

JOHANNES GUTENBERG UNIVERSITY



Doctoral thesis

# “Investigating allosteric inhibition of flaviviral NS2B-NS3 proteases”

*submitted in fulfilment of the regulations for attaining the degree of*

*"Doktor der Naturwissenschaften (Dr. rer. nat)"*

*in the subject of Pharmaceutical and Medicinal Chemistry*

*at the Faculty of Chemistry, Pharmacy, Geography and Geoscience*

*at the Johannes Gutenberg University of Mainz*

Franziska Freifrau v. Hammerstein-Gesmold, nee Ohrloff

born in Bonn, NRW, Germany

Mainz 2020



**Submitted at the Faculty of Chemistry, Pharmacy, Geography and Geoscience:**

**Dean:**

████████████████████

**Name of 1st reviewer:**

██

**Name of 2nd reviewer:**

████████████████████

**Date of doctoral examination:**

*D77 (Dissertation Universität Mainz)*



## **Declaration of Authorship**

I, Franziska v. Hammerstein, declare that this thesis titled, “**Investigating allosteric inhibition of flaviviral NS2B-NS3 proteases**” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

---

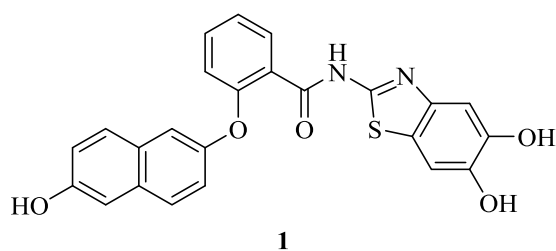
**Date**

**Signature**



## Abstract

Flaviviruses are RNA viruses with positive-sense single stranded RNA genomes (Chambers *et al.*, 1990). The viral *Flaviviridae* family contains over 70 viruses including dengue virus (DENV) and zika virus (ZIKV) (Kuno *et al.*, 1998). Due to increased global travel and climate change, flavivirus-based disease outbreaks have been reported around the world and are no longer restricted to defined regions (Rios, 2009). Of great concern are arthropod-borne flaviviruses (aboviruses, transmission via mosquitos or ticks) such as DENV, ZIKV and tick-borne encephalitis virus (TBEV). This was demonstrated for instance by the ZIKV outbreak 2015/16 in Latin America, where ignorance of the hazardousness of the disease paired with lack of available medical treatment led to an uncontrolled and rapid viral spread (Ai, Zhang and Zhang, 2016). Several decades of research have focused on the development of anti-flaviviral treatments, however up to now there is no specific medical treatment or vaccination for these quickly emerging and highly threatening diseases available. The flaviviral NS2B-NS3 protease is well known as a drug target and thus extensively investigated. However, the unusual narrow and little distinct active site of this serine protease causes problems in drug discovery. The development of allosteric inhibitors could bypass this demanding challenge.



Previous studies in the working group of [REDACTED] have identified **1** as a highly potent, non-competitive DENV NS2B-NS3 protease inhibitor, with  $IC_{50}$  values of 4.2  $\mu$ M and 1.0  $\mu$ M against DENV 2 and DENV 3 respectively (Wu *et al.*, 2015).

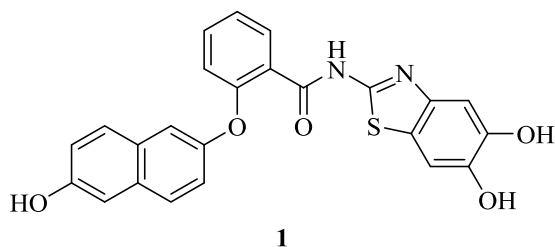
The present work shows that **1** binds to the hydrophobic allosteric binding site, located behind the active site, of the DENV NS2B-NS3 protease. A metabolic study revealed that the dimethylated precursor compound (pro-**1**) acts as a prodrug, being metabolised to **1** in cells. Consistent with this, a structure-activity relationship (SAR) study confirmed that the dihydroxyl substituted benzothiazole moiety is the structural driving force of inhibition. A flaviviral phylogenetic study suggested that the specific allosteric site is a common feature for flaviviral NS2B-NS3 proteases. In this context it could be shown that **1** also inhibits the NS2B-NS3 proteases of ZIKV, TBEV, West Nile virus (WNV), spondweni virus (SWV), Japanese encephalitis virus (JEV) and yellow-fever virus (YFW) and hence, leading to the forward guidance that approaching this well-defined allosteric binding site could lead to the development of a PAN-flaviviral inhibitor.





## Zusammenfassung

Flaviviren sind behüllte RNA Viren mit positiver Polarität des einzelsträngigen RNA Genoms (Chambers *et al.*, 1990). Mit über 70 weiteren Vertretern gehören die Flaviviren der Familie *Flaviviridae* an und beinhalten unter anderem das Dengue Virus (DENV) und Zika Virus (ZIKV) (Kuno *et al.*, 1998). Die immense Zunahme des weltweiten Reisens und die globale Erwärmung haben dazu beigetragen, dass flavivirale Krankheitsausbrüche nicht mehr nur lokal begrenzt auftreten, sondern ein weltweites Problem darstellen (Rios, 2009). Große Sorgen bereiten die flaviviralen Arboviren (Viren, die hauptsächlich durch Arthropoden, wie Moskitos und Zecken, übertragen werden), welches beispielsweise der ZIKV Ausbruch 2015/16 in Lateinamerika verdeutlicht. Die Unwissenheit über die Gefährlichkeit des Zikafiebers, gepaart mit dem Fehlen von Impfstoffen und spezifischen Arzneimitteln, führten zu der unkontrollierten und schnellen Verbreitung des ZIKV (Ai, Zhang and Zhang, 2016). Trotz jahrzehntelanger Forschung an anti-flaviviralen Medikamenten gibt es bis zum heutigen Tag noch keine spezifischen Therapeutika gegen diese sich schnell verbreitenden und gefährlichen Krankheiten. Die flavivirale NS2B-NS3 Protease ist ein bekanntes und bereits umfangreich erforschtes pharmazeutisches Gebiet. Erhebliche Schwierigkeiten in der Arzneistoffentwicklung der flaviviralen NS2B-NS3 Protease bereitet das schmale und wenig ausgeprägte aktive Zentrum der Protease, welches durch die Entwicklung von allosterischen Inhibitoren umgangen werden könnte.



Vorherige Arbeiten in der Arbeitsgruppe von [REDACTED] [REDACTED] haben **1** als einen vielversprechenden, nichtkompetitiven DENV NS2B-NS3 Proteaseinhibitor, mit  $IC_{50}$  Werten von 4.2  $\mu$ M (DENV 2) und 1.0  $\mu$ M (DENV 3), identifiziert (Wu *et al.*, 2015).

Die vorliegende Arbeit zeigt, dass **1** an der hydrophoben allosterischen Bindestelle auf der Rückseite des aktiven Zentrums der DENV NS2B-NS3 Protease bindet und dass die dimethylierte Vorstufe (pro-**1**) als Prodrug fungiert. Im Einklang dazu zeigte eine Struktur-Wirkungsbeziehungsstudie, dass der zweifach hydroxylierte Benzothiazolsubstituent das entscheidende Strukturmotiv dieses Inhibitors ist. Eine phylogenetische Studie von Flaviviren deutete an, dass die allosterische Bindetasche ein verbreitetes Strukturmotiv aller flaviviralen NS2B-NS3 Proteasen ist. Es konnte gezeigt werden, dass **1** die NS2B-NS3 Protease von ZIKV, TBEV, West Nile Virus (WNV), Spondweni Virus (SWV), Japanese Encephalitis Virus (JEV) und Yellow-fever Virus (YFW) hemmt. Vor diesem Hintergrund könnte die Entwicklung eines übergreifenden flaviviralen NS2B-NS3 Proteaseinhibitors möglich sein.



## Acknowledgements

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]



## Table of contents

Declaration of Authorship.....	I
Abstract.....	III
Zusammenfassung.....	V
Acknowledgements.....	VII
Abbreviations and symbols.....	XIII
Table of figures.....	XV
Table of Compounds.....	XIX
1. Introduction.....	1
1.1. Phylogeny of flaviviruses.....	1
1.2. Flaviviruses as human pathogens.....	1
1.3. Transmission and distribution of arthropod-borne flaviviral diseases.....	1
1.4. Clinical pictures and treatments of dengue and zika fever.....	3
1.5. Flaviviral vaccines and prevention methods.....	4
1.6. Flaviviral structure and replication.....	6
1.7. Fundamentals of the flaviviral NS2B-NS3 protease.....	8
1.7.1. Commonly used constructs of flaviviral NS2B-NS3 protease.....	8
1.7.2. Structure of flaviviral NS2B-NS3 protease.....	8
1.7.3. Enzymatic reaction of flaviviral NS2B-NS3 protease.....	9
1.7.4. Allosteric inhibition of DENV NS2B-NS3 protease.....	10
2. Objectives.....	13
3. Material and methods.....	15
3.1. Computational analyses of flaviviral NS2B-NS3 protease.....	15
3.1.1. Molecular docking on DENV 2 NS2B-NS3 protease.....	16
3.1.2. Virtual screening of a purchasable chemical database of small molecules on DENV 2 NS2B-NS3 protease.....	18
3.1.3. Homology model of ZIKV NS2B-NS3 protease.....	18
3.1.4. Molecular docking on ZIKV NS2B-NS3 protease.....	18

3.2.	Synthesis of allosteric flaviviral NS2B-NS3 inhibitors containing the benzothiazol moiety .....	19
3.2.1.	Synthesis of 5,6-dimethoxybenzothiazol-2-amine .....	19
3.2.2.	Synthesis of 2-(phenylthio)benzoic acids.....	19
3.2.3.	HBTU coupling reaction of 5,6-dimethoxybenzothiazol-2-amine with benzoic acid derivates .....	20
3.2.4.	Deprotection of the methoxy groups .....	21
3.3.	Protein production of the flaviviral NS2B-NS3 protease .....	22
3.3.1.	Plasmid construct design of DENV 2 and ZIKV NS2B-NS3 unlinked protease via Gibson Assembly.....	22
3.3.2.	Construct design of flaviviral NS2B-NS3 linked proteases .....	29
3.3.3.	Site-directed mutagenesis .....	32
3.3.4.	Recombinant expression and purification of flaviviral NS2B-NS3 .....	36
3.4.	Analytical methods used to examine the flaviviral NS2B-NS3 protease .....	38
3.4.1.	Fluorometric enzyme assay .....	38
3.4.2.	Microscale thermophoresis (MST).....	39
3.4.3.	Metabolism study .....	41
3.4.4.	NMR Spectroscopy .....	42
3.4.5.	Protein crystallisation .....	42
3.4.6.	Cleavage site study of ZIKV NS2B-NS3 protease.....	43
3.4.7.	Lipid sedimentation assay of ZIKV NS2B-NS3 protease .....	44
3.4.8.	Phylogenetic study of flaviviral proteases.....	45
4.	Results and discussion.....	47
4.1.	DENV NS2B-NS3 protease .....	47
4.1.1.	Structure-activity relationship studies (SAR).....	47
4.1.2.	Metabolic study of pro-1 and pro-16.....	51
4.1.3.	Structural investigations of the DENV 2 NS2B-NS3 protease .....	57
4.1.4.	Summary and discussion of the DENV NS2B-NS3 protease results .....	65
4.2.	ZIKV NS2B-NS3 protease .....	67
4.2.1.	Homology model and molecular docking of ZIKV NS2B-NS3 protease .....	67

4.2.2.	Protein preparation and autocatalytic cleavage of ZIKV NS2B-NS3 protease.....	69
4.2.3.	Inhibition of the ZIKV NS2B-NS3 protease.....	72
4.2.4.	Protein-lipid interaction.....	73
4.2.5.	Structural investigations of the ZIKV NS2B-NS3 protease.....	74
4.2.6.	Summary and discussion of the ZIKV NS2B-NS3 protease results .....	76
4.3.	Other flaviviral NS2B-NS3 proteases .....	78
4.3.1.	Phylogenetic study .....	78
4.3.2.	Protein preparation and enzymatic analysis .....	81
4.3.3.	Summary and discussion of the other flaviviral NS2B-NS3 protease results .....	84
5.	Conclusion and perspectives .....	85
6.	Appendix .....	87
6.1.	Plasmid maps .....	87
6.1.1.	Vector: pET-15b.....	87
6.1.2.	Vector: pET-11a.....	88
6.1.3.	Vector: pACYCDuet-1.....	89
6.1.4.	Vector: pCOLADuet-1 .....	90
6.2.	Sequences and protein physicochemical properties .....	91
6.2.1.	Dengue NS2B-NS3 protease constructs.....	91
6.2.2.	ZIKV NS2B-NS3 protease constructs.....	93
6.2.3.	DENV1 NS2B-NS3 linked protease (p266).....	95
6.2.4.	DENV3 NS2B-NS3 linked protease (p258).....	96
6.2.5.	DENV4 NS2B-NS3 linked protease (p257).....	97
6.2.6.	JEV NS2B-NS3 linked protease (K/A) (p253) .....	98
6.2.7.	SWV NS2B-NS3 linked protease (R/A) (p255).....	99
6.2.8.	TBEV NS2B-NS3 linked protease (p251) .....	100
6.2.9.	UTV NS2B-NS3 linked protease (K/A) (p254) .....	101
6.2.10.	WNV NS2B-NS3 linked protease (K/A) (p252).....	102
6.2.11.	YFV NS2B-NS3 linked protease (p256).....	103

6.3.	Synthesised compounds.....	104
6.3.1.	4-Chloro- <i>N</i> -(5,6-dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)benzamide [10] .....	104
6.3.2.	<i>N</i> -(5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)-3-hydroxybenzamide [11] .....	105
6.3.3.	2-((5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)carbonyl)phenyl acetate [12].....	106
6.3.4.	<i>N</i> -(5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)-2-mercaptobenzamide [13].....	107
6.3.5.	<i>N</i> -(5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)-3,5-dimethylbenzamide [14] .....	108
6.3.6.	<i>N</i> 1, <i>N</i> 3-Bis(5,6-dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)isophthalamide [15].....	109
6.3.7.	<i>N</i> -(5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)-2-(phenylthio)benzamide [16].....	110
6.3.8.	<i>N</i> -(5,6-Dihydroxybenzo[ <i>d</i> ]thiazol-2-yl)-2-( <i>p</i> -tolylthio)benzamide [17] .....	112
6.3.9.	<i>N</i> -(5,6-Dimethoxybenzo[ <i>d</i> ]thiazol-2-yl)-2-((6-methoxynaphthalen-2-yl)oxy)- benzamide [Pro-1] .....	113
6.4.	NMR spectrum of the DENV 2 NS2B-NS3 linked protease.....	114
6.5.	Homology model of ZIKV NS2B-NS3 .....	115
6.6.	Mass spectrometry of WT ZIKV NS2B-NS3 protease .....	116
6.7.	MST measurements .....	118
6.7.1.	Linked ZIKV NS2B-NS3 (R95A) protease with 1 .....	118
6.7.2.	Linked ZIKV NS2B-NS3 (S135A) protease with 1 .....	119
6.8.	Crystallisation of the linked ZIKV NS2B-NS3 (R95A) protease .....	120
6.9.	Phylogenetic trees.....	121
6.9.1.	Evolution of flaviviral polyprotein .....	121
6.9.2.	Evolution of flaviviral NS2B-NS3 protease.....	122
6.9.3.	Alignment of additionally selected flaviviral constructs .....	123
7.	Bibliography.....	125
	Curriculum vitae.....	135



## Abbreviations and symbols

ADE	Antibody-dependent enhancement
AIC	Akaike Information Criteria
AMC	Amino-4-methylcoumarin
BMI	<i>N</i> -benzylmaleimide
C	Capsid
CPD	Compounds
DCM	Dichloromethane
DENV	Dengue virus
DF	Dengue fever
DHF	Dengue haemorrhagic fever
DMF	<i>N,N</i> -Dimethylformamid
DNS	Deoxyribonucleic acid
DSS	Dengue shock syndrome
E	Envelope
<i>E.coli</i>	Escherichia coli
EMI	<i>N</i> -ethylmaleimide
ER	endoplasmic reticulum
FDA	Food and Drug Administration
GER	Germany
GIS	Geographic information system
HBTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl)uronium hexafluorophosphate
IGR	Insect growth regulator
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-D-thiogalactoside
JEV	Japanese encephalitis virus
$K_d$	dissociation constants
$K_M$	Michaelis constant
LB	Lysogeny broth
LC-MS	Liquid chromatography–mass spectrometry
LE	Ligand efficiency
MCS	Multiple cloning site
MST	Microscale thermophoresis
NADPH	Nicotinamide adenine dinucleotide phosphate
NCR	Noncoding regions

Ni-NTA	Ni <sup>2+</sup> -nitriloacetic acid
NMM	<i>N</i> -Methylmorpholine
NS	Non-structural
NTD	Neglected tropical diseases
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
PDB	Protein data base
prM	Membrane precursor
prM	Membrane precursor
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SAR	Structural-activity relationship
SDS	Sodium dodecyl sulfate
SIT	Sterile insect technique
SMILES	Simplified molecular-input line-entry system
SWV	Spondweni virus
TBEV	Thick-borne encephalitis virus
TEV	Tobacco Etch virus
T <sub>m</sub>	Melting temperature
US	United States
UTV	Usutu virus
UV-VIS	Ultraviolet-visible
ViPR	Virus Pathogen Resource
WNV	West Nile virus
YFV	Yellow fever virus
ZF	Zika fever
ZIKV	Zika virus
ΔG	Gibbs free energy

## Table of figures

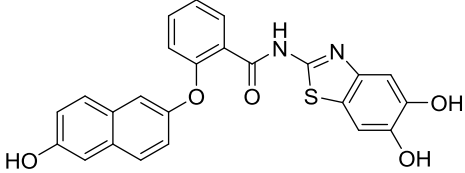
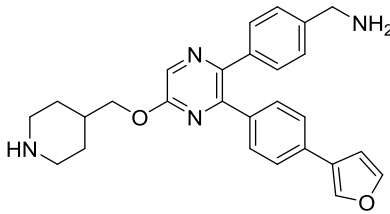
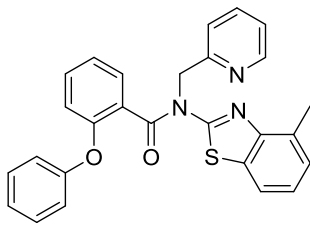
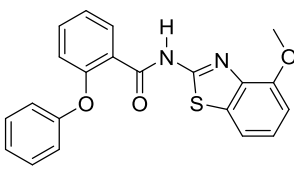
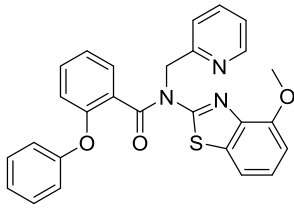
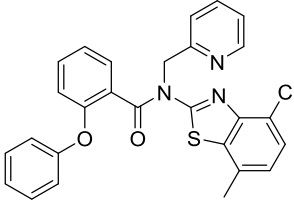
<b>Figure 1.1:</b> Summary of the global distribution of West Nile virus (WNV), dengue virus (DENV), tick-borne encephalitis virus (TBEV) and Japanese encephalitis virus (JEV). The map is drawn on the basis of Ishikawa <i>et al.</i> , 2014. ....	2
<b>Figure 1.2:</b> Schematics of a flavivirus particle. The flaviviral positive-sense single-stranded RNA genome is enclosed by the capsid protein (C), which further is enveloped by a lipid membrane. On the surface of the virus the membrane protein (M) and envelope protein (E) are found. The illustration is drawn on the basis of Slon Campos, Mongkolsapaya and Screaton, 2018. ....	6
<b>Figure 1.3:</b> A.) Viral genome depicted with the open reading frame (ORF), the 5' and 3' noncoding regions (NCR) are indicated. B.) Polyprotein membrane topology and its processing indicated with arrows. C.) “Open“ (PDB: 2FOM) and “closed“ (PDB: 3U1I) conformation of the DENV NS2B (yellow) co-factor and NS3 (green) protease domain. The active-site residues (His-51, Asp-75, Ser-135) are highlighted as red sticks. D.) Flaviviral NS2B-NS3 protease constructs used throughout this work. ....	7
<b>Figure 1.4:</b> General scheme of the reaction mechanism of a serine protease. All starts with the nucleophilic attack from the Ser-135 on the carbonyl of the peptide (1) and is followed immediately by the collapse of the stabilised tetrahedral intermediate (stabilised by the oxyanion hole) to the acyl enzyme intermediate, releasing the amine (2). The final deacylation is initiated from the nucleophilic attack of water (3) to form another stabilised tetrahedral intermediate complex (4) and is fulfilled with the collapse of the tetrahedral intermediate releasing the carboxyl product and regeneration of the catalytic triad (5) (Hedstrom, 2002). ....	9
<b>Figure 1.5:</b> DENV NS3 protease domains are coloured in different shades of grey and the DENV NS2B co-factor in red hue. A.) Alignment of the DENV NS2B-NS3 protease structure in the “closed” conformation (PDB: 3U1I), including the peptidic active site (His-51, Asp-75 and Ser-135 in green) inhibitor (yellow) and the two “open” conformations (PDB: 2FOM and the disputable PDB: 6MO0), including the reportedly allosteric inhibitor 2 (magenta). B.) Structure alignment representing the two “open” conformations of the DENV NS2B-NS3 protease (PDB: 2FOM and the disputable PDB: 6MO0), in which the C-terminus region of the co-factor orientates away from the active site of the NS3 domain, including the allosteric inhibitor 2 (magenta, PDB: 6MO0). ....	12
<b>Figure 3.1:</b> Protein sequence alignment of DENV3 NS2B-NS3 (PDB: 3U1I) with the DENV 2 NS2B-NS3 sequence used for enzymatic testing. DENV NS2B co-factor is deposited in grey and NS3 protease in red. Catalytic triad is marked in cyan and the amino acids contributing to the allosteric site are highlighted in magenta. Amino acids contributing to the allosteric site that were mutated prior docking are underlined. ....	16

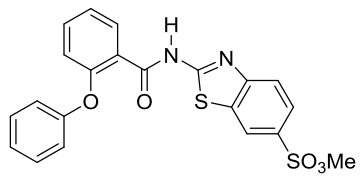
<b>Figure 3.2:</b> Schematic representation of the flaviviral NS2B-NS3 unlinked protease constructs. Both gene fragments were transformed in separate vectors (A). The NS3 protease domain was placed into a pACYCDuet-1 vector highlighted in green and possessed a His-tag including a TEV cleavage site. The co-factor NS2B coloured in yellow was integrated into a pCOLADuet-1 vector. Additionally both gene fragments were introduced into a pACYCDuet-1 vector (B). .....	23
<b>Figure 3.3:</b> Sequence of the DENV 2 NS2B-NS3 linked protease construct (above) and a schematic representation of the essential parts (below). NS2B co-factor is coloured in grey, the NS3 part in red and the active site in cyan. ....	29
<b>Figure 3.4:</b> Extraction of alignment of DENV 2 (NS2B red, NS3 grey) with ZIKV (NS2B yellow, NS3 green), catalytic triad is highlighted in cyan. ....	30
<b>Figure 3.5:</b> Sequence of the ZIKV NS2B-NS3 linked protease and a schematic scheme. Co-factor is highlighted in yellow, NS3 protease domain in green, active site in cyan, His-tag in blue and the TEV cleavage site is underlined. ....	31
<b>Figure 3.6:</b> Schematic representation of the principle of site-directed mutagenesis via quick change PCR. Oligonucleotides are represented by blue arrows and the point mutation by a red asterisk. The sketch is drawn on the basis of <a href="https://blog.addgene.org/site-directed-mutagenesis-by-pcr">https://blog.addgene.org/site-directed-mutagenesis-by-pcr</a> . ....	32
<b>Figure 3.7:</b> Hydrolysis of Boc-Gly-Arg-Arg-AMC by the <i>flaviviral</i> NS2B-NS3 protease. ....	38
<b>Figure 3.8:</b> MST dilution scheme of the sample preparation for the ZIKV S135A with 1 experiment. The final inhibitor concentrations for each capillary are listed below. ....	40
<b>Figure 3.9:</b> Approximate distribution of phospholipids in the endoplasmic reticulum of mammalian cells. The pie charts is drawn on the basis of Andreyev <i>et al.</i> , 2010. ....	44
<b>Figure 4.1:</b> Progress curves of substrate hydrolysis in presence of the purchased chemicals selected by virtual screening results tested against DENV 2 NS2B-NS3 protease in the fluorometric enzyme assay. ....	49
<b>Figure 4.2:</b> Scaffold hopping of 1. The benzothiazole side (framed in light grey) was left unchanged, while the diaryl ether moiety (circuited in dark grey) was altered. ....	49
<b>Figure 4.3:</b> Structures and molecular weights of pro-1 and pro-16. ....	51
<b>Figure 4.4:</b> LC-MS-spectrum of the metabolite pro-1 at 0 min. ....	52
<b>Figure 4.5:</b> LC-MS-spectrum of the metabolite pro-1 at 240 min. ....	52
<b>Figure 4.6:</b> Single demethylation of pro-1 and its structural isomers. ....	53
<b>Figure 4.7:</b> Tandem-MS of pro-1 ( $m/z = 487.3$ ) and the single demethylated metabolite derivate of pro-1 ( $m/z = 473.2$ ). ....	53
<b>Figure 4.8:</b> Structural isomers of the single demethylated pro-1 and their individual fragment sizes. ....	54
<b>Figure 4.9:</b> Fragmentations of the single demethylated derivate of pro-1. ....	54
<b>Figure 4.10:</b> LC-MS-spectrum of the metabolite pro-16 at 0 min. ....	55
<b>Figure 4.11:</b> LC-MS-spectrum of the metabolite pro- 17 at 240 min. ....	55

<b>Figure 4.12:</b> Tandem-MS of pro-16 metabolite. ....	56
<b>Figure 4.13:</b> Metabolic products of pro-16, aromatic hydroxylated derivate (left) and the single demethylated, aromatic hydroxylated compound (right). ....	56
<b>Figure 4.14:</b> NMR spectrum overlay of the DENV 2 NS2B-NS3 linked protease in red and the DENV 2 NS2B NS3 unlinked protease in blue.....	58
<b>Figure 4.15:</b> DENV 2 NS2B-NS3 protease (PDB code 2M9P) and the location of the introduced point mutations. NS2B co-factor is highlighted in red and the NS3 domain coloured in grey. Active site (Ser 135) is presented in green, allosteric site mutations in blue and the offsite position (Gly 14) in yellow. ....	60
<b>Figure 4.16:</b> Bar chart representing the $K_M$ values of DENV 2 NS2B-NS3 linked protease and its single mutants. ....	62
<b>Figure 4.17:</b> Graphical representation of the enzyme activities of the DENV 2 cysteine mutants irreversibly modified via <i>N</i> -benzylmaleimide (BMI) and <i>N</i> -ethylmaleimide (EMI) respectively...	63
<b>Figure 4.18:</b> $IC_{50}$ values of 1 for inhibition of the different irreversibly modified cysteine residues. ....	64
<b>Figure 4.19:</b> Sequence alignment of DENV (NS2B in red, NS3 in grey) and ZIKV (NS2B in yellow, NS3 in green) NS2B-NS3 protease constructs highlighting the active (cyan) and allosteric site (magenta). * (asterisk) stands for a fully conserved residue, : (colon) and . (period) indicate conservation between groups of strongly, respectively weakly similar properties. ....	67
<b>Figure 4.20:</b> Comparison of the molecular docking poses of 1 in the allosteric site of the DENV (above) and ZIKV (below) NS2B-NS3 protease. For DENV the co-factor is coloured in red and the NS3 domain in grey. The ZIKV NS2B region is highlighted in yellow and the NS3 in green. Amino acids contributing to the allosteric site are visualised in magenta and 1 is represented in cyan. ....	68
<b>Figure 4.21:</b> SDS-Page of the unlinked (lane 1) and linked (lane 3) ZIKV NS2B-NS3 constructs. The linked (lane 3) protease is autocatalytically cleaved. Inactive linked S135A enzyme (lane 4) and linked R95A point mutated ZIKV NS2B-NS3 protease (lane 5) proved not to be cleaved. Like the wild type (WT), the mutation R29A was equally cleaved. ....	70
<b>Figure 4.22:</b> Illustration of the ZIKV NS2B-NS3 linked protease construct highlighting the molecular weights of the separate fragments in kDa. ....	70
<b>Figure 4.23:</b> $IC_{50}$ determination by fitting ZIKV NS2B-NS3 linked (R95A) protease activity [%] against the concentration of inhibitor 1 [ $\mu$ M]. ....	72
<b>Figure 4.24:</b> SDS-page of the lipid sedimentation assay for ZIKV NS2B-NS3 linked R95A protease. ER lipid composition mimics the rough endoplasmic reticulum and PC-PE is the positively charged lipid mixture composition. ....	73

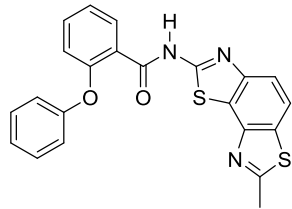
<b>Figure 4.25:</b> Annotated weblogo of the flaviviral NS2B-NS3 proteases. The motif sequence above the bars represents the situation in the Zika virus protease. Red bars indicate the active sites of the proteases (His-51, Asp-75, Ser1-35) and blue bars represent the allosteric sites.....	79
<b>Figure 4.26:</b> The evolutionary history of the additionally selected flaviviruses. ....	80
<b>Figure 4.27:</b> Molecular docking from the ZIKV protease with 1 from chapter 4.2.1 (above) in combination with an extraction of the sequence alignment of the selected flaviviruses (below). Amino acids contributing to the allosteric site are highlighted in magenta and the serine of the catalytic triad in cyan. ....	81
<b>Figure 6.1:</b> NMR spectrum of the DENV 2 NS2B-NS3 linked protease. Black signals represent the DENV 2 NS2-NS3 protease without inhibitor and the red signals the protease after addition of 1 (protease/inhibitor, 1:3). ....	114
<b>Figure 6.2:</b> Structure alignment of the ZIKV NS2B-NS3 protease homology model with the template 2M9P resulting in a RMSD value of 0.880. ....	115
<b>Figure 6.3:</b> Phi-Psi Plot of the ZIKV NS2B-NS3 protease homology model. ....	115
<b>Figure 6.4:</b> SDS-page (12%) from 28.02.2016. Protein band extraction for mass spectrometry analysis of WT ZIKV NS2B-NS3 protease experiencing autocatalytic cleavage highlighted in magenta. ....	116
<b>Figure 6.5:</b> Mass spectrum of the ~ 26 kDa protein band (see Figure 6.4). ....	116
<b>Figure 6.6:</b> Mass spectrum of the ~ 10 kDa protein band (see Figure 6.4). ....	117
<b>Figure 6.7:</b> Picture was taken from the documentation system Formulatrix Rock Imager™, MarXtal University Marburg. Conditions: 35% PEG 200, 100 mM MES, pH 9.....	120
<b>Figure 6.8:</b> The evolutionary history of <i>flaviviruses</i> based on the entire polyprotein. ....	121
<b>Figure 6.9:</b> The evolutionary history of <i>flaviviruses</i> based on the NS2B-NS3 protease sequence. ....	122
<b>Figure 6.10:</b> Sequence alignment of the newly designed flaviviral NS2B-NS3 protease constructs highlighting the active (cyan) and allosteric (magenta) site.....	123

**Table of Compounds**

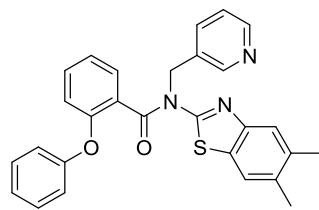
Compound	No.
	1
	2
	3
	4
	5
	6



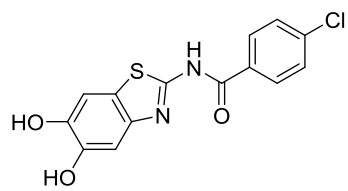
7



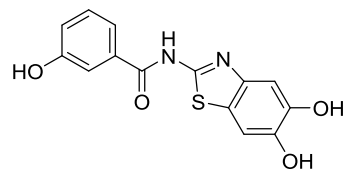
8



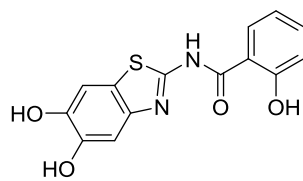
9



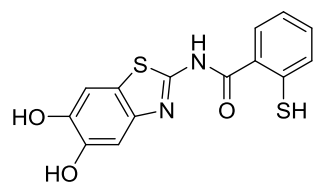
10



11



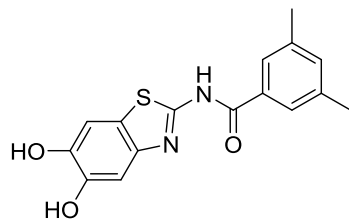
12



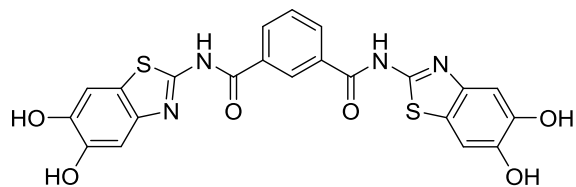
13



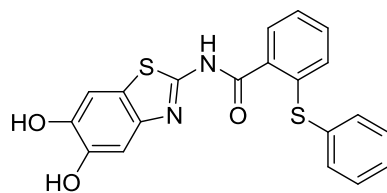
Table of Compounds



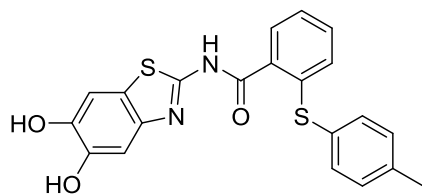
14



15



16



17



## **1. Introduction**

### **1.1. Phylogeny of flaviviruses**

Flaviviruses belong to the family *Flaviviridae*, which also includes hepaciviruses and pestiviruses. Difficulties in flaviviral classification derive mainly through their extensive geographic distribution and the tremendous diversity of arthropod vectors or vertebrates hosts associated with viral transmission. Kuno *et al.*, 1998 classified the viruses into clusters, clades, and species. Their phylogenetic study revealed that two clusters (non-vector and vector-borne viruses) evolve from a putative ancestor and the vector-borne virus cluster further branches into tick-borne and mosquito-borne viruses (illustrated in Figure 6.8). All flaviviruses of human pathogens belong to the last two clusters.

### **1.2. Flaviviruses as human pathogens**

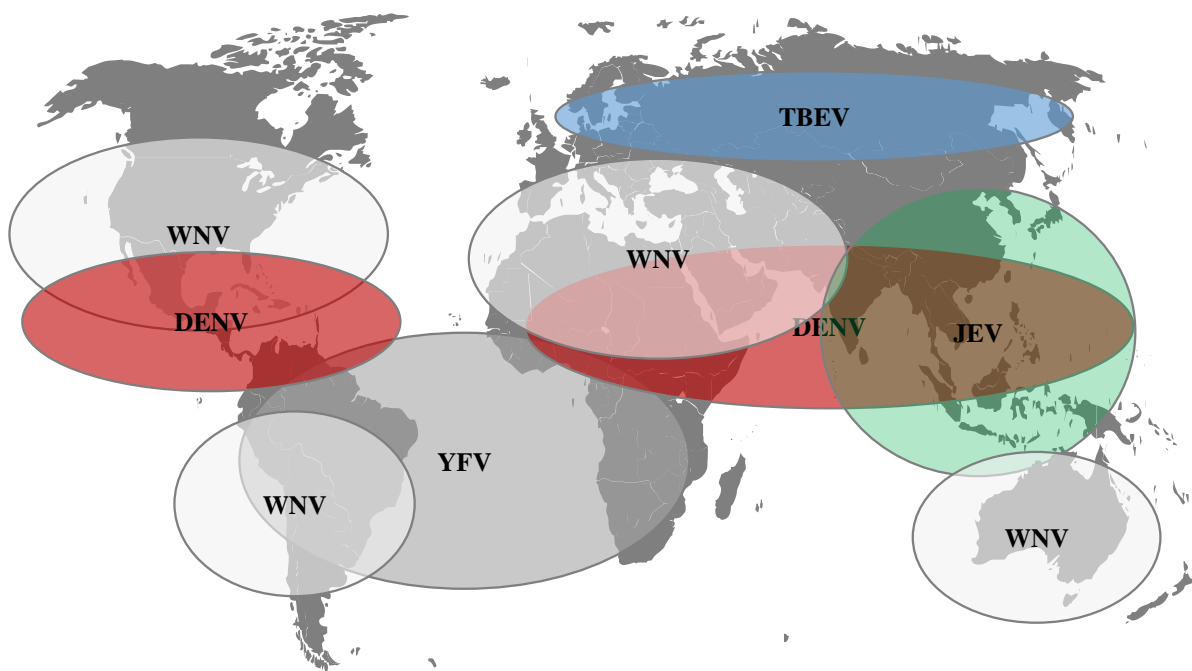
Arthropod-borne viruses (abbr.: arboviruses), that are transmitted to vertebrates through infected mosquitoes or ticks, belonging to the family *Flaviviridae* (genus *Flavivirus*) typically manifest in haemorrhagic fever, encephalitis, paralysis, and icterus. These virally provoked human diseases are often unpredictable in their severity of clinical manifestation, long-term persistence and further degree of geographical spread, thus making them extremely hazardous to human beings (Gould and Solomon, 2008). The flavivirus dengue for example is classified as one of the few neglected tropical diseases (NTDs) present in over 128 countries and effecting more than 40% of the world's population (Brady *et al.*, 2012).

### **1.3. Transmission and distribution of arthropod-borne flaviviral diseases**

Two flaviviruses of specific human importance are the dengue virus (DENV) and zika virus (ZIKV). Both viruses are mosquito-borne viruses and are mainly transmitted via *Aedes aegypti* and *Aedes albopictus*. These *Aedes* species have been identified to be common vectors for transmitting several tropical fever viruses. The mosquitos are active during daytime and only the female bites for blood to ensure egg maturation (Ferreira-De-Lima and Lima-Camara, 2018). In urban areas, viral transmission to humans occurs commonly through the bite of an infected mosquito, while mosquitos are infected by ingesting blood from a viremic host. After a few days, the blood meal leads to the rise of infectious viral particles in the salivary secretions of the mosquito and thus renders the insect capable of transmission to a new vertebrate host. If the viral concentration in the human blood is high enough, flaviviruses can even be passed on via blood transfusions or other forms of blood exchange.

Amongst the arthropod-borne flaviviruses, ZIKV takes an exceptional position for transmission of the virus. On the basis of the outbreak 2015/16 in Brazil it was identified that ZIKV can be passed on by sexual intercourse as well as from a mother to its unborn child (Sakkas *et al.*, 2018), with potentially serious neurological consequences as described in chapter 1.4.

Arboviruses belonging to the family *Flaviviridae* are found all around the world and can approximately be classified by the habitat of their transmission-vector. Figure 1.1 exemplarily summarises the global distribution of selected flaviviruses (Ishikawa, Yamanaka and Konishi, 2014).



**Figure 1.1: Summary of the global distribution of West Nile virus (WNV), dengue virus (DENV), tick-borne encephalitis virus (TBEV) and Japanese encephalitis virus (JEV). The map is drawn on the basis of Ishikawa *et al.*, 2014.**

## Introduction

Many of the mosquito-borne flaviviral diseases, such as dengue, zika and yellow-fever, are found in tropical and subtropical regions of the world, directly correlating with the transmission vector habitat of mosquitos of the *Aedes* and *Culex* families. An expansion of the viral distribution area, as for instance shown for the West Nile virus (WNV) in Figure 1.1, indicates additional transmission vector reservoirs such as birds. The opposite is seen for the Japanese encephalitis virus (JEV), where infected humans do not develop sufficient viremia to infect feeding mosquitoes and hence for transmission an additional viral reservoir is required in close human contact, i.e. pigs in rural regions. In comparison, the tick-borne encephalitis virus (TBEV) is transmitted via ticks as reflected by a different habitat compared to that of mosquitos-borne diseases.

