.4 Structure-Based Drug Design (SBDD) 22

1.4.5 Molecular Docking 23

CHAPTER 2: METHODOLOGY **25**

2.1 Materials 25

2.1.1 Software and Hardware 25

2.1.2 Reagents and Other Consumable Materials 25

2.1.3 Instruments 27

2.2 Methods 27

2.2.1 Molecular modeling 27

2.2.2 Synthesis 28
2.2.3 Synthesis of naphthoquinone derivatives 29
2.2.4 Synthesis of cinnamic acid derivatives 31
2.2.4(a) Synthesis of 4-benzylideneoxazol-5-ones: 31
2.2.4(b) Synthesis of α-(Aroylamino) cinnamic acids: 32
2.2.5 Enzyme inhibition assay 32

CHAPTER 3: RESULTS AND DISCUSSION 34
3.1 Selection of natural product scaffolds for the design of NS2B-NS3 protease inhibitors 34
3.1.1 Naphthoquinone as scaffold for NS2B-NS3 protease inhibitors 34
3.1.2 α-(Arylamino) cinnamic acid as scaffold for NS2B-NS3 protease inhibitors 49
3.2 Synthesis of designed compounds 63
3.2.1 Synthesis of naphthoquinone Mannich base derivatives and binaphthoquinone derivatives 63
3.2.2 Synthesis of α-(arylamino) cinnamic acid derivatives 71
3.3 DEN2 NS2B-NS2 protease inhibition assay 77

CHAPTER 4: CONCLUSION 86
4.1 Future studies 86
REFERENCES 88
APPENDICES
PUBLICATIONS AND PRESENTATIONS

LIST OF TABLES

		Page
Table 1.1	The activity of peptides against DENV2 NS2B-NS3 protease	11
Table 1.2	Active modified peptides having inhibition against NS2B-NS3 protease.	12
Table 1.3	Natural product inhibitors of DENV2 NS2B-NS3 protease	16
Table 1.4	Synthetic Inhibitors of DENV2 NS2B-NS3 protease	18
Table 3.1	Binding affinity of quinone based ligands from literature	36
Table 3.2	Details of interactions of naphthoquinone based compounds	40
Table 3.3	The structures of naphthoquinone derivatives and their affinity towards NS2B-NS3 protease	43
Table 3.4	Binding energies of α-(arylamino) cinnamic acid based compounds	55
Table 3.5	Details of interactions of α-(arylamino) cinnamic acid based compounds towards NS2B-NS3 protease	55
Table 3.6	Binding energies of all designed α-(arylamino) cinnamic acid based compounds	61
Table 3.7	Summarized binding potential of all compounds towards NS2B-NS3 protease	80

LIST OF FIGURES

		Page
Figure 1.1	Schematic diagram of the dengue virus genome and polyprotein.	3
Figure 1.2	Life cycle of Dengue virus.	4
Figure 1.3	Schechter and Berger system of nomenclature	7
Figure 1.4	The structure of NS2B-NS3 complex.	8
Figure 1.5	The charge relay system	9
Figure 1.6	Mechanism of serine protease.	10
Figure 1.7	Cyclic peptides which are active against NS2B-NS3 protease	14
Figure 1.8	Peptide conjugates.	15
Figure 3.1	Flowchart of naphthoquinone based molecular design	35
Figure 3.2	Binding mode of naphthoquinone based inhibitors	37
Figure 3.3	Scaffold hopping of pyranoquinone to lawsone	39
Figure 3.4	The docked poses of designed naphthoquinone based inhibitors.	46
Figure 3.5	Binding modes of α-(arylamino) cinnamic acid based designed inhibitors.	52
Figure 3.6	Flowchart of α-(arylamino) cinnamic acid based molecular design	57
Figure 3.7	Binding poses of the designed compounds based on α-(arylamino) cinnamic acid based derivatives	58
Figure 3.8	Reaction mechanism of amino naphthoquinone formation	60

Figure 3.9	Proton NMR spectra of SS-5	63
Figure 3.10	Reaction mechanism of amino binaphthoquinone formation	69
Figure 3.11	Proton NMR spectra of SS-10	69
Figure 3.12	Reaction mechanism of amino α-(arylamino) cinnamic acid derivatives formation	72
Figure 3.13	Proton NMR spectra of SS-13	73
Figure 3.14	Binding poses of SS-5, SS-9 and SS-12	84

LIST OF SCHEMES

		Page
Scheme 2.1	Synthesis of binaphthoquinones	29
Scheme 2.2	Synthesis of amino-naphthoquinones	30
Scheme 2.3	Synthesis of cinnamic acid derivatives	31
Scheme 3.1	Synthesis of naphthoquinone Mannich base derivatives	63
Scheme 3.2	Synthesis of binaphthoquinone derivatives	68
Scheme 3.3	Synthesis of α-(arylamino) cinnamic acid derivatives	71

LIST OF ABBREVIATION& SYMBOLS

AUC	Area Under Curve
Arg	Arigine
Asn	Asparagine
Asp	Aspartic acid
br	Broad
C	Capsid
˚C	Degree centigrade
DHF	Dengue Hemorrhagic Fever
DENV	Dengue Virus
dd	Doublet
dt	Doublet of Triplet
E	Envelope
e.g.	*exempli gratia*
ER	Endoplasmic reticulum
FTIR	Fourier Transform Infrared
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
hr	Hour
IC_{50}	Inhibitory $Concentration_{50}$
Ile	Isoleucine
Kg	Kilogram
K_i	Inhibition constant
µg/mL	Microgram per milliliter
µL	Microliter
mg	Milligram
mL	Milliliter
mm	Millimeter
min	Minute
M	Molar
m	Multiplet
nm	Nanometer
NS	Non structural
NMR	Nuclear Magnetic Resonance

Phe	Phenylalanine
KBr	Potassium Bromide
prM	preMembrane
RFU	Relative Fluorescence Unit
RMSD	Root mean square deviation
rpm	Revolution per minute
Sec	Second
Ser	Serine
s	Singlet
SD	Standard deviation
Thr	Threonine
td	Triplet of doublet
Tyr	Tyrosine
UV	Ultra violet
UV	Ultraviolet
USFDA	United States Food and Drug Administration
Val	Valine

REKA BENTUK BERDASARKAN STRUKTUR DAN SINTESIS PERENCAT PROTEASE NS2B-NS3 BERPOTENSI VIRUS DENGI

ABSTRAK

Denggi disebarkan melalui gigitan nyamuk yang dijangkiti oleh salah satu daripada empat serotype virus denggi. Antara sasaran terapeutik yang mantap sepanjang penemuan drug anti-denggi, NS2B-NS3 protease adalah agak popular oleh kerana peranannya dalam pengeluaran virus matang dan memainkan peranan penting dalam mengekalkan kebolehjangkitan. Kerja-kerja di dalam tesis ini memerihalkan kajian komputer dan eksperimen dalam reka bentuk dan sintesis NS2B-NS3 protease baharu daripada struktur semulajadi seperti lawsone dan asid ferulik. Dalam kes naphthoquinone kelas sebatian, adalah difahamkan dari dok pengajian yang penubuhan molekul fleksibiliti boleh membawa kepada lebih ketat mengikat sebatian. Pada masa yang sama, rasional scaffold mengurangkan pendekatan telah dilaksanakan bertujuan untuk mengurangkan saiz molekul untuk membangunkan sebatian berat molekul yang rendah juga terbukti berhasil di laluan reka bentuk daripada anthrakuinon naptokuinon. Dok molekul yang terperinci dan analisis interaksi membawa kepada ramalan bahawa sabatian vanilin (VAN) dan lawsone (LAW) boleh menjadi aktif secara eksperimental. Cerakin enzim kedua-dua sebatian tersebut terhadap NS2B-NS3 protease kemudiannya membuktikan bahawa vanilin dan lawsone adalah aktif dengan IC_{50} masing-masing sebanyak 108 µM dan 878 µM. Reka bentuk tambahan berasaskan dok molekul membawa kepada sintesis lapan belas sebatian. Sebatian yang terpilih adalah milik tiga kelas kimia berbeza iaitu binaptokuinone, aminonaptokuinone dan asid α-(arilamino) sinamik. Semua sebatian yang disintesis telah dicirikan oleh FTIR, NMR ^{1}H dan ^{13}C dan Spektroskopi Jisim

(MS). Akhir sekali, kajian perencatan enzim membawa kepada pengenalpastian sebatian SS-2 dan SS-5 yang menunjukkan IC_{50} masing-masing sebanyak 37.92 µM dan 0.0065 µM. Secara keseluruhan, pendekatan reka bentuk berasaskan struktur dan pengesahan lanjut melalui eksperimen dalam kerja ini berjaya membuktikan konsep reka bentuk drug yang rasional.

STRUCTURE BASED DESIGN AND SYNTHESIS OF POTENTIAL DENGUE VIRUS NS2B-NS3 PROTEASE INHIBITORS

ABSTRACT

Dengue is transmitted by the bite of a mosquito infected with one of the five dengue virus serotypes. As far as anti-dengue drug discovery is concerned, among few established therapeutic targets, NS2B-NS3 protease is quite popular due to its role in production of mature viruses and plays a key role in maintaining infectivity. The work in this thesis described the computational and experimental studies in the design and synthesis of novel NS2B-NS3 protease from natural scaffolds such as lawsone and ferulic acid. In the case of naphthoquinone class of compounds, it was understood from docking studies that incorporation of molecular flexibility could lead to tighter binding compounds. Simultaneously, rational scaffold minimizing approach was undertaken with an aim to reduce molecule size to develop low molecular weight compounds also proved fruitful in the design pathway from anthraquinone to naphthoquinone. Detailed molecular docking and interaction analysis led to the prediction that compounds vanillin (VAN) and lawsone (LAW) could be experimentally active. Subsequent enzymatic assay of these two compounds against NS2B-NS3 protease proved that vanillin and lawsone were active with IC_{50} of 108 µM and 878 µM, respectively. Further design based on molecular docking studies led to the synthesis of derivatives of lawsone and vanillin. The selected compounds belonged to three distinct chemical classes, binaphthoquinones, aminonaphthoquinones and α-(arylamino) cinnamic acids. All the synthesized compounds were characterized by FTIR, proton and ^{13}C NMR, as well as mass spectroscopy. Finally, enzymatic inhibition studies led to the identification compound SS-2 (binaphthoquinone scaffold) and SS-5 (amino naphthoquinone)

which showed IC_{50} of 37.92 µM and 0.0065 µM respectively. Overall, the structure based design approach and subsequent experimental validation through the present work successfully proved the concept of computer aided drug design.

CHAPTER 1: INTRODUCTION

1.1 Statement of the Problem

Arthropod borne infection like dengue is a major concern to the world and a leading cause of illness and death in the tropical and subtropical regions of the world. It is estimated that 3.9 billion people in 128 countries are at risk of infection with dengue viruses (Brady et al. (2012)). Although the scale of dengue infection has not reached pandemic proportion, in Malaysia, from 10 to 16 January 2016, there were 3,500 cases of dengue reported (Hemisphere, (2016)). In Malaysia, there were 43347 cases of which 13081 were serologically confirmed (Omar et al. (2016)). Whilst in neighboring country, Indonesia, there were 500,000 patients being hospitalized in 2014 (Karyanti et al. (2014)). Between 2006 and 2012, India reported an annual average of 20,474 dengue cases (Shepard et al. (2014)) .

Dengue cases have increased drastically in the last 50 years due to the growth of population and uncontrolled urbanization. This further compounded by the improved transportation and ease of travelling both locally and internationally which have in one way or the other, facilitated the spread of dengue virus as well (Murray et al. (2013)).

Thus, there is indeed an urgent need to give due attention and put sufficient research fund in the development of effective antiviral agent for the treatment of dengue.

1.2 Objectives

Based on the problem statement mentioned above, the aim of this study is to develop novel inhibitors of DENV2 NS2B-NS3 protease. Specifically, the objective of this work are to:

1. Investigate the binding mode of important inhibitors of DENV2 NS2B-NS3 protease using molecular docking studies,

2. Design novel protease inhibitors through results obtained from molecular docking studies,

3. Synthesize the designed protease inhibitor molecules,

4. Conduct bioassay studies to evaluate the inhibition potential of the synthesized compounds against protease.

1.3 The Structure and Genome of Dengue Virus

Dengue fever is caused by infection from any one of the five serotype of the virus (DENV1-5) that may be transmitted to human by a day-biting female mosquito *Aedes aegypti* (Henchal & Putnak, (1990)). The virus is a sphere-shaped virus with an envelope. It comprises of a single, positive strand RNA genome that is packed by three structural proteins and a lipid bilayer. The structural proteins are named as capsid, C; membrane associated, M and an envelope, E (Kuhn et al. (2002); Rey, (2003)). The RNA of dengue virus consists of a 5'-untranslated region (UTR), a single open reading frame (ORF) and 3'-untranslated region (UTR) (Dong, Zhang, &

Shi, (2008)). These three structural proteins are encoded by 5'terminus of the open reading frame of the genome. The remainder of the ORF encodes seven nonstructural proteins i.e.NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (Fig.1.1).

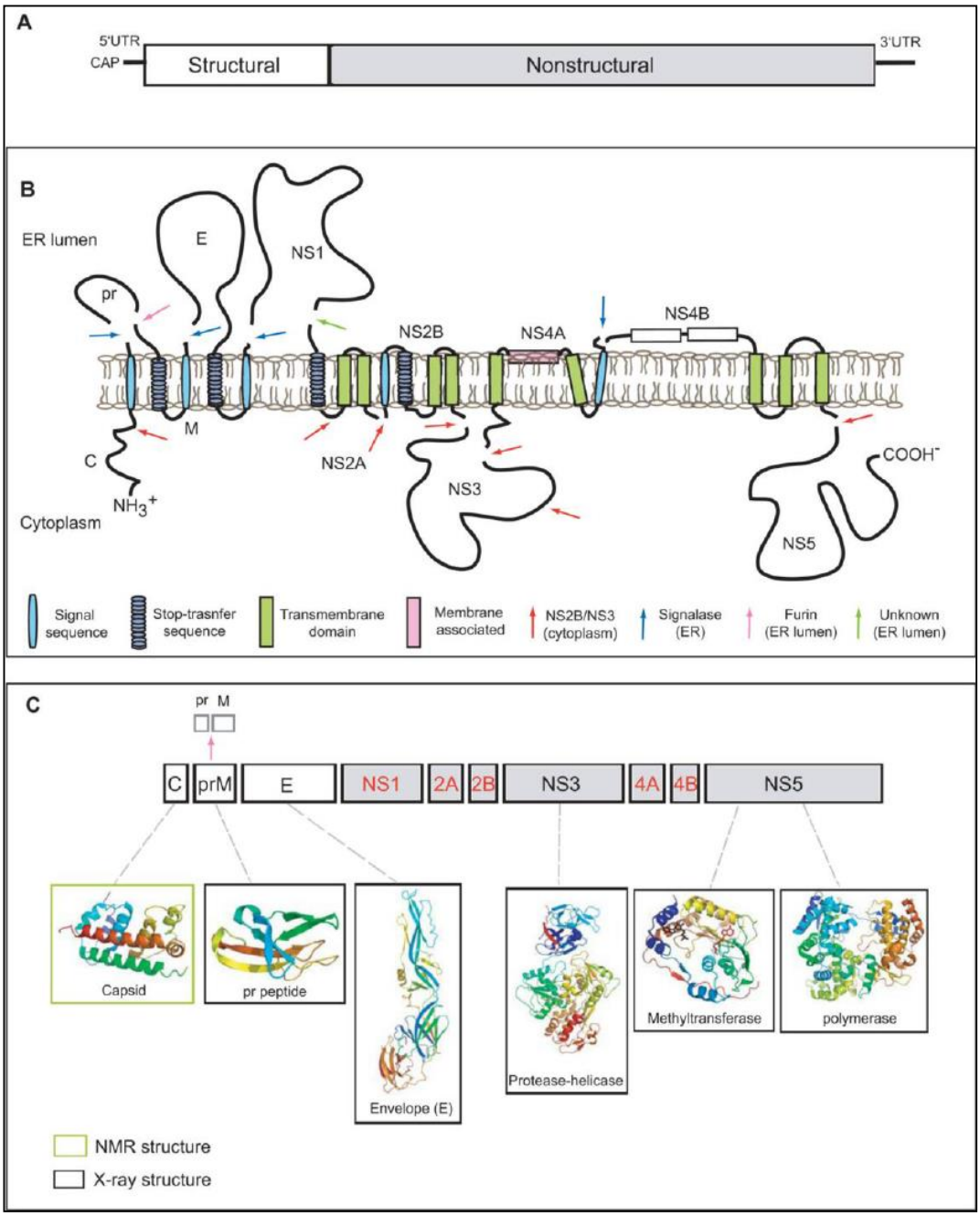


Figure 1.1 Schematic diagram of the dengue virus genome and polyprotein (Aruna, 2014).