In recent years the continuous increase in human travel and global warming has led to the whole world being faced with several different, quickly emerging flaviviral diseases (Rios, 2009).

### **1.4. Clinical pictures and treatments of dengue and zika fever**

The DENV incubation time is found to be between 3 to 14 days. Occasionally, a DENV infection can proceed asymptotically, however more frequently an infection goes along with flu-like symptoms, such as muscle and joint pains, headache, fever and conjunctivitis. Dengue fever (DF) often displays some slightly more characteristic symptoms such as a skin rash similar to that of measles (Kalayanarooj, 2011). Importantly, even when unambiguously identified, there is currently no treatment available against DENV infections. Upon treatment of the symptoms, DF typically abates after one week. Unfortunately, this means that DENV infections are often not correctly identified thus leading to a high number of unreported cases. Estimates assume that annually 390 million DENV infections occur worldwide. In 2 - 5% of the cases the disease displays a more severe form, the dengue haemorrhagic fever (DHF) or even a dengue shock syndrome (DSS) of which 1% result in mortality (Bhatt *et al.*, 2013). Overcoming a DENV infection, leads to a lifelong immunisation against that specific serotype. Currently there are four human pathogen serotypes (DENV 1 – DENV 4) known and a putative fifth serotype (DENV 5) emerging (Mustafa *et al.*, 2015). However, perhaps counterintuitively, an antibody-dependent enhancement (ADE) makes the patient more vulnerable against the other serotypes in future. In general, the immune response caused by a virus infection evokes the formation of virus-specific antibodies and plays an essential role in the virus control. Unfortunately the presence of virus-specific antibodies formed by a previous DENV infection can be advantageous to the entry DENV of a different serotype. Hence a secondary infection of DENV, with a different serotype, often lead to a severer course (DHF or DSS) of the disease (Katzelnick *et al.*, 2017).

Similar to what is observed for DF, a ZIKV infection often shows mild flu-like symptoms, however, in contrast to DENV infections, more severe disease phenotypes are not typically associated with zika fever (ZF). For this reason ZF remained generally underestimated for a long time (Ai, Zhang and Zhang, 2016). Today it is known that a ZIKV infection can provoke severe neurological ailments, such as the Guillain-Barré syndrome and fetal disorders (i.e. microcephaly) (Rasmussen *et al.*, 2016; Baud *et al.*, 2017).

In general, many flaviviral tropical diseases display uncharacteristic disease manifestations and can thus only be confirmed conclusively by microbiological laboratory testing, such as virus isolation in cell cultures and nucleic acid detection via polymerase chain reaction (PCR). These tests however require high standards of laboratory equipment and are expensive, which is often not attainable in the regions of the world where these diseases commonly occur (Figure 1.1). Thus, the number of unreported cases is presumably very high.

Once infected by DENV and ZIKV no anti-flaviviral treatments are currently available. The symptoms can only be palliated by pain relievers (Nitsche, 2019).

### **1.5. Flaviviral vaccines and prevention methods**

Successful vaccines have been approved for yellow fever virus (YFV), TBEV and JEV and are of high aspiration for further flaviviral vaccine developments (Collins and Metz, 2017). There are many vaccine candidates for DENV and ZIKV in clinical development and in 2015, a DENV vaccine (Dengvaxia<sup>®</sup> from Sanofi Pasteur) was released in Mexico, followed by Brazil, El Salvador, Paraguay, and the Philippines. The tetravalent DENV vaccine Dengvaxia<sup>®</sup> displayed different efficiencies against the DENV serotypes and was thus controversially evaluated (Capeding *et al.*, 2014; Vannice *et al.*, 2018). Nevertheless, Dengvaxia<sup>®</sup> was authorised by the European Union in 2018 for use in European endemic areas, worth mentioning that this authorisation came alongside with strict restrictions. Only patients in the age range of 9 to 45 years with an unambiguously verified, via antibody testing, previous DENV infection may be vaccinated (Dengvaxia, *European Medicines Agency*, 2020). In 2019 the United States (US) Food and Drug Administration (FDA) restrictively approved Dengvaxia<sup>®</sup> for individuals aged 9 to 16 living in parts of the US where the DENV is endemic (*First FDA-approved vaccine for the prevention of dengue disease in endemic regions, FDA*, 2020).

## Introduction

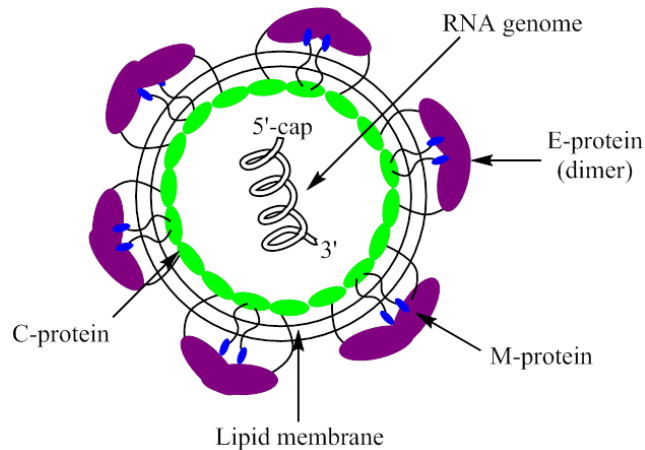
An important issue for embanking flaviviral mosquito-borne diseases are vector-prevention and -control methods. A detailed overview was lately constructed by Rather *et al.*, 2017 and is briefly summarised in Table 1.1.

**Table 1.1: Summary of vector-prevention and -control methods by Rather *et al.*, 2017.**

<b>Categories</b>	<b>Sub classifications</b>
Physical Control	<b>Monitoring and geographic information system (GIS) mapping of DENV infections</b> Efficient for quickly spotting an increase in dengue cases, in order to respond with diverse preventive strategies and making prevention a measurable quantity.
	<b>Determination of mosquito breeding areas</b> Aiming to detect and reduce the population density of dengue vectors.
	<b>Education of control and prevention strategies</b> Training individuals to identify and resolute act with vector habitats.
Biological Control	<b>Paratransgenesis (<i>wolbachia</i> bacteria)</b> Introduction of genetically modified symbiotic bacteria, which are of great harm to the main vector <i>A. aegypti</i> and hence suppressing the population. A very effective agent identified is <i>Wolbachia</i> (Jeffery <i>et al.</i> , 2009).
	<b>Genetic modification of vector species</b> Suppression of the vector population and its replacement / transformation of the genetically modified vector species, providing an effector gene for reduction and inhibition of disease transmission. A genetically modified mosquito species showed an 85% decline in <i>A. aegypti</i> population in Brazil (Pan America Health Organization (PAHO), 2014).
	<b>Sterile insect technique (SIT),</b> Releasing sterilised male vectors in the target population and hence controlling the vector density (Oliva <i>et al.</i> , 2012).
	<b>Usage of <i>larvivorous</i> fish and crustaceans</b> Larvivorous fish, also known as mosquitofish, are an effective strategy to reduce the population of the vector mosquito. The larvae of <i>A. aegypti</i> reside in open water and are eaten by the larvivorous fish.
Chemical Control	<b>Usage of insecticides</b> Usually first choice of strategy, highly effective and time-tested. Problematic is the development of resistance in the target vector population and the harmfulness to the environment.
	<b>Usage of insect growth regulator (IGR)</b> Several chemical compounds are known to hinder the growth and development of vermin, but resistance can occur.
	<b>Usage of pheromones</b> The strategy behind the usage of pheromones is to attract or repel the vector mosquitos. An environment friendly approach is the “attract-and-kill” strategy, where the insect is drawn to the killing agent, rather than the lethal lures released (Ong and Jaal, 2015).

## 1.6. Flaviviral structure and replication

Flaviviruses are small enveloped viruses of approximately 50 nm in diameter, with icosahedral and spherical geometry. The capsid surface consists of E (envelope) and prM (membrane precursor) proteins protecting the nucleocapsid, consisting of C (capsid) protein and the viral genomic ribonucleic acid (RNA) as illustrated in Figure 1.2 (Lindenbach, Thiel and Rice, 2007; Slon Campos, Mongkolsapaya and Screaton, 2018).



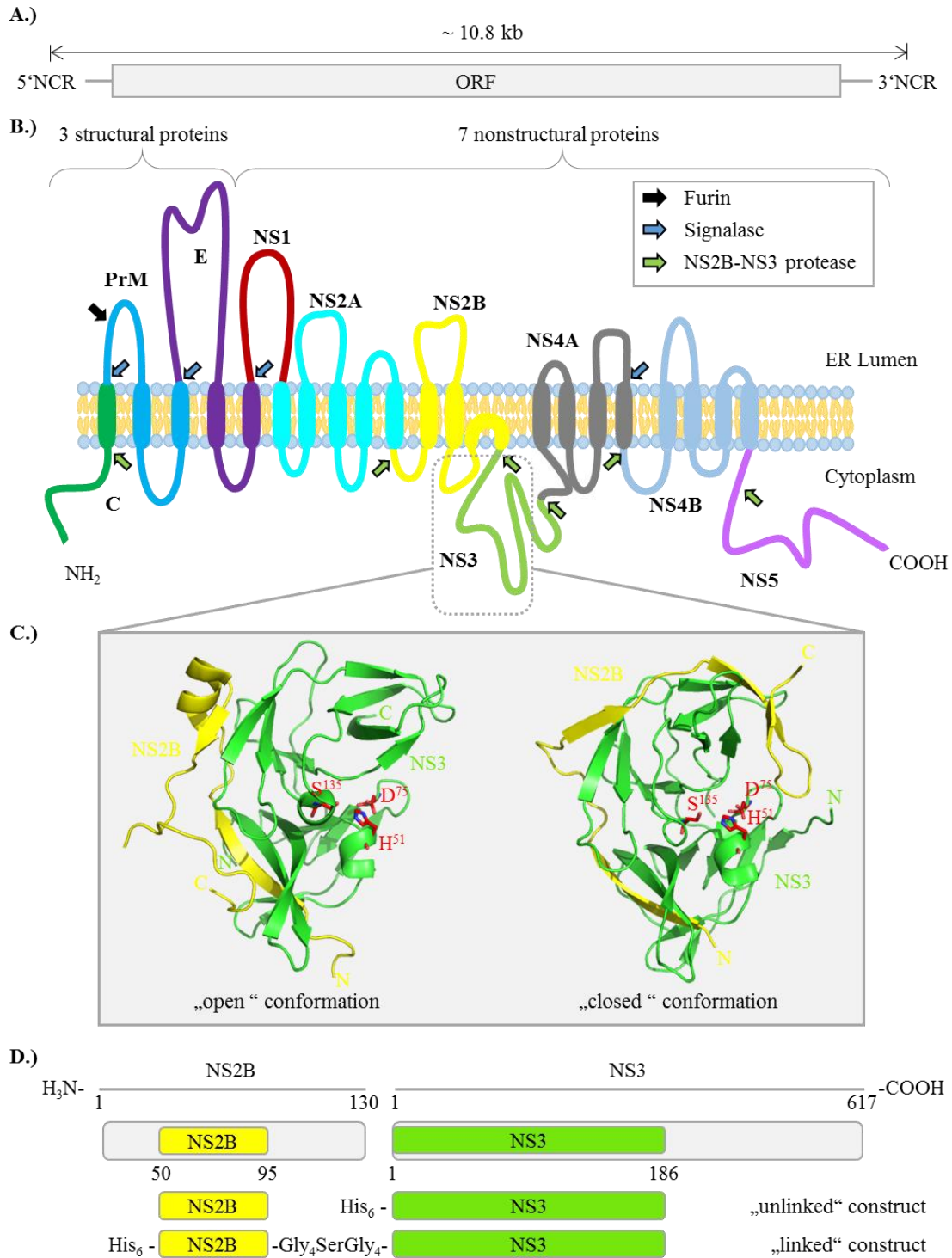
**Figure 1.2: Schematics of a flavivirus particle. The flaviviral positive-sense single-stranded RNA genome is enclosed by the capsid protein (C), which further is enveloped by a lipid membrane. On the surface of the virus the membrane protein (M) and envelope protein (E) are found. The illustration is drawn on the basis of Slon Campos, Mongkolsapaya and Screaton, 2018.**

The viral particles enter the (host) cells by receptor-mediated endocytosis. The envelope fuses with the endosomal membrane and releases the nucleocapsid into the cell cytoplasm. The nucleocapsid opens to uncoat the viral genome (Figure 1.3, A). The viral genome consists of a single-stranded positive-sense ribonucleic acid (RNA), containing about 11 k nucleotide bases. The viral RNA is translated at the ribosomes on the host's rough endoplasmic reticulum (ER) into a polyprotein spanning across the ER. The viral polyprotein consists of three structural (C, PrM and E) and seven non-structural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5) proteins (Figure 1.3, B) (Barrows *et al.*, 2018). The polyprotein precursor is processed by host cell proteases, such as furin and signalase and a specific viral protease, the heterodimeric NS2B-NS3 protease.

The non-structural proteins play an essential role in the viral replication and hence the reproduction of the virus. For this reason, the complete proteolytic processing of the viral precursor polyprotein is essential for the viral replication and hence making the viral NS2B-NS3 protease a highly promising point of attack for drug development (Chambers *et al.*, 1990; Natarajan, 2010; Nitsche, 2019).



## Introduction



**Figure 1.3:** A.) Viral genome depicted with the open reading frame (ORF), the 5' and 3' noncoding regions (NCR) are indicated. B.) Polyprotein membrane topology and its processing indicated with arrows. C.) “Open“ (PDB: 2FOM) and “closed“ (PDB: 3U1I) conformation of the DENV NS2B (yellow) co-factor and NS3 (green) protease domain. The active-site residues (His-51, Asp-75, Ser-135) are highlighted as red sticks. D.) Flaviviral NS2B-NS3 protease constructs used throughout this work.

## 1.7. Fundamentals of the flaviviral NS2B-NS3 protease

The NS3 protease domain arises from the N-terminal third of the multifunctional NS3 protein and requires the short hydrophilic NS2B co-factor for functionality (Figure 1.3, B and C). NS2B ensures proper folding and contributes to the substrate recognition region of the protease (Lindenbach, Thiel and Rice, 2007). In agreement with MEROPS (*MEROPS, the Peptidase Database*, 2020), NS3 is assigned to the peptidase family S7 (*flavivirin* family) within the subclan PA(S) (serine endopeptidase).

### 1.7.1. Commonly used constructs of flaviviral NS2B-NS3 protease

For *in-vitro* studies, such as functional assays, minimalistic constructs consisting of a shortened NS2B co-factor (~ 45 aa, approximately aa 48-95) in conjunction with the core NS3 protease domain (~ 185 aa, approximately aa 1–186) are established throughout literature (Erbel *et al.*, 2006). Either these shortened flaviviral NS2B-NS3 proteins can be expressed binary (“unlinked” constructs) or connected via a polyglycine linker (“linked” constructs) (Figure 1.3, D).

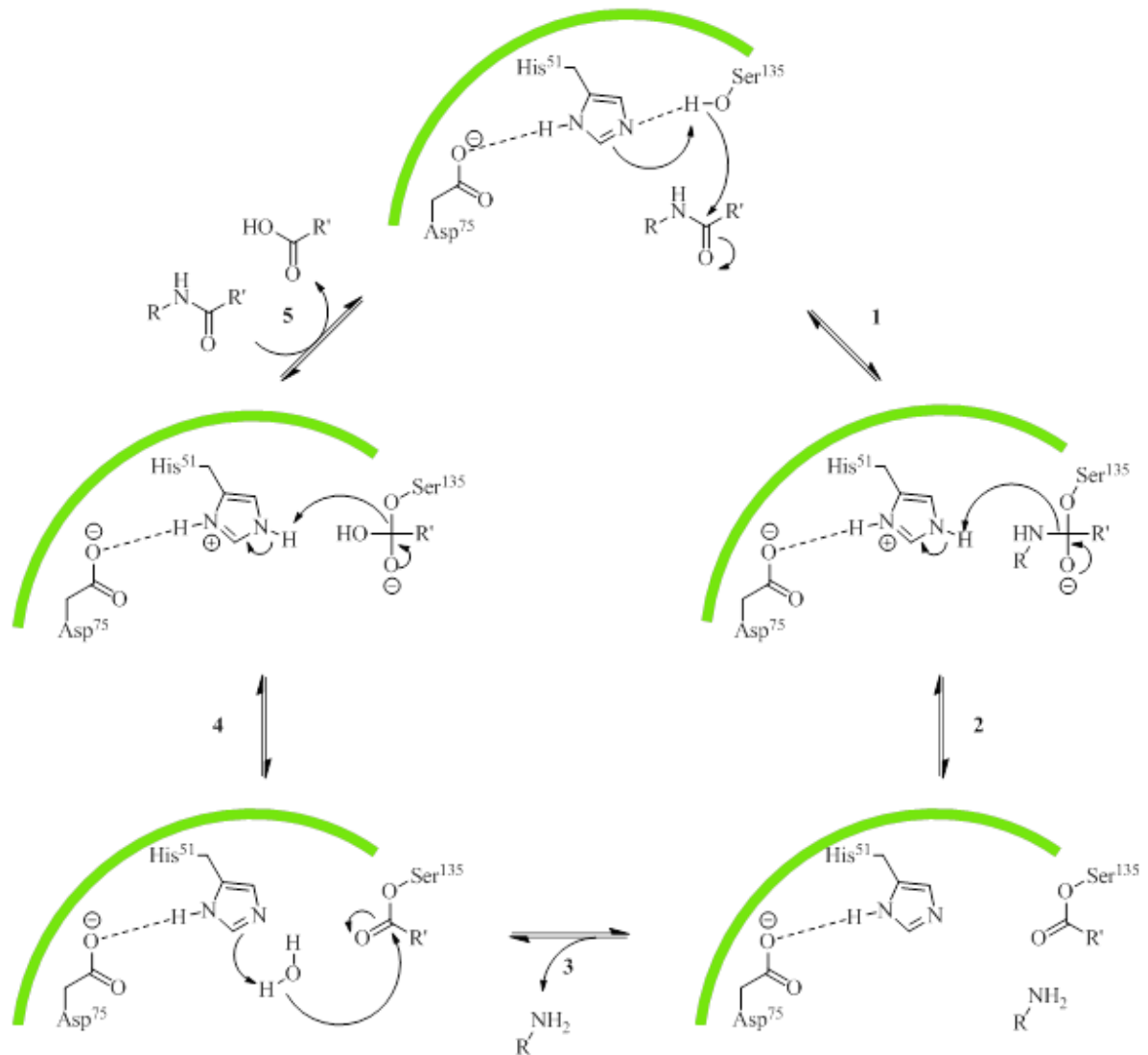
For ZIKV, DENV and several other flaviviral NS2B-NS3 proteases, these shortened constructs were shown to be relatively stable, enzymatically active and are commonly used to study structure and functionality (Kim *et al.*, 2013; Nitsche *et al.*, 2014; Kuiper *et al.*, 2017). It is worth mentioning that some (i.e. ZIKV NS2B-NS3) of these artificial flaviviral linked NS2B-NS3 constructs are affected by autocatalytical cleavage, which can be avoided by the introduction of specific site-directed mutagenesis (Phoo *et al.*, 2016; von Hammerstein *et al.*, 2019).

### 1.7.2. Structure of flaviviral NS2B-NS3 protease

Flaviviral NS2B-NS3 proteases are serine proteases with a trypsin-like fold. The NS2B co-factor is wrapped around the NS3 protease domain and there are currently crystal structures of two distinct conformations, the “open” and “closed” states, available (Figure 1.3, C). Both states have previously been examined extensively via X-ray and NMR studies and differ in the contribution of the co-factor participating on the active site formation (Zhu *et al.*, 2015; Li *et al.*, 2018). The “open” state displays a situation in which the C-terminus region of the co-factor orientates away from the active site of the NS3 domain, whereas the “closed” state is represented by the contribution of the co-factor, forming a  $\beta$ -hairpin covering the active site. The closed state is thought to resemble the active conformation in solution (Piccirillo *et al.*, 2016; Mahawaththa *et al.*, 2017).

### 1.7.3. Enzymatic reaction of flaviviral NS2B-NS3 protease

The NS2B-NS3 protease possesses a catalytic triad consisting of His-51, Asp-75, and Ser-135. The proteolytic reaction is initiated by a nucleophilic attack from the Ser-135 on the carbonyl of the peptide and is promoted by the spatially close His-51, whose stability as a protonated residue in turn is increased by the formation of a hydrogen bond to Asp-75 (Figure 1.4, step 1).



**Figure 1.4: General scheme of the reaction mechanism of a serine protease. All starts with the nucleophilic attack from the Ser-135 on the carbonyl of the peptide (1) and is followed immediately by the collapse of the stabilised tetrahedral intermediate (stabilised by the oxyanion hole) to the acyl enzyme intermediate, releasing the amine (2). The final deacylation is initiated from the nucleophilic attack of water (3) to form another stabilised tetrahedral intermediate complex (4) and is fulfilled with the collapse of the tetrahedral intermediate releasing the carboxyl product and regeneration of the catalytic triad (5) (Hedstrom, 2002).**

Generally, the flaviviral NS2B-NS3 proteases display substrate specificity for two basic amino acid residues in P1 and P2. Subsequently, cleavage of the viral polyprotein precursor via NS2B-NS3 typically occurs at positions harbouring such dibasic motifs, including the NS2A/NS2B, NS2B/NS3, NS3/NS4A and NS4B/NS5 junctions, as well as additional internal cleavage sites in C, NS2A, NS3, and NS4A (Lindenbach, Thiel and Rice, 2007).

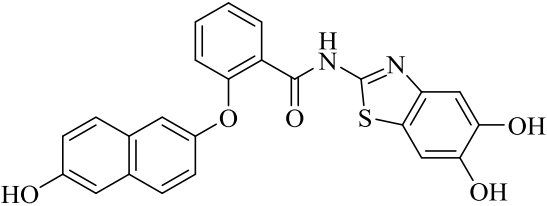
Substrate profiling for flaviviral NS2B-NS3 protease has led to well established functional enzyme assays conditions (Steuer *et al.*, 2009; Nitsche *et al.*, 2014). Several flaviviral NS2B-NS3 protease inhibitors addressing the active site have been identified, but due to poor pharmacokinetics and lack of cellular activities none of them have entered clinical trials yet (Leysen, Clercq and Neyts, 2000; Nitsche, 2019). Compared with other serine proteases such as trypsin, the flaviviral active sites are relatively shallow, rendering rational drug design difficult.

#### 1.7.4. Allosteric inhibition of DENV NS2B-NS3 protease

Scanning the DENV NS2B-NS3 protease surface via cysteine mutagenesis and subsequent chemical modification led to the identification and validation of a region susceptible to allosteric inhibition (Yildiz, 2013). This allosteric site is located behind the active site, centred around alanine 125 and reveals a deep, mostly hydrophobic binding pocket (Yildiz, 2013). Since then several attempts have been made to discover allosteric inhibitors for the DENV NS2B-NS3 protease (Wu *et al.*, 2015; Brecher *et al.*, 2017; Roy *et al.*, 2017; Shiryayev *et al.*, 2017; Nitsche *et al.*, 2019).

Wu *et al.*, 2015 identified **1** (Table 1.2) as a highly potent allosteric DENV 2 and DENV 3 NS2B-NS3 protease inhibitor, with  $IC_{50}$  values of 4.2  $\mu$ M and 1.0  $\mu$ M against DENV 2 and DENV 3 respectively. However, in order to further improve allosteric inhibition, information about the allosteric binding modes was missing.

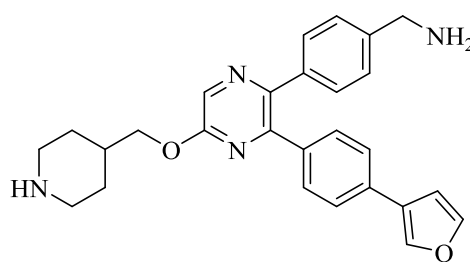
**Table 1.2: Compound 1, highly potent allosteric DENV 2 and DENV 3 NS2B-NS3 protease inhibitor (Wu *et al.*, 2015).**

Structure	DENV 2 ( $IC_{50}$ [ $\mu$ M])	DENV 3 ( $IC_{50}$ [ $\mu$ M])
 <p style="text-align: center;"><b>1</b></p>	4.2 $\pm$ 0.4	1.0 $\pm$ 0.1

## Introduction

Yao *et al.*, 2019 identified a drug-like, allosteric flaviviral NS2B-NS3 protease inhibitor (compound **2**, Table 1.3) and allegedly validated its binding mode by protein crystallisation (Figure 1.5).

**Table 1.3: Compound 2, an allosteric flaviviral NS2B-NS3 protease inhibitor identified by Yao et al, 2019 and its enzymatic data respectively.**

	<b>ZIKV</b>	<b>DENV 2</b>	<b>DENV3</b>	<b>WNV</b>
	$IC_{50}$ [ $\mu\text{M}$ ] ( $EC_{50}$ [ $\mu\text{M}$ ])	$IC_{50}$ [ $\mu\text{M}$ ]	$IC_{50}$ [ $\mu\text{M}$ ]	$IC_{50}$ [ $\mu\text{M}$ ]
<b>2</b>	$0.20 \pm 0.01$ ( $0.30\text{--}0.60$ )	$0.59 \pm 0.02$	$0.52 \pm 0.06$	$0.78 \pm 0.02$

Reportedly the X-ray study revealed for the first time, that the binding of the allosteric inhibitor forces the DENV NS2B-NS3 protease in a catalytically inactive “open” conformation (Figure 1.5, B), preventing NS2B to participate at the active site formation and hence prohibiting binding of the substrate. However, the scientific community discusses this crystal structure (PDB: 6MO0) highly controversial (Green, 2019; Grinter, 2019). Particularly due to the fact that the electron density cannot be clearly assigned to the ligand or other parts of the C-terminus region of the protein, makes the assessment of the structure highly questionable.

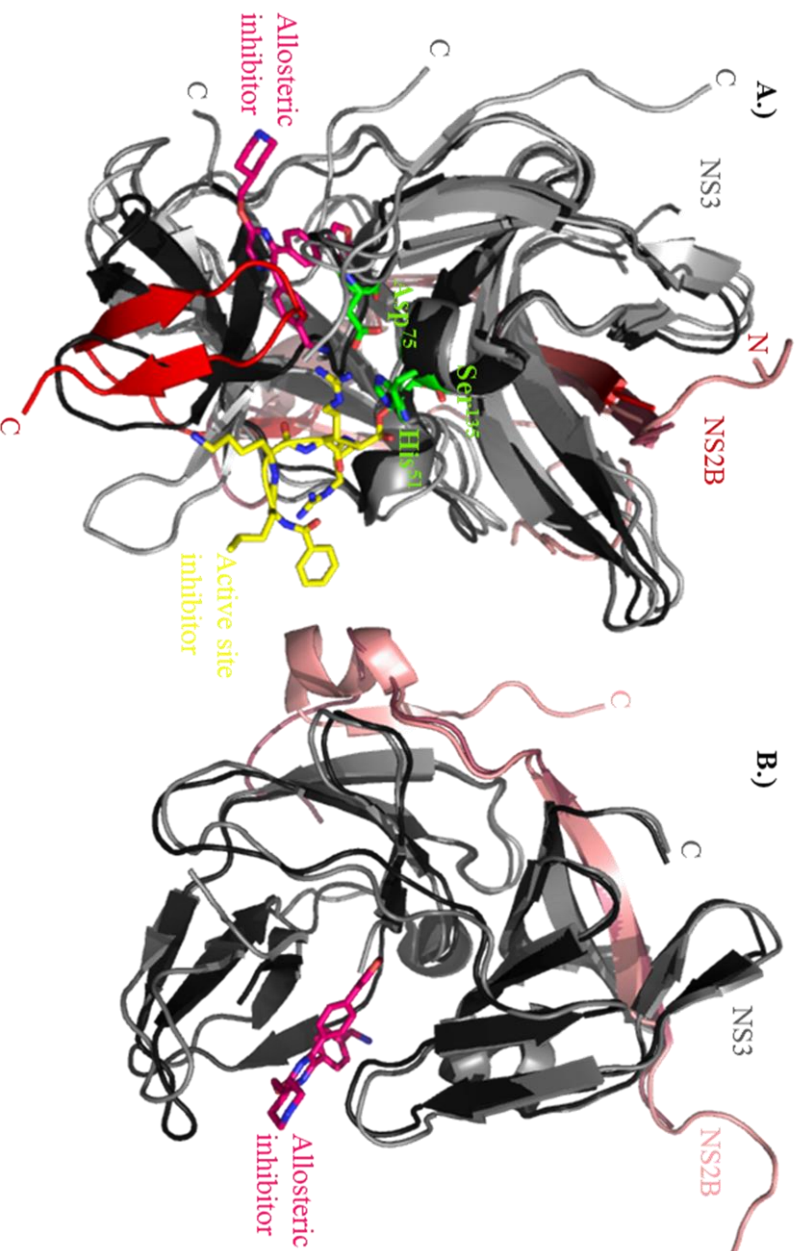


Figure 1.5: DENV NS3 protease domains are coloured in different shades of grey and the DENV NS2B co-factor in red hue. A.) Alignment of the DENV NS2B-NS3 protease structure in the “closed” conformation (PDB: 3U1D), including the peptidic active site (His-51, Asp-75 and Ser-135 in green) inhibitor (yellow) and the two “open” conformations (PDB: 2FOM and the disputable PDB: 6MO0), including the reportedly allosteric inhibitor 2 (magenta). B.) Structure alignment representing the two “open” conformations of the DENV NS2B-NS3 protease (PDB: 2FOM and the disputable PDB: 6MO0), in which the C-terminus region of the co-factor orientates away from the active site of the NS3 domain, including the allosteric inhibitor 2 (magenta, PDB: 6MO0).

## 2. Objectives

Due to the increasing global emergence of flaviviruses as severe human pathogens, the development of anti-flaviviral treatments is of high and urgent demand. This is emphasised by the fact that two-thirds of the world's population are living in areas of high risk for infections with the dengue flavivirus, but to this day no anti-flaviviral treatments are available. The flaviviral NS2B-NS3 protease has been widely described as a promising drug target.

Prior to this work, compound **1** was identified to be a highly potent DENV 2 and DENV 3 NS2B-NS3 protease inhibitor. As highlighted in Table 1.2, compound **1** displays inhibition values in the low, single-digit micromolar range, and follows a non-competitive inhibition mode (Wu *et al.*, 2015).

To promote the development of further allosteric DENV NS2B-NS3 protease inhibitors derived from inhibitor **1**, the aim of this work was to investigate the inhibitor binding region and mode respectively for flaviviral NS2B-NS3 proteases (with a focus on DENV and ZIKV) and to examine the impact of inhibitor **1** in the broader context of other flaviviral NS2B-NS3 proteases.





### 3. Material and methods

All computational programs used are listed in Table 3.1.

**Table 3.1: Software used in this thesis**

Software	Version / Source	Usage
ApE	ApE v2.0.55	Plasmid Editor
BLAST	NCBI ( <i>BLAST: Basic Local Alignment Search Tool</i> , 2020)	Biological sequence alignment tools
ChemDraw	ChemBioDraw Ultra 13.0	Chemical structure drawing
Clustal Omega	EMBL-EBI ( <i>Clustal Omega, Multiple Sequence Alignment, EMBL-EBI</i> , 2020)	Sequence alignment tool
GenBank	NCBI ( <i>GenBank Overview</i> , 2020)	Genetic sequence database
LeadIT	LeadIT-2.1.6	Docking
Microsoft Office	Microsoft Office 2010	Data publication, representation
MOE	MOE 2015.1001	Database handling

#### 3.1. Computational analyses of flaviviral NS2B-NS3 protease

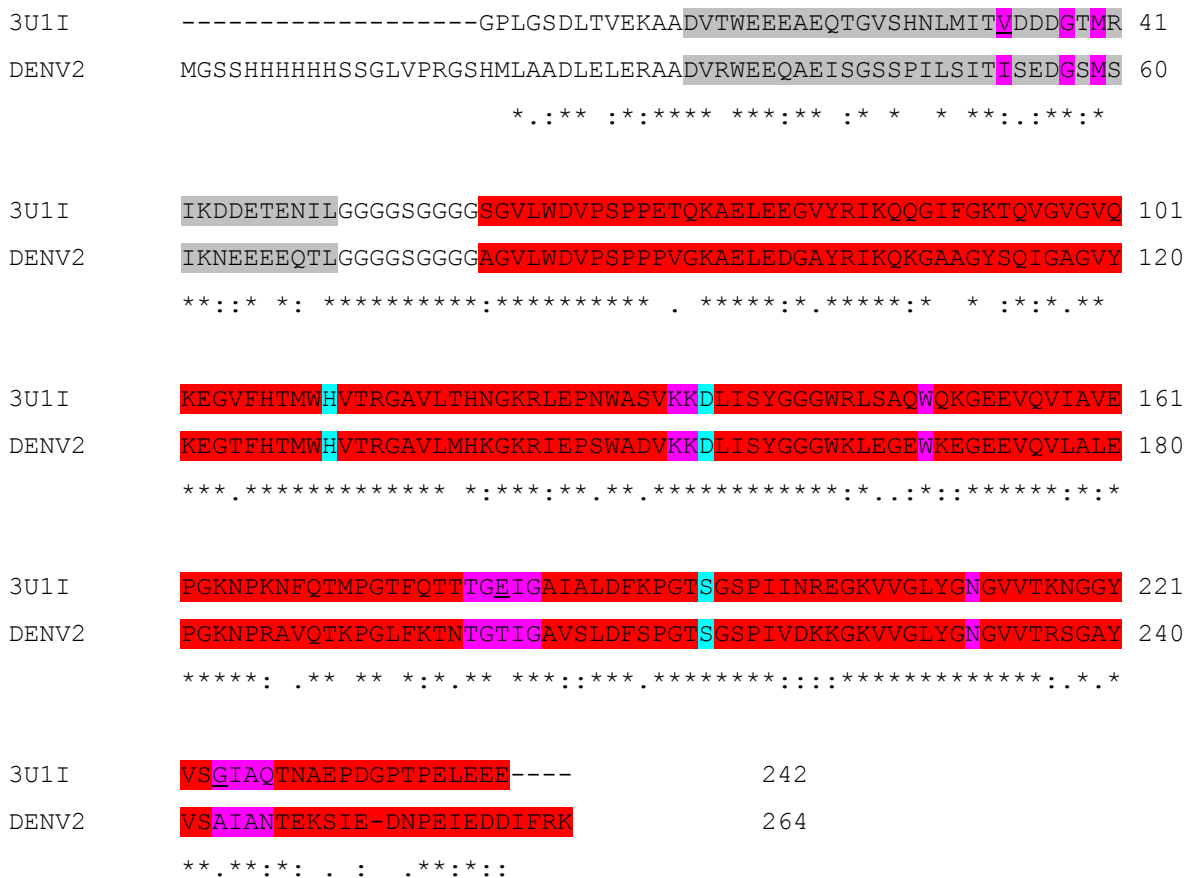
All structures used were freely available in the RCSB Protein Databank (PDB, Berman *et al.*, 2000). Table 3.2 comprises the PDB accession codes of the NS2B-NS3 proteases used.

**Table 3.2: Summary of the PDB accession codes of the flaviviral NS2B-NS3 proteases used.**

PDB code		Virus	Structure type	Conformation
2FOM	(Erbel <i>et al.</i> , 2006)	DENV 2	Crystal structure	Open
2M9P	(Gibbs <i>et al.</i> , 2018)	DENV 2	NMR structure	Closed
3U1I	(Noble <i>et al.</i> , 2011)	DENV 3	Crystal structure	Closed
5LC0	(Lei <i>et al.</i> , 2016)	ZIKV	Crystal structure	Closed
6MO0	(Yao <i>et al.</i> , 2019)	DENV 2	Crystal structure	Closed

### 3.1.1. Molecular docking on DENV 2 NS2B-NS3 protease

Prior to this work, molecular docking of **1** on the crystal structure of DENV3 NS2B-NS3 (PDB: 3U1I) was performed by Wu *et al.*, 2015. Consistent with this, the structure of DENV3 NS2B-NS3 (PDB: 3U1I) was used throughout this work, if not explicitly pointed out otherwise. Prior to molecular docking, mutated residues in the allosteric site found in the DENV2 NS2B-NS3 expression construct used (see chapter 3.3.1) were changed appropriately using MOE2015.1001 (Molecular Operating Environment (MOE), 2015.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite 910, Montreal, QC, Canada, H3A 2R7, 2017.) and are illustrated via underlined residues (figure 3.1).



**Figure 3.1: Protein sequence alignment of DENV3 NS2B-NS3 (PDB: 3U1I) with the DENV 2 NS2B-NS3 sequence used for enzymatic testing. DENV NS2B co-factor is deposited in grey and NS3 protease in red. Catalytic triad is marked in cyan and the amino acids contributing to the allosteric site are highlighted in magenta. Amino acids contributing to the allosteric site that were mutated prior docking are underlined.**

## Material and methods

Further receptor preparations were performed within the LeadIT-2.1.6 (www.biosolveit.com, BioSolveIT GmbH, Sankt Augustin, Germany, 2017.) worksuite (Rarey *et al.*, 1996). The amino acids contributing to the proposed allosteric site were selected manually and residue protonation was done automatically via ProToss. Water molecules were removed and all ligands were energy minimized using the MMFF94 force field (Halgren, 1996) within MOE. The docking protocol was performed under default parameters within LeadIT using the hybrid approach (enthalpy/entropy) for ligand placement.

For docking analysis in LeadIT, HYDE™ Visual Affinities and PoseView™ 2D Pose Visualization were used. HYDE is a scoring function based on a description of hydrogen bond and dehydration energies in protein–ligand complexes. Values are directly related to protein-ligand affinities and are derived from the Gibbs-Helmholtz equation. (Schneider *et al.*, 2013)

### 3.1.2. Virtual screening of a purchasable chemical database of small molecules on DENV 2 NS2B-NS3 protease

A similarity search for *N*-(benzothiazol-2-yl)-2-phenoxybenzamide using SciFinder<sup>®</sup> (<https://scifinder.cas.org/scifinder>) revealed 324 similar compounds. Retaining the commercially available compounds reduced the number to 249 compounds. These 249 buyable compounds were drawn with ChemBioDraw Ultra 13.0 (PerkinElmer, Inc.) and converted into their corresponding simplified molecular-input line-entry system (SMILES) codes. The newly constructed database was transferred into MOE and energy minimized using MMFF94. The self-assembled database was docked with LeadIT using the structure of DENV NS2B-NS3 (PDB: 3U1I) as the protein receptor. The same receptor preparations were used as described in chapter 3.1.1. Docking solutions were stored in FlexX solutions database and the obtained scores were exported in Microsoft Excel containing the scoring function terms: match, lipophilicity, ambiguity, clash and rotatable bonds. Classifications of the results were rendered via comparison of Gibbs free energy ( $\Delta G$ ) and ligand efficiency (LE).

### 3.1.3. Homology model of ZIKV NS2B-NS3 protease

Since there were no crystal structures of the ZIKV NS2B-NS3 protease available when the investigations into this field started, a homology model using MOE was constructed. An automatic PDB search by MOE was performed and revealed that the ZIKV NS2B-NS3 protease sequence and the sequence of the structure DENV 2 NS2B-NS3 (PDB: 2M9P) were highly similar. Hence, the homology model was built on the DENV 2 NS2B-NS3 (PDB: 2M9P) NMR structure. The homology model was assessed using the protein geometry application, for which a Ramachandran plot (see appendix, chapter 6.5) was generated in order to check outliers in terms of backbone dihedral angles. The homology model was validated by alignment of the reference structure and analysed using root mean square deviation (RMSD, equation 1).

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (1)$$

It was calculated by averaging over *N* pairs of atoms and the distance ( $\delta_i$ ) between the two atoms in the *i*-th pair, given in Å. (Kufareva and Abagyan, 2012)

### 3.1.4. Molecular docking on ZIKV NS2B-NS3 protease

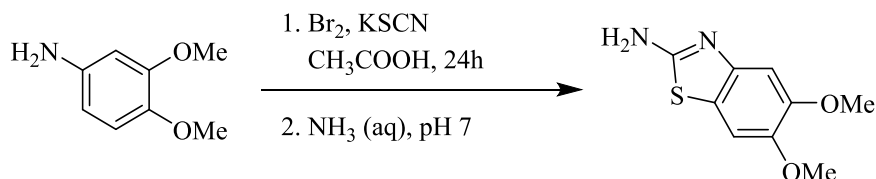
Molecular docking was first performed on the self-constructed homology model described in section 3.1.3. After the first crystal structure of ZIKV NS2B-NS3 (PDB: 5LC0) protease was released (Lei *et al.*, 2016), this crystal structure was used for further computational analysis. The receptor preparation was performed as described in chapter 3.1.1.

### 3.2. Synthesis of allosteric flaviviral NS2B-NS3 inhibitors containing the benzothiazol moiety

All chemicals were purchased from Sigma Aldrich (Darmstadt, GER), VWR (Darmstadt, GER) and J&K Scientific (Lommel, BEL).

#### 3.2.1. Synthesis of 5,6-dimethoxybenzothiazol-2-amine

(Trapani *et al.*, 2001)

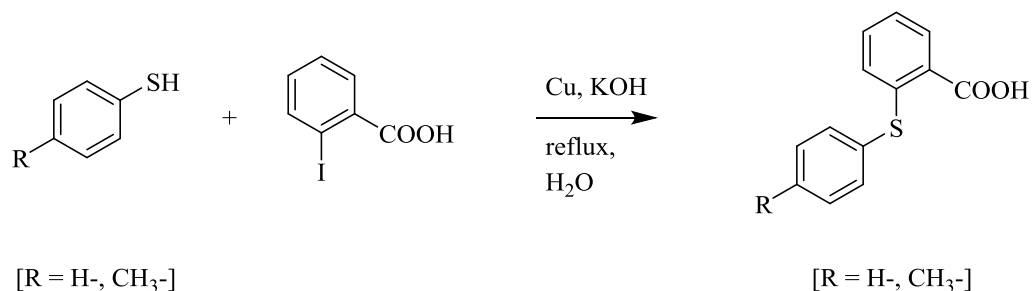


#### Reaction scheme 1: Synthesis of 5,6-dimethoxybenzothiazol-2-amine.

Ammonium thiocyanate (0.025 mol) was added to a solution of the 3,4-dimethoxyaniline (0.005 mol) in glacial acetic acid (50 ml). Bromine (0.005 mol) in acetic acid (10 ml) was added dropwise to this reaction mixture. The exothermic reaction was cooled down with a water bath and stirred at room temperature for 12 – 18 h. Finally, the reaction mixture was neutralised using aqueous ammonia. The product was isolated via filtration and the solid residue was washed thoroughly with distilled water.

#### 3.2.2. Synthesis of 2-(phenylthio)benzoic acids

(Hellwinkel and Bohnet, 1987)

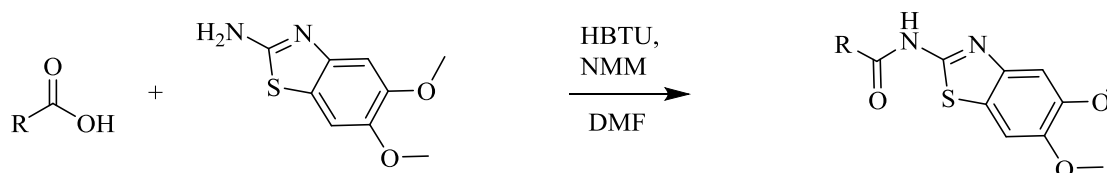


#### Reaction scheme 2: Synthesis of 2-(phenylthio)benzoic acids.

Catalytic amounts of copper (0.05 g) and zinc powder (0.01 g) were added to 20 ml water containing the appropriate thiophenol (0.008 mol). The mixture was heated for 10 min before the 2-iodobenzoic acid (0.008 mol) was added and the reaction mixture allowed refluxing for 20 h. The reaction mixture was filtered hot, to remove the unsolved compounds and cooled down before adding concentrated hydrochloric acid. The precipitate was isolated via filtration and washed several times with water. Recrystallisations in ethanol lead to colourless needles.

### 3.2.3. HBTU coupling reaction of 5,6-dimethoxybenzothiazol-2-amine with benzoic acid derivatives

(Montalbetti *et al.*, 2005)



**Reaction scheme 3: General procedure for the coupling reaction of 5,6-dimethoxybenzothiazol-2-amine with benzoic acid derivatives using HBTU.**

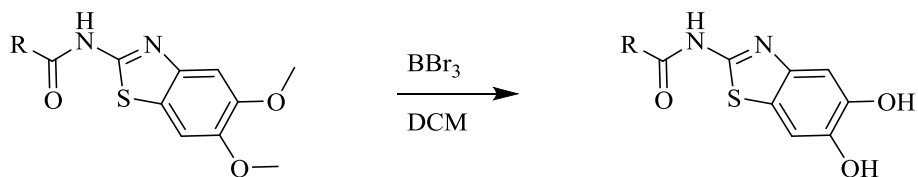
**Table 3.3: Register of the purchased benzoic acids.**

Structure	Name	M [g/mol]	CAS number	End product
	4-Chlorobenzoic acid	156.57	74-11-3	10
	Salicylic acid	138,12	69-72-7	12
	Acetylsalicylic acid	180,16	50-78-2	13
	Thiosalicylic acid	154.19	147-93-3	14
	3,5-Dimethylbenzoic acid	150.17	499-06-9	15
	Isophthalic acid	166.13	121-91-5	16

To a solution of the appropriate benzoic acid (1.0 mmol), *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU, 1.1 mmol) and *N*-methylmorpholine (NMM, 3,5 mmol) in *N,N*-dimethylformamid (DMF) under argon atmosphere 5,6-dimethoxybenzothiazol-2-amine was added and stirred at room temperature for 12 hours. The reaction was quenched via the addition of water and the reaction mixture was extracted with dichloromethane (DCM). DCM was dried over sodium sulphate, filtrated and the solvent was evaporated under reduced pressure. The products were purified using column chromatography.

### 3.2.4. Deprotection of the methoxy groups

(McOmie, Watts and West, 1968)



#### Reaction scheme 4: Demethylation of the 5,6-dimethoxybenzothiazol-2-amine using boron tribromide.

The intermediate product (1 mmol) was solved in 5 mL DCM and cooled down to  $-78\text{ }^\circ\text{C}$ . Under argon atmosphere boron tribromide (1M in DCM, 6 mmol) was added dropwise to the solution. After the addition of boron tribromide was completed, the reaction mixture was slowly brought to room temperature and allowed to stir for 12 hours. The reaction mixture was quenched with water and extracted with DCM. DCM was dried over sodium sulphate, filtrated and the solvent was evaporated under reduced pressure. The products were purified using column chromatography.

### **3.3. Protein production of the flaviviral NS2B-NS3 protease**

Throughout the biochemistry part, if not specifically mentioned otherwise, all chemicals and enzymes were obtained from the following companies: Roth (Karlsruhe, GER), Sigma-Aldrich (München, GER), New England Biolabs (Frankfurt a.M., GER) and VWR (Darmstadt, GER).

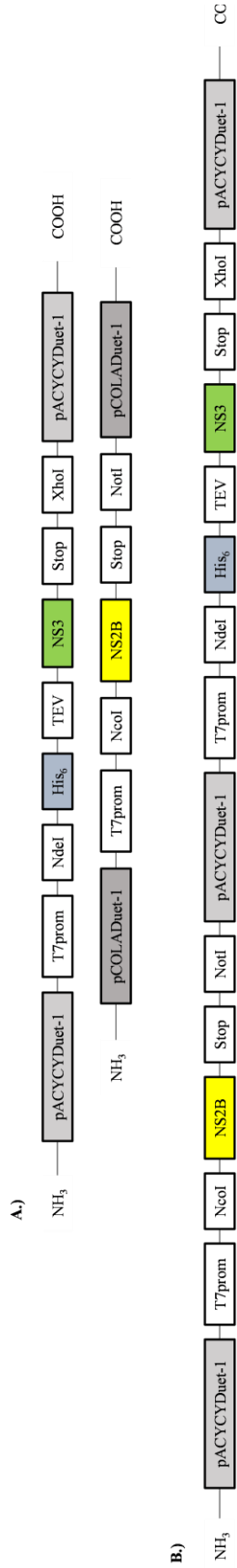
#### **3.3.1. Plasmid construct design of DENV 2 and ZIKV NS2B-NS3 unlinked protease via Gibson Assembly**

To create the DENV 2 and ZIKV NS2B-NS3 unlinked protease constructs as described for DENV 2 by (Kim *et al.*, 2013) the Gibson Assembly cloning method was used to generate plasmids encoding these proteins. Gibson Assembly is a highly effective cloning method for the easy assembly of multiple linear deoxyribonucleic acid (DNA) fragments. Multiple overlapping DNA fragments can be joined in a single isothermal reaction, regardless of fragment length or end compatibility (Gibson *et al.*, 2009).

Two vectors, pACYCDuet-1 and pCOLADuet-1, for co-expression in *Escherichia coli* (*E. coli*) were available. Two gene fragments were designed for each viral protease in order to be transformed in separate vectors (Figure 3.2, A) and additionally, both in one for co-expression (Figure 3.2, B). Table 3.4 summarises the different DENV 2 and ZIKV constructs and their vectors at hand.

The two DNA inserts needed for each unlinked construct were designed in the way that they were usable for all three desired variants. To enable this, the NS2B insert was additionally provided with the sequence between the two multiple cloning sites. The NS3 fragment was supplied with an N-terminal His-tag followed by a TEV (Tobacco Etch virus) protease cleavage site for tag removal. The designed gene fragments listed in Table 3.5 were obtained by Eurofins Genomics (Ebersberg, GER).





**Figure 3.2: Schematic representation of the flaviviral NS2B-NS3 unlinked protease constructs. Both gene fragments were transformed in separate vectors (A). The NS3 protease domain was placed into a pACYCYDuet-1 vector highlighted in green and possessed a His-tag including a TEV cleavage site. The co-factor NS2B coloured in yellow was integrated into a pCOLADuet-1 vector. Additionally both gene fragments were introduced into a pACYCYDuet-1 vector (B).**

**Table 3.4: Summary of the plasmids containing the “linked” and “unlinked” NS2B-NS3 protease constructs of DENV 2 and ZIKV used throughout this work. The “linked” plasmids were purchased and the “unlinked” were made via Gibson Assembly.**

Flaviviral NS2B-NS3 protease	“Linked” construct	“Unlinked” construct	
	His-NS2B-G <sub>4</sub> SG <sub>4</sub> -NS3	NS2B	His-NS3 NS2B / His-NS3
DENV 2	pET-15b ( <i>p092</i> )	pCOLADuet-1 ( <i>p094</i> )	pACYCYDuet-1 ( <i>p093</i> )
ZIKV	pET-11a ( <i>p119</i> )	pCOLADuet-1 ( <i>p122</i> )	pACYCYDuet-1 ( <i>p121</i> ) pACYCYDuet-1 ( <i>p120</i> )

**Table 3.5: Sequences of the purchased DNA fragments for Gibson Assembly of the unlinked constructs.**

<b>Gene fragment</b>	<b>Sequence 5' → 3'</b>
DENV 2 NS2B	ATACCATGGACCTGGAACCTGGAACGTGCGGCAGACGTGCGTTGGGAAGAACA AGCAGAAATCAGCGGTAGCAGCCCGATCCTGTCAATTACGATCTCGGAAGAT GGTAGCATGTCTATTA AAAACGAAGAAGAACAAGACAGACCCTGTAAGCGGCCG CATAATGCTTAAGTCGAACAGAAAGTAATCGTATTGTACACGGCCGCATAAT CGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCC ATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATG
DENV 2 His-TEV-NS3	AGAAGGAGATATACATATGGGTAGCAGCCATCATCATCATCACGGTAGC GAAAACCTGTACTTCCAAGGCTCCGACGGCGTGCTGTGGGATGTTCCGTCTCC GCCGCCGGTGGGTAAAGCAGAACTGGAAGACGGCGCTTATCGTATTAACAG AAGGGTGCCGCGGGCTATTCACAAATCGGTGCGGGCGTTTACAAAGAAGGTA CCTTTCATACCATGTGGCACGTCACGCGTGGTGCAGTGCTGATGCATAAAGGC AAGCGCATTGAACCGTCATGGGCTGATGTCAAAAAGGACCTGATCTCGTACG GCGGTGGCTGGAACTGGAAGGCGAATGGAAGGAAGGCGAAGAAGTTCAGG TCCTGGCGCTGGAACCGGGTAAAAACCCGCGTGCCGTTCAAACCAAACCGGG CCTGTTTAAAGACCAATACGGGTACCATTGGCGCGGTTAGCCTGGATTTCTCTC CGGGTACGAGTGGCTCCCCGATCGTCGACAAAAAGGGTAAAGTGGTTGGCCT GTATGGTAATGGTGTGCTGACCCGTAGTGGTGCCTACGTGTCCGCGATTGCCA ACACGGAAGTCCATTGAAGATAACCCGGAATTGAAGATGACATTTTTTCG TAAGTAACTCGAGTCTGGTAAA
ZIKV NS2B	TATACCATGGATATGTATATTGAACGCGCGGGCGATATTACCTGGGAAAAAG ATGCGGAAGTGACCGGCAACAGCCCGCGCCTGGATGTGGCGCTGGATGAAAG CGGCGATTTTAGCCTGGTGAAGATGATGGCCCGCCGATGCGCTAAGCGGCC GCATAATGCTTAAGTCGAACAGAAAGTAATCGTATTGTACACGGCCGCATAA TCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCC CATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATG
ZIKV His-TEV-NS3	AGAAGGAGATATACATATGGGTAGCAGCCATCATCATCATCACGGTAGC GAAAACCTGTACTTCCAAGGCTCCAGCGGCGCGCTGTGGGATGTGCCGGCCG CGAAAGAAGTGAAAAAAGGCGAAACCACCGATGGCGTGTATCGCGTGATGA CCCGCCGCTGCTGGGCAGCACCCAGGTGGGCGTGGGCGTGATGCAGGAAGG CGTGTTCATACCATGTGGCATGTGACCAAAGGCAGCGCGCTGCGCAGCGGC GAAGGCCGCTGGATCCGTATTGGGGCGATGTGAAACAGGATCTGGTGAAGT ATTGCGGCCCGTGGAACTGGATGCGGCGTGGGATGGCCATAGCGAAGTGCA GCTGCTGGCGGTGCCGCCGGGCGAACGCGCGCGCAACATTCAGACCCTGCCG GGCATTTTTTAAACCAAAGATGGCGATATTGGCGCGGTGGCGCTGGATTATC CGGCGGGCACCGCGCAGCCCGATTCTGGATAAATGCGGCCGCGTGATTGG CCTGTATGGCAACGCGTGGTGAATAAAAACGGCAGCTATGTGAGCGCGATT ACCCAGGGCCCGCGGAAGAAGAAACCCCGTGGAAATGCTTTGAACCGAGCA TGCTGAAAAATAA ACTCGAGTCTGGTAAA

## Material and methods

For plasmid amplification via polymerase chain reaction (PCR) five primers were designed and purchased from Sigma Aldrich (Table 3.6). The multiple cloning sites (MCS) of both plasmids were identical (see appendix, chapter 6.1) consequently only the NS2B\_rev primer, reaching into the differing part, had to be varied.

**Table 3.6: Oligonucleotides used as primers for the unlinked protease constructs of DENV and ZIKV.**

Name	N	T <sub>m</sub> [°C]	Sequence 5' → 3'
NS3_for ( <i>P8</i> )	22	64	TAACTCGAGTCTGGTAAAGAAACCG
NS3_rev ( <i>P10</i> )	41	60	CATATGTATATCTCCTTCTTATACTTAACTAATATACTAAG
NS2B_for ( <i>P9</i> )	26	61	AGAAGGAGATATACATATGGCAGATC
NS2B_rev ( <i>P7</i> )	30	61	TTCCAGGTCCATGGTATATCTCCTTATTAAGTTAAAC
ZIKV NS2B_rev	42	64	CAATATACATATCCATGGTATATCTCCTTATTAAGTTAAAC

The reaction mixture of 25 µL (elucidated in Table 3.7) using the KAPA HiFi HotStart PCR Kit (Roche, Mannheim, GER) was for each amplification prepared.

**Table 3.7: General composition of PCR mixtures.**

Component	Volume [µL]
5x HiFi buffer (fidelity)	5
dNTP Mix	0.75
KAPA HiFi polymerase	1
Forward primer	1.5
Reverse primer	1.5
DNA template	0.5
Milli-Q <sup>®</sup> H <sub>2</sub> O	14.75

The Primus 25 thermo cycler (MWG Biotech, Ebersberg, GER), with the adjusted parameters of programmed cycle settings (exemplified in Table 3.8) was used for amplification.

**Table 3.8: PCR cycle parameter settings.**

Step	Temp. [°C]	Duration	Cycles
Initial denaturation	95	3 min	1
Denaturation	98	20 s	1
Annealing	T <sub>m</sub> – 5 °C	15 s	30
Elongation	72	60 s/kbp	1
Final elongation	75	7 min	1

The PCR product was checked for size and yield on a 1% (w/v TAE buffer) agarose gel. After the agarose gel electrophoresis run, the gel was stained for 20 min in an ethidium bromide bath (0,033% v/v) and the plasmid DNA was visualized with ultraviolet light detection using the Quantum gel documentation system (Peqlab, Erlangen, GER). To eliminate the methylated template DNA a DpnI digestion was carried out (composition described in Table 3.9).

**Table 3.9: General DpnI digestion reaction mixture.**

<b>Component</b>	<b>Volume [<math>\mu</math>L]</b>
10x CutSmart buffer	5
Milli-Q <sup>®</sup> H <sub>2</sub> O	22
DpnI	1
PCR reaction mixture	22

The Gibson assembly master mix was prepared (according to Table 3.10) and placed into a heating block set at 50 °C. 50 ng of the vector and each insert were diluted up to 5  $\mu$ L with autoclaved Milli-Q<sup>®</sup> H<sub>2</sub>O and added to the Gibson assembly master mix.

**Table 3.10: Gibson Assembly reaction mixtures.**

<b>5x Isothermal reaction mix:</b>	<b>Assembly master mix:</b>
3 mL 1 M Tris-HCl (pH 7.5), sterile	320 $\mu$ L 5x Isothermal Master Mix
300 $\mu$ L 1 M MgCl <sub>2</sub> , sterile	0.64 $\mu$ L 10 U/uL T5 exonuclease
600 $\mu$ l dNTP Mix 10 mM	20 $\mu$ L 2U/uL Phusion DNA Pol
300 $\mu$ L 1M DTT (50 mM)	0.16 $\mu$ L 40 U/uL Taq DNA Ligase
1.5 g PEG-8000	860 $\mu$ L MQH <sub>2</sub> O
300 $\mu$ L 100 mM NAD	
MQH <sub>2</sub> O to top up to 6 ml	

## Material and methods

After 1 hour, competent DH5 $\alpha$  *E. coli* bacteria were transformed with the DNA (Table 3.11). Throughout this work, competent cells were cultivated from the commercial strains and made competent using CaCl<sub>2</sub>.

**Table 3.11: *Escherichia coli* strains in stock.**

Bacterial strain	Genotype
BL21 (DE3) Gold (Agilent Technologies)	B F <sup>-</sup> ompT hsdS (rB <sup>-</sup> mB <sup>-</sup> ) dcm <sup>+</sup> Tetr gal $\lambda$ (DE3) endA Hte
DH5 $\alpha$ (Agilent Technologies)	F <sup>-</sup> $\Phi$ 80lacZ $\Delta$ M15 $\Delta$ (lacZYA-argF) U169 <i>recA1 endA1 hsdR17</i> (rK <sup>-</sup> , mK <sup>+</sup> ) <i>phoA supE44</i> $\lambda$ - <i>thi-1 gyrA96 relA1</i>

For each transformation, 30 ng of DNA were combined with chemically competent *E. coli* bacteria (Table 3.11) and placed on ice for 30 min. The cells were heat-shocked at 42 °C for 45 s and subsequently incubated with 300  $\mu$ L lysogeny broth medium (LB-medium) at 37 °C at 225 rpm for 60 min. Afterwards cells were plated onto antibiotic containing LB-agar plates (Table 3.12) and incubated over night at 37 °C. From each LB-agar plate 10 clones were picked and overnight cultures with DH5 $\alpha$  cells in 5 mL LB-medium, containing appropriate amounts of antibiotics (Table 3.12) were prepared.

**Table 3.12: Concentration of antibiotics.**

Antibiotics	Stock solution	Desired final concentrations
Ampicillin	100 mg/mL in Milli-Q <sup>®</sup> H <sub>2</sub> O	100 $\mu$ g/mL
Chloramphenicol	30 mg/mL in ethanol	30 $\mu$ g/mL
Kanamycin	30 mg/mL in Milli-Q <sup>®</sup> H <sub>2</sub> O	30 $\mu$ g/mL

To ensure that the selected clones contained the desired insert a test PCR, amplifying the insert part between the restriction sites, was carried out using 1  $\mu$ L of the overnight culture. PCR was performed as described in Table 3.7 and Table 3.8 using the oligonucleotides in Table 3.13.

**Table 3.13: Oligonucleotides for test PCR.**

Name	N	Tm (°C)	Sequence 5' $\rightarrow$ 3'
NS2(NcoI)for (P2)	41	68	GGAATTCCCATGGGCAGCAGCCATCATCATC
NS2(NotI)rev (P4)	43	70	GCGGCCGCTCACCAGGGTCTGTCTTCTTCTTC
NS3(NdeI)for (P5)	38	64	GGAATTCCATATGGCAGGCGTGCTGTGGG
NS3(XhoI)for (P6)	39	64	GAAGATGACATTTTTTCGTAAGTGACTCGAG

Cells of the overnight cultures were yielded via centrifugation at 3900 rpm for 10 min at 4 °C (Centrifuge 5810 R, Eppendorf) and only the promising plasmid DNA was cleaned using the E.Z.N.A Plasmid Mini Kit I (Omega Bio-Tek, see protocol I of instruction manual). DNA concentrations were determined via absorption measurements at  $\lambda = 260$  nm with a Thermo Scientific NanoDrop 2000c UV-VIS spectrophotometer.

Verification of successfully cloned and modified DNA constructs was provided by GENterprise GENOMICS (StarSeq GmbH, Mainz, GER). Sequencing mixtures were prepared with ~ 500 ng plasmid DNA and 1  $\mu$ L of the appropriate sequencing primer (10  $\mu$ M stock, Table 3.14) and topped up to a final volume of 7  $\mu$ L by adding autoclaved Milli-Q® H<sub>2</sub>O.

**Table 3.14: Oligonucleotides for DNA sequencing.**

Name	N	Sequence 5' → 3'
T7_promoter	20	TAATACGACTCACTATAGGG
T7_terminator	19	GCTAGTTATTGCTCAGCGG
JW8	22	TGCGACTCCTGCATTAGGAAAT

Generally, for DNA sequencing of pET-11a and pET-15b derived plasmids, the T7\_promoter and T7\_terminator primers were used. For sequencing of pACYCDuet-1 and pCOLADuet-1 derived plasmids, the JW8 and the T7\_terminator primers were used (Table 3.14).

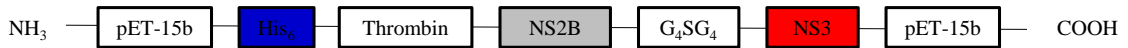
**Table 3.15: Expression plasmids obtained via Gibson Assembly of DENV 2 and ZIKV NS2B-NS3 unlinked protease.**

Construct	Vector	Restriction sites	Resistance
DENV 2 NS2B His-TEV-NS3	pACYCDUET-1	NcoI, NotI/NdeI, XhoI	chloramphenicol
DENV 2 His-TEV-NS3 alone	pACYCDUET-1	NdeI, XhoI	chloramphenicol
DENV 2 NS2B alone	pCOLADUET-1	NcoI, NotI	kanamycin
ZIKV NS2B His-TEV-NS3	pACYCDUET-1	NcoI, NotI/NdeI, XhoI	chloramphenicol
ZIKV His-TEV-NS3 alone	pACYCDUET-1	NdeI, XhoI	chloramphenicol
ZIKV NS2B alone	pCOLADUET-1	NcoI, NotI	kanamycin

### 3.3.2. Construct design of flaviviral NS2B-NS3 linked proteases

The flaviviral NS2B-NS3 linked protease constructs were designed according to the expression construct pET-15b encoding DENV 2 NS2B-NS3 linked protease (represented in Figure 3.3), obtained from the working group of [REDACTED].

MGSSHHHHHHSSGLVPRGSHMLAADLELERAADVRWEEQAEISGSSPILSITISEDGSMSEIKNEEEEQTL  
 GGGGSGGGGAGVLWDVPSPPPVGKAELEDGAYRIKQKGAAGYSQIGAGVYKEGTFHTMWHVTRGAVLMHK  
 GKRIEPSWADVKKDLISYGGGWKLEGWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNTGTIGAVSLDFS  
 PGTSGSPIVDKKGKVVGLYNGVVTRSGAYVSAIANTEKSIEDNPEIEDDIFRK



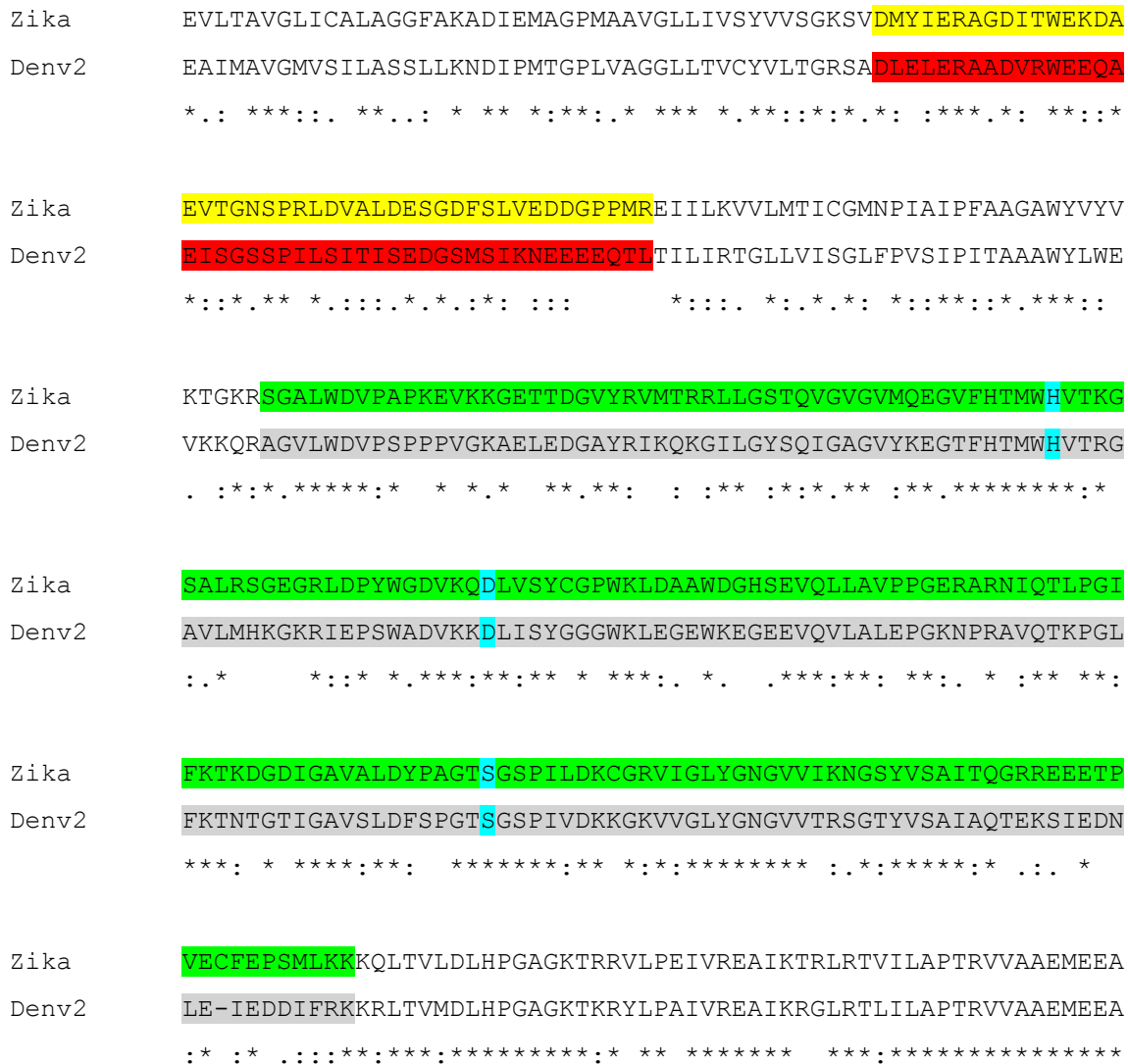
**Figure 3.3:** Sequence of the DENV 2 NS2B-NS3 linked protease construct (above) and a schematic representation of the essential parts (below). NS2B co-factor is coloured in grey, the NS3 part in red and the active site in cyan.

All genome data were taken from the GenBank<sup>®</sup>, as listed in Table 3.16 with their corresponding identification number.

**Table 3.16:** GenBank<sup>®</sup> identification number of polyprotein sequences.

Organism	GenBank ID	References
Dengue virus 1 (DENV 1)	AEV54606	(Lambrechts <i>et al.</i> , 2012)
Dengue virus 2 (DENV 2)	AB189124	(Sasmono <i>et al.</i> , 2015)
Dengue virus 3 (DENV 3)	ABY73735	(Sun <i>et al.</i> , 2011)
Dengue virus 4 (DENV 4)	AAK58017	(Blaney <i>et al.</i> , 2001)
Japanese encephalitis virus (JEV)	ABU94627	(Fulmali <i>et al.</i> , 2011)
Spondweni virus (SWV)	ABI54480	(Moureau <i>et al.</i> , 2010)
Tick-borne encephalitis virus (TBEV)	AHF27215	(Frey <i>et al.</i> , 2014)
Usutu virus (UTV)	AGP50649	(Nikolay <i>et al.</i> , 2013)
West Nile virus (WNV)	ACV90471	(Pybus <i>et al.</i> , 2012)
Yellow fever virus (YFV)	AGO04419	(Bonaldo <i>et al.</i> , 2017)
Zika virus (ZIKV)	KJ776791	(Baronti <i>et al.</i> , 2014)

Selected flaviviral polyprotein sequences were aligned, using the multiple sequence alignment tool Clustal Omega (EMBL-EBI), with the highly explored DENV 2 polyprotein and the relevant ZIKV NS2B and NS3 regions were identified and isolated (Figure 3.4).



**Figure 3.4: Extraction of alignment of DENV 2 (NS2B red, NS3 grey) with ZIKV (NS2B yellow, NS3 green), catalytic triad is highlighted in cyan.**

All flaviviral NS2B-NS3 linked protease constructs were designed analogously to the DENV 2 NS2B-NS3 linked protease construct, only differing in the tag cleavage site. As the TEV protease has a higher specificity for its recognition site, compared to thrombin (LVPR/GS), a TEV cleavage site (ENLYFQ/G) was introduced instead (see Figure 3.3 and Figure 3.5).



## Material and methods

MGSS HHHHHHGSENL<sup>Y</sup>FQGS DMYIERAGDITWEKDAEVTGNSPRLDVALDESGDFSLVEDDGPPMRGGG  
 GSGGGG SGALWDVPAPKEVKKGETTDGVYRVMTRRLLGSTQVGVGMQEGVFHTMWHVTKGSALRSGEG  
RLDPYWGDVKQDLVSYCGPWKLDAAWDGHSEVQLLAVPPGERARNIQTLPGIFKTKDGDIGAVALDYPA  
GTSGSPILDKCGRVIGLYGNGVVIKNGSYVSAITQGRREEETPVECFEPSMLKK



**Figure 3.5: Sequence of the ZIKV NS2B-NS3 linked protease and a schematic scheme. Co-factor is highlighted in yellow, NS3 protease domain in green, active site in cyan, His-tag in blue and the TEV cleavage site is underlined.**

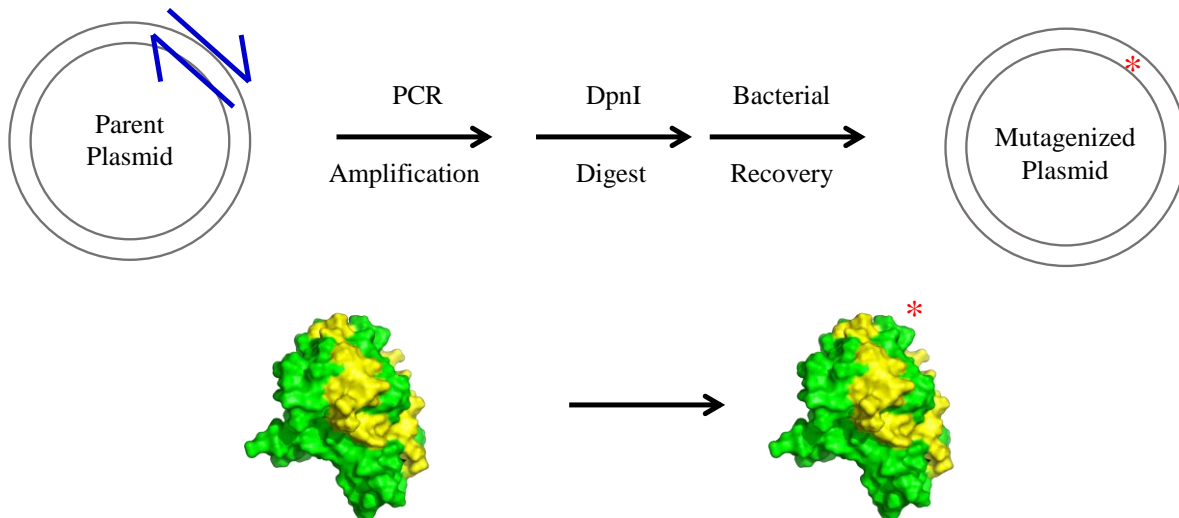
Due to autocatalytic self-cleavage problems with the ZIKV NS2B-NS3 protease construct (see chapter 4.2.2), specific point mutations at a potential cleavage sites in Japanese Encephalitis virus (JEV), Spondweni virus (SWV), Usutu virus (UTV) and West Nile virus (WNV) were introduced (Table 3.17).

**Table 3.17: Introduced point mutations at potential cleavage sites in JEV, SWV, UTV and WNV.**

	JEV	SWV	UTV	WNV
<b>Mutagenesis</b>	K95A	R95A	K95A	K95A

### 3.3.3. Site-directed mutagenesis

To study the interaction of inhibitors to the DENV 2 and the ZIKV NS2B-NS3 protease, point mutations were introduced via quick change PCR (Figure 3.6).



**Figure 3.6:** Schematic representation of the principle of site-directed mutagenesis via quick change PCR. Oligonucleotides are represented by blue arrows and the point mutation by a red asterisk. The sketch is drawn on the basis of <https://blog.addgene.org/site-directed-mutagenesis-by-pcr>.

All oligonucleotides were self-designed and purchased from Sigma-Aldrich (Darmstadt, GER). For the primer design, the relevant DNA regions were identified and isolated. Using the information of Table 3.18, the codons were changed appropriately. Attentions were drawn to the minimum amount of codon changes and a high average usage frequency (%) of the newly introduced codon in *E.coli* genes. The following general conditions were tried to be fulfilled: 1. primers should be between 30 and 40 base pairs long, 2. overlap should be at least 10 base pairs long, 3. melting temperature should be between 70 and 80 °C and 4. oligonucleotides should end on G or C.

**Table 3.18: Codon usage in *E. coli* genes. % represents the average usage frequency. The table is drawn on the basis of (*Codon usage*, 1996).**

	Codon	Amino acid	%	Codon	Amino acid	%	Codon	Amino acid	%	Codon	Amino acid	%
<b>U</b>	UUU	F	1.9	UCU	S	1.1	UAU	T	1.6	UGU	C	0.4
	UUC	F	1.8	UCC	S	1	UAC	T	1.4	UGC	C	0.6
	UUA	L	1	UCA	S	0.7	UAA	STOP	0.2	UGA	STOP	0.1
	UUG	L	1.1	UCG	S	0.8	UAG	STOP	0.03	UGG	W	1.4
<b>C</b>	CUU	L	1	CCU	P	0.7	CAU	H	1.2	CGU	R	2.4
	CUC	L	0.9	CCC	P	0.4	CAC	H	1.1	CGC	R	2.2
	CUA	L	0.3	CCA	P	0.8	CAA	Q	1.3	CGA	R	0.3
	CUG	L	5.2	CCG	P	2.4	CAG	Q	2.9	CGG	R	0.5
<b>A</b>	AUU	I	2.7	ACU	T	1.2	AAU	H	1.6	AGU	STOP	0.7
	AUC	I	2.7	ACC	T	2.4	AAC	H	2.6	AGC	STOP	1.5
	AUA	I	0.4	ACA	T	0.1	AAA	K	3.8	AGA	R	0.2
	AUG	M	2.6	ACG	T	1.3	AAG	K	1.2	AGG	R	0.2
<b>G</b>	GUU	V	2	GCU	A	1.8	GAU	D	3.3	GGU	G	2.8
	GUC	V	1.4	GCC	A	2.3	GAC	D	2.3	GGC	G	3
	GUA	V	1.2	GCA	A	2.1	GAA	E	4.4	GGA	G	0.7
	GUG	V	2.4	GCG	A	3.2	GAG	E	1.9	GGG	G	0.9
	<b>U</b>			<b>C</b>			<b>A</b>			<b>G</b>		

Melting temperatures ( $T_m$ ) were calculated using equation (2).

$$T_m [^{\circ}\text{C}] = 81,5 + 0,41 * (\%GC) - \frac{675}{N} - (\%mismatch) \quad (2)$$

% GC = GC content, % mismatch = amount of mismatched bases, N = total number of comprised nucleotides

**Table 3.19: Oligonucleotides for DENV 2 quick change PCR.**

<b>Name</b>	<b>N</b>	<b>T<sub>m</sub> [°C]</b>	<b>Sequence 5' → 3'</b>
DENVII_S135C_for	27	82	CCGGGTACGTGTGGCTCCCCGATCGTC
DENVII_S135C_rev	30	82	CGGGGAGCCACACGTACCCGGAGAGAAATC
DENVII_N152L_for	32	77	CTGTATGGTCTTGGTGTCTCGTGACCCGTAGTGG
DENVII_N152L_rev	32	79	CTACGGGTACAGACACCAAGACCATAACAGGCC
DENVII_T120I_for	38	78	CTGTTTAAGACCAATATCGGTACCATTGGCGCGGTTAG
DENVII_T120I_rev	34	76	CCGCGCCAATGGTACCGATATTGGTCTTAAACAG
DENVII_G14C_for	30	77	CCGCCGGTGTGCAAAGCAGAACTGGAAGAC
DENVII_G14C_rev	28	80	CTGCTTTGCACACCCGGCGGGAGACGG
DENVII_T118A_for	32	78	CAAACCAAACCGGGCCTGTTTAAGGCCAATAC
DENVII_T118A_rev	29	76	CGCCAATGGTACCCGTATTGGCCTTAAAC
DENVII_T120A_for	33	80	CAAACCGGGCCTGTTTAAGACCAATGCGGGTAC
DENVII_T120A_rev	30	80	CTAACCGCGCCAATGGTACCCGCATTGGTC
DENVII_T120D_for	35	75	CAAACCGGGCCTGTTTAAGACCAATGATGGTACC
DENVII_T120D_rev	32	74	GGCTAACCGCGCCAATGGTACCATCATTGGTC
DENVII_T118D_for	33	75	CAAACCAAACCGGGCCTGTTTAAGGACAATACG
DENVII_T118D_rev	32	76	CGCCAATGGTACCCGTATTGGCCTTAAAC
DENVII_T118I_for	34	75	CAAACCAAACCGGGCCTGTTTAAGATCAATACGG
DENVII_T118I_rev	36	77	CTAACCGCGCCAATGGTACCCGTATTGATCTTAAAC
DENVII_S135A_for	30	78	CTGGATTCTCTCCGGGTACGGCTGGCTCC
DENVII_S135A_rev	32	79	CTTTTTGTGACGATCGGGGAGCCAGCCGTAC
DENVII_T118C_for	36	78	GTTTAAGTGCAATACGGGTACCATTGGCGCGGTTAG
DENVII_T118C_rev	36	77	CGTATTGCACTTAAACAGGCCCGGTTTGGTTTGAAC
DENVII_T120C_for	33	77	CCAATTGCGGTACCATTGGCGCGGTTAGCCTGG
DENVII_T120C_rev	35	76	GGTACCGCAATTGGTCTTAAACAGGCCCGGTTTGG
DENVII_T122C_for	32	79	CAATACGGGTTGCATTGGCGCGGTTAGCCTGG
DENVII_T122C_rev	35	79	CGCCAATGCAACCCGTATTGGTCTTAAACAGGCC
DENVII_A164C_for	38	76	CGTGTCTGCATTGCCAACACGGAAAAGTCCATTGAAG
DENVII_A164C_rev	33	77	GCAATGCAGGACACGTAGGCACCACTACGGGTC
DENVII_A166C_for	39	78	CGATTTGCAACACGGAAAAGTCCATTGAAGATAACCCGG
DENVII_A166C_rev	31	77	GTGTTGCAAATCGCGGACACGTAGGCACCAC
DENVII_N152Q_for	32	79	GTATGGTCAGGGTGTCTCGTGACCCGTAGTGGTG
DENVII_N152Q_rev	34	79	GACACCCTGACCATAACAGGCCAACCCTTTACCC

As a negative control for enzymatic analysis, the serine 135 contributing to the catalytic triad was mutated into alanine. All the cysteine mutations were integrated for the purpose of labelling. It is worth mentioning that the DENV 2 NS2B-NS3 protease construct naturally does not possess cysteine residues.