1.3.1 Life cycle of dengue virus and role of important proteins

The life cycle of the dengue virus starts with the attachment and binding of the virions on the host cell surface, followed by entry into the cell by endocytosis and endosome formation (Fig.1.2).

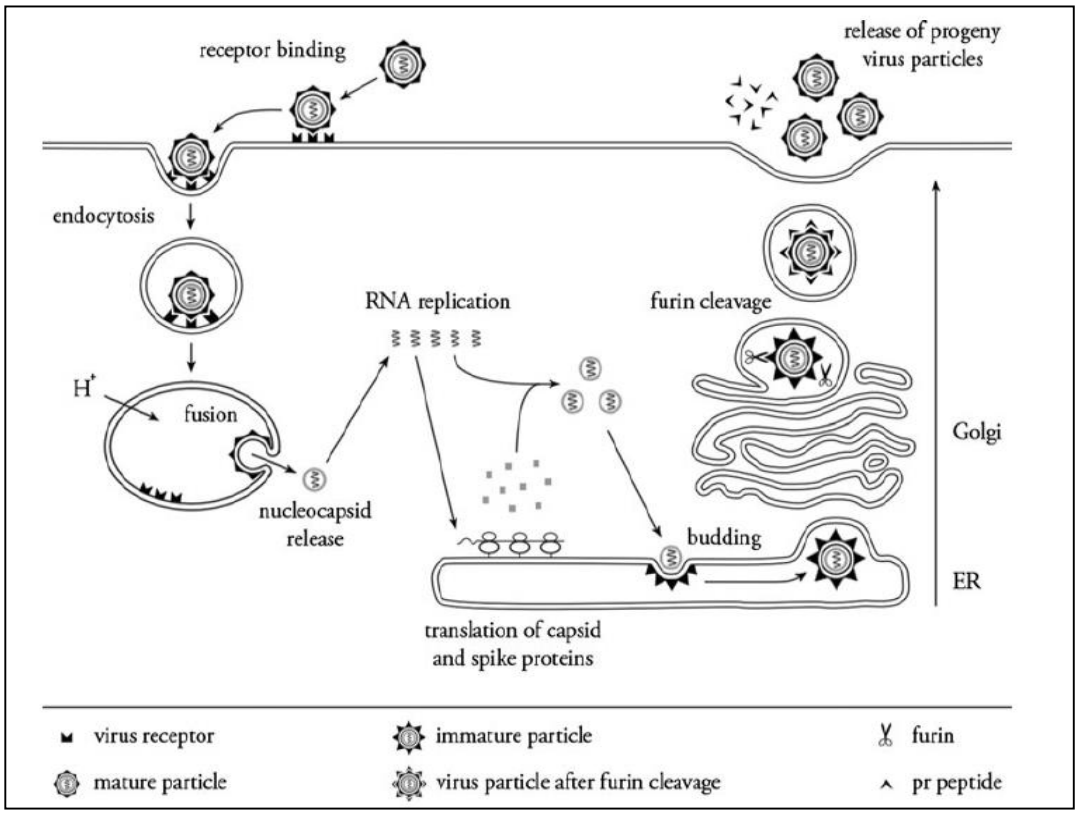


Figure 1.2 Life cycle of Dengue virus. (Aruna, 2014)

After entry of virus cell and subsequent uncoating of the nucleocapsid, the RNA molecule is translated into a single polyprotein (Miller, Kastner, Krijnse-Locker, Bühler, & Bartenschlager, (2007)). During this process, the signal and stop-transfer sequences of the polyprotein direct its back-and-forth translocation across the endoplasmic reticulum (ER) membrane. The polyprotein is processed by proteases into three structural proteins (C, prM, and E) and seven nonstructural (NS) proteins

(NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). The E protein is glycosylated at amino acid residue Asn67 and Asn153 to obtain appropriate folded conformation of the protein (Chambers et al. (1995)). Other potential N-linked glycosylation sites are located in prM at position 7, 31, and 52 and within NS1 at position 130 and 207 (Luo et al. (2008); Wu et al. (2003)). Once the proteins are translated and folded, the non-structural proteins initiate replication of the viral genome (Miller et al. (2007)).

The newly synthesized RNA is subsequently packed by the C protein to form a nucleocapsid. The prM and E proteins form heterodimers that are oriented into the lumen of the ER which will then be associated into trimers. These oligomeric interactions are believed to induce a curved surface lattice, which helps in virion budding (Chambers, (1990); Falgout et al. (1991)).

Interestingly, encapsulation of nucleocapsid during virus assembly is not pivotal as the genesis of capsid-less subviral particles (Bera et al. (2007)); Clum et al., (1997)); Gorbalenya et al., (1993); Shiryaev et al. (2007)). Detailed structural analysis of newly formed virions has revealed that a single virion consist of 180 prM/E heterodimers which protrude outward from the surface of the virus as 60 trimeric spikes (Aleshin et al., 2007; Erbel et al. 2006; Wu et al., (2005)). The immature particles formed in the ER continue to mature during their journey through the secretory pathway. The slightly acidic pH (*5.8–6.0) of the trans-Golgi network (TGN) leads to dissociation of the prM/E heterodimers, which further initiates the formation of 90 dimers that lie flat on the viral surface, with prM capping the fusion

peptide of the E protein. This structural reorganization of the glycoproteins leads to cleavage of prM by the cellular endoprotease furin (Xu et al. (2005); Yamashita et al. (2008); Yao et al. (1997)). Furin cleavage happens at a Arg-X-(Lys/Arg)-Arg (where X is any amino acid) recognition sequence and formation of membrane-associated M and a "pr" peptide takes place. A recent study has shown that the pr peptide remains attached with the virion until the virus is secreted to the extracellular environment (Xu et al. (2005)). Both prM protein and pr peptide are believed to act as chaperones that stabilizes the E protein during passage through the secretory pathway. This in turn, prevents premature conformational changes of the E protein that may lead to membrane fusion. After dissociation of the pr peptide, mature virions are formed which then released to infect new cells.

1.3.2 Dengue virus NS2B-NS3 protease

The proteases class of enzymes carry out protein hydrolysis at an accelerated rate although with absolute specificity which ensures prevention of uncontrolled proteolysis of the proteins for any organism. Proteases can be further divided into serine, cysteine, aspartic, threonine and metallo proteases. The catalytic nucleophile (hydroxyl group) of serine in serine protease active site helps in the degradation of proteins. Protease binds with the substrate in a specific manner that is represented by Schechter and Berger system of nomenclature (Berger A. (1967)) (Fig. 1.3).