## Material and methods

Glycine 14 in the offsite position, being far away from the active and allosteric site where reactions and/or ligand binding are taking place, was chosen for selective labeling with fluorophore dyes for single molecular spectroscopy experiments. Threonine 118, 120 and 122 and alanine 164 and 166 contributing to the allosteric site were made to be addressed covalently with the maleimide compounds *N*-benzylmaleimide (BMI) and *N*-ethylmaleimide (EMI). Further mutations of Thr-118 and Thr-120 into alanine, aspartic acid and isoleucine were performed to possibly disrupt interactions between allosteric site and allosteric inhibitors such as 1. Asparagine 152 is highly conserved throughout the flaviviral NS2B-NS3 proteases. It belongs to the allosteric site and was termed to be the “molecular switch” between the open and closed conformation (Piccirillo *et al.*, 2016). Mutagenesis on this amino acid had the purpose of exploring the possible introduction of conservative changes, i.e. leucine and glutamine and determination of the enzymatic kinetics.

**Table 3.20: Oligonucleotides for ZIKV quick change PCR.**

Name	N	T <sub>m</sub> [°C]	Sequence 5' → 3'
ZIKV_S135A2_for	32	80	GGCACCGCCGGTAGCCCGATTCTGGACAAATG
ZIKV_S135A2_rev	29	80	CTACCGGCGGGTGCCCGCCGGATAATCCAG
ZIKV_D120A_for	31	77	GTAAAGGCGGACCTGGTTAGCTACTGCGGTC
ZIKV_D120A_rev	30	76	CAGGTCCGCCTTAACATCACCCCAATACGG
ZIKV_D122A2_for	25	78	GACGGCGCTATTGGTGCGGTTGCGC
ZIKV_D122A2_rev	32	78	CCAATAGCGCCGTCCTTGGTTTTGAAGATACC
ZIKV_D120T_for	33	80	CCCTGCCGGGTATCTTCAAACCAAGACCGGCG
ZIKV_D120T_rev	30	78	GCGCAACCGCACCAATATCGCCGGTCTTGG
ZIKV_D122T_for	35	76	CCGGGTATCTTCAAACCAAGGACGGCACCATTGG
ZIKV_D122T_rev	32	75	GATAATCCAGCGCAACCGCACCAATGGTGCCG
ZIKV_R95A_for	27	79	CCGATGGCTGGCGGTGGTGGTAGCGGC
ZIKV_R95A_rev	28	80	CACCGCCAGCCATCGGCGGACCATCGTC
ZIKV_R29A_for	32	80	GACCCGTGCTCTGCTGGGTAGCACCCAGGTT
ZIKV_R29A_rev	29	81	CAGCAGAGCACGGGTCATCACACGGTAAACACCG

rev, revers; for, forward

The quick change PCRs on the ZIKV NS2B-NS3 constructs were carried out as described in chapter 3.3.1 using PCR mixtures of Table 3.7 and PCR cycling parameters of Table 3.8

S135A, R29A and R95A were introduced for investigations into the autocatalytic cleavage of the ZIKV linked NS2B-NS3 construct. Mutations of D122 and D124 to Ala and Thr, respectively, had the purpose of investigating the interactions of the inhibitor by imitating the DENV situation.

### 3.3.4. Recombinant expression and purification of flaviviral NS2B-NS3

Ni<sup>2+</sup>-nitriloacetic acid (Ni-NTA) resin was obtained from Qiagen (Hilden, GER), all other chemicals were obtained from Carl Roth (Karlsruhe, GER) unless indicated otherwise. Purchased chemicals were at least analytical grade or better.

Expression was carried out in *E. coli* strain BL21-Gold (DE3) (see Table 3.11). Cells were grown in LB medium (see Table 3.22) with the corresponding antibiotic (see Table 3.12) at 37 °C to an OD<sub>600 nm</sub> of ~ 0.5 and induced with 1 mM isopropyl-D-thiogalactoside (IPTG). Cells were then grown at 20 °C for 12 - 14 h. After harvest, cells were shock frozen in liquid nitrogen and stored at -20 °C until further use.

For purification, cells were resuspended in lysis buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 0.1 v/v Triton X-100, RNase, DNase, lysozyme, 1 mM PMSF, 1 mM DTT, 1 mM benzamidine) and lysed by sonication (Sonoplus, Bandelin). The lysate was cleared by centrifugation and the protein purified by immobilised metal affinity chromatography (IMAC). After extensive washing (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole), the protein was eluted (250 mM imidazole). The His-tag was removed by TEV cleavage during a dialysis step (20 mM Tris-HCl pH 8, 150 mM NaCl). After reverse IMAC, the protein was concentrated using Vivaspin-6 or Vivaspin-20 centrifugal filters (GE Healthcare) with a molecular weight cut-off of 10 kDa and finally a gelfiltration step (HiLoad 16/60 Superdex 75 or 200 column, GE Healthcare) was carried out (50 mM Tris-HCl pH 8, 150 mM NaCl). The entire protein expression and purification was carried out at 4 °C, if not pointed out otherwise.

The protein concentrations throughout purification were determined by absorption measurements at  $\lambda = 280$  nm using a UV-VIS spectrophotometer (Thermo Scientific NanoDrop 2000c). Actual protein concentration ( $c$ ) results from the measured value of Abs<sub>280 nm</sub> divided by the corresponding molar absorption coefficient  $\epsilon_{280 \text{ nm}}$ , which was calculated for each protein based on the amino acid composition using the online tool ExpASy – ProtParam (<https://web.expasy.org/protparam>).

$$c \left[ \frac{\text{mol}}{\text{L}} \right] = \frac{\text{Abs}_{280 \text{ nm}}}{\epsilon_{280 \text{ nm}} \left[ \frac{\text{L}}{\text{M} \cdot \text{cm}} \right] * d [\text{cm}]} \quad (3)$$

Throughout all steps, the sample purity was assessed through a discontinuous Tris-glycine SDS-polyacrylamide gel electrophoresis system. The gels were prepared with the two SDS-page buffer formulations described in Table 3.21. As running buffer 1% (w/w) sodium dodecyl sulfate (SDS), 192 mM glycine and 25 mM Tris base was used. Protein and peptide bands were visualised via staining with staining dye (45% (v/v) ethanol, 10% (v/v) AcOH, and 514 mg/L Coomassie Brilliant Blue) and destaining with 10% (v/v) AcOH. All protein and peptide samples were mixed

## Material and methods

with 4x SDS loading dye (200 mM Tris-HCl pH 6.8, 400 mM DTT, 8% SDS (w/w), 0.4% bromophenol blue (w/w) and 40% glycerol (v/v)).

**Table 3.21: SDS-page buffer formulations.**

Ingredients	Stacking gel	Running gel	
		12%	15%
H <sub>2</sub> O	3.4 mL	6.6 mL	4.6 mL
1.5 mM Tris-HCl pH 8.8	-	5 mL	5 mL
1.0 mM Tris-HCl pH 6.8	0.63 mL	-	-
30% acrylamide mix	0.83 mL	8 mL	10 mL
SDS	50 µL	200 µL	200 µL
APS	50 µL	200 µL	200 µL
TEMED	5 µL	20 µL	20 µL

<sup>15</sup>N-labeled protein was created by bacteria growth in M9-Media. M9 medium is a minimal medium consisting of nine ingredients listed in Table 3.22. In this medium the only nitrogen source was <sup>15</sup>N-NH<sub>4</sub>Cl, which was purchased from Eurisotope (Saarbrücken, GER). The cultures were grown and expressed analogously to the unlabeled protein.

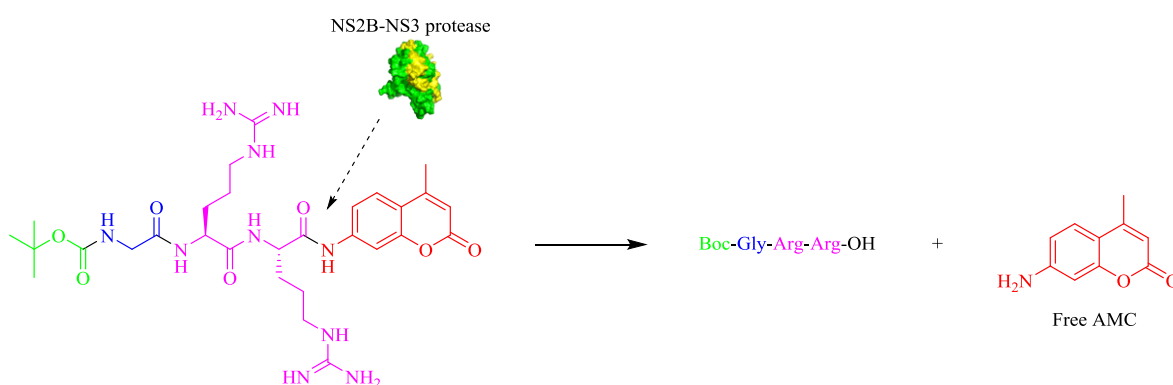
**Table 3.22: Composition of used bacterial growth media.**

Medium	Ingredients
Luria Broth (LB)	5 g/L yeast extract, 10 g/L NaCl, 10 g/L peptone
Terrific Broth (TB)	23,6 g/L yeast extract, 11.8 g/L peptone, 9.4 g/L K <sub>2</sub> HPO <sub>4</sub> , 2.2 g/L KH <sub>2</sub> PO <sub>4</sub>
M9 medium	15.0 g/L KH <sub>2</sub> PO <sub>4</sub> , 33.9 g/L Na <sub>2</sub> HPO <sub>4</sub> and 2.5 g/L NaCl 2 mM MgCl <sub>2</sub> , 10 µM FeCl <sub>3</sub> , 1 mL magic mix, 10 mL trace elements mix 0.75 g/L [ <sup>15</sup> N]-NH <sub>4</sub> Cl Eurisotope 2 g/L glucose
Magic mix	Centrum® vitamin tablet dissolved in 20 mL MPH <sub>2</sub> O
Trace elements mix	0.2 g/L CaCl <sub>2</sub> -dihydrate 0.2 g/L ZnSO <sub>4</sub> -heptahydrate 0.2 g/L MnSO <sub>4</sub> -monohydrate 5 g/L thiamine and niacin 0,1 g/L biotin

### 3.4. Analytical methods used to examine the flaviviral NS2B-NS3 protease

#### 3.4.1. Fluorometric enzyme assay

The fluorometric enzyme assay was used to evaluate the enzymatic kinetics of the different purified flaviviral NS2B-NS3 proteases and to investigate the interactions with allosteric inhibitors. All fluorometric enzyme assays were performed on a M200 infinite<sup>®</sup> pro or Spark<sup>®</sup> microplate reader using Boc-Gly-Arg-Arg-7-amino-4-methylcoumarin (Boc-GRR-AMC, Bachem, Bad-Neuenahr-Ahrweiler, GER) as substrate. Released AMC was excited at a wavelength of 380 nm and detected at 460 nm. Figure 3.7 outlines the substrate hydrolysis. The linear correlation of the increase of the fluorescence intensity with time and the enzyme activity was used for quantifying enzyme kinetics.



**Figure 3.7: Hydrolysis of Boc-Gly-Arg-Arg-AMC by the *flaviviral* NS2B-NS3 protease.**

The enzyme assay was carried out in a Tris-HCl buffer (50 mM Tris-HCl pH 9.0, 20% glycerol, 1 mM Chaps) at 25 °C in a final volume of 200  $\mu$ L (see Table 3.23) using white 96-well plates.

**Table 3.23: Pipetting scheme for fluorometric enzyme assay.**

Ingredients	V [ $\mu$ L]
Assay buffer	180
Enzyme	5
Inhibitor (or DMSO)	10
Substrate (4nM in DMSO)	5



## Material and methods

$IC_{50}$  values were determined by varying the inhibitor concentration in a range from 0.1 to 100  $\mu\text{M}$  and calculated from the slopes of the progress curves against the inhibitor concentrations using the nonlinear four-parameter logistic fit (equation (4)) of Grafit 5.0.13 (Leatherbarrow, 2010).

$$y = \frac{\text{Range}}{1 + \left(\frac{[I]}{IC_{50}}\right)^s} + \text{Background} \quad (4)$$

Range = difference in fluorescence signal between 0% and 100% activity, Background = fluorescence signal of 0% activity, [I] = inhibitor concentration and s = slope factor.

In general, the  $IC_{50}$  values were only determined of the inhibitors which proved an inhibition of over 50% at a concentration of 20  $\mu\text{M}$ .

The kinetic data were determined in duplicates for 30 min with varying substrate concentrations from 50 to 1000  $\mu\text{M}$ . The Michaelis constant ( $K_M$ ) calculated from the slopes of the initial 10 min of the reactions using the nonlinear fit of the Michaelis-Menten equation (equation (5)) is the substrate concentration at which the reaction rate is half of the maximum reaction rate at saturation ( $v_{max}$ ).

$$v = \frac{v_{max} * [S]}{K_m + [S]} \quad (5)$$

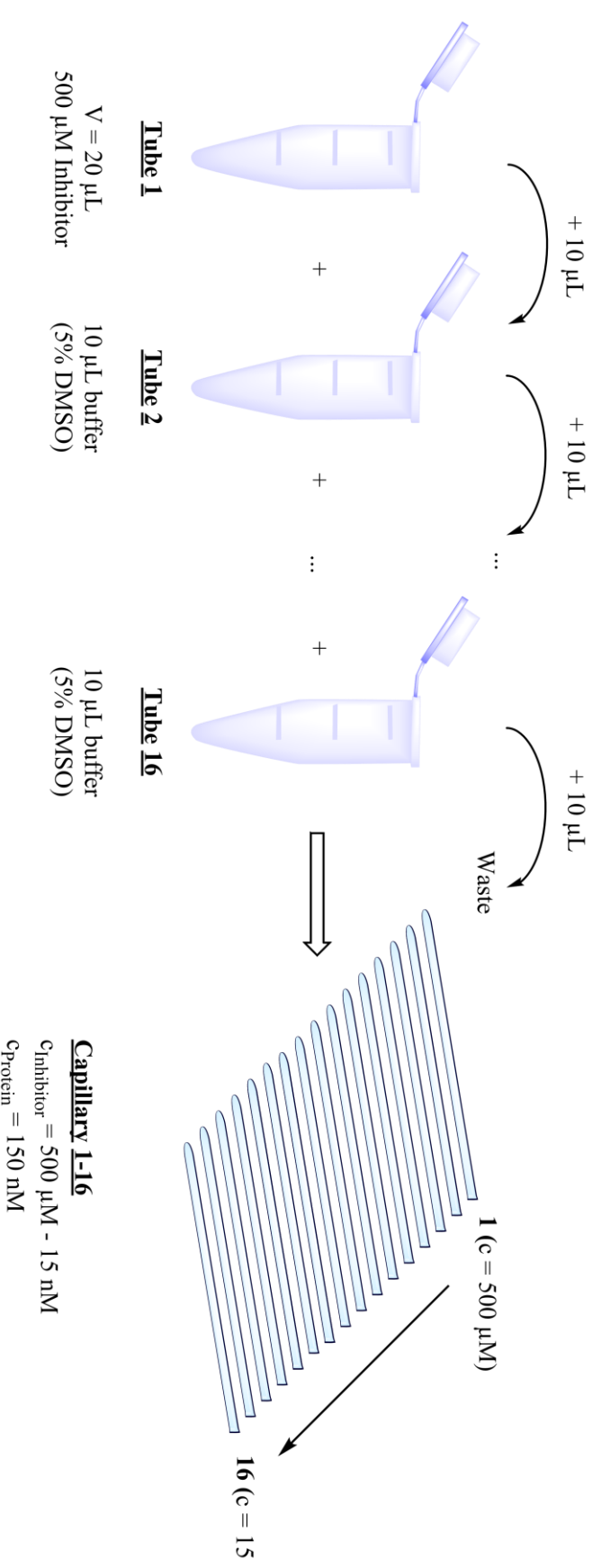
$v$  = reaction rate and [S] = substrate concentration

### 3.4.2. Microscale thermophoresis (MST)

To determine the dissociation constants ( $K_d$ ), reflecting the interaction between the inhibitors and the ZIKV R95A protease, MST was used. The purified protease was labeled with Monolith NT Protein Labeling kit RED-NHS from NanoTemper Technologies (München, GER), containing NT-647-NHS as fluorescent dye for primary amines, e.g. lysine residues. The protein labeling was carried out according to the vendor provided labeling protocol. 100  $\mu\text{L}$  of a 20  $\mu\text{M}$  protein solution were labeled with 5 fold of the provided labeling dye.

10  $\mu\text{L}$  of the 300 nM labeled protein stock solution was combined with 16 different concentrations (ranging from ~ 500  $\mu\text{M}$  to ~ 15 nM) of the inhibitors ( $V = 10 \mu\text{L}$ ) (see Figure 3.8).

The compounds were diluted with a buffer containing 1 mM CHAPS, 50 mM Tris and 20% glycerol at pH 9. To ensure equal DMSO concentrations, DMSO was added to a final concentration of 5% (v/v). The measurement was carried out on a NanoTemper® Monolith NT.115 instrument (MST power of 10%, LED power of 50%, 30 s laser-on time and 5 s laser-off time) with the use of standard capillaries provided in the MST kit.



Capillary no.	$c$ (1, R95A) [ $\mu\text{M}$ ]	$c$ (1, S135A) [ $\mu\text{M}$ ]	Capillary no.	$c$ (1, R95A) [ $\mu\text{M}$ ]	$c$ (1, S135A) [ $\mu\text{M}$ ]
1	625.000	500.000	9	2.441	1.953
2	312.500	250.000	10	1.221	0.977
3	156.250	125.000	11	0.610	0.488
4	78.125	62.500	12	0.305	0.244
5	39.063	31.250	13	0.153	0.122
6	19.531	15.625	14	0.076	0.061
7	9.766	7.813	15	0.038	0.031
8	4.883	3.906	16	0.019	0.015

Figure 3.8: MST dilution scheme of the sample preparation for the ZIKV S135A with 1 experiment. The final inhibitor concentrations for each capillary are listed below.

### 3.4.3. Metabolism study

The metabolism study was carried out under my supervision by a student [REDACTED] in the context of his bachelor thesis.

Pooled male rat liver microsomes were purchased by Corning® Gentest™ and characterised by the company (Table 3.24).

**Table 3.24: Enzyme assay results characterised by Corning® Gentest™ of the purchased pooled male rat liver microsomes.**

Enzyme measured	Assay	Enzyme Activity
Total P450	“Omura and Sato”	650 pmole/mg
OR	Cytochrome c reductase	310 nmole (mg * min)
Cytochrome b <sub>5</sub>	Spectrophotometric	400 pmole/mg
CYP3A	Testosterone 6β-hydroxylase	8700 pmole/(mg*min)
CYP2C	Testosterone 16β-hydroxylase	5300 pmole/(mg*min)
CYP1A	7-Ethoxyresorufin O-Deethylase	150 pmole/(mg*min)

The nicotinamide adenine dinucleotide phosphate (NADPH) generating system was prepared via incubation at 37 °C for 10 minutes of 2-amino-2-(hydroxymethyl)-1,3-propanedio (Trizma®) buffer (790 µL, 100 mM, pH 7.4), MgCl<sub>2</sub> (50 µL, 80 mM), glucose-6-phosphate (50 µL, 100 mM), NADP disodium salt (50 µL, 20 mM) and glucose-6-dehydrogenase (50 µL, 100 IU/mL). Additions of microsomes (50 µL, 20 mM) were carried out thorough homogenisation and further 10 minutes of incubation. The metabolism reaction was started by addition of the inhibitor.

Aliquots of 50 µL were taken at 0, 5, 10, 20, 30, 60, 90, 150 and 240 minutes and added to 200 µL of ice-cold acetonitrile to quench the reaction. 50 µL of an internal standard (naphthaline, 0.12 mM) were added to each aliquot and centrifuged at 4 °C and 16000 rpm for 10 minutes. The supernatant was removed and the aliquots were stock frozen at -20 °C until LC-MS analysis was performed.

LC-MS/MS-system: Agilent Poroshell 120 EC-C18 150x2.10 mm 4 µM column; mobile phase: 55% acetonitrile, 35% H<sub>2</sub>O, 10% of a 0.1% solution of formic acid in water. Ion chromatograms were obtained using electronic filters for the ions of interest.

### 3.4.4. NMR Spectroscopy

The NMR-experiments were carried out by [REDACTED], University of Mainz.

Standard 1D  $^1\text{H}$ - and 2D  $^1\text{H}$ -,  $^{15}\text{N}$ - NMR spectra (Bruker, Karlsruhe) were recorded for selectively  $^{15}\text{N}$ -labeled DENV NS2B-NS3 prepared from cells grown in minimal medium supplemented with  $^{15}\text{N}$  ammonium chloride as the sole nitrogen source. Protein purification was carried out as described above (chapter 3.3.4). Purified protease was buffer-exchanged to NMR buffer (20 mM HEPES (pH 7.3), 2 mM DTT). The DENV 2 NS2B-NS3 linked sample was concentrated to 170  $\mu\text{M}$  and the unlinked sample to 110  $\mu\text{M}$ . The NMR-experiments were carried out at 298 K and measured at a Bruker 800 MHz spectrometer equipped with cryogenic triple resonance probe. HSQC spectra were recorded with standard pulse sequences and processed with TopSpin 2.1 (Bruker, Karlsruhe).

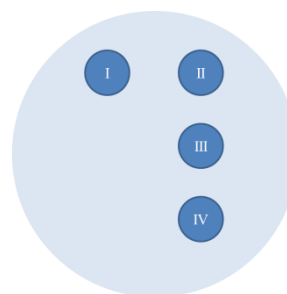
### 3.4.5. Protein crystallisation

Robot assistant protein crystallisation experiments were carried out by the working group of [REDACTED]. DENV 2 NS2B-NS3 protease crystallisations trials, including co-crystallisation with compound **1**, were mainly carried out by [REDACTED]. For protein crystallization of ZIKV NS2B-NS3 protease [REDACTED] was assisting. The University of Marburg has a crystallisation laboratory (MarXtal) equipped with the two crystallisation robots, Digilab Honeybee 963<sup>TM</sup> and Cartesian Microsys<sup>TM</sup> SQ 4000. The documentation system Formulatrix Rock Imager<sup>TM</sup> was used to observe crystal growth.

The crystallisation experiment of the purified ZIKV NS2B-NS3 protease was partly done at the University of Mainz using Crystal Screen and Crystal Screen 2 purchased from Hampton research (Aliso Viejo, CA, USA). Four stock solutions were prepared, two containing exclusively the ZIKV NS2B-NS3 protease and the other two containing ZIKV NS2B-NS3 protease along with the inhibitor **1** (Table 3.25).

**Table 3.25: Concentrations of the stock solutions and pipetting scheme for crystallisation experiment.**

Stock solution	[ZIKV R65A] (mg/mL)	[ <b>1</b> ] (mM)
Solution I	20	/
Solution II	10	/
Solution III	20	5
Solution IV	10	10



## Material and methods

Crystallisation screens were performed with the hanging drop vapour diffusion method using ComboPlate™ 24-Well Protein Crystallization Plates (Greiner Bio-One). The reservoir was filled with 500 µL of in the Crystal Screen provided solutions and the crystallisation drops consisting of 1 µL protein solution plus 1 µL reservoir were pipetted as shown in the scheme of Table 3.25. The trays were stored at 20 °C and observed over a period of 6 months.

### 3.4.6. Cleavage site study of ZIKV NS2B-NS3 protease

Due to protein cleavage problems, while purifying the ZIKV NS2B-NS3 linked protease, a cleavage site experiment was carried out. This study was carried out under my supervision by a student [REDACTED] in the context of her research module. In order to identify whether the resulting cleavage results from autocatalytic cleavage following a *cis*- or a *trans*-mechanism, the enzymatically active ZIKV NS2B-NS3 (R95A) linked protease was incubated with the enzymatically inactive ZIKV NS2B-NS3 (S135A) linked protease. As the experiment was observed via 15% SDS-pages (Table 3.21), the ZIKV NS2B-NS3 (R95A) linked protease was kept in a concentration below the SDS-Page detection level. The ZIKV cleavage site study was set up as described in Table 3.26.

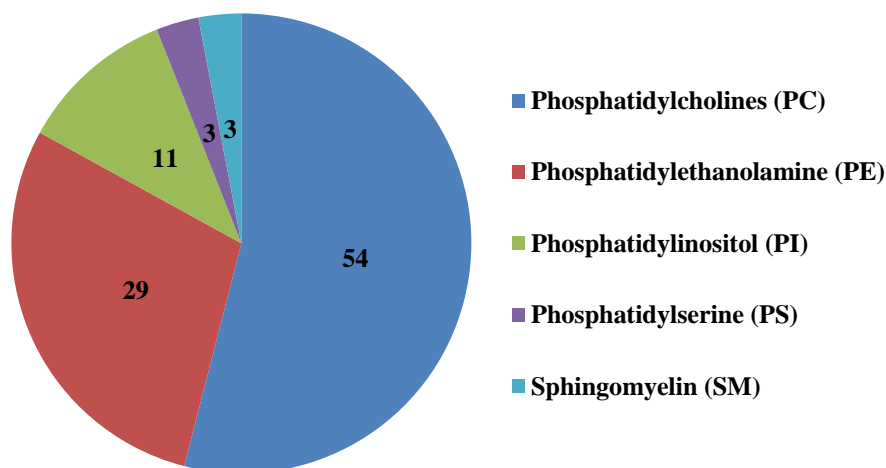
**Table 3.26: Experimental setup for the autocatalytic cleavage site study.**

Procedure	T (°C)	[S135A] (mg/mL)	[R95A] (mg/mL)	Ratio (S/A : R/A)
Experiment 1	RT	0.136	0.008	1 : 0.05
Experiment 2	37	0.136	0.008	1 : 0.05
Experiment 3	37	2.9	2.9	1 : 1
Experiment 4	37	1.36	0.08	1 : 0.05

1 µL samples were taken in different time intervals from each experiment and immediately quenched by adding 4x SDS-dye. At first, the samples were taken in a high frequency of every 5 minutes and at the end broadening the period to 30 minutes/sample. The SDS-page samples were shock frozen in liquid nitrogen and stored at -20 °C before being analysed.

### 3.4.7. Lipid sedimentation assay of ZIKV NS2B-NS3 protease

The lipid sedimentation assay was used to study the putative interactions of ZIKV NS2B-NS3 (R95A) linked protease with lipids and was carried out under my supervision by a student [REDACTED] in the context of her research module. As the ZIKV NS2B-NS3 protease is a membrane bound protein, one lipid composition was chosen to mimic the endoplasmic reticulum (ER-mix) according to Figure 3.9 and the other without the negatively charged lipids (see Table 3.27).



**Figure 3.9:** Approximate distribution of phospholipids in the endoplasmic reticulum of mammalian cells. The pie charts is drawn on the basis of Andreyev *et al.*, 2010.

**Table 3.27:** Liposome composition of the sedimentation assay mixtures.

Lipid composition	PC % (w/w)	PE % (w/w)	PG % (w/w)
“ER-mix”	54	29	17
“positive charged-mix”	62.5	37.5	0

Lipid solutions of 25 mg/mL PC, PE and PG in chloroform (Avanti Lipids) were mixed to obtain the two lipid mixture compositions of Table 3.27. The organic solvent was removed by nitrogen gas aeration and dried in a desiccator over-night. The lipid sediment mixtures were then suspended in buffer (50 mM Tris-HCl pH 8, 150 mM NaCl), resulting in a concentration of 4 mg/mL. To obtain multilamellar liposomes the lipid mixture was extruded 15 times using the mini-extruder set (Avanti Polar Lipids) in combination with a 400 nm membrane. 20 µL of the ZIKV NS2B-NS3 (R95A) linked protease (50 µM) was mixed with 780 µL of the 4 mg/mL lipid stock solutions and incubated at 4 °C under mild shaking for 1 hour. Before centrifugation, a SDS-PAGE sample was taken, to ensure the presences of protein and to obtain an initial concentration. Both mixtures were

## Material and methods

then centrifuged for 1 hour at 22,500 rpm (Optima MAX-XP, Beckmann Coulter) and 4 °C. To avoid resuspension the supernatant was rapidly separated from the liposome pellet, which was then resuspended in 80 µL buffer solution. SDS-PAGE samples were taken from the supernatant and the suspended pellet. Protein solutions without liposomes were used as negative controls. The samples were run on 15% SDS-PAGE gels.

### 3.4.8. Phylogenetic study of flaviviral proteases

The phylogenetic study was carried out by [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]

#### 3.4.8.1. Domain architecture of the ZIKV polyprotein

Using Pfam (El-Gebali *et al.*, 2019) the ZIKV polyprotein sequence was entered as reference to compare the feature architecture of all other polyproteins against it. 14 Pfam-A matches to this search were found and the data was downloaded from Virus Pathogen Resource (ViPR). Applied filters included: complete genomes, include polyprotein, remove identical proteins and name filter–polyprotein, leading to 7,748 obtained sequences (download date: 6. Dec. 2016). For taxon composition the sequences were extracted by the non-redundant list of organisms as defined by the 'organisms:' flag in the fasta header. This revealed a list of 311 organisms, which was further reduced manually to 143 taxa of interest. Features were annotated and feature architecture similarity computed using default settings. For each organism a representative strain with the highest FAS score was selected. The sub-selected organisms and polyproteins that had a FAS score of  $\geq 0.8$  were used for tree reconstruction and finally comprised the list down to 99 sequences. The polyproteins were aligned with muscle (Edgar 2004). To enhance the phylogenetic signal, columns where more than 50% of the sequences were represented by a gap, were removed. ProtTest (Abascal *et al.* 2005) using the Akaike Information Criteria (AIC) was run to determine the best fitting evolutionary model for the data. This resulted in LG+G+F as best fitting model (Le and Gascuel, 2008). The phylogenetic tree was inferred with [<https://github.com/stamatak/standard-RAxML>][raxml] v. 8.2.4 (Stamatakis, 2006) enabling the rapid bootstrapping option (Stamatakis, Hoover and Rougemont, 2008). 100 bootstrap replicates were performed. Tree display was done with FigTree and is illustrated in chapter 6.9.1.

### 3.4.8.2. Extraction of the flaviviral protease

For the identification of the flaviviral proteases the shortened construct sequence commonly used in literature (as described in chapter 1.7.1) was used. This comprised the flaviviral sequences to two Pfam domains, the N-terminal NS2B and the C-terminal Peptidase\_S7 (NS3) Pfam domain. 87 protease sequences adhere to the canonical domain architecture and were extracted on this basis. The phylogenetic tree was generated and is displayed in chapter 4.3.1.

The flaviviral NS2B-NS3 protease sequences were aligned to analyse the conservation of the C-terminal region using Jalview (Waterhouse *et al.*, 2009). Based on the aligned sequences a weblogo was generated with WebLogo (Crooks *et al.*, 2004) and subsequently annotated with Inkscape<sup>®</sup> (see chapter 4.3.1, Figure 4.25).



## 4. Results and discussion

### 4.1. DENV NS2B-NS3 protease

DENV NS2B-NS3 protease is a promising drug target and thus extensively investigated in this regard. The unusual narrow and structurally relatively flat active site of the serine protease causes problems in drug discovery (see chapter 1.7.4). Development of allosteric inhibitors can bypass this challenge. Previous studies have identified **1** (Table 1.2) as a highly potent allosteric DENV NS2B-NS3 protease inhibitor, with  $IC_{50}$  values of 4.2  $\mu\text{M}$  and 1.0  $\mu\text{M}$  against DENV 2 and DENV 3 respectively (Wu *et al.*, 2015).

This chapter summarises the development of allosteric inhibitors evolved from compound **1** and represents investigations into the allosteric site of the DENV 2 NS2B-NS3 protease.

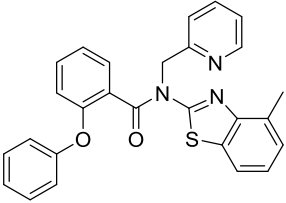
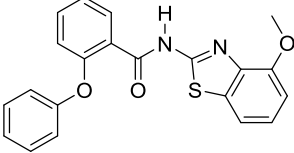
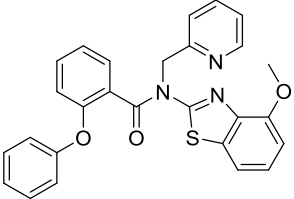
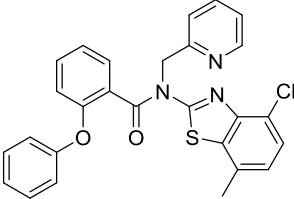
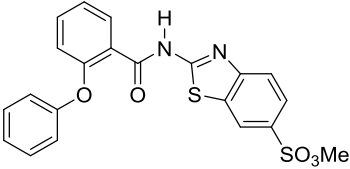
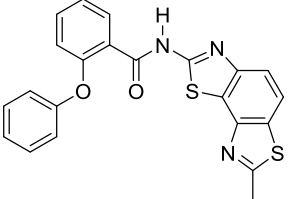
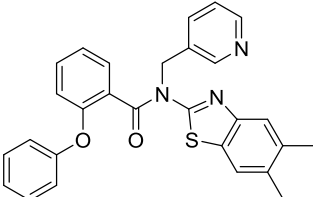
#### 4.1.1. Structure-activity relationship studies (SAR)

##### 4.1.1.1. SAR study of purchasable putative inhibitors using virtual screening

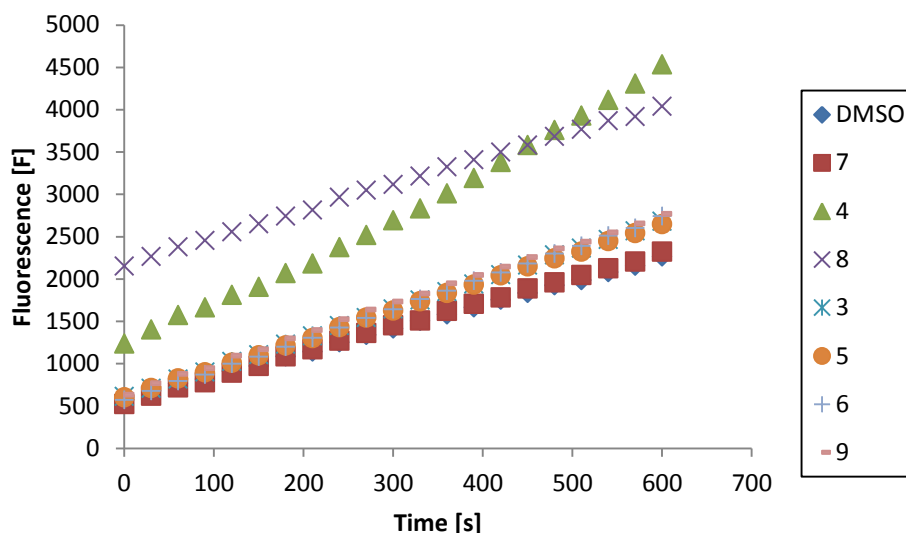
Virtual screening results of the commercially purchasable chemical library (carried out as described in chapter 3.1.1) resulted in ranking from highly potent to less viable compounds for DENV NS2B-NS3 protease inhibition. For SAR study purposes two highly potent, two passably and three less promising compounds were selected, purchased from Ambinter (Orléans, FR) and tested against the DENV 2 NS2B-NS3 protease in the fluorometric enzyme assay (described in 3.4.1). Table 4.1 summarises the results of the virtual screening and Figure 4.1 presents the enzymatic results of the purchased chemicals tested on the DENV 2 NS2B-NS3 protease.

Via virtual screening, compounds **3** and **4** were identified to be the most potent compounds for inhibition of the DENV NS2B-NS3 protease. Both compounds obtained a highly negative value of  $\Delta G$  and a LE value of 0.21. Representing the middle of the virtual screening results, compounds **5** and **6** were chosen. Both compounds still proved a negative  $\Delta G$  of -26 kJ/mol and -23 kJ/mol and with LE values of 0.18 and 0.16 these compounds represented the mean. **7**, **8** and **9** had the least prospective regarding the virtual screening results. The  $\Delta G$  values proved to be around 0 kJ/mol and equally LE values of zero.

**Table 4.1: Virtual screening results of the commercially available compound library obtained by a similarity search using SciFinder. The table includes the compound identification number from Ambinter in square brackets. Compounds were ranked by the screening outcome using  $\Delta G$  and LE as indicator.**

Compound identity	Structure	$\Delta G$ [kJ/mol]	LE
3 [Amb 3283285]		-29	0.21
4 [Amb 3422386]		-25	0.21
5 [Amb 3283341]		-26	0.18
6 [Amb 3283752]		-23	0.16
7 [Amb 6398800]		-4	0.03
8 [Amb 1262978]		0	0.00
9 [Amb 3277965]		2	0.00

## Results and discussion

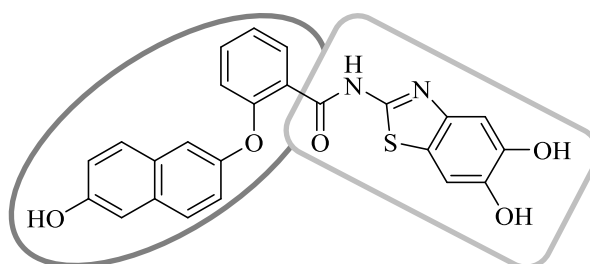


**Figure 4.1: Progress curves of substrate hydrolysis in presence of the purchased chemicals selected by virtual screening results tested against DENV 2 NS2B-NS3 protease in the fluorometric enzyme assay.**

Fluorometric enzyme assay results are graphically presented in Figure 4.1. Exceptionally of compound **4**, the progress curves of substrate hydrolysis showed that the compounds had no inhibitory effects on the DENV NS2B-NS3 protease. Compounds **4** and **8** proved interferences with the fluorometric enzyme assay.