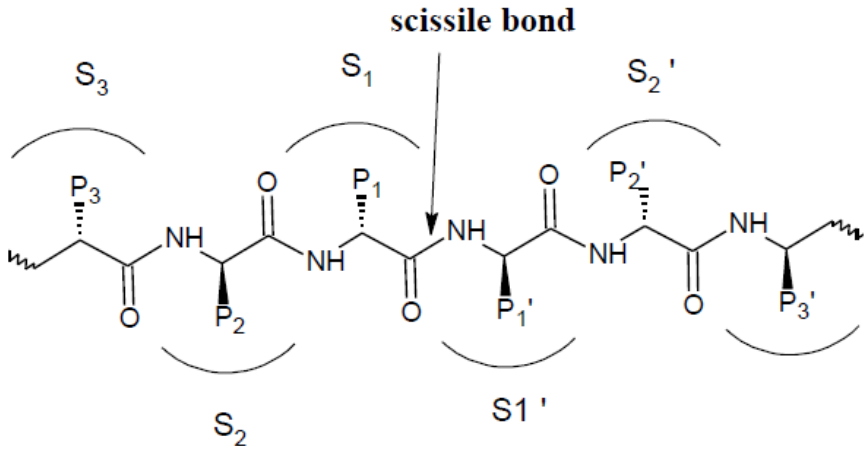


Fig 1.3 Schehter and Berger system of nomenclature.

In this method of nomenclature, the amino acid residues (or side chains) of substrate are labeled from the N to C terminus as Pn,.....,P3,P2,P1,P1',P2',P3',.....Pn' and the corresponding binding sub sites of the enzyme are labeled as Sn,.....,S3, S2, S1, S1',S2',S3',.....Sn'. The peptide bond hydrolysis is carried out between P1,P1' and the bond is called the scissile bond (Fig. 1.3).

In this case, the P1 residue of the substrate is called the primary specificity residue. The specificity and selectivity of a protease primarily depends on the nature of the P1 residue. Serine proteases are further classified as elastase-like (small hydrophobic residues at P1), trypsin-like (positively charged residues/basic residues at P1) or chymotripsin-like (large hydrophobic residues at P1).

The N-terminal domain of NS3 (aa 1–169) is a chymotrypsin like serine protease that cleaves the viral polyprotein (Chambers et al. (1990)) (Fig.1.4). The NS3 protease requires the NS2B cofactor to function as an active enzyme (Falgout et al,

1991). NS2B is a14 kDa protein that consists of three domains: two trans-membrane regions located at the N and C terminals and a central region of 47 amino acids (amino acids 49–96) which acts as a protein cofactor of the NS3 protease (Clum et al. (1997)). The dengue virus NS3 protein is neither soluble nor catalytically active as a protease *in vitro*, suggesting that it does not fold properly without the NS2B protein(Kim et al. (2013); T. Xu et al. (2005)).

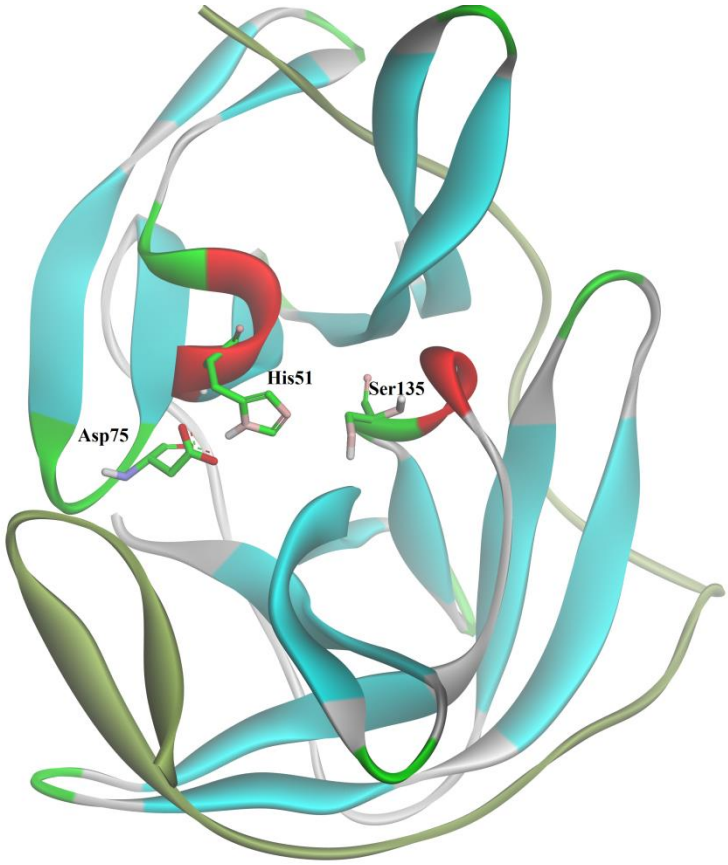


Figure 1.4 The structure of NS2B-NS3 complex and orientation of active site residues: Asp75, His51 and Ser135 (Wichapong, (2010)).

1.3.3 Proteolytic mechanism of NS2B-NS3 protease

DENV NS2B-NS3 protease is a member of serine protease family. These proteases are a class of proteolytic enzymes, which are characterized by the presence of

reactive serine residues. It has a catalytic triad containing histidine, serine and aspartic acid residues. As shown in Fig. 1.4, the active site's geometry as well as the proteolytic activity is governed by hydrogen bonding interaction of the catalytic residues. The hydrogen bonding networks generate a very reactive oxyanion species on serine that attacks the carbonyl group of the peptide bond in the substrate which leads to the formation of a tetrahedral intermediate, in which the carbonyl carbon of the substrate changes from sp^2 to sp^3. This process converts serine to alkoxide (ionised serine) through a charge relay system (Fig. 1.5) (Melino et al., 2007).

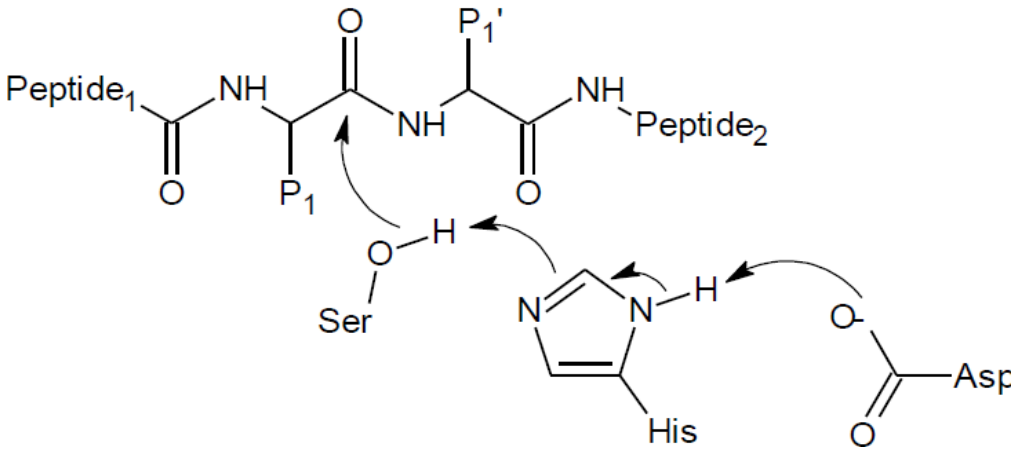


Fig. 1.5 Charge relay system for activation of an active site serine residue.

The active site of serine (activated nucleophile) then attacks the carbonyl group of the scissile amide bond. This intermediate undergoes rearrangement to release N-terminal peptide leaving C-terminal attached to serine. Subsequent hydrolysis by the water molecule near the active site releases C-terminal peptide from serine which regenerates the catalytic triad for its next cycle of proteolysis (Fig. 1.6).

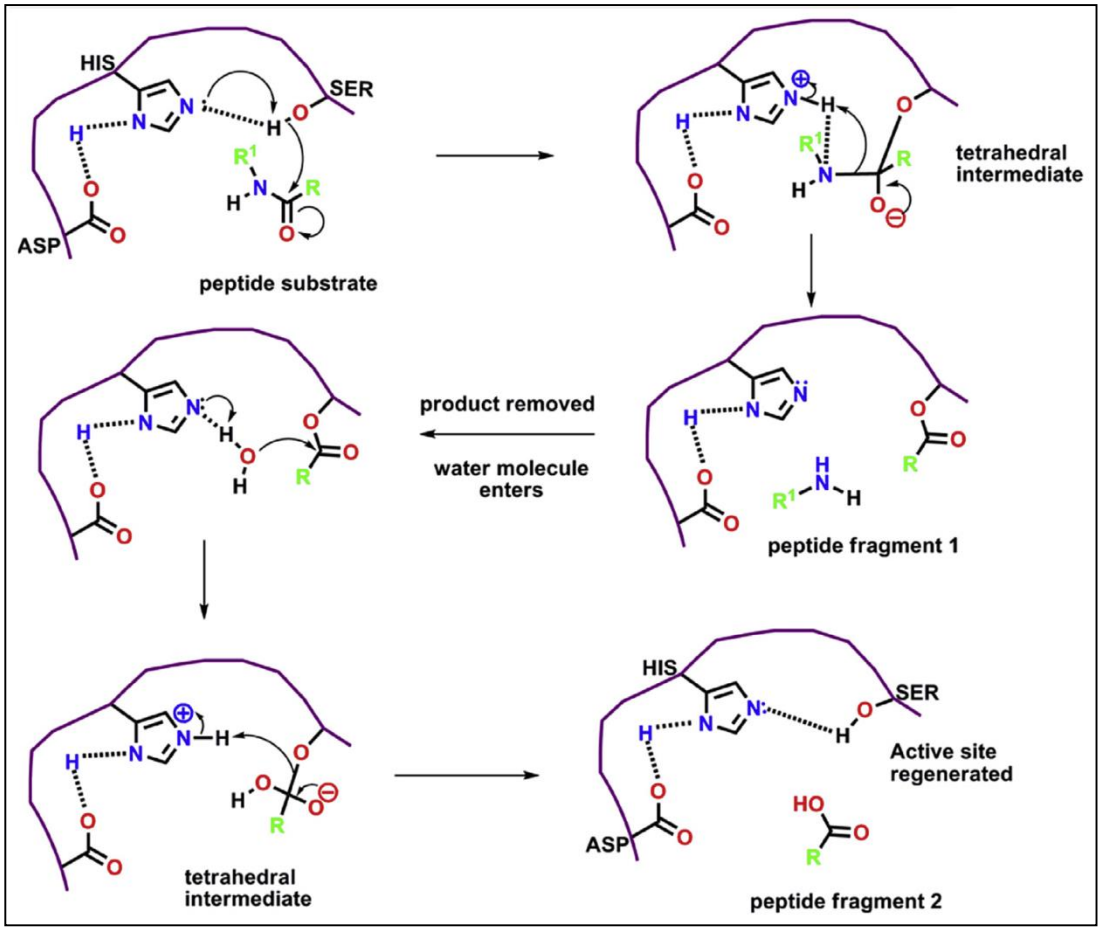


Figure 1.6 Mechanism of serine protease (Timiri, Sinha, & Jayaprakash, (2016)).

1.4 NS2B-NS3 Protease Inhibitors

1.4.1 Peptide inhibitors

The discovery of dengue NS2B-NS3 protease inhibitors were initially based on the information obtained from kinetic profiling of non-prime substrates. (Li et al. (2005)) (Frecer et al., (2010), Yang et al. (2011), Katzenmeier et al., (2004)).These efforts were quite challenging as the enzyme has a solvent-exposed, shallow active site and is dependent on the selectivity for substrates containing basic amino acids (arginine and lysine) at P1 and P2 positions (Cregar-Hernandez et al. (2011)). Nevertheless, many peptidic inhibitors have been discovered to have NS2B-NS3 protease

inhibition activities (Behnam et al. (2016); Nitsche et al.(2012), Nitsche et al. (2013), Rothan et al. (2012), Prusis et al. (2013), Xu et al. (2012)).

Li et al. (2005) designed synthetic peptides from the substrate kinetic profiling attached to electrophilic war-heads of serine proteases inspired by earlier work (Powers, et al., (2002)). Starting from two tetra-peptides (Bz-Nle-Lys-Arg-Arg-OH and Bz-Nle-Lys-Thr-Arg-OH)* sequences a series of ten inhibitors with different electrophilic warhead were synthesized. Peptides having boronic acid warhead (Bz-NIe-Lys-Arg-Arg-$B(OH)_2$) was shown to have an excellent inhibition at a concentration of Ki of 43nM(Yin et al. (2006)).

Tables 1.1 and 1.2 summarize some sequence and activity of the key peptide inhibitors which are being reported in literature.

Table 1.1 The activity of peptides against DENV2 NS2B-NS3 protease

Peptide	**Activity (μM)**
Bz-Nle-Lys-Arg-Arg-H	K_i= 5.8
Bz-Nle-Lys-Arg-Arg-$B(OH)_2$	K_i= 0.043
Bz-Nle-Lys-Arg-Arg-CF_3	K_i=0.85
Bz-Phe-Lys-Arg-Arg-H	K_i= 6.8
Bz-Lys-Arg-Arg-H	K_i = 1.5
Ph-Ac-Lys-Arg-Arg-H	IC_{50} = 6.7
2-cyano-N-cyclopropylbut-2-eanmide-Arg-Lys-Nle-NH_2	K_i= 4.9
Ac-Phe-Ala-Ala-Gly-Arg-Arg-H	K_i = 16.3
Retrocyclin-1	IC_{50}= 21.4
Protegrin-1	K_i= 5.85
Latarcin 1	IC_{50} = 12.7

Prusis and co-workers implemented a rational step-wise substitution/deletion of charged amino acids from prime site in order to achieve drug-likeness. A peptide inhibitor, WYCW-NH_2 showed significant protease inhibition with an IC_{50} of 4.8μM (Prusis et al. (2013)). Molecular modeling studies revealed that the *N*-terminal of the inhibitors interacted with Asp75 residue of the active site. The first two amino acids of the peptide interacted with S3 and S4 pockets, respectively while the fourth amino acid was found to established pi-pi stacking interaction with Tyr161 of S1 pocket.

Table 1.2 Active modified peptides having inhibition against NS2B-NS3 protease.

Inhibitor	**Ki value (μM)**
Ac-Phe-Ala-Ala-Gly-Arg-Arg-CHO	16 ± 3
Ac- Phe-Ala-Ala-Gly-Arg-Arg -αketo-Ser-Leu-$CONH_2$	47 ± 3
Ac-Thr-Thr-Ser-Thr-Arg-Arg -αketo-SL-$CONH_2$	220 ± 55
Ac- Thr-Thr-Ser-Thr -Arg-Arg -αketo-GTGN-$CONH_2$	368 ± 47
Ac-Arg-ThrSer-Lys -Lys-Arg -$CONH_2$	12.14
Ac- Lys -Lys-Arg -$CONH_2$	22.31
Ac- Phe-Ala-Ala-Gly-Arg- Lys -$CONH_2$	25.87
Ac- Thr-Thr-Ser-Thr-Arg-Arg -$CONH_2$	45.96
Ac-EV Lys -Lys-Gln Arg -$CONH_2$	66.68
Ac-Lys-Arg-$CONH_2$	121.53
Ac-Gly-Lys-Arg-$CONH_2$	152.30
Ac-Ser-Lys -Lys-Arg-$CONH_2$	187.60

Bioavailability and metabolic stability are two major challenges that have to be addressed in the development of linear peptide as an inhibitor. The problems were addressed by designing cyclic peptides which have the advantage of good permeability and metabolic stability when compared with linear peptides (Horton, et al., (2002)).

Various cyclic peptides were already being evaluated against DENV NS2B-NS3 protease and in support this case, four examples could be highlighted (Fig.1.7). Kalata B1 is a plant-based cyclopeptide (Fig.1.7a) and its derivatives (Fig. 1.7a(i) and (ii)) were synthesized and evaluated against dengue protease (Gao, et al., (2010)). They were found to inhibit DENV2 protease with an IC_{50} of 4.3 µM and 9.3 µM respectively.

In a different study, retrocyclin-1 (Fig. 1.7b) was tested for DENV NS2B-NS3 protease activity at different temperatures. The (IC_{50}) values at 40^{o}C, 37^{o}C and 28^{o}C were found to be 14.1, 21.4 and 46.1 µM respectively. Retrocyclin-1 also inhibited DENV2 replication in Vero cells (Rothan et al. (2012)).