### 4.1.1.2. Self-established screening library of eight compounds on the basis of **1**

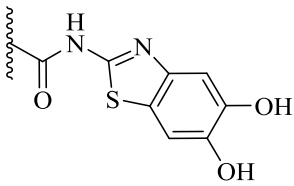
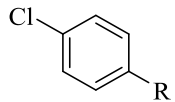
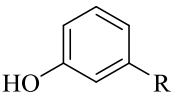
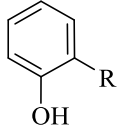
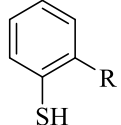
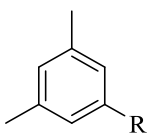
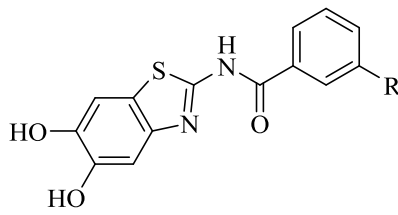
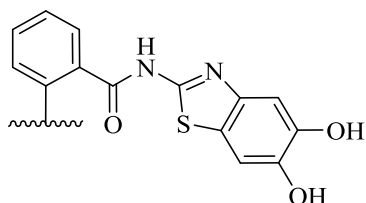
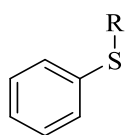
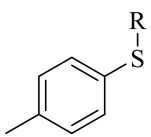
Having identified compound **1** as lead structure for allosteric inhibition of DENV 2 and DENV 3 NS2B-NS3 proteases, the replacement of the naphthalene side leaving the benzothiazole part unchanged was investigated (see Figure 4.2).



**Figure 4.2: Scaffold hopping of 1. The benzothiazole side (framed in light grey) was left unchanged, while the diaryl ether moiety (circuited in dark grey) was altered.**

The following compounds (CPD) were synthesised as described in chapter 3.2. Inhibitory data, using fluorometric enzyme assays (described in 3.4.1) were obtained in order to establish structure-activity relationships.

**Table 4.2: Summary of the synthesised and enzymatically analysed CPDs.  $IC_{50}$  values of inhibition of DENV 2 protease were only determined if inhibition at 20  $\mu$ M inhibitor concentration was better than 50%.**

Backbone	Substitute	CP D	Inhibition at 20 $\mu$ M [%]	$IC_{50}$ [ $\mu$ M]
		10	88	$4.3 \pm 1.0$
		11	13	n.d.
		12	25	n.d.
		13	53	$20,2 \pm 0,8$
		14	50	$23.0 \pm 0.0$
		15	100	$1.2 \pm 0.2$
		16	100	$5.1 \pm 0.3$
		17	100	$5.9 \pm 0.8$

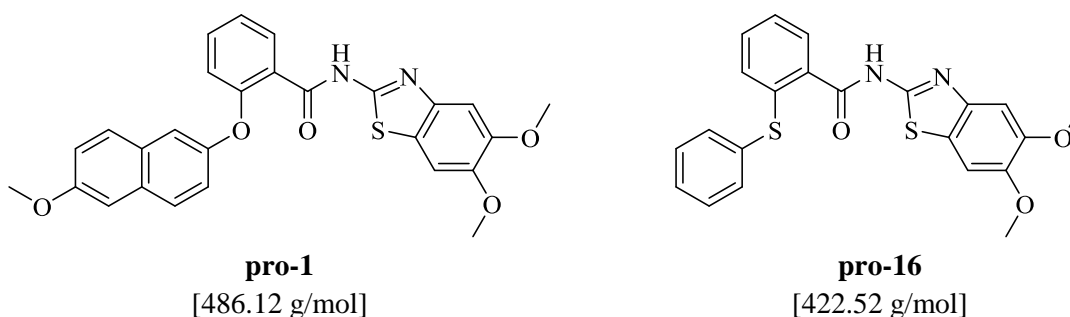
## Results and discussion

Substitution of the phenyl ring at para position with chlorine (**10**) as a functional group, achieved an  $IC_{50}$  value similar to that of inhibitor **1**. An OH- substitution in ortho- (**11**) and meta- (**12**) position of the phenyl group diminished the inhibitor effect. Only 13% and 25% inhibition at a concentration of 20  $\mu\text{M}$  respectively were determined. Introducing the thiol functional group in the meta position (**13**) sustained a  $IC_{50}$  value of 20.3  $\mu\text{M}$ . Double methyl substitutions at both ortho positions of the phenyl group (**14**) led to a comparable  $IC_{50}$  value of 23.0  $\mu\text{M}$ . **15**, the largest synthesised molecule, proved an  $IC_{50}$  value of 1.2  $\mu\text{M}$ . Maintaining the diaryl bridging, however substituting the ether bridge with a thioether led to  $IC_{50}$  values of 5.1  $\mu\text{M}$  and 5.9  $\mu\text{M}$  for **16** and **17** respectively.

### 4.1.2. Metabolic study of pro-1 and pro-16

Wu *et al.*, 2015 reported that the dimethoxy prodrugs were inactive against the DENV 2 NS2B-NS3 protease in the fluorometric enzyme assay, but showed antiviral activity against Vero cells infected with DENV 1 and DENV 2. These findings led to the assumption that the dimethoxy compounds acted as prodrugs, being demethylated in cells yielding the more potent dihydroxy derivatives.

In the fluorometric enzyme assay this phenomenon was also noticed for the synthesised inhibitor **16** and its precursor molecule, **pro-16**. Hence a metabolic study was carried out to identify the metabolic products. The metabolic study was performed for **pro-1** and **pro-16**, represented in Figure 4.3 and analysed via LC-MS/MS as described in chapter 3.4.3. Naphthalene was used as an internal standard, as it is not ionisable it was not visible in the LC-MS-spectrum.



**Figure 4.3: Structures and molecular weights of pro-1 and pro-16.**

#### 4.1.2.1. Metabolic study of pro-1

The LC-MS/MS-spectrum of the metabolite **pro-1** at 0 min (Figure 4.4) represented the zero point of the metabolism study.

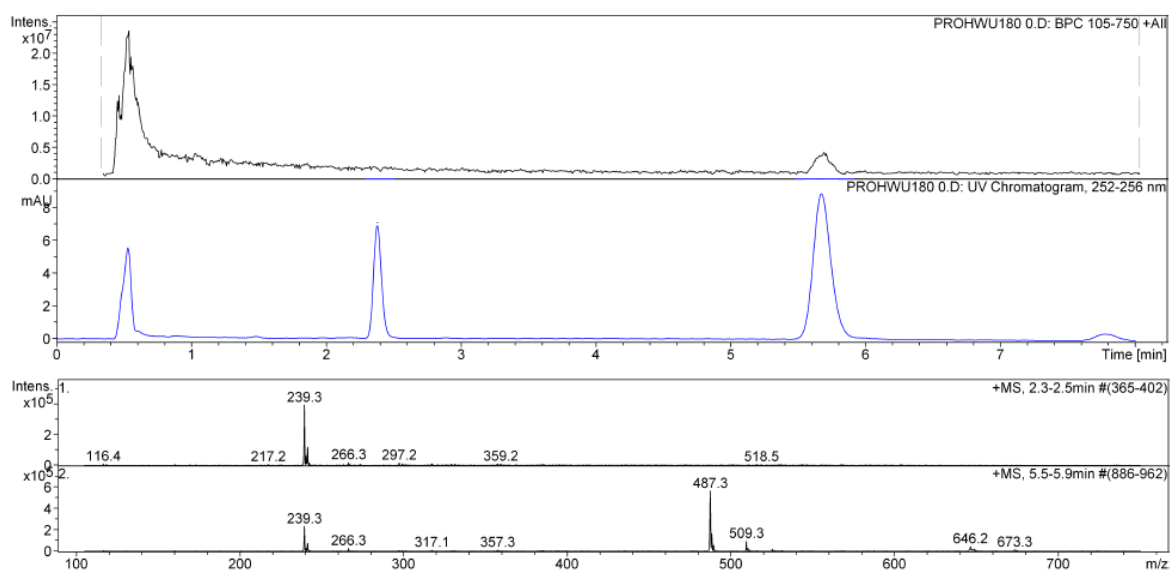


Figure 4.4: LC-MS-spectrum of the metabolite **pro-1** at 0 min.

The chromatogram (Figure 4.4, above) visualised at 0.5 min retention time the Trizma® buffer, at 2.4 min the internal standard naphthalene and at 5.6 min **pro-1**. The MS-spectra (Figure 4.4, below) correlated with the chromatogram as the m/z ratio of 487.3 represented **pro-1**. Throughout all MS-spectra a contamination at m/z ratio of 239.3 were visible and throughout this work negligible.

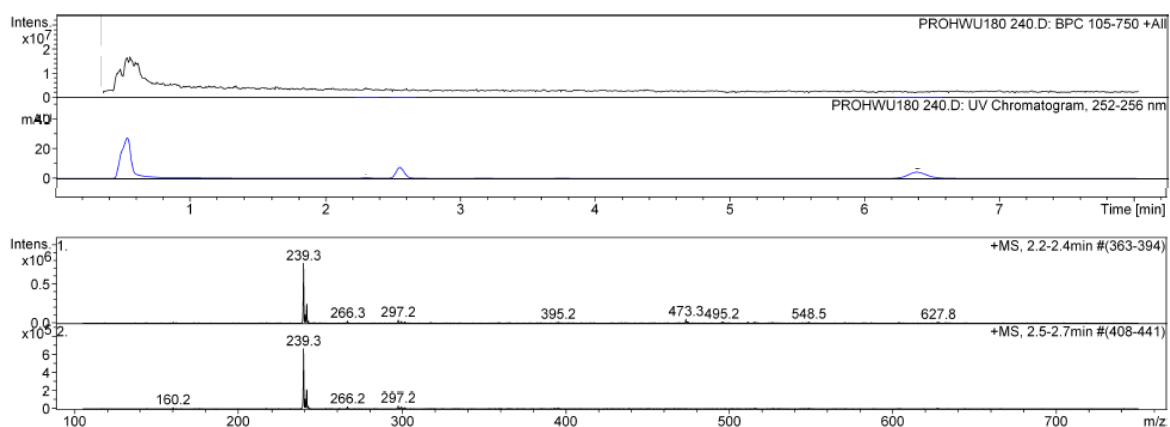
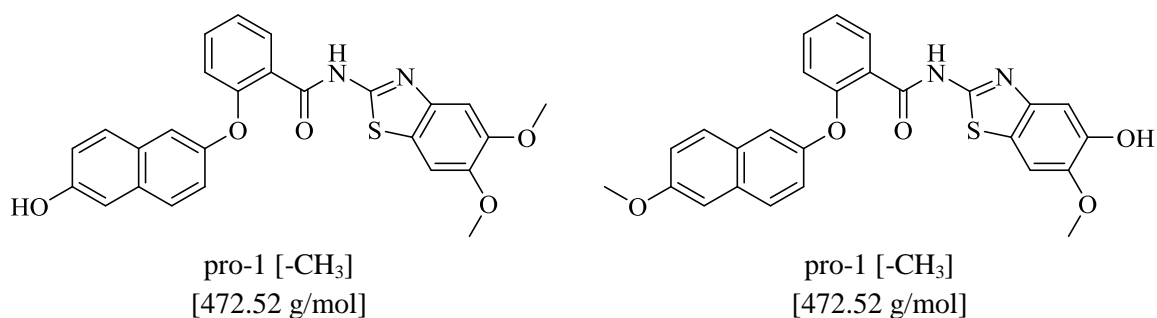


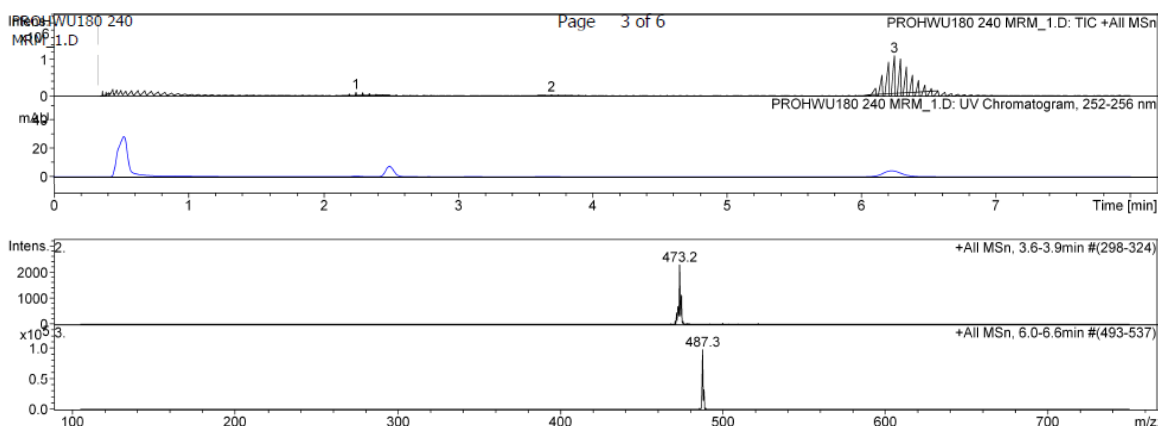
Figure 4.5: LC-MS-spectrum of the metabolite **pro-1** at 240 min.

## Results and discussion

Figure 4.5 represents the end-point of the metabolic study after 240 min. The chromatogram showed that the **pro-1** signal had lost significantly in intensity and two signals at retention times of 2.15 min and 3.8 min occurred. The mass spectrum revealed a new signal at  $m/z$  ratio 473.3, which correlated with the single demethylated **pro-1** derivate. Two signals at different retention times of the chromatogram agreed with one peak in the MS-spectrum if the single demethylation occurred at two different positions as highlighted in Figure 4.6.

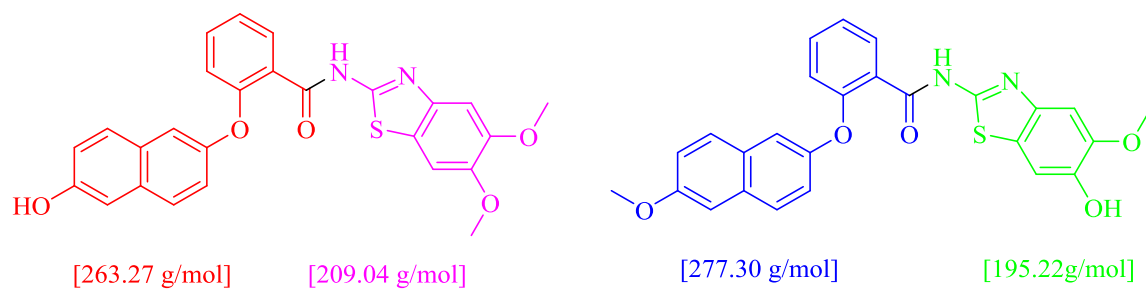


**Figure 4.6: Single demethylation of pro-1 and its structural isomers.**

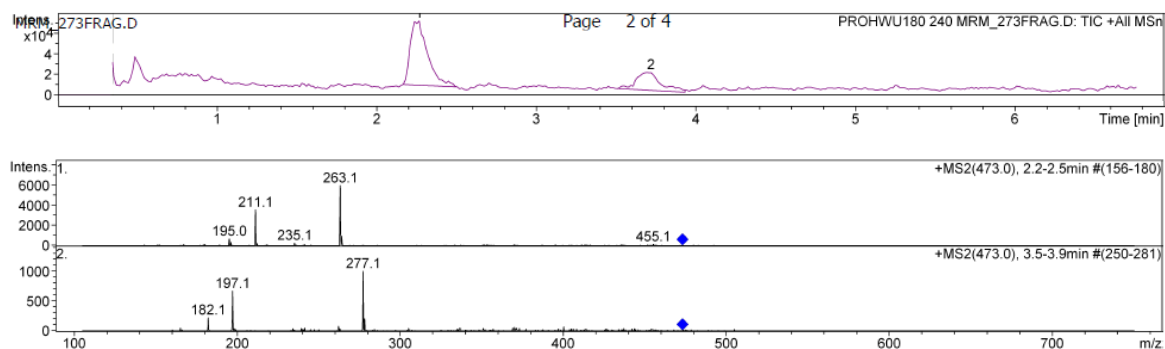


**Figure 4.7: Tandem-MS of pro-1 ( $m/z = 487.3$ ) and the single demethylated metabolite derivate of pro-1 ( $m/z = 473.2$ ).**

The tandem-MS spectrum clarified that the two newly obtained signals in the chromatogram (highlighted in Figure 4.7 with 1 and 2) had exactly the same  $m/z$  ratio of 473.2 and matched with a single demethylated product of **pro-1**.



**Figure 4.8: Structural isomers of the single demethylated pro-1 and their individual fragment sizes.**



**Figure 4.9: Fragmentations of the single demethylated derivate of pro-1.**

Fragmentation of the newly obtained signals in the chromatogram revealed, that all four fragments of the two possible isomers were detectable. The  $m/z$  ratios of 263.1 and 211.1 belonged to the derivate, which had been demethylated at the naphthalene side and 277.1 and 197.1 to the isomer being demethylated at the benzothiazole moiety shown in Figure 4.8. Unfortunately it could not be determined which one of the two methoxy groups on the benzothiazole moiety was demethylated.



## 4.1.2.2. Metabolic study of pro-16

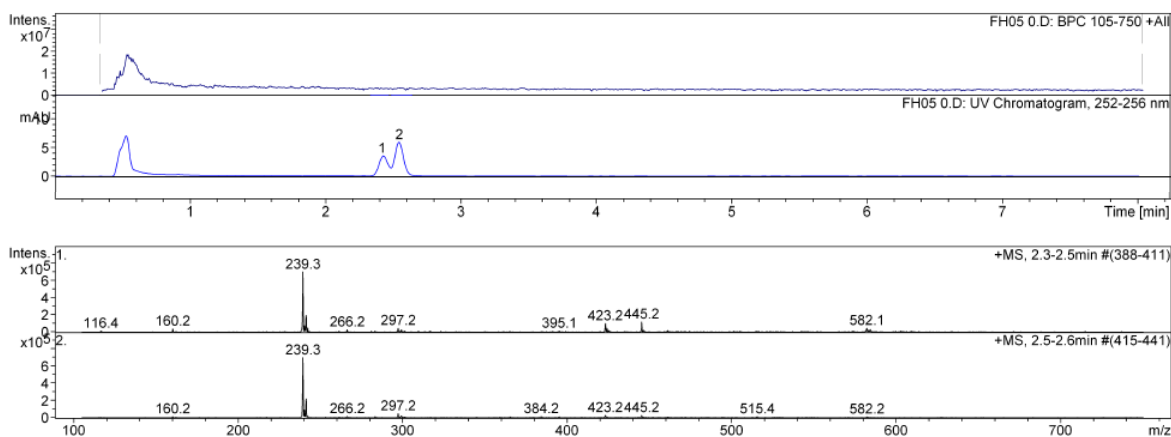


Figure 4.10: LC-MS-spectrum of the metabolite pro-16 at 0 min.

Figure 4.10 demonstrates the starting point of the metabolic study of **pro-16**. The signals of naphthalene and **pro-16** overlap. The signal with the smaller retention time could be identified as the **pro-16** and the other signal representing the internal standard naphthalene. The  $m/z$  ratio of 423.2 in the MS-spectrum confirmed this assumption.

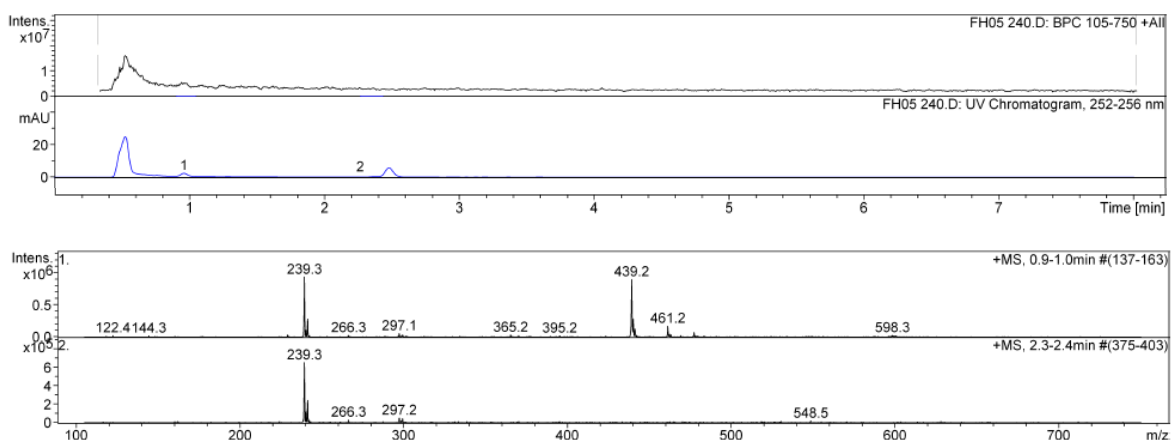
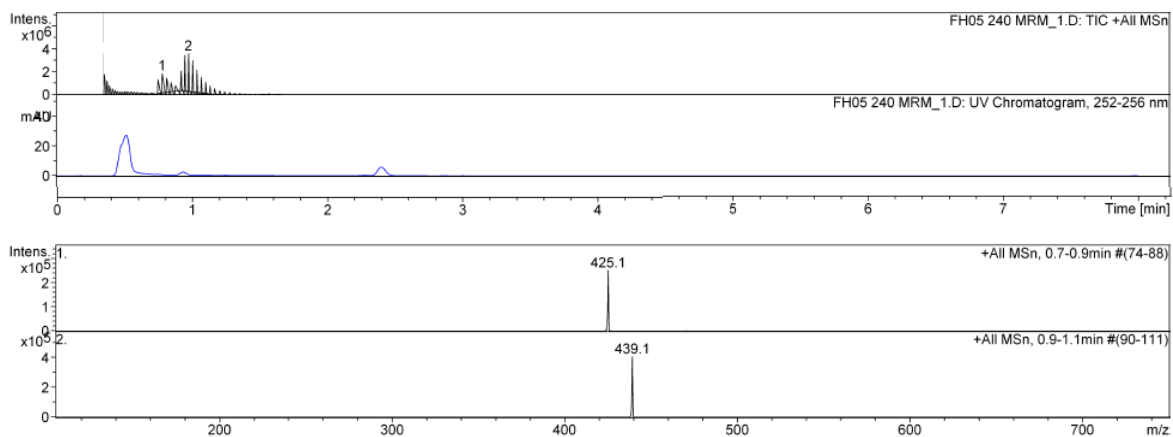


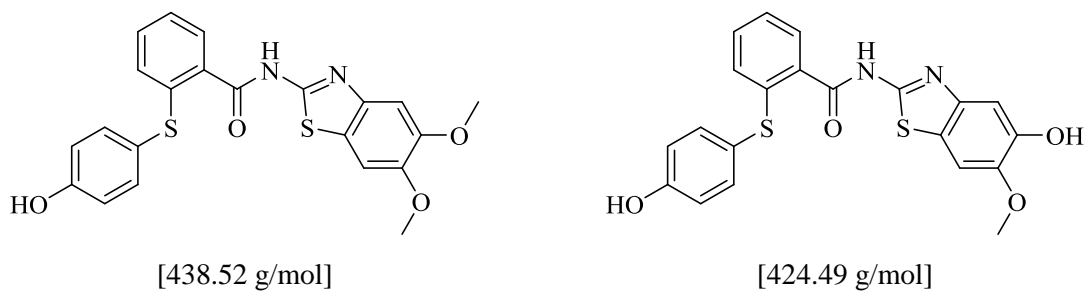
Figure 4.11: LC-MS-spectrum of the metabolite pro-17 at 240 min.

The chromatogram after 240 min (shown in Figure 4.11) pointed out that the **pro-16** signal had completely disappeared and a new signal with a shorter retention time was detected. In the MS spectra a peak at  $m/z$  ratio of 439.2 was identified. The increase in  $m/z$  ratio could be explained by the introduction of an aromatic hydroxyl group.



**Figure 4.12: Tandem-MS of pro-16 metabolite.**

The tandem-MS revealed two peaks in the MS-spectra, 425.1 and 439.1, indicating a combination of aromatic hydroxylation and demethylation at the benzothiazole side (see Figure 4.13).



**Figure 4.13: Metabolic products of pro-16, aromatic hydroxylated derivate (left) and the single demethylated, aromatic hydroxylated compound (right).**

Equally to the metabolite of pro-1 it could not be determined which one of the two methoxy groups, on the benzothiazole moiety, was demethylated (see 4.1.2.1).

### 4.1.3. Structural investigations of the DENV 2 NS2B-NS3 protease

#### 4.1.3.1. NMR spectroscopy of the linked and unlinked DENV 2 NS2B-NS3 protease

The NMR spectroscopy experiments were carried out in order to visualise the protein-ligand interactions of the DENV 2 NS2B-NS3 protease and compound **1**. Applying the procedure described in chapter 3.3.4 the protease was isotopically  $^{15}\text{N}$ -labeled and via NMR spectroscopy examined as described in chapter 3.4.4.

At first, the DENV 2 NS2B-NS3 linked protease construct was examined using NMR spectroscopy. The obtained NMR spectrum is shown in the appendix, Figure 6.1. The NMR spectrum revealed that the DENV 2 NS2B-NS3 protease was successfully  $^{15}\text{N}$ -labeled and that the addition of the inhibitor **1** caused a significant number of signals to shift. However, signal overlies most likely caused by the glycine linker made an assignment impossible and demonstrated the necessity to use the DENV NS2B-NS3 unlinked construct instead.

The unlinked construct was made via Gibson assembly method described in chapter 3.3.1 and sustained by following general expression and purification protocol elucidated in chapter 3.3.4. Enzymatic kinetics of the DENV 2 NS2B-NS3 unlinked protease was determined as described in chapter 3.4.1 and compared with the linked construct.

**Table 4.3: Experimentally determined enzymatic kinetics of the DENV2 NS2B-NS3 linked and unlinked protease in comparison with the literature values (Kim *et al.*, 2013).**

DENV 2 construct	Experimental				References		
	$K_M$ [ $\mu\text{M}$ ]	$k_{cat}$ [ $\text{s}^{-1}$ ]	$k_{cat}/K_M$ [ $\text{s}^{-1}\text{M}^{-1}$ ]	$IC_{50}$ (1) [ $\mu\text{M}$ ]	$K_M$ [ $\mu\text{M}$ ]	$k_{cat}$ [ $\text{s}^{-1}$ ]	$k_{cat}/K_M$ [ $\text{s}^{-1}\text{M}^{-1}$ ]
Linked	74.8	0.30	4051.5	$4.2 \pm 0.4$	31.5	0.19	5970
Unlinked	88.0	0.47	2681.7	$5.6 \pm 0.5$	53.7	0.18	330

The enzyme kinetic data represented in Table 4.3 of DENV 2 NS2B-NS3 unlinked and linked proteases proved that the two proteases exhibited similar catalytic activities determined via Boc-GRR-AMC hydrolysis. This is consisted with the results published by Kim *et al.*, 2013 tracing Bz-nKRR-AMC cleavage under similar fluorometric enzyme assay conditions.

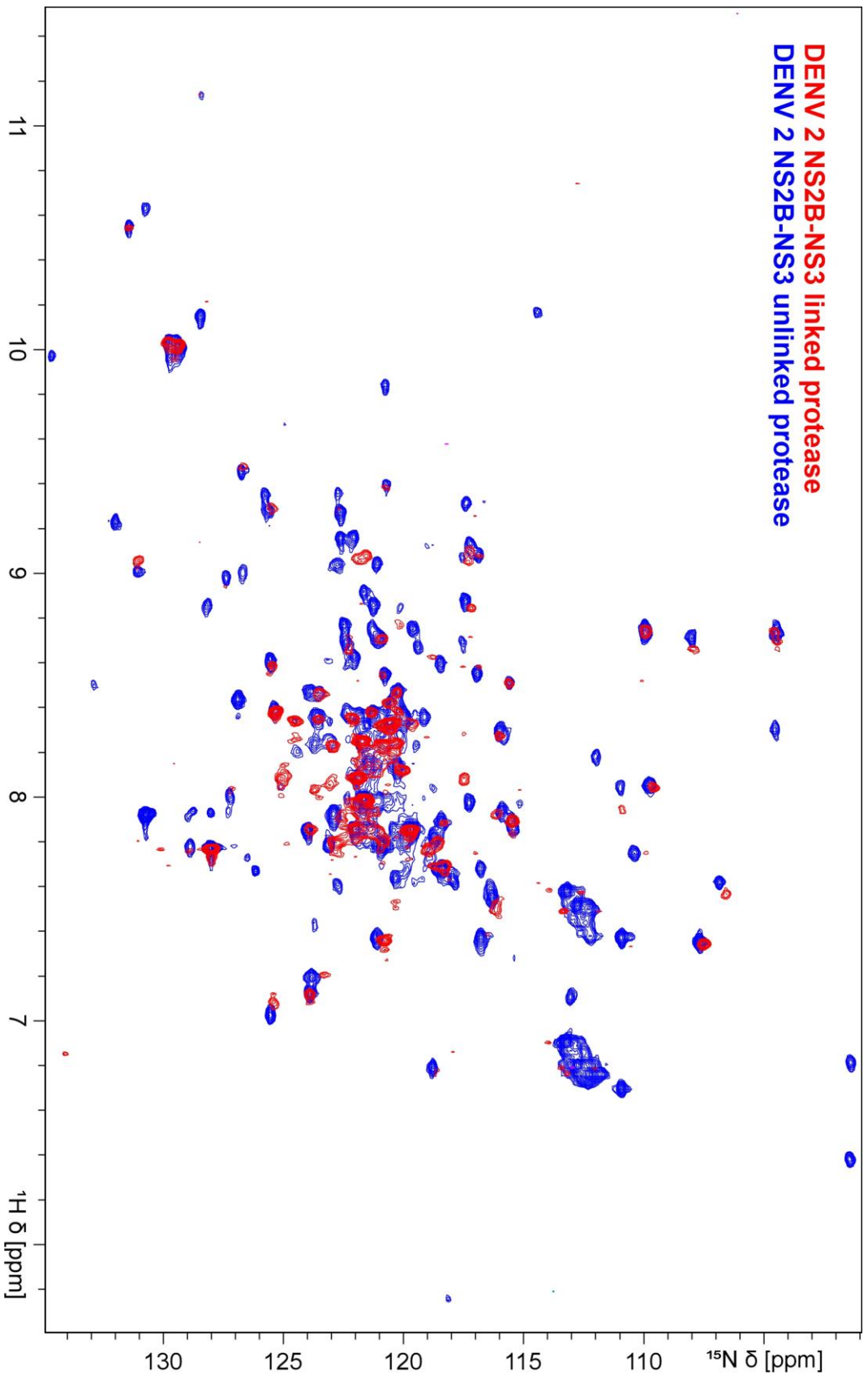


Figure 4.14: NMR spectrum overlay of the DENV 2 NS2B-NS3 linked protease in red and the DENV 2 NS2B NS3 unlinked protease in blue.

Figure 4.14 illustrates that the unlinked protease was successfully isotopically  $^{15}\text{N}$ -labeled and that the NMR-spectra of the linked and unlinked proteases were similar but differing in some regions. There were more signals for the unlinked construct detected and the signals were dispersed, both being favourable for further investigations using NMR spectroscopy. NMR peak assignment of the obtained spectra in dependence on the published work of Kim *et al.*, 2013 was not possible.

#### 4.1.3.2. Protein crystallisation of DENV 2 NS2B-NS3 protease

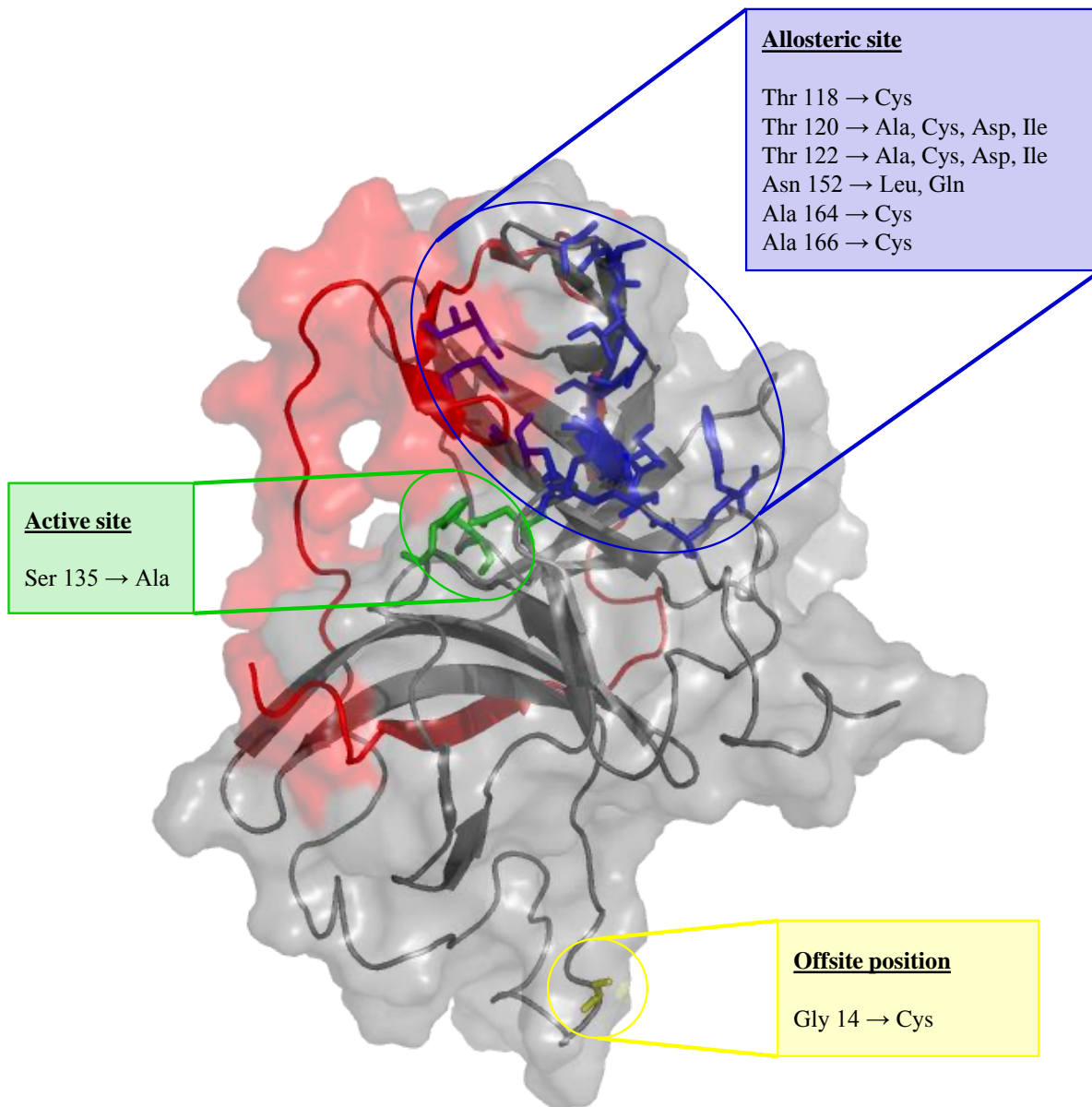
For crystallisation experiments, using the equipment described in chapter 3.4.5, the DENV 2 NS2B-NS3 protease including the inhibitor **1** was sent to the working group of [REDACTED].

Crystals of the DENV 2 NS2B-NS3 protease could be observed by [REDACTED]. These crystals were examined via X-ray and did not include the allosteric inhibitor **1**. Empty obtained crystals were used for soaking methods, but no crystals containing the **1** could be obtained so far.

#### 4.1.3.3. Site directed mutagenesis DENV 2 NS2B-NS3 protease

As there were no crystal structures of DENV NS2B-NS3 protease co-crystallised with an allosteric inhibitor available, molecular binding of **1** to the allosteric pocket was uncertain. Molecular docking of **1** into the proposed allosteric pocket gave an impression of possible binding modes which were investigated using site directed mutagenesis illustrated in Figure 4.15.

The point mutations were introduced via quick change PCR as described in chapter 3.3.3 using primers listed in Table 3.19. The sequence-checked plasmids were expressed and purified after the standard procedure elucidated in chapter 3.3.4. The obtained DENV 2 NS2B-NS3 proteases were examined using the fluorometric enzyme assay exemplified in chapter 3.4.1. Table 4.4 summarises the determined enzymatic data of the different DENV 2 NS2B-NS3 protease mutants.

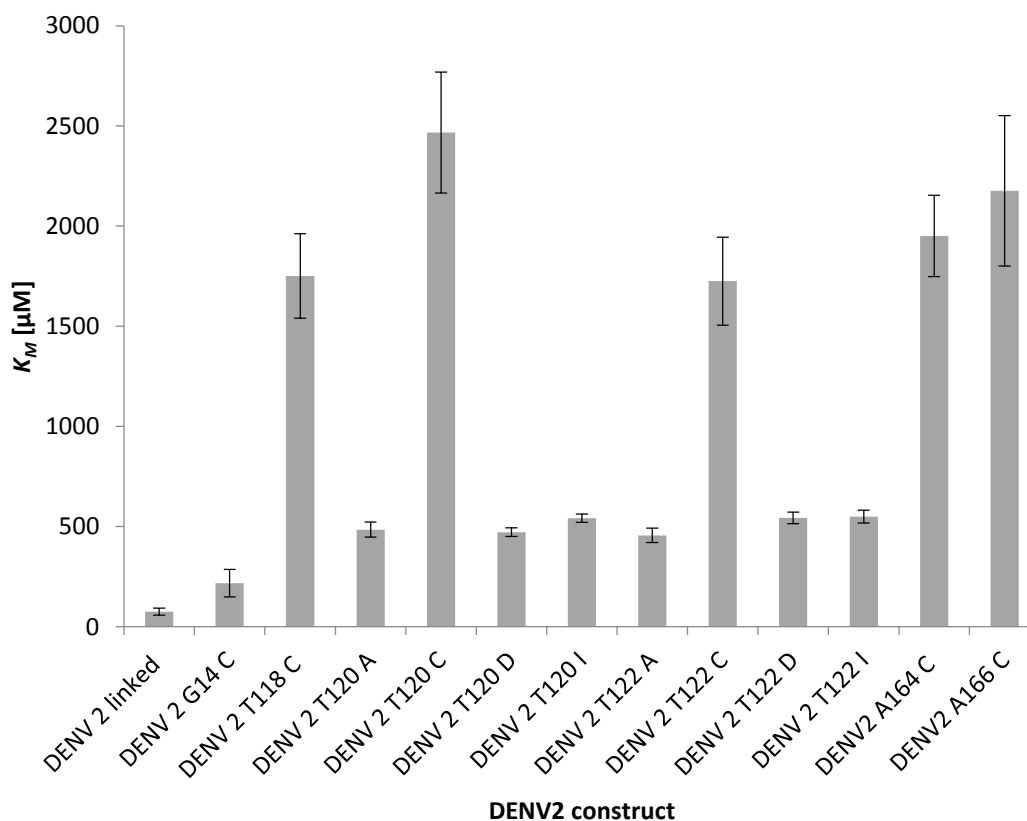


**Figure 4.15: DENV 2 NS2B-NS3 protease (PDB code 2M9P) and the location of the introduced point mutations. NS2B co-factor is highlighted in red and the NS3 domain coloured in grey. Active site (Ser 135) is presented in green, allosteric site mutations in blue and the offsite position (Gly 14) in yellow.**

Table 4.4: Kinetic data of the single mutated DENV 2 NS2B-NS3 proteases determined fluorometrically via hydrolysis of Boc-GRR-AMC.