The third example, conotoxin MrIA (Fig. 1.7c), is a peptide from *Conus marmoreus* which was found to inhibit DENVNS2B-NS3 protease significantly (Xu et al. 2012). Synthetic Protegrin-1 (Fig. 1.7d) is another cyclic peptide which was shown to have activity against DENV NS2B-NS3 protease. It was found to inhibit the replication of DENV2 in MK2 cells and DENV2 NS2B-NS3 protease with an IC_{50} of 11.7 µM (Rothan et al. (2012)). In that study, three concentrations of PG-1 were used to test the peptide stability and its capability to reduce dengue viral replication in MK2 cell lines. The results clearly explain that the viral copy number got reduced with increasing concentrations of protegrin-1. In addition to that, the highest inhibition percentage was observed when the PG-1 concentration was 12.5 µM at 24, 48 and 72 hrs.

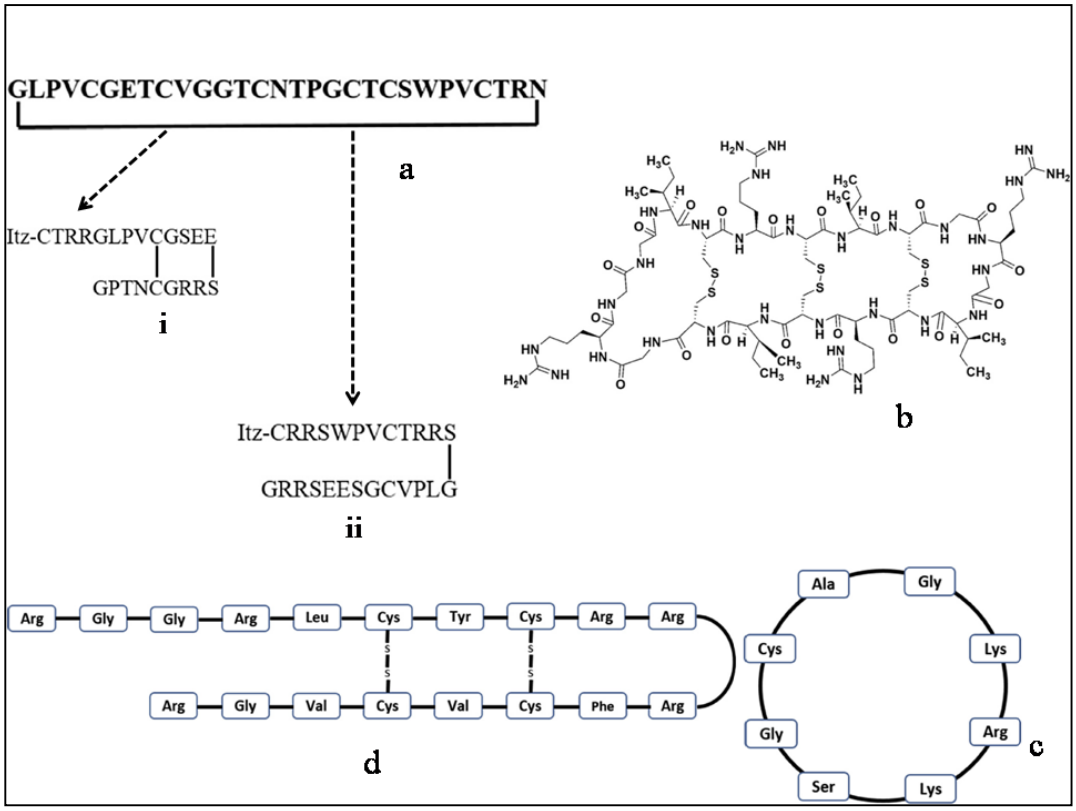


Figure 1.7 Cyclic peptides which are active against NS2B-NS3 protease.

1.4.2 Peptide-conjugates

Peptides attached to small organic molecules could imitate the natural substrates and may be a potential inhibitor for a targeted protein. The strategy behind this approach was to combine the drug-likeness of small molecules with the specificity of substratemimicking peptides. Thiazolidine and rhodanine (cap) scaffolds were conjugated with substrate mimicking peptides (Arg-Lys-Nle-NH_2) to get *N*-substituted 5-arylidenethiazolidinone-(di/tri)-peptide hybrid constructs (cap-Arg-Lys-Nle-NH_2). These compounds were evaluated for their inhibitory activity on DENV NS2B-NS3 protease as presented in Fig.1.8a. Molecular modeling evaluation revealed the location of cap region (rhodanine and thiazolidinone) near the

hydrophobic region of NS2B cofactor whilst Nle, Lys and Arg residues were located in S10, S1 and S2 pockets, respectively (Nitsche et al. (2013)).

In a second example as illustrated in Fig.1.8b, a new class of dipeptides with a core moiety of methionine-prolineanilides were developed as potential DENV NS2B-NS3 protease inhibitor. It was revealed that for an optimum activity, it is necessary to have *p*-nitrophenyl group, proline with *S*-configuration and methionine with free amino- and thioether in non-oxidized form (Zhou et al. (2013)).

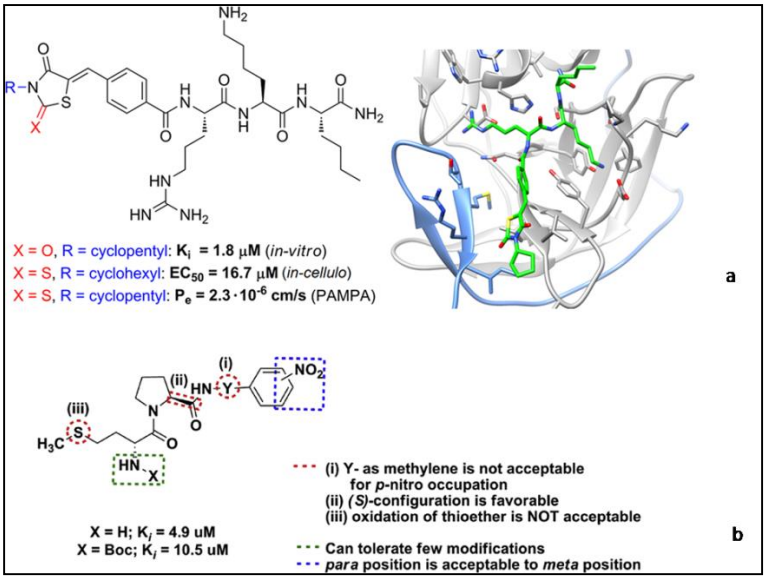


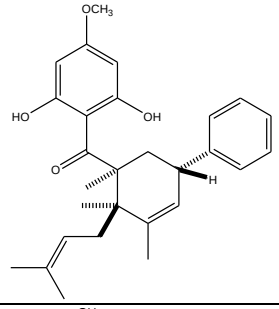
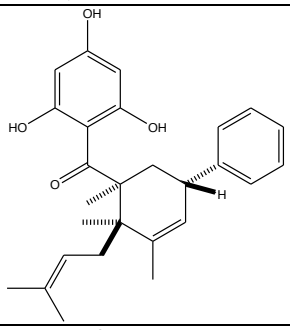
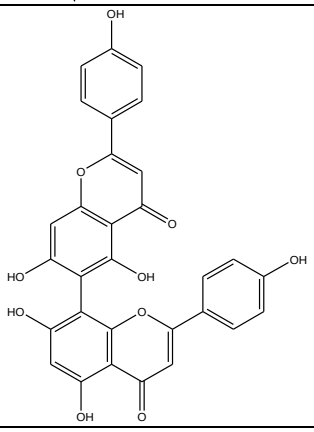
Figure 1.8 Peptide conjugates active against NS2B-NS3 protease. *a* (Nitsche et al. (2013)) and b (Timiri et al. (2016))

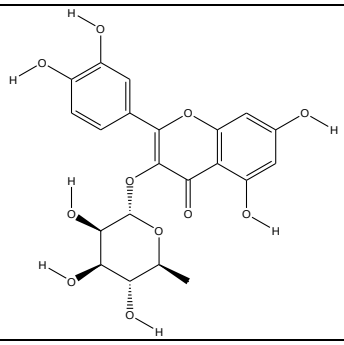
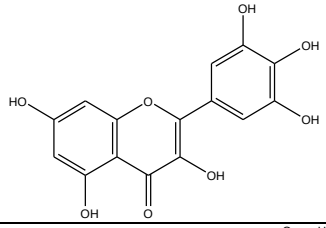
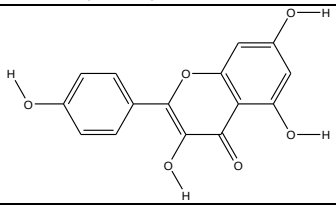
1.4.3 Small molecule inhibitors (natural and synthetic)

Natural products such as 4-hydroxypanduratin A and panduratin A (Table 1.3) were also reported to inhibit DENV protease with a K_i of 21 μM and 25 μM against DENV NS2B-NS3 protease (Kiat, Pippen, Yusof, & Ibrahim, (2006)).