DENV2 constructs	Enzyme conc. [nM]	$K_M$ [ $\mu\text{M}$ ]	SE $K_M$	$v_{\text{max}}$ [RFU $\text{s}^{-1}$ ]	SE $v_{\text{max}}$	$v_{\text{max}}$ [ $\mu\text{M s}^{-1}$ ]	SE $v_{\text{max}}$	$k_{\text{cat}}$ [ $\text{s}^{-1}$ ]	SE $k_{\text{cat}}$	$k_{\text{cat}}/K_M$ [ $\text{M}^{-1} \text{s}^{-1}$ ]	Error $k_{\text{cat}}/K_M$ [ $\text{M}^{-1} \text{s}^{-1}$ ]
Unlinked	143.0	88	19.1	115.5	8.7	$3.4 \times 10^{-2}$	$2.5 \times 10^{-3}$	0.236	$1.3 \times 10^{-4}$	2681.7	127.8
Linked	50.0	75	17.5	52	6.2	$1.5 \times 10^{-2}$	$1.81 \times 10^{-3}$	0.304	$3.62 \times 10^{-5}$	4062.3	222.8
G14C	14.3	218.1	68.7	2.4	0.1	$7.0 \times 10^{-4}$	$2.9 \times 10^{-5}$	0.049	$2.0 \times 10^{-6}$	224.8	70.8
S135A	240.1	/	/	/	/	/	/	/	/	/	/
T118C	14.3	1751.3	210.3	20.7	1.8	$6.0 \times 10^{-3}$	$5.3 \times 10^{-4}$	0.422	$3.7 \times 10^{-5}$	240.9	28.9
T120A	6.7	484.9	37.4	5.3	0.2	$1.5 \times 10^{-3}$	$5.8 \times 10^{-5}$	0.232	$8.7 \times 10^{-6}$	478.1	36.9
T120C	14.3	2466.5	301.9	60.8	5.7	$1.8 \times 10^{-2}$	$1.7 \times 10^{-3}$	1.239	$1.2 \times 10^{-4}$	502.5	61.5
T120D	1.6	472.2	22.3	4.8	0.1	$1.4 \times 10^{-3}$	$2.9 \times 10^{-5}$	0.905	$1.9 \times 10^{-5}$	1916.1	90.5
T120I	9.5	541.6	21.4	12.4	0.2	$3.6 \times 10^{-3}$	$5.8 \times 10^{-5}$	0.383	$6.2 \times 10^{-6}$	706.4	27.9
T122A	11.2	455.9	36.3	7.4	0.3	$2.2 \times 10^{-3}$	$8.8 \times 10^{-5}$	0.192	$7.8 \times 10^{-6}$	421.9	33.6
T122C	23.3	1725.0	219.2	39.1	3.5	$1.1 \times 10^{-2}$	$1.0 \times 10^{-3}$	0.490	$4.4 \times 10^{-5}$	283.8	36.1
T122D	12.8	543.3	28.7	10.3	0.3	$3.0 \times 10^{-3}$	$8.8 \times 10^{-5}$	0.235	$6.9 \times 10^{-6}$	433.3	22.9
T122I	17.0	549.5	31.4	5.8	0.3	$1.7 \times 10^{-3}$	$8.8 \times 10^{-5}$	0.100	$5.2 \times 10^{-6}$	181.4	10.4
N152L	155.2	/	/	/	/	/	/	/	/	/	/
N152Q	180.0	/	/	/	/	/	/	/	/	/	/
A164C	81.0	1950.4	202.8	34.4	2.6	$1.0 \times 10^{-2}$	$7.6 \times 10^{-4}$	0.124	$9.4 \times 10^{-6}$	63.6	6.6
A166C	17.7	2176.8	375.3	30.2	3.9	$8.8 \times 10^{-3}$	$1.1 \times 10^{-3}$	0.498	$6.4 \times 10^{-5}$	228.6	39.4

As no free AMC could be detected with the S135A mutant, the assay showed that the S135A mutation leads to an inactive protease. Equally, the mutations of the highly conserved asparagine 152 into leucine or glutamine, respectively, led to a complete loss of enzyme activity.



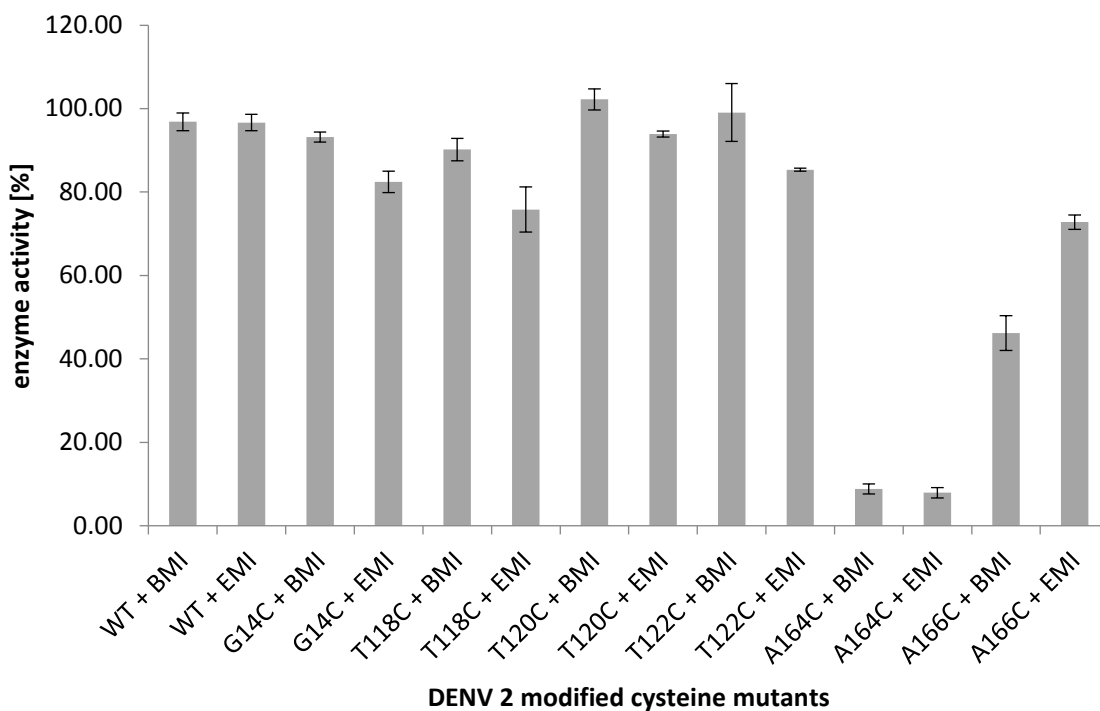
**Figure 4.16:** Bar chart representing the  $K_M$  values of DENV 2 NS2B-NS3 linked protease and its single mutants.

All point mutations of the DENV 2 linked protease had an effect on the substrate-enzyme affinity represented by the  $K_M$  values shown in Figure 4.16. G14C had the smallest change in  $K_M$  value compared with the DENV 2 linked protease. The newly introduced cysteine for G14 is far away from the active site and differs only in the thiol group. Hence, this mutation has a feasible small impact on the substrate affinity. Comparing the  $K_M$  value of the DENV 2 linked protease with the mutants introduced in the allosteric site it was clearly seen that the  $K_M$  values increased immensely depending on the location of the exchanged amino acid. All allosterically introduced cysteine mutants proved an increase of the  $K_M$  value by 20 - 30 -fold. The mutations of either threonine or alanine respectively to cysteine were significant. Threonine has a hydroxyl group and is a more polar amino acid. Alanine is the smallest amino acid with a hydrophobic side chain and differs in size and charge immensely respectively to cysteine. The other mutations revealed an increase of the  $K_M$  values of around 7 -fold indicating a less pronounced effect on the substrate affinity.



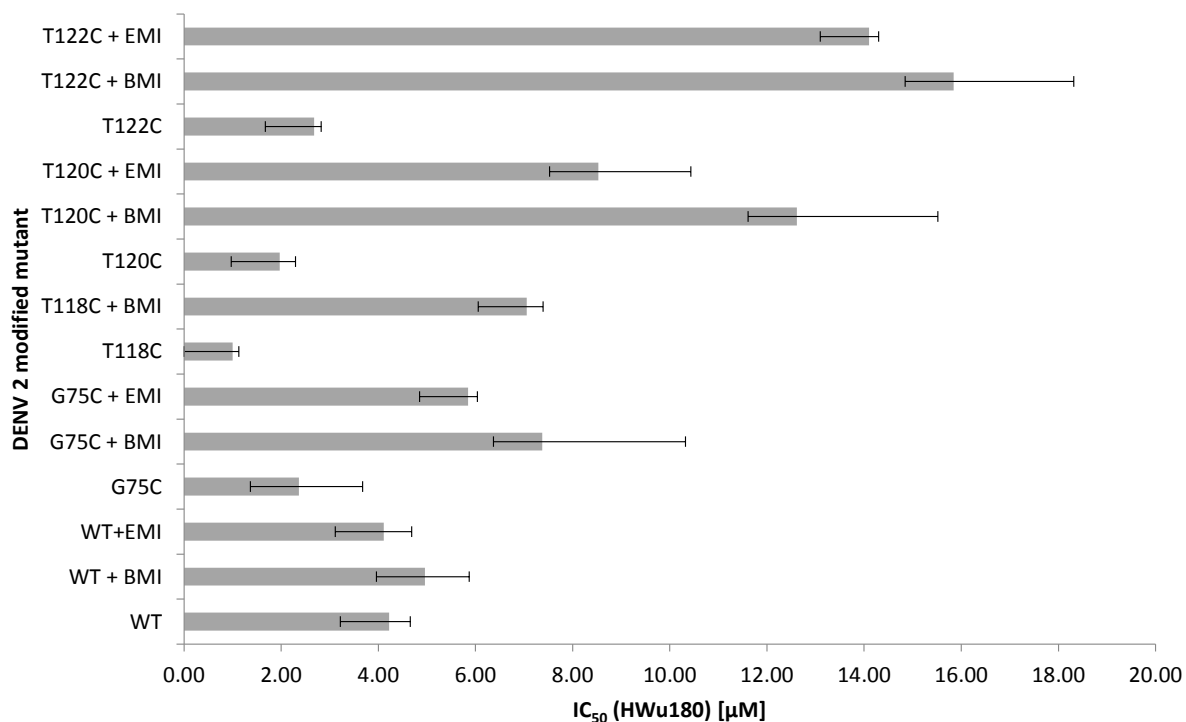
## Results and discussion

The newly introduced cysteine residues were irreversibly modified with *N*-benzylmaleimide (BMI) and *N*-ethylmaleimide (EMI), respectively. Both maleimide compounds reacted covalently with the cysteine residues in the allosteric site and were enzymatically analysed via the hydrolysis reaction of the AMC substrate as described in chapter 3.4.1.



**Figure 4.17:** Graphical representation of the enzyme activities of the DENV 2 cysteine mutants irreversibly modified via *N*-benzylmaleimide (BMI) and *N*-ethylmaleimide (EMI) respectively.

Figure 4.17 illustrated that the modified mutants A164C + BMI and A164C + BMI had almost completely lost enzymatic activity. Reduced activity was also determined for A166C + BMI, A166C + EMI and T118C + EMI. All other maleimide-modified enzymes showed enzymatic activities higher than 80% compared to the wildtype and were therefore declared to be only little affected by the covalent bonding of a small molecule in the allosteric site and thus could be used for further investigations.



**Figure 4.18:**  $IC_{50}$  values of **1** for inhibition of the different irreversibly modified cysteine residues.

The mutants containing modified cysteine residues were exposed to **1** and the  $IC_{50}$  values were determined. Figure 4.18 brings out that the  $IC_{50}$  values of inhibition of the modified mutant enzymes T122C (+ BMI and + EMI), T120C (+ BMI and + EMI) and T118C + BMI were significantly higher than without the modifications, pointing out that the modifications interrupt the enzyme-ligand binding.

#### 4.1.4. Summary and discussion of the DENV NS2B-NS3 protease results

The similarity search of *N*-(benzothiazol-2-yl)-2-phenoxybenzamide by SciFinder led to a virtual screening of 249 compounds into the proposed allosteric site, resulting in a ranking of highly potent to less promising putative inhibitors. Seven commercially available molecules were purchased from Ambinter and tested fluorometrically against the DENV 2 NS2B-NS3 protease. None of the identified compounds proved any inhibitory effect on the viral enzyme. Conspicuously none of the purchased substances included the 5,6-dihydroxybenzothiazol moiety with the focus lying on the two contiguous hydroxyl groups at the benzothiazole moiety.

Another SAR study approach of synthesising eight compounds by leaving the 5,6-dihydroxyl-2-aminobenzothiazole moiety of the lead structure unchanged and altering the naphthalene side led to the identification of inhibitory compounds. All synthesised molecules revealed to have some action against the viral protease of which **10**, **15**, **16** and **17** proved to have analogue  $IC_{50}$  values compared with the lead structure **1**. This demonstrated that the *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)benzamide moiety is the essential structural feature for DENV 2 NS2B-NS3 protease inhibitors.

Taking both SAR study results into account it can be said that the *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)benzamide moiety is the inhibitory driving force with a high focus lying on the hydroxyl substituted benzothiazole. This finding was consistent with Wu *et al.*, 2015 which reported that the dimethoxy compounds were inactive against the DENV 2 NS2B-NS3 protease in the fluorometric enzyme assay. Enzymatically, this was also confirmed for the synthesised compounds **10** – **17** and their dimethoxy precursors, respectively.

Wu *et al.*, 2015 suggested that the dimethoxy compounds acted as prodrugs and can be demethylated in cells yielding the more potent dihydroxy derivatives. In this work this assumption was investigated for **pro-1** and **pro-16** by a metabolic study using pooled male rat liver microsomes and analysing the metabolites via LC-MS. For **pro-1** a demethylation of one of the three methoxy groups could be detected and hence the assumption could be verified. Most likely, **pro-1** is fully demethylated in cells, however the microsomal testing system used seemed exhausted and only a single methoxy group could be deprotected. In comparison, **pro-16** showed two metabolic products. One metabolite could be identified to be the aromatic hydroxylated derivate and the other the single demethylated product. Equally for **pro-16** it can be assumed that it would be completely demethylated and hydroxylated at the aromatic unit in cells. In summary, the metabolic study verified that both precursor substances were metabolised and therefore probably act as prodrugs.

Given that molecular docking and hence following lead structure optimisation did not have the expected success, validation of the actual binding site of **1** was inevitable. Two approaches, being NMR-spectroscopy and site directed mutagenesis, were consulted.

The successfully <sup>15</sup>N-labeled DENV 2 NS2B-NS3 linked protease was examined via NMR spectroscopy and proved to be unserviceable as the linker caused signal overlay. It could be shown that the addition of inhibitor **1** caused signal shifts and confirmed that NMR spectroscopy is an adequate analysis method to study this protein-ligand interaction. The signal overlay problem could be overcome by using the unlinked construct as described by Kim *et al.*, 2013. The unlinked viral protease, made via Gibson assembly, successfully expressed and purified, was fluorometrically analysed, revealing that the newly made construct showed comparable enzymatic kinetics as described by Kim *et al.*, 2013. It was possible to obtain the <sup>15</sup>N-labeled unlinked enzyme for NMR spectroscopy purposes and a NMR spectrum overlay of the unlinked and linked proteases showed that the unlinked protease signals were more dispersed and hence to a greater extent useful for examination purposes. Unfortunately, signal assignment using the published data (Kim *et al.*, 2013) was not possible and hence a full signal assignment would have to be carried out. Due to high costs and unpredictable time consumption, this was postponed. In addition, the low solubility of **1** in aqueous solutions was in need of improvement for protein-ligand NMR analysis.

A set of point mutations were introduced and kinetically examined. Mutating the serine in the catalytic triad expectedly led to an inactive enzyme and therefore this mutant could be used as a negative control in further investigations. Changing the highly conserved asparagine (N152) (illustrated in Figure 4.15) equally led to a loss of active NS2B-NS3 protease and demonstrated the significance of this amino acid. Introduction of a conservative mutation in the periphery (G14C) had no effect on the enzyme substrate binding and for this reason is ideal for single-molecule spectroscopy as aimed for. All other point mutations were positioned into the allosteric site and had an impact on the enzyme kinetics, emphasising the presence and relevance of this allosteric site. The newly introduced cysteine mutants were irreversibly modified with *N*-benzylmaleimide (BMI) and *N*-ethylmaleimide (EMI), respectively. After inspection of the enzymatic activities, *IC*<sub>50</sub> values of the inhibitor **1** for each modified protease were determined and revealed that binding of the inhibitor to the allosterically modified proteases were significantly disrupted. Hence, it was validated for the first time that binding of the inhibitor **1** takes place in the proposed allosteric site.

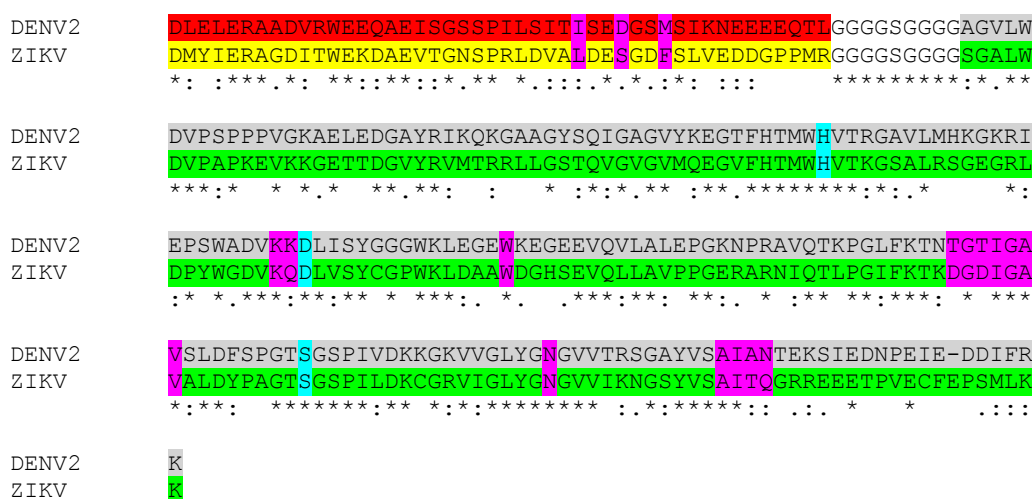
## 4.2. ZIKV NS2B-NS3 protease

During the research on the DENV NS2B-NS3 protease, the world was faced with a dramatic outbreak of the ZIKV. Amongst others, this was the reason for the expansion of the research field. Both, DENV and ZIKV genomes encode for polyproteins containing the viral NS2B-NS3 proteases, which are essential for polyprotein procession and viral replication as in detail described in chapter 1.7.

The ZIKV NS2B-NS3 protease was consulted to study the protein-ligand interactions of a viral NS2B-NS3 protease with respect to the allosteric inhibitor **1**, using the natural given differences and similarities as auxiliary tool. The results are elucidated in this chapter.

### 4.2.1. Homology model and molecular docking of ZIKV NS2B-NS3 protease

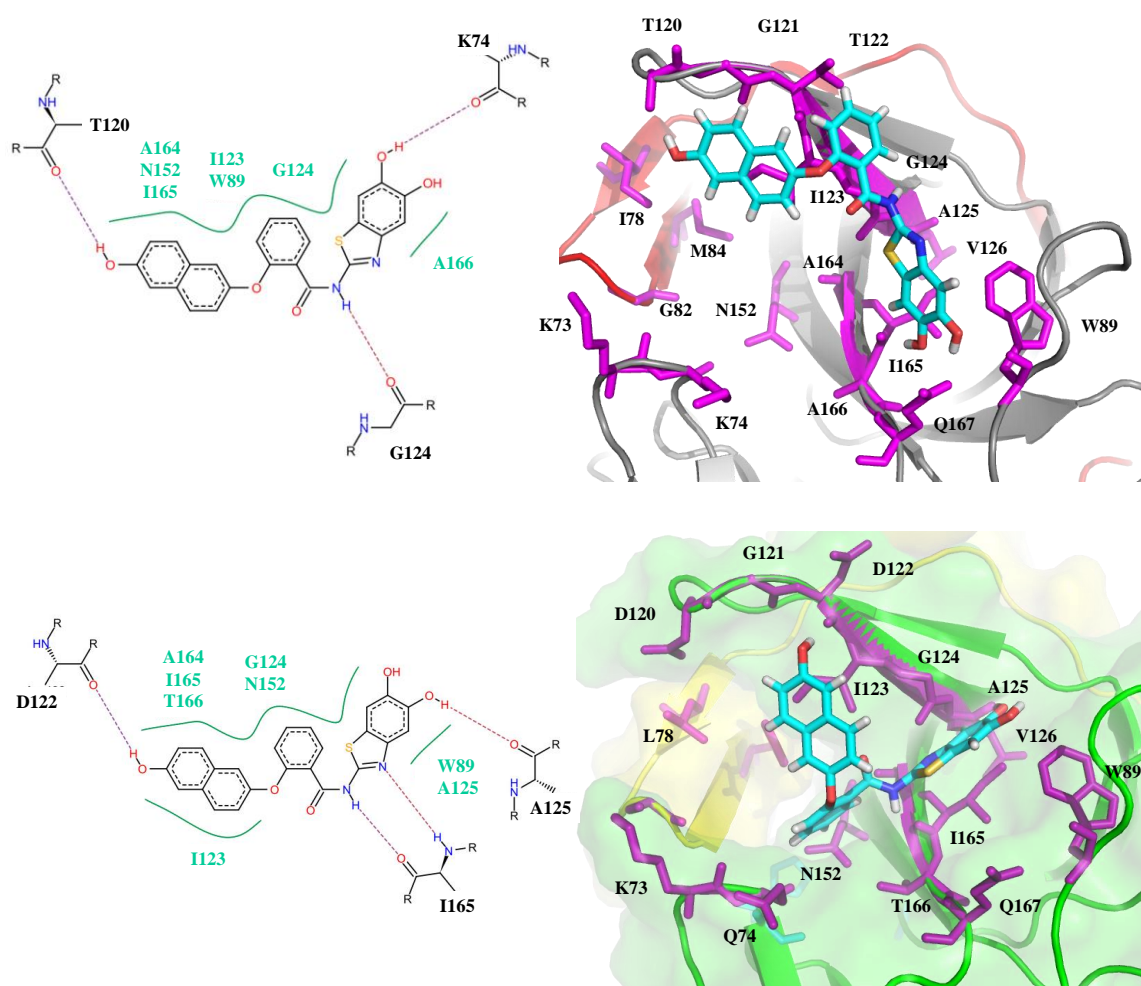
Investigations into the ZIKV NS2B-NS3 protease started before the first crystal structure by Lei *et al.*, 2016 was released. In order to perform molecular docking a homology model was constructed. The homology model of the ZIKV NS2B-NS3 protease was prepared as described in chapter 3.1.3. Identified using MOE PDB search function, the NMR derived DENV 2 NS2B-NS3 protease structure (PDB: 2M9P) was used as structural template. The homology model of the ZIKV NS2B-NS3 protease obtained an RMSD value of 0.880 and the geometry was examined using the Phi-Psi Plot to eliminate outliers (see chapter 6.5). Based on a sequence alignment (Figure 4.19) and the homology model, the DENV analogue allosteric site of the ZIKV NS2B-NS3 protease was identified.



**Figure 4.19:** Sequence alignment of DENV (NS2B in red, NS3 in grey) and ZIKV (NS2B in yellow, NS3 in green) NS2B-NS3 protease constructs highlighting the active (cyan) and allosteric site (magenta). \* (asterisk) stands for a fully conserved residue, : (colon) and . (period) indicate conservation between groups of strongly, respectively weakly similar properties.

Observations of the amino acids contributing to the allosteric site showed 56% identity and 89% similarity. Only minor conservative changes were seen in the NS2B co-factor region, however significant differences were determined for the threonine residues 120 and 122 in the NS3 protease domain of DENV. In ZIKV NS2B-NS3 protease, these residues are substituted to aspartic acids, which vary significantly in size and charge compared to their DENV NS2B-NS3 counterparts. Threonine has a polar but uncharged side chain and in comparison, the aspartic acid side chain is negatively charged. Molecular docking studies were performed to investigate these differences.

Amongst others, molecular docking applying the homology model was carried out on the allosteric DENV NS2B-NS3 protease inhibitor **1** investigated in chapter 4.1.



**Figure 4.20:** Comparison of the molecular docking poses of **1** in the allosteric site of the DENV (above) and ZIKV (below) NS2B-NS3 protease. For DENV the co-factor is coloured in red and the NS3 domain in grey. The ZIKV NS2B region is highlighted in yellow and the NS3 in green. Amino acids contributing to the allosteric site are visualised in magenta and **1** is represented in cyan.

Comparing the docking poses obtained for the two viral NS2B-NS3 proteases shown in Figure 4.20, it was noticeable that in both cases the ligand followed the same orientations. Both poses indicate that the hydroxyl group at the naphthalene moiety were facing and interacting with the amino acids 120 and 122 respectively. A threonine in DENV (Figure 4.20, above) and aspartic acid in ZIKV (Figure 4.20, below) were found in these positions and traced back to backbone interactions. In both cases, the benzothiazole part reached out into the direction of isoleucine 165. Disparities were noticeable for the benzamide moiety. The ligand **1** orientated in the ZIKV protease binding pocket inwards (Figure 4.20, below), towards N152 and in DENV outwards (Figure 4.20, above), towards residue asparagine 122.

The structural similarity of the predicted allosteric site of the ZIKV NS2B-NS3 protease according with the sustained docking poses led to the hypothesis that the ZIKV NS2B-NS3 protease would also be inhibited by **1** in a non-competitive manner (see chapter 4.2.2). Site-directed mutagenesis and enzymatic analysis could reveal the disparity between threonine and aspartic acid in position 120 and 122 (see chapter 4.2.5.3).

### **4.2.2. Protein preparation and autocatalytic cleavage of ZIKV NS2B-NS3 protease**

During expression and purification of the wild type (WT) ZIKV NS2B-NS3 linked protease protein cleavage as seen in Figure 4.21 (lane 3) occurred. The cleaved fragments were in size equal to that of the unlinked protease and the cleaved protease proved to be enzymatically active as shown in Table 4.5. In order to eliminate external parameters that could cause the protein cleavage, a point mutation at the active site serine (Ser135) was introduced. As demonstrated in Figure 4.21 (lane 4) the S135A mutant was not cleaved, hence, external factors causing the decomposition were excluded and conclusively, the protein cleavage must occur autocatalytically.

Mass spectrometry (MS) was consulted in order to determine the exact cleaved fragments. The MS results shown in chapter 6.6 indicated that the cleavage occurred in the region of 7.26 kDa. Taking into account that the protease prefers basic amino acids in P1, the cleavage was identified to take place just after arginine 95 in the NS2B co-factor region (see Figure 6.6).

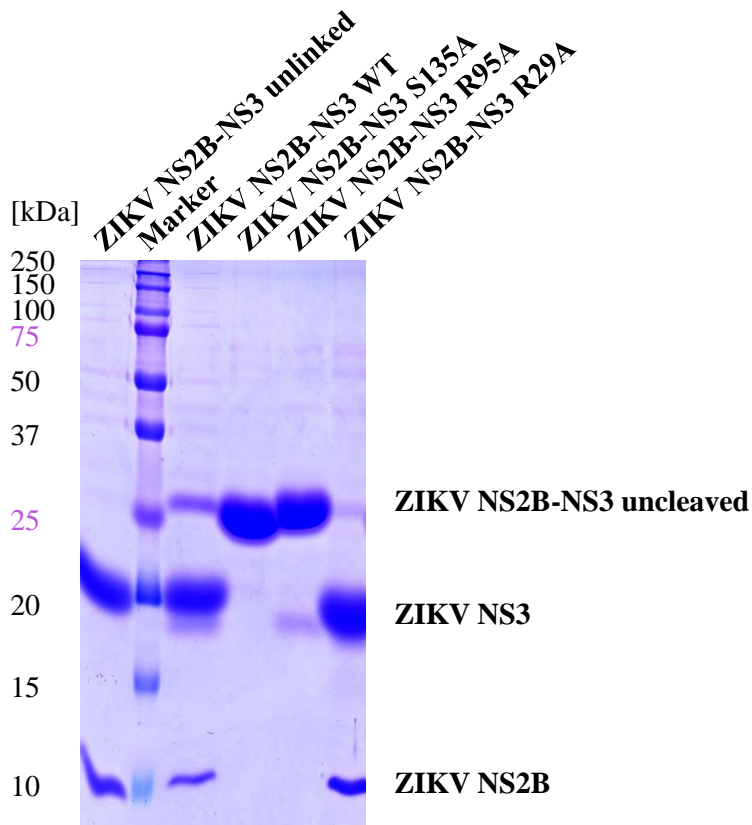


Figure 4.21: SDS-Page of the unlinked (lane 1) and linked (lane 3) ZIKV NS2B-NS3 constructs. The linked (lane 3) protease is autocatalytically cleaved. Inactive linked S135A enzyme (lane 4) and linked R95A point mutated ZIKV NS2B-NS3 protease (lane 5) proved not to be cleaved. Like the wild type (WT), the mutation R29A was equally cleaved.

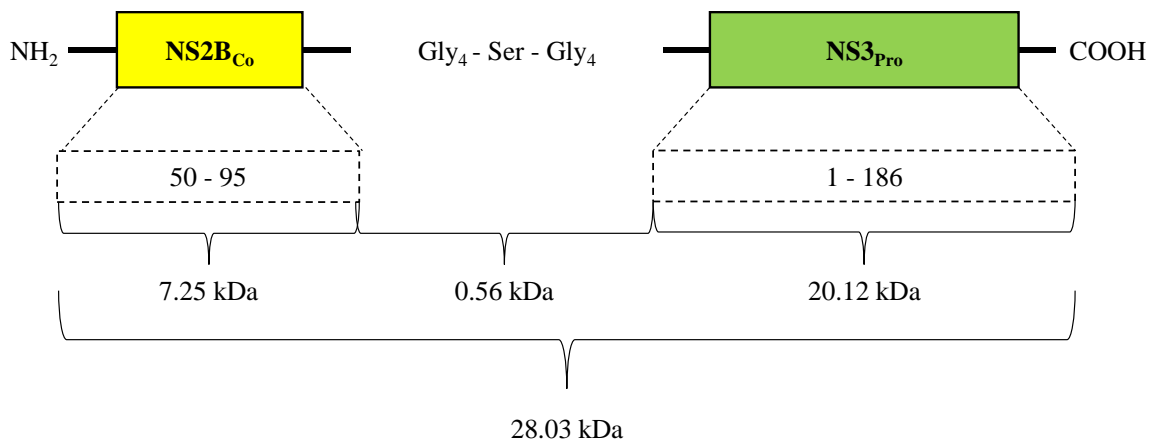


Figure 4.22: Illustration of the ZIKV NS2B-NS3 linked protease construct highlighting the molecular weights of the separate fragments in kDa.



## Results and discussion

Figure 4.21 (lane 5) displays that the single mutation R95A of ZIKV NS2B-NS3 protease prohibits the autocatalytic cleavage and leads to an enzymatically active enzyme as described in Table 4.5.

There were two possible variants of a cleavage mechanism, either the protein cleaves itself resulting in a so-called *cis*-cleavage or each macromolecule is cleaved by other proteases following a *trans*-cleavage mechanisms. Investigations of the phenomenon led to the development of the cleavage site study described in chapter 3.4.6. Incubation of the S135A mutant with the WT linked protease showed that the S135A mutant was not cleaved by the WT linked protease. Conclusively the WT linked protease followed a *cis*-cleavage mechanism.

Furthermore, in literature a double mutation of the amino acids (NS2B-) R95 and (NS3-) R29 is commonly used to bypass autocatalytic cleavage problems (Lei *et al.*, 2016). In the present work, it was shown that only the R95A mutation was necessary to prohibit cleavage, as R29A had no influence on the autocatalytic cleavage as shown in Figure 4.21 (lane 6).

The expressed and purified ZIKV NS2B-NS3 proteases were examined using the fluorometric enzyme assay described in chapter 3.4.1. Table 4.5 represents the enzymatic kinetics of the ZIKV NS2B-NS3 unlinked and linked constructs and the single-point mutated ZIKV linked proteases S135A, R95A and R29A.

**Table 4.5: Enzymatic kinetics of unlinked and linked ZIKV NS2B-NS3 constructs and the linked point mutated S135A, R95A and R29A proteases.**

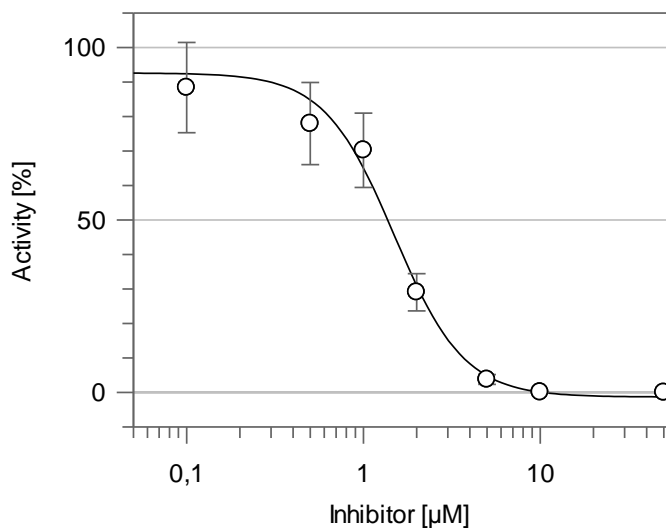
ZIKV construct	$K_M$ [ $\mu\text{M}$ ]	$k_{cat}$ [ $\text{s}^{-1}$ ]	$k_{cat}/K_M$ [ $\text{s}^{-1}\text{M}^{-1}$ ]
Unlinked	$584.3 \pm 133.4$	$0.20 \pm 2.3 \times 10^{-02}$	$339.6 \pm 86.7$
WT linked	$557.7 \pm 61.5$	$0.03 \pm 1.4 \times 10^{-03}$	$47.2 \pm 5.8$
S135A	na	na	na
R95A	$367.5 \pm 32.5$	$0.06 \pm 2.2 \times 10^{-03}$	$162.4 \pm 15.6$
R29A	$539.5 \pm 61.8$	$0.20 \pm 1.1 \times 10^{-02}$	$365.2 \pm 46.5$

na, not active

The kinetic data of the different enzymes represented in Table 4.5 pointed out that the substrate affinities reflected by the  $K_M$  values for the unlinked constructs (unlinked, WT linked and R29A) were almost equal, indicating that the mutation had no influence on the substrate binding. The substrate affinity of the linked R95A protease was smaller and hence proved an even higher affinity towards the same substrate in comparison with the unlinked proteases. The substrate turn-over rates however, suggested that the unlinked protease cleaved the substrate twice as quickly as the linked protease.

### 4.2.3. Inhibition of the ZIKV NS2B-NS3 protease

According to enzymatic analysis of the DENV NS2B-NS3 protease, investigations into the ZIKV NS2B-NS3 protease were carried out using the linked R95A mutant. The fluorometric enzymatic assay and the  $IC_{50}$  value determination of **1** were performed as described in chapter 3.4.1.



**Figure 4.23:**  $IC_{50}$  determination by fitting ZIKV NS2B-NS3 linked (R95A) protease activity [%] against the concentration of inhibitor **1** [ $\mu\text{M}$ ].

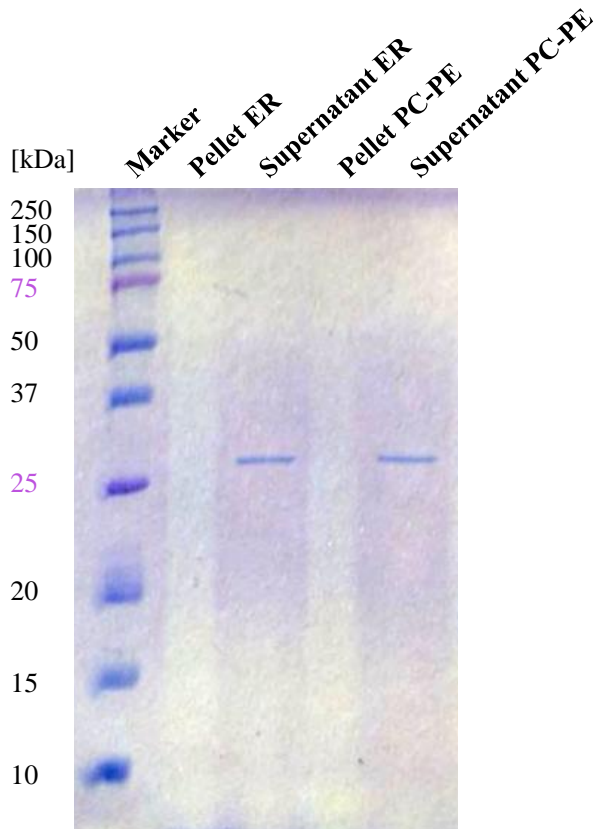
The experimentally determined (Figure 4.23)  $IC_{50}$  value of **1** on the ZIKV NS2B-NS3 linked (R95A) protease was calculated to  $1.5 \pm 0.2 \mu\text{M}$ . This result was in accordance with the values for the DENV NS2B-NS3 protease published by Wu *et al.*, 2015 elucidated in chapter 4.1.

Validation of the fluorometrically determined enzymatic data was obtained by microscale thermophoresis (MST). The MST experiments were carried out as described in chapter 3.4.2. The  $K_d$  value for the ZIKV NS2B-NS3 linked R95A protease with **1** was  $1.7 \mu\text{M}$  (see appendix 6.7.1). As the  $IC_{50}$  and  $K_d$  values were highly similar, both analytical methods were concluded to be reliable.

The enzymatically inactive ZIKV NS2B-NS3 linked (S135A) protease was labeled and analysed by means of MST in the presence of **1**. The experimental MST data is found in chapter 6.7 and the  $K_d$  value was determined to be  $1.4 \mu\text{M}$ . This highly similar result indicated that the S135A protease is folded correctly, including the fold of the allosteric site.

#### 4.2.4. Protein-lipid interaction

The viral NS2B-NS3 protein is a membrane bound protein (see chapter 1.7). Therefore investigations of protein-lipid interactions of the artificial linked protease construct were of interest. Examinations were carried out by the use of the lipid-sedimentation assay described in chapter 3.4.7.



**Figure 4.24: SDS-page of the lipid sedimentation assay for ZIKV NS2B-NS3 linked R95A protease. ER lipid composition mimics the rough endoplasmic reticulum and PC-PE is the positively charged lipid mixture composition.**

The pellet lanes (Figure 4.24, lane 2 and 4) showed no band and hence, illustrated the absence of protein. All protein was found in the supernatant (Figure 4.24, lane 3 and 5) and revealed that there was no protein-lipid interaction between the ZIKV NS2B-NS3 linked (R95A) protease and the two different lipid compositions.

## **4.2.5. Structural investigations of the ZIKV NS2B-NS3 protease**

### **4.2.5.1. NMR spectroscopy**

The NMR spectroscopy experiments were carried out in order to determine the structure of the ZIKV NS2B-NS3 protease and additionally examine their interactions between the two proteases and **1**.

Applying the general procedure (described in chapter 3.4.4) the ZIKV NS2B-NS3 linked and unlinked proteases were successfully isotopically <sup>15</sup>N-labeled. The obtained NMR spectra showed that the linked and unlinked ZIKV NS2B-NS3 proteases were both folded. The glycine linker proved to have an influence on the dynamic, which was seen in the heterogenic line width of the ZIKV NS2B-NS3 linked protease spectrum. The obtained NMR spectra showed that **1** binds to the ZIKV NS2B-NS3 protease.

### **4.2.5.2. Crystallisation**

The manual method (see chapter 3.4.5) of crystallisation experiments for the ZIKV NS2B-NS3 protease in absence and presence of **1** were carried out at the University of Mainz. The experiments did not lead to any protein crystals.

More promising was the crystallisation done at MarXtel (see chapter 3.4.5), which resulted in crystal growth (see appendix 6.8, Figure 6.7). Isolation of the small crystals was impossible, for which reason X-ray examination was not feasible.

### **4.2.5.3. Site directed mutagenesis**

As highlighted in chapter 4.2.1 the DENV 2 and ZIKV NS2B-NS3 proteases contain significant differences in the amino acids 122 and 124. DENV 2 exhibits at these positions two threonines, whereas ZIKV two aspartic acids. Site directed mutagenesis was carried out (as described in chapter 3.3.3) and the mutated proteases were enzymatically examined (as described in chapter 3.4.1).

## Results and discussion

**Table 4.6: Enzymatic kinetics of the point-mutated ZIKV NS2B-NS3 linked proteases and the  $IC_{50}$  values of **1**, respectively. Only the R95A mutated derivate were not cleaved autocatalytically.**

<b>ZIKV linked construct</b>	<b><math>K_M</math> [<math>\mu</math>M]</b>	<b><math>v_{max}</math> [RFU/s]</b>	<b><math>IC_{50}</math> (<b>1</b>) [<math>\mu</math>M]</b>
D122A	280.9	13.1	$2.6 \pm 0.6$
D122T	227.2	10.2	$3.5 \pm 0.2$
R95A, D122T	185.9	5.2	$4.3 \pm 0.4$
D124A	236.4	7.5	$3.3 \pm 0.5$
D124T	240.8	6.7	$3.6 \pm 0.4$
R95A, D124T	150.7	15.2	$4.5 \pm 0.5$

The results listed in Table 4.6 coherently indicated that the affinity of the substrate towards the linked (R95A) protease is higher than that of the unlinked. Conversely, the substrate turn-over rate of the unlinked was higher than that of the linked (R95A), being consistent with the enzymatic results determined in chapter 4.2.2. The obtained  $IC_{50}$  values of **1** with respect to the different mutated enzymes were almost equal. Even the introduction of the small alanine did not have any significant effect on the protein-ligand interaction.

#### 4.2.6. Summary and discussion of the ZIKV NS2B-NS3 protease results

The established homology model suggested that the allosteric site was likewise present in the ZIKV NS2B-NS3 protease and enabled the identification of the amino acids contributing to the predicted allosteric binding pocket. The allosteric sites of the two proteases were highly similar with mainly conservative mutations, but significant differences in the residues 122 and 124 were found. DENV accommodate threonine residues, whereas the ZIKV protease harbours aspartic acids. Molecular docking indicated that the orientation of the allosteric inhibitor **1** was similar between the DENV and ZIKV NS2B-NS3 protease and hence the inhibitor **1** should equally inhibit the ZIKV protease. This hypothesis was confirmed and it was shown that **1** had almost the same  $IC_{50}$  value for ZIKV (1.5  $\mu\text{M}$ ) in comparison to DENV 2 (4.2  $\mu\text{M}$ ) and DENV3 (1.0  $\mu\text{M}$ ) (Wu *et al.*, 2015).

Protein preparation of the WT ZIKV NS2B-NS3 linked protease revealed a cleavage problem. The linked protease was cleaved into two fragments, similar in size to that of the unlinked protease and still enzymatically active. External factors causing the cleavage could be eliminated by preparing the inactive S135A mutant and showed that the protease was cleaved autocatalytically. Mass spectrometry indicated that the cleavage occurred at around the R95 residue and a mutation into alanine supported the hypothesis that R95 is the relevant cleavage site for ZIKV NS2B-NS3. A double mutation of R95 and R29, commonly used in literature, was thus identified to be unnecessary. It was shown that the autocatalytic protein cleavage occurred following a *cis*-cleavage mechanism.

Fluorometric enzymatic analysis of the ZIKV NS2B-NS3 protease showed that the linked enzyme showed higher affinity towards the substrate in comparison to the unlinked protease, however the substrate turn-over rates were higher for the unlinked construct. These results are consistent with the findings of Kuiper *et al.*, 2017. Via MST a  $K_d$  value of 1.7  $\mu\text{M}$  for the binding of inhibitor **1** was obtained and hence, the protein-ligand interaction could be confirmed using an independent analytical method. The MST measurement of the S135A mutant with **1** emphasized by a  $K_d$  value of 1.4  $\mu\text{M}$  that both analytical methods were meaningful.

The protein-lipid sedimentation study pointed out that for the two chosen lipid compositions no interaction occurred. This was a surprising result, as the NS2B-NS3 protease is known to be a membrane bound protein. However, for the laboratory tests an artificial linked construct is used. The linked protease is reduced in size to its absolute minimum (see chapter 1.7.1), hence, potentially explaining the absence of the protein-lipid interaction and emphasising that the frequently used shortened constructs only mimic extracts of the reality.

## Results and discussion

For structural investigations, NMR spectroscopy and crystallisation experiments were consulted. Unfortunately it was not possible to grow a protein crystal before the structure was firstly solved by Lei *et al.*, 2016. However, both ZIKV NS2B-NS3 protease constructs (linked and unlinked) were successfully isotopically <sup>15</sup>N-labeled and the obtained spectra proved that they were both folded. During works in progress Mahawaththa *et al.*, 2017 released a solved NMR structure of the linked ZIKV NS2B-NS3 protease, for which reason the intention of the structure determination of ZIKV NS2B-NS3 protease was abandoned. Equally to the NMR experiments carried out for the DENV 2 NS2B-NS3 protease (described in chapter 4.1.3), the unlinked ZIKV NS2B-NS3 protease construct is favourable for further investigations in the interplay of protease and allosteric inhibitor. Using NMR spectroscopy the determination of the binding region of **1** on the ZIKV NS2B-NS3 protease was unsuccessful. Due to the already in chapter 4.1 mentioned solubility problem of **1**, further investigations were postponed.

Site-directed mutagenesis highlighted that there was no significant difference for the inhibition by **1** found between threonine and aspartic acid in position 122 and 124. Molecular docking predicted the interactions between ligand and protein to be backbone interactions and hence possibly explaining why exchange of the amino acids had no significant effect. This result is in accordance with the earlier findings that the *N*-(5,6-dihydroxybenzo[d]thiazol-2-yl)benzamide moiety is the inhibitory driving force with a high focus lying on the hydroxyl substituted benzothiazole.

### 4.3. Other flaviviral NS2B-NS3 proteases

Having identified **1** as an allosteric protease inhibitor of DENV and ZIKV NS2B-NS3 proteases, further investigations into other flaviviral NS2B-NS3 proteases were of substantial interest and these results are summarised in this chapter.

As there were no well-founded phylogenetic studies on the flaviviral NS2B-NS3 proteases available, an evolutionary history of flaviviruses was done by [REDACTED]. This was used to identify additional flaviviral NS2B-NS3 proteases of higher and lower similarity in order to investigate the allosteric inhibition of compound **1** further.

#### 4.3.1. Phylogenetic study

The phylogenetic study was carried out as described in chapter 4.3.1. The evolutionary history of the flaviviral polyprotein and the NS2B-NS3 protease are illustrated in chapter 6.9, Figure 6.8 and Figure 6.9. The phylogenetic tree based on the NS2B-NS3 protease (Figure 6.9) captured the same groups as the tree based on the flaviviral polyprotein (Figure 6.8). The first subdivision from the putative ancestor into two clusters, mosquito-borne and tick-borne viruses, is clearly seen in both trees. The mosquito-borne viruses can further be separated into their main transmission vectors (i.e. *Aedes* and *Culex* respectively). Viruses transmitted via *Aedes* mosquitoes exemplarily belong to the YFV, DENV and SWV groups, whereas JEV, UTV and WNV are associated with the *Culex* mosquitoes. Both trees underlined that DENV and ZIKV polyprotein and NS2B-NS3 protease are closely related, but deriving from different clans.



A sequence alignment of the NS3 proteases was used to identify the variation of the amino acids in the allosteric site of the phylogenetically analysed flaviviruses, shown in Figure 4.25.

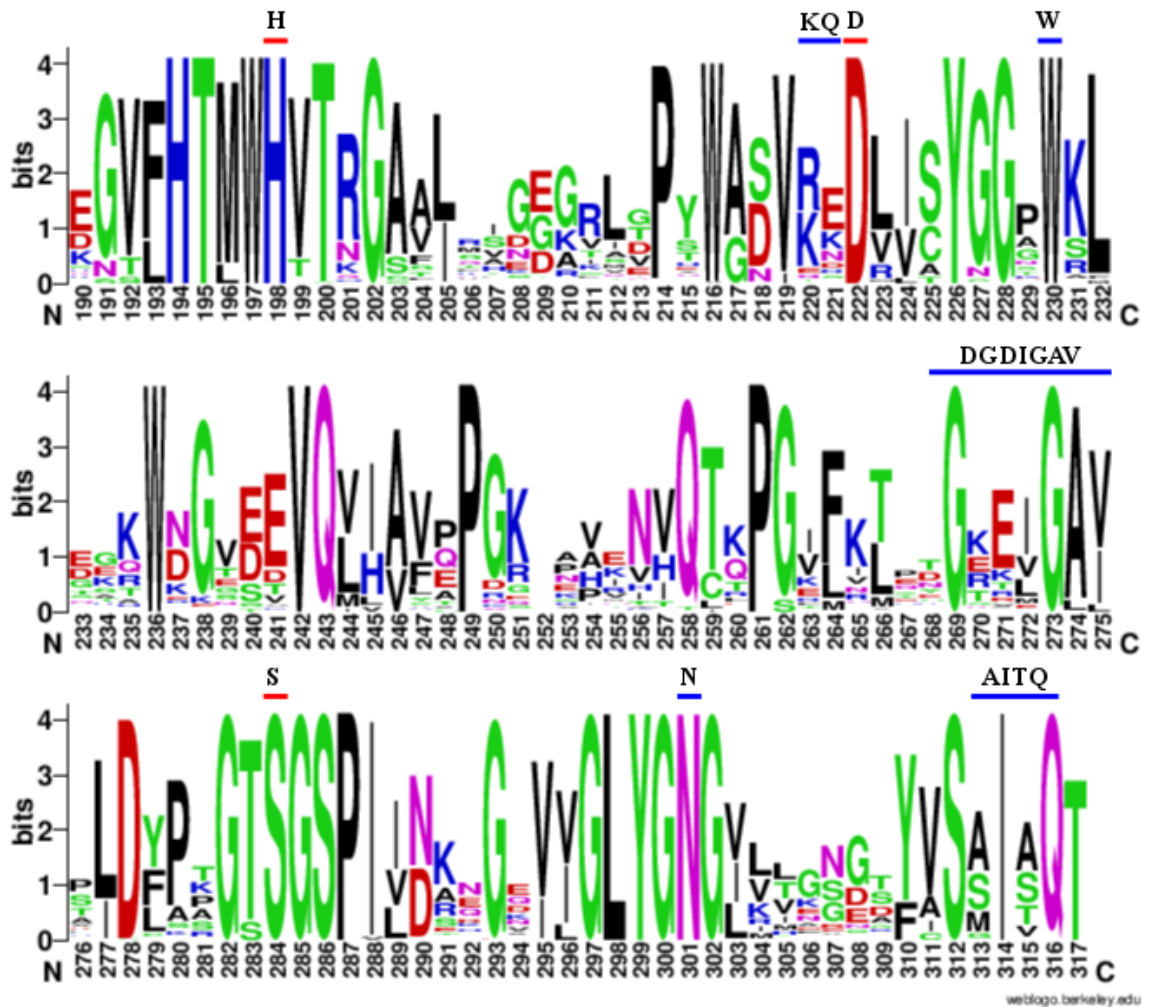
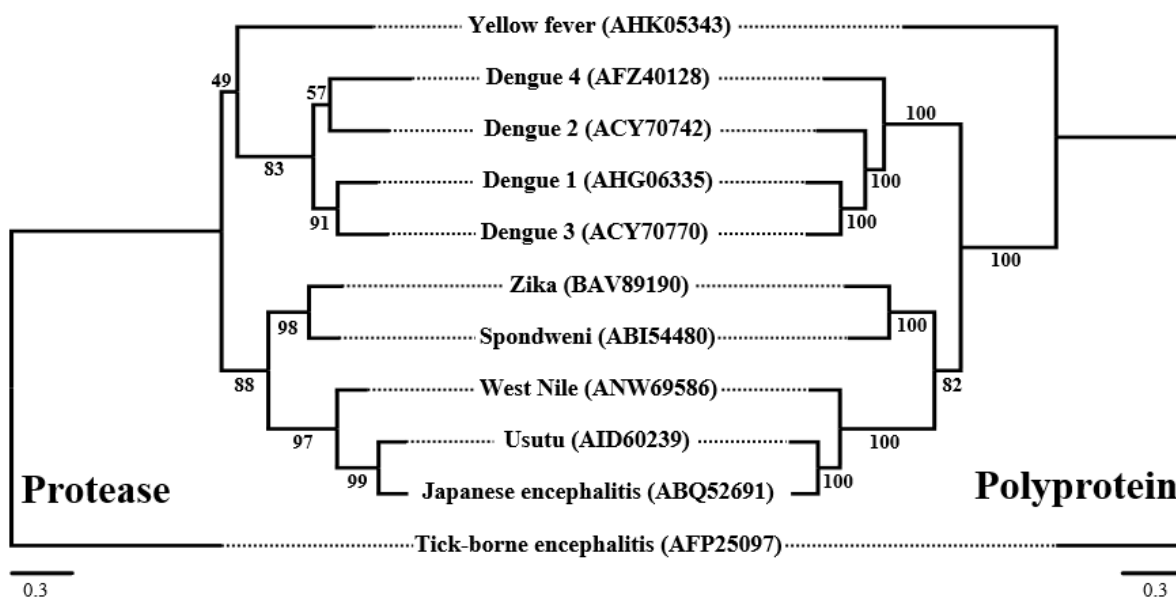


Figure 4.25: Annotated weblogo of the flaviviral NS2B-NS3 proteases. The motif sequence above the bars represents the situation in the Zika virus protease. Red bars indicate the active sites of the proteases (His-51, Asp-75, Ser-35) and blue bars represent the allosteric sites.

W83, G121, G124, N152, I165 and Q167 were identified as highly conserved amino acids contributing to the allosteric site. All other involved amino acids seemed to be variable to some degree. Taking the results obtained from the molecular docking (see chapter 4.2.1) into account, the additional flaviviral NS2B-NS3 proteases outlined in Figure 4.26 were selected for further analysis purposes.



**Figure 4.26: The evolutionary history of the additionally selected flaviviruses.**

The degree of kinship from the newly selected flaviviruses was highly dispersed. Closely related and derived from the same group are DENV 1 to 4 and the yellow-fever virus (YFV) on the one hand, ZIKV and Spondweni virus (SWV), Usutu virus (UTV) and Japanese encephalitis virus (JEV) with the West Nile virus (WNV) on the other hand. The tick-borne encephalitis virus (TBEV) was shown to be part of a different group and hence had less relation to the other proteases. A full NS2B-NS3 protease alignment of the newly designed NS2B-NS3 protease constructs is found in chapter 6.9.3.

Figure 4.27 highlights the docking results (above) from the homology model of the ZIKV NS2B-NS3 protease with **1** (see chapter 4.2.1) in combination with a sequence alignment of the selected flaviviruses (below). The main interactions of the benzothiazole moiety seemed to derive from throughout the flaviviruses highly conserved G124, N152 and I165 residues, whereas the hydroxyl substituted naphthalene moiety reached out towards the highly variable residues at position 120, 122 and 124.

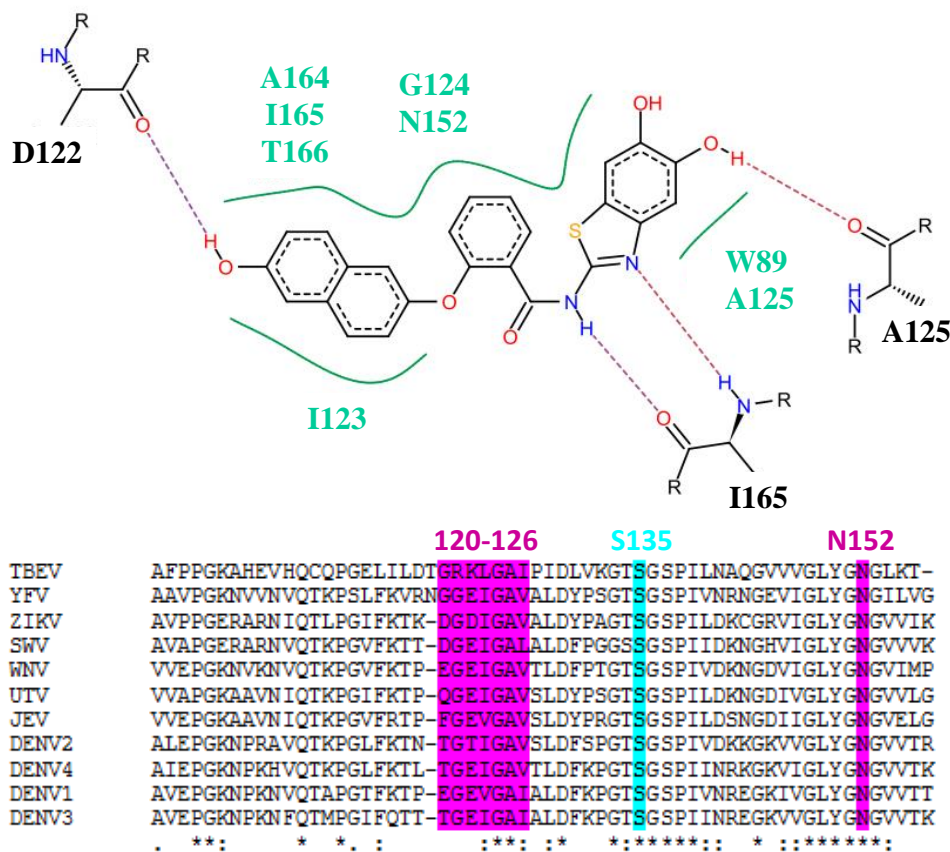


Figure 4.27: Molecular docking from the ZIKV protease with 1 from chapter 4.2.1 (above) in combination with an extraction of the sequence alignment of the selected flaviviruses (blow). Amino acids contributing to the allosteric site are highlighted in magenta and the serine of the catalytic triad in cyan.

### 4.3.2. Protein preparation and enzymatic analysis

The additionally selected flaviviral NS2B-NS3 proteases were constructed as described in chapter 3.3.2, the plasmids were purchased and checked in sequence. Expression and purification was executed as elucidated in chapter 3.3.4. The flaviviral proteases were examined using the standard fluorometric enzyme assay as outlined in detail in chapter 3.4.1. Table 4.7 summarises the determined enzymatic kinetic data of the new flaviviral NS2B-NS3 protease constructs.

Table 4.7: Enzymatic kinetic data of the newly established flaviviral NS2B-NS3 proteases and their  $IC_{50}$  values of 1, respectively.

Flaviviral constructs	Enzyme conc. [nM]	$K_M$ [ $\mu$ M]	SE $K_M$	$v_{max}$ [ $\mu$ M s <sup>-1</sup> ]	SE $v_{max}$	$k_{cat}$ [s <sup>-1</sup> ]	SE $k_{cat}$	$k_{cat}/K_M$ [M <sup>-1</sup> s <sup>-1</sup> ]	Error $k_{cat}/K_M$ [M <sup>-1</sup> s <sup>-1</sup> ]	$IC_{50}$ (1) [ $\mu$ M]	SE $IC_{50}$
DENV1	78.8	435.7	25.4	1.2	2.9 x10 <sup>-2</sup>	0.015	3.7 x10 <sup>-4</sup>	34.9	2.9	7.4	0.6
DENV3	32.2	269.9	33.1	0.7	2.9 x10 <sup>-2</sup>	0.022	9.1 x10 <sup>-4</sup>	80.8	13.3	6.7	1.0
DENV4	25.9	258.2	25.6	2.5	1.2 x10 <sup>-1</sup>	0.096	4.5 x10 <sup>-3</sup>	370.9	54.2	5.4	2.9
TBEV	59.8	293.1	58.6	1.0	8.8 x10 <sup>-2</sup>	0.016	1.5 x10 <sup>-3</sup>	55.0	16.0	4.4	1.0
WNV	13.5	158.8	17.9	1.5	5.8 x10 <sup>-2</sup>	0.112	4.3 x10 <sup>-3</sup>	706.9	106.9	5.3	0.7
SWV	3179.6	484.1	35.7	4.0	1.5 x10 <sup>-1</sup>	0.001	4.6 x10 <sup>-5</sup>	2.6	0.3	na	na
UTV	1822.5	na	na	Na	na	na	na	na	na	na	na
JEV	513.6	254.1	40.1	1.1	2.9 x10 <sup>-2</sup>	0.002	5.7 x10 <sup>-5</sup>	8.3	1.5	12.4	7.8
YFV	660.0	80.8	20.4	3.6	8.8 x10 <sup>-2</sup>	0.005	1.3 x10 <sup>-4</sup>	67.9	18.8	7.1	0.1

SE: standard error, na: not applicable

## Results and discussion

Apart from the UTV, all other newly established flaviviral NS2B-NS3 protease constructs were enzymatically active and hydrolysed the Boc-GRR-AMC substrate sufficiently. The determined  $K_M$  values were comparable with that of DENV and ZIKV. The different enzymatic efficiencies represented by the  $k_{cat}/K_M$  values varied significantly amongst the examined proteases. WNV and DENV4 exhibited a 10 times higher value as the other proteases. All functional enzymes were inhibited by **1** with highly similar  $IC_{50}$  values.

In order to eliminate the fact that **1** was an unselective inhibitor it was tested against some other proteases frequently used in the laboratory. Table 4.8 summarises the percentage of inhibition of the different proteases by **1** in a final concentration of 20  $\mu\text{M}$  and their  $IC_{50}$  values respectively.

**Table 4.8: Percentage of inhibition of the different proteases by **1** in a final concentration of 20  $\mu\text{M}$  and their  $IC_{50}$  values respectively. (n.d. = not determined)**

<b>Protease</b>	<b>Protease inhibition [%]</b>	<b><math>IC_{50}</math> (<b>1</b>) [<math>\mu\text{M}</math>]</b>
$\alpha$ -Chymotrypsin	80	12.3 $\pm$ 9.4
Proteasome chymotrypsin-like activity	100	17.2 $\pm$ 13.2
Proteasome postglutamyl/caspase-like activity	100	4.6 $\pm$ 2.0
Proteasome trypsin-like activity	100	17.2 $\pm$ 5.1
Cathepsin B	0	na
Cathepsin L	0	na

na: not applicable

### 4.3.3. Summary and discussion of the other flaviviral NS2B-NS3 protease results

A phylogenetic study on the arboviruses of the family *Flaviviridae*, focusing mainly on the NS2B-NS3 protease, was established and revealed that the putative ancestor equally branches into two clusters, mosquito-borne and tick-borne viruses. The phylogenetic tree based on the flaviviral NS2B-NS3 protease was consulted to identify the degree of relatedness in comparison to the diversity of the amino acids contributing to the allosteric site (Figure 4.25). Thereupon nine more flaviviral NS2B-NS3 protease (Figure 4.26) constructs were designed analogously to the well-established DENV NS2B-NS3 proteases described in chapter 1.7.1.

After successful protein expression and purification of the various flaviviral NS2B-NS3 proteases, enzymatic analysis showed that all proteases, impart from the UTV, were enzymatically active and could be examined by following the Boc-GRR-AMC hydrolysis using the standard DENV fluorometric assay described in chapter 3.4.1.

Even though the selected constructs represented a broad spectrum of the flaviviral NS2B-NS3 proteases and hence, included a significant amount of natural given differences (Figure 6.10), **1** inhibited all of them. The determined  $IC_{50}$  values for **1** amongst the tested flaviviral NS2B-NS3 proteases were in the range from 4 – 12  $\mu$ M and hence, highly similar. These results indicate that **1** is a broad applicable flaviviral NS2B-NS3 protease inhibitor and suggests that the allosteric binding site is present throughout all flaviviruses.

Selectivity of **1** was examined by testing the inhibitor against other proteases frequently examined in the laboratory. **1** inhibits the human proteasome, with the three distinct catalytic activities of chymotrypsin-like, caspase-like and trypsin-like summarized in Table 4.8. The proteasome chymotrypsin-like activity exhibits the hydrolysis of peptides after hydrophobic amino acids, whereas the caspase-like and trypsin-like activities after acidic and basic amino acids, respectively. The inhibition of the proteasome maybe explained by the extremely small size of **1** in comparison with the giant (~ 2000 kDa), multiple functional protein and hence, leading to unspecific binding of **1** in the cylindrical 20S core. Alpha-chymotrypsin is a digestive enzyme and preferentially cleaves hydrophobic, aromatic amino acids in P<sub>1</sub> position of the substrate. The enzyme is inhibited by **1**, presumably resulting from **1** binding to the hydrophobic binding sites. Cathepsin B and cathepsin L are both cysteine proteases, exhibiting cleavage after amino acids with positive charged side chains. Given that they are not being inhibited by **1**, shows that binding of **1** is not arbitrary.

## 5. Conclusion and perspectives

The demand for an anti-flaviviral treatment has never been greater. The DENV NS2B-NS3 protease has been identified as a drug target and is heavily investigated. Due to the unusual narrow and little distinctive active site of the serine protease, problems in drug discovery arise. Remedy could provide the development of allosteric inhibitors, such as **1** (Wu *et al.*, 2015).

Having identified compound **1** as lead structure for allosteric DENV 2 and DENV 3 NS2B-NS3 protease inhibitors, the investigations summarised in this work showed that the *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)benzamide moiety is the inhibitory driving force, with a high focus lying on the twice hydroxyl substituted benzothiazole part. Here it was validated that the dimethoxy precursor compounds derived from **1** acted as prodrugs and are metabolised to the hydroxyl groups in cells. By the means of site-directed mutagenesis this work confirmed for the first time, that binding of the inhibitor **1** actually takes place in the proposed allosteric site of the DENV 2 NS2B-NS3 protease. However, due to low solubility of **1** in aqueous solution, crystallisation and NMR experiments carried out on the DENV 2 NS2B-NS3 protease were unsuccessful.

On this account optimisations of **1** are inevitable. The focus should lie on altering the naphthalene side leaving the benzothiazole part unchanged. For further structural elucidations, such as crystallisation and NMR experiments, polarity and hence solubility in aqueous solution of **1** necessarily must be increased. According to the correctness of the structure published by Yao *et al.*, 2019 (discussed in chapter 1.7.4 ) of the DENV 2 NS2B-NS3 protease containing the allosteric inhibitor **2** (PDB: 6MO0), this newly obtained information should be used for computational analysis. Redocking of **2** and docking of **1** and its derivatives into DENV 2 NS2B-NS3 protease containing the allosteric inhibitor (PDB: 6MO0) should provide further distinctions of the binding mode of **1**. Yao *et al.* indicated that an allosteric inhibitor forces the DENV protease to embrace the enzymatically inactive “open” conformation. As it is still uncertain that the “open” conformation arises from a crystallisation artefact and hence, could well differ in solution, the dynamic investigations of the NS2B co-factor, in absence and presence of the allosteric inhibitors, using NMR spectroscopy experiments should be continued. This work points out, that the unlinked DENV NS2B-NS3 protease construct should be used for future experiments in this field.

The presented examination of the allosteric binding region led to the preparation of the linked and unlinked ZIKV NS2B-NS3 protease constructs. Due to autocatalytic cleavage of the WT linked ZIKV NS2B-NS3 protease construct, the introduction of a double mutation (R95A and R29G) is commonly used in literature. In this work it was shown that the autocatalytic protein cleavage occurred following a *cis*-cleavage mechanism and that only the single mutation at position R95A is essential to prohibit autocatalytic cleavage. For the first time it was shown that **1** inhibits the ZIKV

NS2B-NS3 protease in an equally efficient manner, represented by the  $IC_{50}$  value of 1.5  $\mu$ M. It was shown via protein-lipid sedimentation assay, based on the linked ZIKV NS2B-NS3 protease construct and a lipid mix mimicking the ER composition, that no protein-lipid interactions were traceable. This result emphasised that the frequently used shortened linked ZIKV NS2B-NS3 construct is highly artificial.

On these grounds it should be well-considered in future which constructs (linked or unlinked) should be used for further examinations. The linked construct is highly artificial and not necessarily required. Experiments on the protein stability of the two constructs should be carried out in order to justify the use of either constructs. According to this, it might be of interest to repeat the protein-lipid sedimentation assay for the unlinked construct. Furthermore this work suggests the use of the minimally modified linked ZIKV NS2B-NS3 protease construct, containing the single mutant (R95A) essential to prohibit autocatalytic cleavage, for future structural and functional studies.

Finally this work presented that **1** inhibits a broad spectrum of other flaviviral NS2B-NS3 proteases. DENV 1, DENV 4, UTV, TBEV, WNV, SWV, JEV and YFV were successfully expressed, purified and enzymatically analysed. Besides UTV, the above mentioned flaviviral NS2B-NS3 proteases were enzymatically active and showed to be inhibited by **1**. The inhibition values of the other flaviviral NS2B-NS3 proteases were equally to DENV and ZIKV found to be in a low micromolar range.

These findings lead to several interesting questions, such as: Do these flaviviral allosteric binding sites have a specific function, for example serve as a “conformation switch” between the “open” and “closed” state? Is there a difference in the allosteric site found in the *in vitro* used artificial constructs and the full-length NS2B-NS3 protease present *in vivo*?

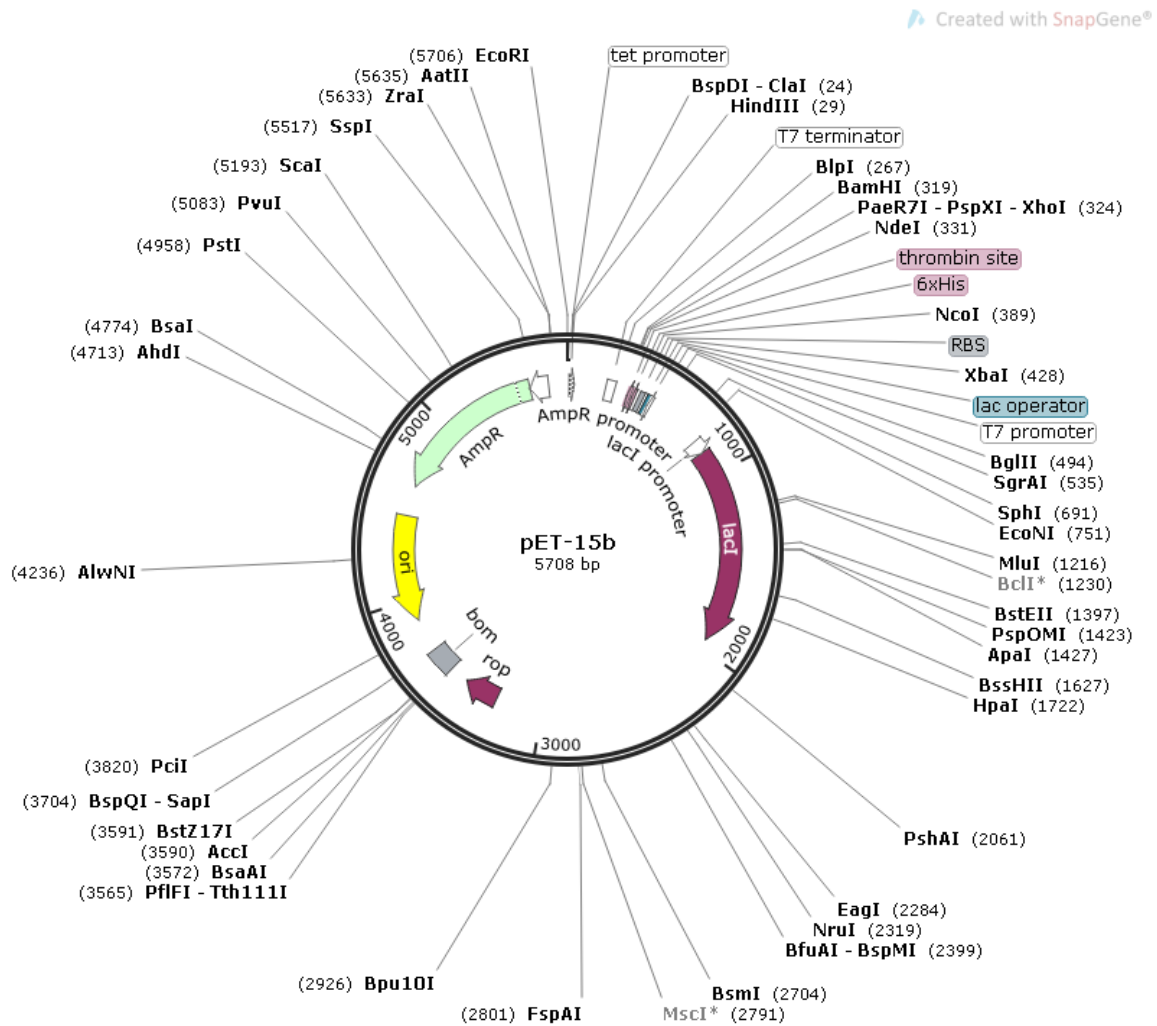
In summary, the results of this work highlight that approaching the flaviviral allosteric binding pocket of the NS2B-NS3 protease could well lead to the development of a PAN-flaviviral inhibitor.



## 6. Appendix

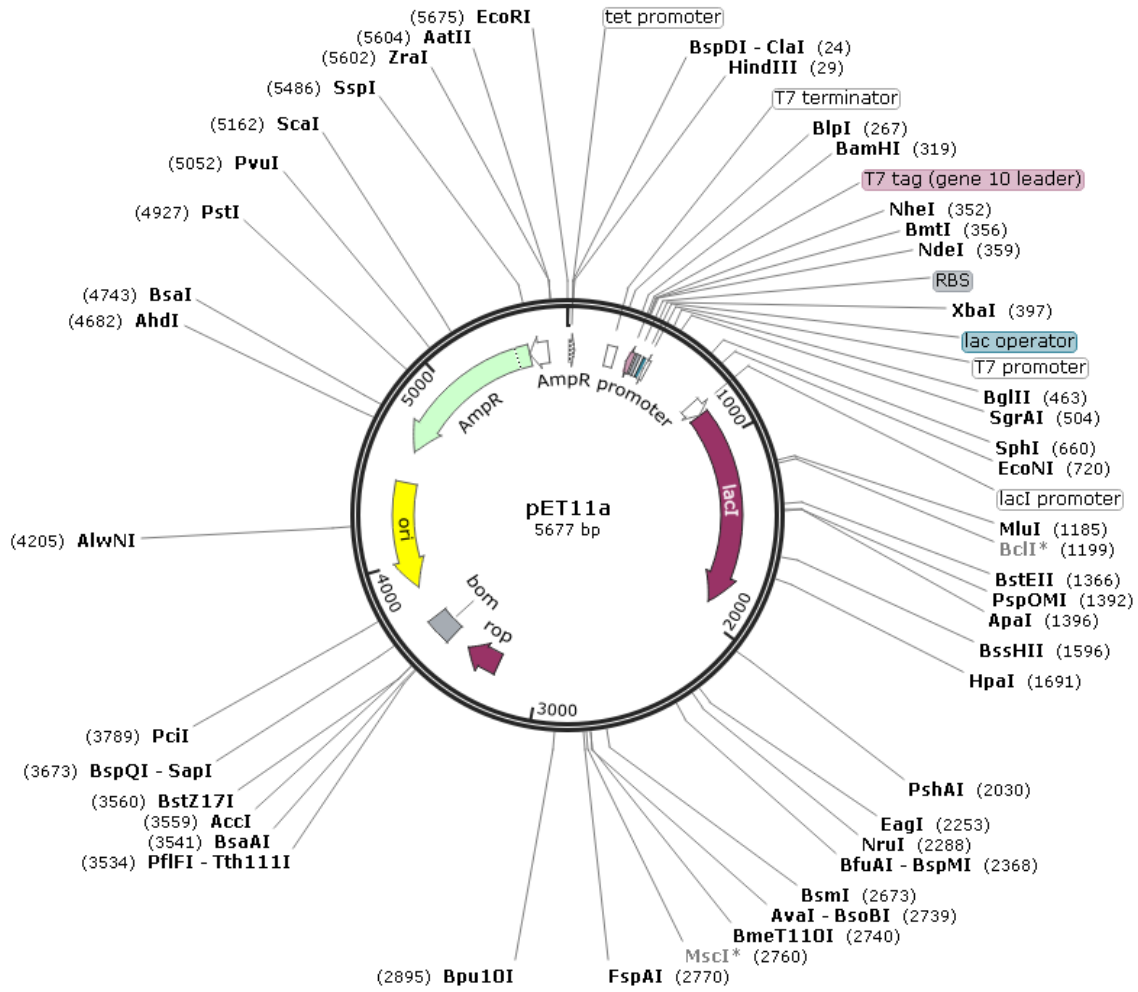
### 6.1. Plasmid maps

#### 6.1.1. Vector: pET-15b



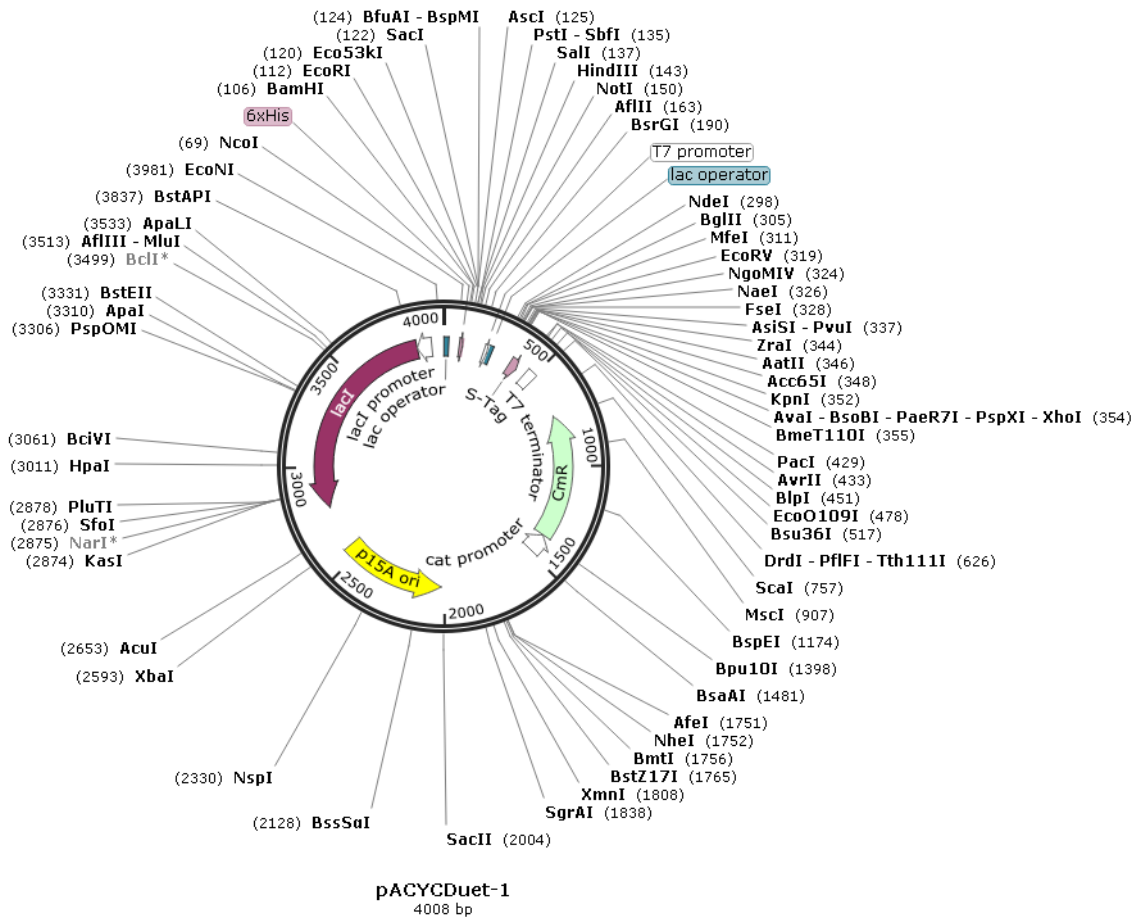
## 6.1.2. Vector: pET-11a

Created with SnapGene®



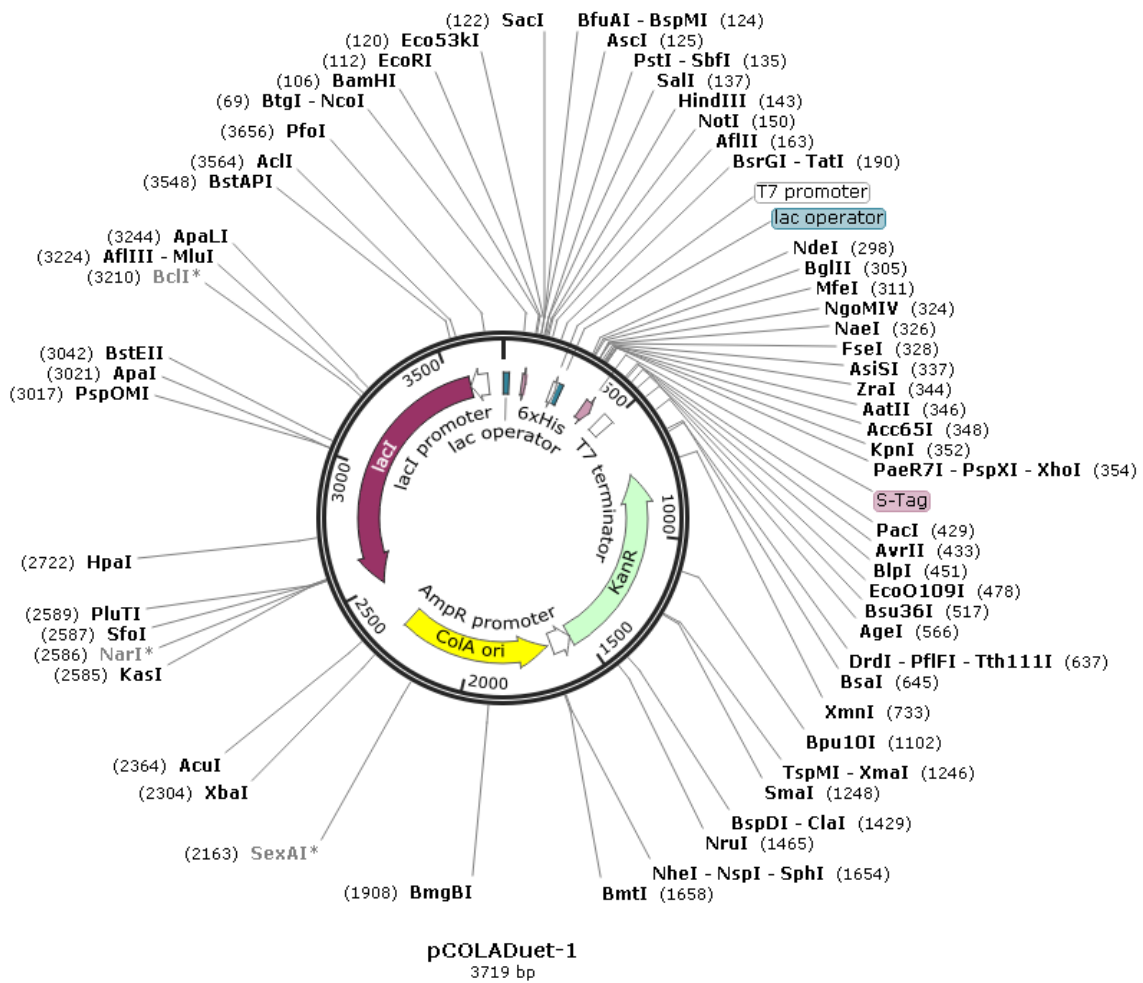
### 6.1.3. Vector: pACYCDuet-1

Created with SnapGene®



## 6.1.4. Vector: pCOLADuet-1

Created with SnapGene®



## 6.2. Sequences and protein physicochemical properties

### 6.2.1. Dengue NS2B-NS3 protease constructs

#### DENVII NS2B alone (p094)

```
atggacctggaactggaacgtgcgccagacgtgcggttgggaagaacaagcagaaatcagc
M D L E L E R A A D V R W E E Q A E I S
ggtagcagcccgatcctgtcaattacgatctcgggaagatggtagcatgtctattaaaaac
G S S P I L S I T I S E D G S M S I K N
gaagaagaagaacagaccctgtaa
E E E E Q T L -
```