Agathisflavone and myricetin were observed to be non-competitive inhibitors of dengue virus serotype 2 NS2B-NS3 protease with Ki values of 11 and 4.7 μM, respectively (de Sousa et al. (2015)). Docking studies propose a binding mode of the flavonoids in a particular allosteric binding site (around Gln88, Gln167, and Gly124) of the enzyme which is near to its catalytic triad (de Sousa et al. (2015)).

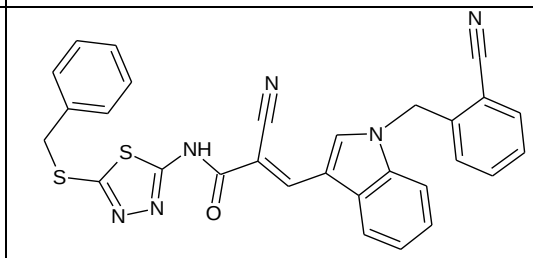
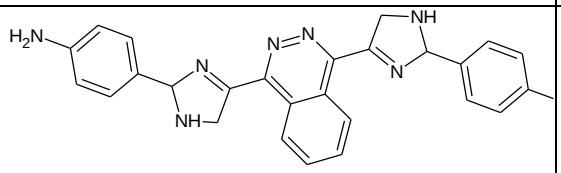
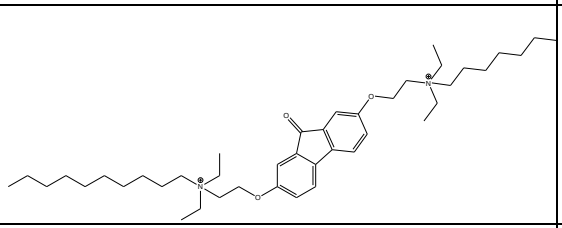
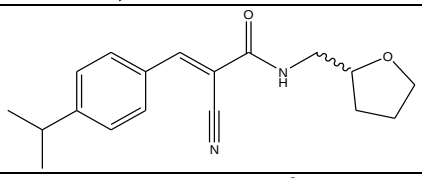
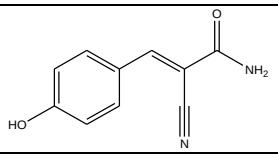
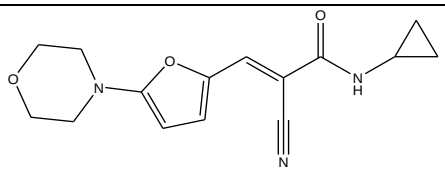
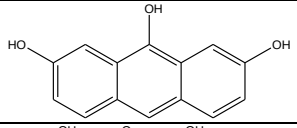
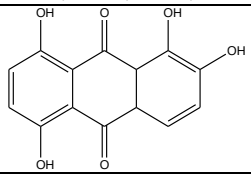
Table 1.3 Natural Product Inhibitors active against NS2B-NS3 protease.

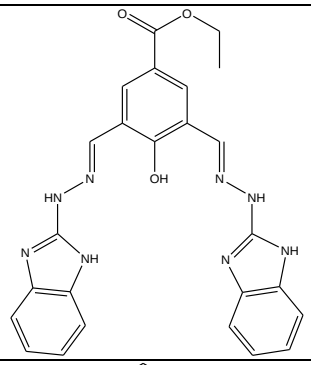
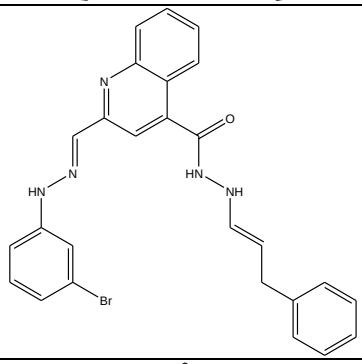
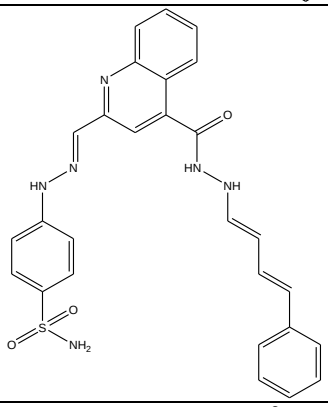
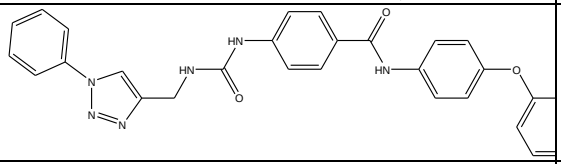
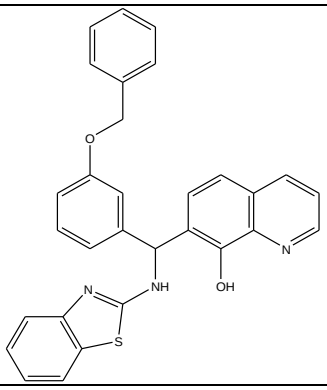
Compound Name	**Chemical class**	**Structure**	**Activity (μM)**		**Source plant**	**References**
			DV2	DV3		
Panduratin A	Cyclohexenyl Chalcone derivative		K_i= 25	NA	*Boesenbergia rotunda* (L.) Mansf.	Kiat et al. 2006
4-hydroxypanduratin A	Cyclohexenyl chalcone derivative		K_i= 21	NA	*Boesenbergia rotunda* (L.) Mansf.	Kiat et al. 2006
Agathisflavone	Biflavonoid		K_i = 11.1	K_i = 15.3	*Cenostigma macrophyllum* Tul.	de Sousa et al. 2015

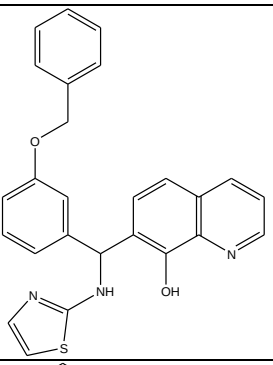
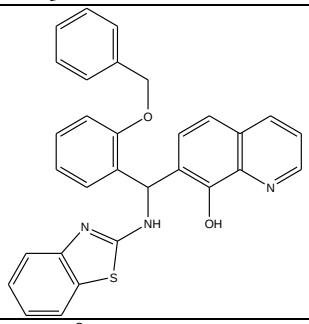
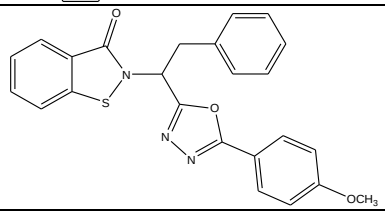
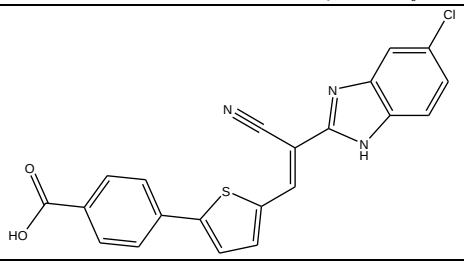
Quercitrin	Flavonoid		K_i = 25.7 ±	K_i = 28.0	*Byrsonima Coccolobifolia* Kunth	de Sousa et al. 2015
Myricetin	Flavonoid		K_i = 4.7	K_i = 25.7	*Byrsonima coccolobifolia* Kunth	de Sousa et al. 2015
Kaempferol	Flavonoid		K_i = 22.3	K_i = 24.2	*Byrsonima Coccolobifolia* Kunth	de Sousa et al. 2015

Arylcyanoacrylamide derivatives (Table 1.4) with electrophilic nitrile trap for Ser135 hydroxyl group was found to be yet another logical approach in designing DENV NS2B-NS3 protease inhibitors. Electron-donating groups like hydroxyloramine, either at *para*- or *meta*-position is essential to have increased potency. It was assumed that aryl and amide terminals were accommodated in S1 and S1' pockets, respectively (Nitsche, Steuer, & Klein, (2011)). The best compound for this effort showed a percentage of inhibition of 46.6 ±1.4 with an IC_{50} of 50 µM.

Table 1.4 Synthetic Inhibitors active against NS2B-NS3 protease

Chemical class	Structure	Activities (μM)	References
Thiadiazoloacrylamide		IC_{50} =2.24	Liu et al. 2014
Imidazophthalazine		IC_{50} = 6.00	Bodenreider et al. 2009
Alkyl ammonium		IC_{50} = 15.4	Yang et al. 2011
Arylcyanoacrylamide		K_i = 98.1 μM*	Nitsche et al.,2011
Arylcyanoacrylamide		K_i = 35.7 μM*	Nitsche et al. 2011
Arylcyanoacrylamide		K_i = 184 μM*	Nitsche et al. 2011
Anthracene		K_i = 3.0	Tomlinson a et al., 2011
Anthracene		K_i = 10.0	Tomlinson et al., 2011

Benziimidazole		IC_{50} = 13.1	Deng et al. 2012
Quinoline		IC_{50} = 9.5	Deng et al. 2012
Quinoline		IC_{50} = 7.5	Deng et al.,2012
Aminobenzamide		52.5% at 25 μM	Aravapali et al. 2012
Quinoline		IC_{50} = 2.4	Lai et al. 2013

Quinoline		IC_{50} = 3.5	Lai et al. 2013
Quinoline		IC_{50} = 5.2	Lai et al. 2013
Benzisothiazole		80% at 25 μM	Lai et al. 2012
Benzimidazole		IC_{50} = 6.0	Raut et al. 2015

In a separate initiative, a novel series of thiadiazoloacrylamide analogues as DENV NS2B-NS3 protease inhibitors was designed based on the structure of the hit obtained through a high-throughput screening of commercially available chemical library containing ~7000 compounds which resulted in the discovery of a compound with $IC_{50}$of 2.24 μM against DENV2 NS2B-NS3 protease (Liu et al. (2014)). The docked conformation of active compounds from this study shows that the nitrile group is involved in hydrogen-bonding interactions with the side chain of Ser83 and the main chain of Met84. The benzyl group in these compounds was predicted to

have hydrophobic interactions with residues, including Val154, Val155, Gly159 and Ile86.

Attempts have also been made to identify potential lead compounds against DENV NS2B-NS3 protease using high-throughput screening (Bodenreider et al. (2009)). The best compound from this study with the central phthalazine ring forms a *pi–pi* interaction with the phenyl group of Tyr161. Both charged imidazoline groups are involved in salt bridge or hydrogen bonds with several residues (Asp129, Gly159, and Asn84). These molecules were found to have good inhibitory effect and has been shown by their experimental inhibition values: (a) Ki of 15.43±2.12 µM (Yang et al. (2011)), (b) IC_{50} of 4.2 µM on DENV2 and IC_{50} of 0.99 µM on DENV3 (Wu et al. (2015)), (c) IC_{50} of 2.0 µM on DENV1, IC_{50} of 2.3 µM on DENV2, IC_{50} of 1.6 µM on DENV3 and IC_{50} of 2.0 µM on DENV4 cells (Bodenreider et al. (2009)). In a separate study, Deng and co workers discovered benziimidazole class of compound was found active against NS2B-NS3 protease (IC_{50} = 13.12 µM) by virtual screening. This study was further continued with scaffold hopping initiatives to discover the quinoline class of compounds with IC_{50} of 7.5 µM. In 2013, Lai and co workers (Lai et al. (2013)) reported 8-hydroxyquinoline (HQ) derivatives for inhibition of DENV2 protease. The kinetic constant (K_i) for the most potent 8-HQ-aminobenzothiazole inhibitor from this study with an IC_{50} value of 0.91 ± 0.05 µM was determined to be 2.36 ± 0.13 µM. The mode of action of compounds in this work is to inhibit the DENV2 NS2B-NS3pro by a competitive mode of inhibition. The best compound was predicted to be forming favorable hydrophobic interaction of the bulkier 8-HQ with three aromatic residues phe130, tyr150 and tyr161 whereas,

another active compound from this work with bulkier benzothiazole moiety form hydrophobic interaction with val155.

1.4.4 Structure-Based Drug Design (SBDD)

Understanding the small-molecule ligands binding with therapeutic drug targets (proteins) plays a key role in discovery of novel drugs (Blaney, (2012)). structure based drug design (SBDD) refers to the rational utilization of structural data from macromolecular targets, obtained from x-ray diffraction or through homology modeling (Mandal et al., (2009)). The end result of this whole exercise is to discover novel molecules with high receptor binding affinity. The availability of three-dimensional macromolecular structures takes care of the binding site topology, including the presence of clefts, cavities and sub-pockets.

Electrostatic properties, such as charge distribution, are also understood. Current SBDD methods enable the design of novel molecules consisting of necessary features for efficient modulation of the protein target. Selective binding of a validated drug target by high affinity ligand which can interfere with specific cellular processes, ultimately leading to the required therapeutic effects (Urwyler, (2011)). SBDD is a iterative method consisting of stepwise information gathering. Starting from a known protein structure, computational studies are conducted to identify potential ligands. These molecular modeling procedures are followed by the experimental synthesis of the most promising candidates (Wilson et al., (2011)). Subsequently, evaluations of biological properties, such as potency, affinity and

efficacy, are carried out using in-vitro and in-vivo studies (Fang et al., (2012)). Provided that active compounds are identified, the crystal structure of the ligand bound receptor complex can be solved. The available structure allows the observation of intermolecular physical forces which actually leads the process of molecular recognition. Structural descriptions of ligand-receptor complexes are utilized for the investigation of binding conformations, characterization of important interactions, exploration of unknown binding sites, mechanism of action studies and the elucidation of small molecule-induced changes in conformation (Kahsai et al. (2011)).

1.4.5 Molecular Docking

Molecular docking is one of the most frequently used methods in SBDD because of its ability to predict the conformation of small-molecule within the appropriate target binding site (Xuan-Yu et al., (2011)). Following the development of the first algorithms in the eighties, molecular docking has gradually being used as an indispensible tool in inhibitor discovery (Lopez-Vallejo et al. (2011)). For example, investigations involving ligand binding modes and the corresponding intermolecular interactions that stabilize the ligand-receptor complex, can be easily carried out (Huang et al., (2010)). In addition to that, molecular docking algorithms execute quantitative predictions of binding energy, providing affinity based ranking order of docked compounds. Molecular docking methodology can be conveniently used to screen virtual libraries of drug-like molecules in order to obtain leads for further drug development.

The identification of the most likely binding conformations is carried out two steps: (i) exploration of a conformational space with different potential binding modes; (ii) accurate prediction of the interaction energy associated with each of the predicted binding conformation (Kapetanovic et al., (2008)). Molecular docking programs perform these functions through a iterative process, in which the ligand conformation is evaluated by specific scoring functions. This entire process is carried out repeatedly until it converges to a solution of minimum energy (Huang et al., (2010); Kapetanovic et al., (2008); Yurievet al., (2011)).

Content of the Thesis

This chapter has reviewed the dengue virus from a perspective of epidemiology and clinical manifestation. The structure of the dengue virus, specific therapeutic targets related to dengue virus, molecular mechanism of dengue virus NS2B-NS3 protease, inhibitor development has been discussed in detail.

In Chapter 2, various methodologies related to inhibitor design, organic synthesis and enzyme assay are discussed. In Chapter 3, results and discussions related to inhibitor design, organic synthesis and bioassay experiments are elaborated.

In Chapter 4, conclusion and future studies has been described to take the current work forward towards a deeper understanding of the mechanism of the developed compounds.