Number of amino acids 47  
Molecular weight 5267.73 g/mol  
Theoretical pI 3.88  
Ext. coefficient (280 nm) 5500 M<sup>-1</sup> cm<sup>-1</sup>

#### DENVII NS3 alone (p093)

```
atgggtagcagccatcatcatcatcacggtagcgaaaacctgtacttccaaggctcc
M G S S H H H H H G S E N L Y F Q G S
gcaggcgtgctgtgggatgttccgtctccgcccggtgggtaaagcagaactggaagac
A G V L W D V P S P P P V G K A E L E D
ggcgcttatcgtattaaacagaaggggtgccgcccggctattcaciaaatcgggtgcccggctt
G A Y R I K Q K G A A G Y S Q I G A G V
taciaaagaaggtacctttcattaccatgtggcagtcacgcggtggtgcagtgctgatgcat
Y K E G T F H T M W H V T R G A V L M H
aaaggcaagcgcattgaaccgtcatgggctgatgtcaaaaaggacctgatctcgtacggc
K G K R I E P S W A D V K K D L I S Y G
gggtggctggaactggaaggcgaatggaagggaaggcgaagaagttcagggtcctggcgctg
G G W K L E G E W K E G E E V Q V L A L
gaaccgggtaaaaaccgcgctgccgttcaaaccaaaccgggacctgtttaagaccaatacg
E P G K N P R A V Q T K P G L F K T N T
ggtaccattggcgcggttagcctggatttctctccgggtacgagtggtcctcccgatcgtc
G T I G A V S L D F S P G T S G S P I V
gacaaaaagggtaaagtgggtggcctgtatggtaatgggtgctcgtgacctgtagtggtgcc
D K K G K V V G L Y G N G V V T R S G A
tacgtgtccgcgattgccaacacggaaaagtccattgaagataaccggaaattgaagat
Y V S A I A N T E K S I E D N P E I E D
gacatcttttcgtaagtaa
D I F R K -
```

Number of amino acids 205  
Molecular weight 22093.83 g/mol  
Theoretical pI 7.14  
Ext. coefficient (280 nm) 37930 M<sup>-1</sup> cm<sup>-1</sup>

### DENVII NS2B-NS3 unlinked protease

Number of amino acids 252  
Molecular weight 27343.54 g/mol  
Theoretical pI 5.42  
Ext. coefficient (280 nm) 43430 M<sup>-1</sup> cm<sup>-1</sup>

### DENV2 NS2B-NS3 linked protease (p092)

```
atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcccagccat
M G S S H H H H H S S G L V P R G S H
atgctggcggcagacctggaactggaactgctgcggcagacgtgcttgggaagaacaagca
M L A A D L E L E R A A D V R W E E Q A
gaaatcagcggtagcagccccgatcctgtcaattacgatctcgggaagatggtagcatgtct
E I S G S S P I L S I T I S E D G S M S
attaaaaacgaagaagaagaacagaccctgggcggtagcggcggtagcggcggtagcggcggtagc
I K N E E E E Q T L G G G G S G G G G A
ggcgtgctgtgggatgttccgtctccgcccgccggtagcggtaaaagcagaactggaagacggc
G V L W D V P S P P P V G K A E L E D G
gcttatcgtattaaacagaagggtagcggcggctattcacaatcggtagcggcggcggtagc
A Y R I K Q K G A A G Y S Q I G A G V Y
aaagaaggtacctttcataccatgtggcagctcacgcgtggtagcagtgctgatgcataaa
K E G T F H T M W H V T R G A V L M H K
ggcaagcgcattgaaccgtcatgggctgatgtcaaaaaggacctgatctcgtacggcggtagc
G K R I E P S W A D V K K D L I S Y G G
ggctggaaactggaaggcgaatggaagggaaggcgaagaagttcaggctcctggcggtagcggaa
G W K L E G E W K E G E E V Q V L A L E
ccgggtaaaaaccgcgtagcggcttcaaaccaaaccgggcctgtttaagaccaatacgggtagc
P G K N P R A V Q T K P G L F K T N T G
accattggcgcggtagcctggatttctctccggtagcagtggtcctcccgatcgtcgcac
T I G A V S L D F S P G T S G S P I V D
aaaaagggtaaagtggtagcctgtatggtaatggtagcgtgacctgtagtggtgcctac
K K G K V V G L Y G N G V V T R S G A Y
gtgtccgcgattgccaacacggaaaagtccattgaagataaccggaaattgaagatgac
V S A I A N T E K S I E D N P E I E D D
attttccgtaagtga
I F R K -
```

Number of amino acids 264  
Molecular weight 28037.31 g/mol  
Theoretical pI 5.71  
Ext. coefficient (280 nm) 41940 M<sup>-1</sup> cm<sup>-1</sup>

## 6.2.2. ZIKV NS2B-NS3 protease constructs

### ZIKB NS2B alone (p122)

atggatatgtatattgaacgcgcggcgatattacctgggaaaaagatgcggaagtgacc  
M D M Y I E R A G D I T W E K D A E V T  
ggcaacagccccgcgcctggatgtggcgctggatgaaagcggcgatTTTTAGCCTGGTGGAA  
G N S P R L D V A L D E S G D F S L V E  
gatgatggccccgcatgcgctaa  
D D G P P M R -

Number of amino acids 47  
Molecular weight 5244.72g/mol  
Theoretical pI 3.88  
Ext. coefficient (280 nm) 6990 M<sup>-1</sup> cm<sup>-1</sup>

### ZIKV NS3 alone (p121)

atgggtagcagccatcatcatcatcacggtagcgaaaaacctgtacttccaaggctcc  
M G S S H H H H H G S E N L Y F Q G S  
agcggcgcgctgtgggatgtgccggcgccgaaagaagtgaaaaaggcgaaaccaccgat  
S G A L W D V P A P K E V K K G E T T D  
ggcgtgtatcgcgtgatgacccgcgcctgctggggcagcaccaggtggcgctggcgctg  
G V Y R V M T R R L L G S T Q V G V G V  
atgcaggaaggcgtgtttcataccatgtggcatgtgaccaaaggcagcgcgctgcgcagc  
M Q E G V F H T M W H V T K G S A L R S  
ggcgaaggccgcctggatccgtattggggcgatgtgaaacaggatctggtgagctattgc  
G E G R L D P Y W G D V K Q D L V S Y C  
ggccccgtggaaactggatgccccgctgggatggccatagcgaagtgcagctgctggcggtg  
G P W K L D A A W D G H S E V Q L L A V  
ccgccgggcaacgcgcgcgcaacattcagaccctgccgggcatttttaaaaccaaagat  
P P G E R A R N I Q T L P G I F K T K D  
ggcgatattggcgcggtggcgctggattatccggcgggcaccagcggcagcccgattctg  
G D I G A V A L D Y P A G T S G S P I L  
gataaatgcggccgcgctgattggcctgtatggcaacggcgtggtgattaaaaacggcagc  
D K C G R V I G L Y G N G V V I K N G S  
tatgtgagcgcgattaccagggccgcgcgaagaagaacccccgggtggaatgctttgaa  
Y V S A I T Q G R R E E E T P V E C F E  
ccgagcatgctgaaaaaataa  
P S M L K K -

Number of amino acids 206  
Molecular weight 22387.28 g/mol  
Theoretical pI 6.75  
Ext. coefficient (280 nm) 38055 M<sup>-1</sup> cm<sup>-1</sup>

### ZIKN NS2B-NS3 unlinked protease

Number of amino acids 253  
Molecular weight 27613.99 g/mol  
Theoretical pI 5.36  
Ext. coefficient (280 nm) 45045 M<sup>-1</sup> cm<sup>-1</sup>

### ZIKV NS2B-NS3 linked protease (p119)

catatgggtagcagccatcatcatcatcaccacggcagcgagaaacctgtactttcaaggc  
H M G S S H H H H H H G S E N L Y F Q G  
agcgacatgtatatcgagcgtgcgggcgacattacctgggaaaaggatgcggaagttacc  
S D M Y I E R A G D I T W E K D A E V T  
ggtaacagcccgcgtctggatggtgcgctggacgaaagcggtagcttcagcctggttgaa  
G N S P R L D V A L D E S G D F S L V E  
gacgatgggtccgccgatgctggcggtggtggttagcggcggtggcggttagcggcgcgctg  
D D G P P M R G G G G S G G G G S G A L  
tgggacgtgccggcgccgaaggaagtgaaaaagggcgaaaccaccgacgggtgtttaccgt  
W D V P A P K E V K K G E T T D G V Y R  
gtgatgaccctgctctgctgggttagcaccagggtgggtggtggtgatgcaggaaggt  
V M T R R L L G S T Q V G V G V M Q E G  
gttttccacaccatgtggcaggttaccagggttagcgcgctgctagcggcgaaggtcgt  
V F H T M W H V T K G S A L R S G E G R  
ctggacccgtattggggtgatgtaagcaggacctggttagctactgcggtccgtggaaa  
L D P Y W G D V K Q D L V S Y C G P W K  
ctggacgcggcggtgggatggtcacagcgaagtgcagctgctggcggttccgccgggtgaa  
L D A A W D G H S E V Q L L A V P P G E  
cgtgcgcgtaacatccagaccctgccgggtatcttcaaaaccaaggacggcgatattggt  
R A R N I Q T L P G I F K T K D G D I G  
gcggtgcgctggattatccggcgggcaccagcggtagcccgattctggacaaatgcggc  
A V A L D Y P A G T S G S P I L D K C G  
cgtgttattggcctgtacggtaacggcgttgtgatcaagaacggtagctacgttagcgcg  
R V I G L Y G N G V V I K N G S Y V S A  
atcaccaggggtcgtcgtgaggaagaaaccccggtggagtgctttgaaccgagcatgctg  
I T Q G R R E E E T P V E C F E P S M L  
aaaaaataa  
K K -

Number of amino acids 261  
Molecular weight 28026,29 g/mol  
Theoretical pI 5.36  
Ext. coefficient (280 nm) 45045 M<sup>-1</sup> cm<sup>-1</sup>



### 6.2.3. DENV1 NS2B-NS3 linked protease (p266)

ggcagcagccaccaccaccaccaccacggcagcagagaatctgtactttcaaggcagcagc  
 G S S H H H H H H G S E N L Y F Q G S D  
 ctgagcctggagaaggcggcggaagtgagctgggaggaagaggcggaacacagcggcgcg  
 L S L E K A A E V S W E E E A E H S G A  
 agccacaacatcctggttgaggtgcaggacgatggtaccatgaagattaaaaacgaagag  
 S H N I L V E V Q D D G T M K I K N E E  
 cgtgacgataccctgggtggcggtggcagcgggtggcggtggcagcggcggttctgtgggac  
 R D D T L G G G G S G G G S G V L W D  
 accccgagcccgcggaagttgagcgtgcggtgctggacgatggtatctaccgtattatg  
 T P S P P E V E R A V L D D G I Y R I M  
 caacgtggcctgctgggtcgtagccaggtggcggttggtgtgttccaagaaaacgttttt  
 Q R G L L G R S Q V G V G V F Q E N V F  
 cacaccatgtggcagcgttaccctggcgcggtgctgatgtaccaaggtaaacgtctggag  
 H T M W H V T R G A V L M Y Q G K R L E  
 ccgagctgggcgagcgtgaagaaagacctgatcagctatggtggcggttggcgtctgcag  
 P S W A S V K K D L I S Y G G G W R L Q  
 ggcagctggaacaccggtgaagaggttcaagtgattgcggttgaaccgggcaagaaccg  
 G S W N T G E E V Q V I A V E P G K N P  
 aaaaacgtgcaaaccgcgcccgggtaccttcaagaccccggaaggcgagggttggtgcgatt  
 K N V Q T A P G T F K T P E G E V G A I  
 gcgctggattttaaaccgggcaccagcggtagcccgatcgtaaacgtgagggcaagatt  
 A L D F K P G T S G S P I V N R E G K I  
 gtgggtctgtacggcaacggtgtggttaccaccagcggtagctatgtgagcgcgattgcg  
 V G L Y G N G V V T T S G T Y V S A I A  
 caggcgaaggcgagccaagaaggcccgtgccggagattgaagacgaagtgttccgcaag  
 Q A K A S Q E G P L P E I E D E V F R K

Number of amino acids     260  
 Molecular weight            27899.90 g/mol  
 Theoretical pI              5.28  
 Ext. coefficient (280 nm)   41940 M<sup>-1</sup> cm<sup>-1</sup>

#### 6.2.4. DENV3 NS2B-NS3 linked protease (p258)

ggtagcagccaccaccaccaccacggcagcgagaacctgtactttcaaggcagcgac  
G S S H H H H H H G S E N L Y F Q G S D  
ctgaccgttgagaaggcggcggtgtgacctgggaggaagaggcggagcagaccggtgtt  
L T V E K A A D V T W E E E A E Q T G V  
agccacaacctgatgatcaccgtggacgatgacggcaccatgcgtatcaaggatgacgaa  
S H N L M I T V D D D G T M R I K D D E  
accgagaacatcctgggtggcgggtggcagcgggtggcagcgggtgttctgtgggat  
T E N I L G G G G S G G G S G V L W D  
gtgccgagcccgcggaaacccaaaaagcggagctggaagaggcggtttaccgtatcaag  
V P S P P E T Q K A E L E E G V Y R I K  
cagcaaggcattttcggtaaaacccaagtggcggttggtgtgcaaaaagagggtgtgttt  
Q Q G I F G K T Q V G V G V Q K E G V F  
cacaccatgtggcatgttaccctggcgggtgctgaccacaacggtaaacgtctggaa  
H T M W H V T R G A V L T H N G K R L E  
ccgaactggcgagcgtaagaagatctgatcagctatggtggcgggtggcgtctgagc  
P N W A S V K K D L I S Y G G G W R L S  
gcgagtggaagggcgaagaggttcaggtgatcgcgggtggaaccgggcaagaacccg  
A Q W Q K G E E V Q V I A V E P G K N P  
aaaaacttccagaccatgccgggcatttttcaaaccaccaccggcgagatcgggtgcgatt  
K N F Q T M P G I F Q T T T G E I G A I  
gcgctggacttcaaaccgggcaccagcggtagcccgatcattaaccgtgaaggcaagggtg  
A L D F K P G T S G S P I I N R E G K V  
gttggctctgtacggcaacgggtgtggttaccaaaaacggcgggttatggttagcgggtattgcg  
V G L Y G N G V V T K N G G Y V S G I A  
cagaccaatgcggaaccggacggcccgaccccggaactggaagaagaatgtttaagaag  
Q T N A E P D G P T P E L E E E M F K K

Number of amino acids 260  
Molecular weight 27974.03 g/mol  
Theoretical pI 5.28  
Ext. coefficient (280 nm) 40450 M<sup>-1</sup> cm<sup>-1</sup>

### 6.2.5. DENV4 NS2B-NS3 linked protease (p257)

ggcagcagccaccaccaccaccaccacggcagcagagaacctgtactttcaaggcagcagac  
 G S S H H H H H H G S E N L Y F Q G S D  
 ctgagcctggagaaggcggcgaatgtgcagtgaggacgagatggcggatatacccggcagc  
 L S L E K A A N V Q W D E M A D I T G S  
 agcccgatcattgaggtgaagcaagacgaagatggtagcttcagcatccgtgacgttgag  
 S P I I E V K Q D E D G S F S I R D V E  
 gaaaccaacatgattgggtggcgggtggcagcgggtggcgggtggcagcggcgcgctgtgggat  
 E T N M I G G G G S G G G G S G A L W D  
 gttccgagcccgggcgacccaagaaagcggcgctgagcaggggtgtttaccgtatcatg  
 V P S P A A T K K A A L S E G V Y R I M  
 cagcgtggcctgttcggtaaaacccaagtggcggttggtattcacatggaaggcgtgttt  
 Q R G L F G K T Q V G V G I H M E G V F  
 cacaccatgtggcagtgaccggtggcagcgttatctgccacgaaaccggtcgtctggaa  
 H T M W H V T R G S V I C H E T G R L E  
 ccgagctggcgagcgttcgtaacgatatgattagctatggtggcgggtggcgtctgggt  
 P S W A D V R N D M I S Y G G G W R L G  
 gacaagtgggataaagaggaagacgtgcaggttctggcgatcgagccgggcaagaaccgg  
 D K W D K E E D V Q V L A I E P G K N P  
 aaacacgtgcaaaccaagccgggtctgttcaaaaccctgaccggcgaaattgggtgcggtt  
 K H V Q T K P G L F K T L T G E I G A V  
 accctggactttaagccgggcaccagcggtagcccgatcattaaccgtaagggcaaagtg  
 T L D F K P G T S G S P I I N R K G K V  
 atcggctctgtacggcaacgggtgtggttacccaaaagcgggtgattatgtagcgcgattacc  
 I G L Y G N G V V T K S G D Y V S A I T  
 caggcggaaacgcattgggtgaaccggactatgaagttgatgaagacatTTTTTCGAAA  
 Q A E R I G E P D Y E V D E D I F R K

Number of amino acids	259
Molecular weight	27987.20 g/mol
Theoretical pI	5.54
Ext. coefficient (280 nm)	41940 M <sup>-1</sup> cm <sup>-1</sup>

### 6.2.6. JEV NS2B-NS3 linked protease (K/A) (p253)

ggtagcagccatcatcatcatcacggtagcgcgagaacctgtactttcaaggtagcgac  
G S S H H H H H H G S E N L Y F Q G S D  
atgtggctggagcgtgcccggacattagctgggaaatggacgcggcgattaccggtagc  
M W L E R A A D I S W E M D A A I T G S  
agccgctcgtctggatgtgaagctggacgatgacggtgatttccacctgatcgatgaccg  
S R R L D V K L D D D G D F H L I D D P  
ggtgttccgtggcgcggtggcagcgggtggcgggtggcgggtggtttctgggac  
G V P W A G G G G S G G G G G V F W D  
accccgagcccgaagccgtgcagcaaaggcgataccaccaccgggtgtttaccgtatcatg  
T P S P K P C S K G D T T T G V Y R I M  
gcgcggtggtattctgggtacctaccaggcgggtgttggcgtgatgtacgaaaacgtgttc  
A R G I L G T Y Q A G V G V M Y E N V F  
cacaccctgtggcacaccaccctgggtgcccgatcatgagcgggtgaaggcaaactgacc  
H T L W H T T R G A A I M S G E G K L T  
ccgtattggggcagcgttaaagaagatcgatcgcgtagcggcgggtccgtggcgtttcgac  
P Y W G S V K E D R I A Y G G P W R F D  
cgtaaatggaacggtaccgatgacggttcaggatcggttgaggcgggtaaagcggcg  
R K W N G T D D V Q V I V V E P G K A A  
gttaacatccaaaccaaacgggtgttttccgtaccccggttggtaagtgggtgcggtt  
V N I Q T K P G V F R T P F G E V G A V  
agcctggactatccgcgtggtaccagcggcagcccgattctggacagcaacggatgatt  
S L D Y P R G T S G S P I L D S N G D I  
atcggcctgtatggtaacggcgttgaactgggtgatggttagctacgttagcgcgattggt  
I G L Y G N G V E L G D G S Y V S A I V  
caaggcgaacgtcaagaggaaccgggtgccggaggcgtataccccgaacatgctgcgtaag  
Q G E R Q E E P V P E A Y T P N M L R K

Number of amino acids 260  
Molecular weight 27953.96 g/mol  
Theoretical pI 5.20  
Ext. coefficient (280 nm) 58900 M<sup>-1</sup> cm<sup>-1</sup>

### 6.2.7. SWV NS2B-NS3 linked protease (R/A) (p255)

ggcagcagccatcatcatcatcatcacggtagcgcgagaatctgtattttccaaggcagcgcac  
 G S S H H H H H H G S E N L Y F Q G S D  
 atgtacatcgagaagggttgcgacattagctgggacaaggatgaggatcaccgggtacc  
 M Y I E K V C D I S W D K D A E I T G T  
 agccccgcgtctggatggtgcgctggacgatagcgggtgatttcagcctgattcaggacgat  
 S P R L D V A L D D S G D F S L I Q D D  
 ggtccgcccaccgcggtggcggtggcagcgggtggcggtggcagcggcgcgatgtgggac  
 G P P T A G G G G S G G G S G A M W D  
 atccccgagccccgcgtgaggttaagaaagggtgaaaccaccgcgggcggtgaccgtattatg  
 I P S P R E V K K G E T T A G V Y R I M  
 acccgtaaactgctgggtagcaccagggttggcgggcggtgatgcacgagggtgttttt  
 T R K L L G S T Q V G A G V M H E G V F  
 cacaccatgtggcatgtgaccaaaggcagcgcgctgcgtagcgggtgaaggccgctctggac  
 H T M W H V T K G S A L R S G E G R L D  
 ccgtactggggtaacgttaaacaagatctgatcagctattgcccccggtggaagctggac  
 P Y W G N V K Q D L I S Y C G P W K L D  
 ggtaaattgggatggcggttagcgcgaggtgcagctgatcgcgggtggcgccgggtgaaactgctg  
 G K W D G V S E V Q L I A V A P G E R A  
 cgtaacgttcaaaccaagccggggtgttcaaaaccaccgatgggtgaaattgggtgctgctg  
 R N V Q T K P G V F K T T D G E I G A L  
 gcgctggattttccgggtggcagcagcggtagcccgatcattgataagaacgggtcacggt  
 A L D F P G G S S G S P I I D K N G H V  
 atcggcctgtacggtaacggcggtggttgtaaaagcggcagctatgtgagcgcgattatg  
 I G L Y G N G V V V K S G S Y V S A I M  
 cagaccgagaagatggaagaaccggcggttgactgctttgaagaggacatgctgcgtaaa  
 Q T E K M E E P A V D C F E E D M L R K

Number of amino acids	260
Molecular weight	27748.97 g/mol
Theoretical pI	5.32
Ext. coefficient (280 nm)	43555 M <sup>-1</sup> cm <sup>-1</sup>

### 6.2.8. TBEV NS2B-NS3 linked protease (p251)

ggcagcagccatcatcatcatcatcacggcagcgcgagaatctgtactttcaaggtagccaa  
G S S H H H H H H G S E N L Y F Q G S Q  
ctggttgcggagtgagcgggttgcgttgagtgccacccggaactggtaacgaggggtggc  
L V A E W S G C V E W H P E L V N E G G  
gaagttagcctgcgtgtgcgtcaggacgcgatgggcaacttccacctgaccgagctggaa  
E V S L R V R Q D A M G N F H L T E L E  
aaagaggagcgtatgggtggcgggtggcagcgggtggcgggtggcagcgcgatctgggtgttcagc  
K E E R M G G G G S G G G S D L V F S  
ggtcaaggtggccgtgagcgtggtgaccgtccgtttgaagttaaggatggtgtgtaccgt  
G Q G G R E R G D R P F E V K D G V Y R  
atcttcagcccggcctgttttggggtcagaaccaagttggtgtgggctatggtagcaaa  
I F S P G L F W G Q N Q V G V G Y G S K  
ggcgttctgcacaccatgtggcatgtgaccctggtgctggcgcgtgagcattgacgatgcg  
G V L H T M W H V T R G A A L S I D D A  
ggtgcggtccgtagctggtggcggatgtgctgaggatgtggtttgctatggtggcgcgtgg  
V A G P Y W A D V R E D V V C Y G G A W  
agcctggaggaaaagtggaaaggtgaaaccgttcaagtgcgatgcgtttccgccgggcaag  
S L E E K W K G E T V Q V H A F P P G K  
gcgcatgaggttcaccagtccaaccgggtgaaactgatcctggacaccggccgtaagctg  
A H E V H Q C Q P G E L I L D T G R K L  
ggtgcatcccgattgatctggtgaaaggcaccagcggtagcccgattctgaacgcgcag  
G A I P I D L V K G T S G S P I L N A Q  
ggcgtggttgtgggtctgtacggcaacggctctgaaaaccaacgaaacctacgtgagcagc  
G V V V G L Y G N G L K T N E T Y V S S  
atcgcgcaaggtgaagcggaaaagagccgccgaatctgccgcaagcggttgtgggtacc  
I A Q G E A E K S R P N L P Q A V V G T  
ggttggaccagcaaa  
G W T S K

Number of amino acids 265  
Molecular weight 28436.58 g/mol  
Theoretical pI 5.79  
Ext. coefficient (280 nm) 54555 M<sup>-1</sup> cm<sup>-1</sup>

### 6.2.9. UTV NS2B-NS3 linked protease (K/A) (p254)

ggtagcagccatcatcatcatcatcatggtagcgcgagaacctgtactttcaaggcagcgcac  
 G S S H H H H H H G S E N L Y F Q G S D  
 ctgtggctggagcgtgcggcggacattacctgggaaagcgatgcggcgcattaccgggtacc  
 L W L E R A A D I T W E S D A A I T G T  
 agccagcgtctggacgtgaagctggatgacgatggcgattttcacctgatcaacgatccg  
 S Q R L D V K L D D D G D F H L I N D P  
 ggtgttccgtggg**gcg**gggtggcgggtggtagcggcgggtgggtgggtgggtggcgtgttctgggat  
 G V P W **A** G G G G S G G G G G V F W D  
 accccggcgcgcgtacacctatccgaagggtgataccagcccgggtgtgtaccgtatcatg  
 T P A P R T Y P K G D T S P G V Y R I M  
 acccgtcgtatcctgggtgcgtaccaaactcggcgttgggtgtgatgtacgaggggtgttctg  
 T R R I L G A Y Q I G V G V M Y E G V L  
 cacaccctgtggcacaccaccctgggtgcggcgattcgtagcgggtgaaggccgtctgacc  
 H T L W H T T R G A A I R S G E G R L T  
 ccgtattgggggttgcgtgaaggaagaccgtattacctacggcgggtccgtggaaactggac  
 P Y W G C V K E D R I T Y G G P W K L D  
 cgtaagtggaaacggctctggacgatgttcaactgatcgtgggttgcgcgggtaaaagcggcg  
 R K W N G L D D V Q L I V V A P G K A A  
 gtgaacatccagaccaaaccgggtatcttcaaaacccgcagggcgaaatcgggtgcgggt  
 V N I Q T K P G I F K T P Q G E I G A V  
 agcctggattaccgcgagcggtagcagcggtagcccgatcctggacaaaaacgggtgacatc  
 S L D Y P S G T S G S P I L D K N G D I  
 gttggtctgtacggtaacggcgttgtgctgggcaacggtagctacgttagcgcgattggt  
 V G L Y G N G V V L G N G S Y V S A I V  
 caaggcgaacgtgaagaagaaccgctgccggatgcgtacaatgcggacatgctgcgtaag  
 Q G E R E E E P L P D A Y N A D M L R K

Number of amino acids	260
Molecular weight	27911.06 g/mol
Theoretical pI	5.55
Ext. coefficient (280 nm)	60390 M <sup>-1</sup> cm <sup>-1</sup>

### 6.2.10. WNV NS2B-NS3 linked protease (K/A) (p252)

ggcagcagccatcatcatcatcacggcagcgagaacctgtactttcaaggcagcgac  
G S S H H H H H H G S E N L Y F Q G S D  
atgtggatcgaacgcaccgcgacattagctgggagagcgatgcggaaattaccggtagc  
M W I E R T A D I S W E S D A E I T G S  
agcgagcgtgtggacggttcgtctggacgatgacggtaacttccagctgatgaacgatccg  
S E R V D V R L D D D G N F Q L M N D P  
ggtgcccgtggcggtggcggtggcggtggcggtggcggtggcggtgctgtgggat  
G A P W A G G G G S G G G G G V L W D  
accccagcccgaaggagtacaagaaagggtgacaccaccaccggcgtttatcgtatcatg  
T P S P K E Y K K G D T T T G V Y R I M  
accctggtctgctgggcagctaccaagcgggtgcccggctgatggttgaagggtgtgtt  
T R G L L G S Y Q A G A G V M V E G V F  
cacaccctgtggcataaccaccaaagggtgctgcgctgatgagcggcgagggctcgtctggac  
H T L W H T T K G A A L M S G E G R L D  
ccgtactgggtagcgtgaaagaagaccgtctgtgctatggtggcccgtggaagctgcag  
P Y W G S V K E D R L C Y G G P W K L Q  
cacaaatggaacggccaggatgaggttcaaataattgtggttgaaccgggcaagaacgtg  
H K W N G Q D E V Q M I V V E P G K N V  
aaaaacggtcaaaccaagccggcggttcaaaccggagggtgaaatcggcgcggtt  
K N V Q T K P G V F K T P E G E I G A V  
accctggactttccgaccggtaccagcggcagcccgattgtggataaaaacggtgacggt  
T L D F P T G T S G S P I V D K N G D V  
attggcctgtacggtaacggcgtgatcatgccgaacggtagctatatcagcgcgattggt  
I G L Y G N G V I M P N G S Y I S A I V  
caaggcgaacgtatggacgaaccgattccggcgggtttgaaccggagatgctgcgtaag  
Q G E R M D E P I P A G F E P E M L R K

Number of amino acids 260  
Molecular weight 28004.15 g/mol  
Theoretical pI 5.14  
Ext. coefficient (280 nm) 55920 M<sup>-1</sup> cm<sup>-1</sup>



### 6.2.11. YFV NS2B-NS3 linked protease (p256)

```

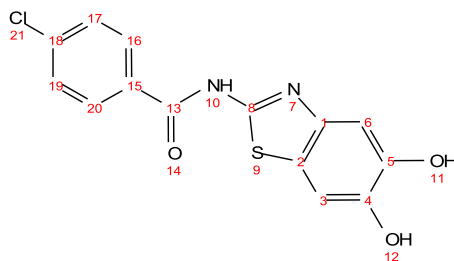
ggtagcagccaccaccaccatcatcacggtagcgcgagaatctgtactttcaaggtagcgggt
G S S H H H H H H G S E N L Y F Q G S G
ctggaactgaagaagctgggcgagggttagctgggaggaagaggcggagatcagcggtagc
L E L K K L G E V S W E E E A E I S G S
agcgcgcggttatgatgtggcgctgagcgcgagcagggcgaattcaagctgctgagcgaagag
S A R Y D V A L S E Q G E F K L L S E E
aaagttccgctgggatgggtggcggtggcagcgggtggcggtggcagcggcgacgtgctgtgg
K V P W D G G G G S G G G G S G D V L W
gatattccgacccccgaagatcattgaagagtgcgagcacctggaagacggtatctatggc
D I P T P K I I E E C E H L E D G I Y G
attttcaaagcacctttctgggtgagcagcagcgtggtgtggcggttgcgcaaggtggc
I F Q S T F L G A S Q R G V G V A Q G G
gtgttccacaccatgtggcacggtaccggtgctgtttctggtgctgtaacggcaagaaa
V F H T M W H V T R G A F L V R N G K K
ctgatcccgcgctgggcgagcgtgaaagaggacctggttgcgtacggtggcagctggaaa
L I P S W A S V K E D L V A Y G G S W K
ctggaaggtcgttgggatggcgaagaggaagtgcagctgattgctggcggttccgggtaaa
L E G R W D G E E E V Q L I A A V P G K
aacgtggttaacgttcaaaccaagccgagcctgttcaaagtgcgtaacggtggcgagatt
N V V N V Q T K P S L F K V R N G G E I
ggtgcggttgcgctggactaccgagcgggtaccagcggcagcccgattgtgaaccgtaac
G A V A L D Y P S G T S G S P I V N R N
ggtgaagttatcggcctgtatggtaacggcatttctggtgggtgataacagctttgtagc
G E V I G L Y G N G I L V G D N S F V S
gcgatcagccagaccgaagtgaagaagagggtaaagaagaactgcaagagattccgacc
A I S Q T E V K E E G K E E L Q E I P T
atgctgaaaaaa
M L K K

```

Number of amino acids	264
Molecular weight	28329.56 g/mol
Theoretical pI	5.09
Ext. coefficient (280 nm)	47440 M <sup>-1</sup> cm <sup>-1</sup>

## 6.3. Synthesised compounds

### 6.3.1. 4-Chloro-*N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)benzamide [10]



Melting point: 351.2 °C

$^1\text{H-NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 7.95 – 7.84 (m, 2H, C16/C20-H), 7.58 – 7.47 (m, 2H, C17/C19-H), 7.36 (s, 1H, C3-H), 7.29 (s, 1H, C6-H).

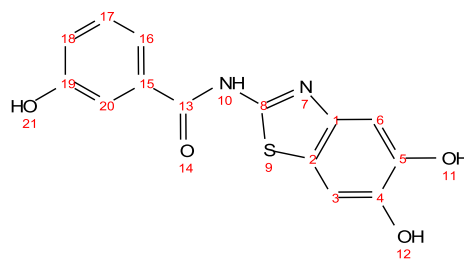
$^{13}\text{C-NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 168.20 (C13), 157.18 (C8), 145.69 (C5), 144.74 (C1), 142.13 (C4), 138.08 (C18), 132.31 (C15), 129.24 (C16/C20), 128.96 (C17/C19), 126.41 (C2), 107.97 (C3), 103.81 (C6).

FT-IR:  $\nu$  [ $\text{cm}^{-1}$ ] = 3356, 3091, 1681, 1592, 1487, 1471, 1448, 1408, 1366, 1311, 1286, 1245, 1213, 1149, 1093, 1036, 1012, 906, 864, 843, 787, 745, 676.

ESI-MS: 321.0 [ $\text{M-H}^+$ ]- (calculated for  $\text{C}_{14}\text{H}_9\text{ClN}_2\text{O}_3\text{S}$ :  $M = 320.0$ )

Purity: 88%

### 6.3.2. *N*-(5,6-Dihydroxybenzo[*d*]thiazol-2-yl)-3-hydroxybenzamide [11]



Melting point: 344.2 °C

<sup>1</sup>H-NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 7.53 – 7.37 (m, 2H, C16/C6-H), 7.37 – 7.26 (m, 3H; C20/C17/C3-H), 6.98 (dd, *J* = 7.8, 2.2 Hz, 1H, C18-H).

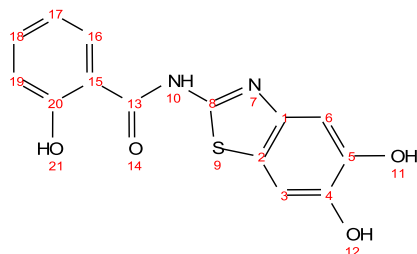
<sup>13</sup>C-NMR (75 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 166.58 (C13), 159.18 (C8), 157.27 (C19), 145.69 (C5), 144.74 (C1), 143.11 (C4), 133.83 (C15), 129.77 (C17), 126.41 (C2), 120.04 (C16/C18), 112.81 (C20), 109.97 (C3), 103.18 (C6).

FT-IR:  $\nu$  [cm<sup>-1</sup>] = 3390, 3265, 3053, 1649, 1599, 1582, 1547, 1495, 1474, 1431, 1364, 1288, 1265, 1223, 1215, 1171, 1148, 1099, 1043, 1035, 1010, 905, 870, 842, 831, 788, 730, 717, 680.

ESI-MS: 303.0 [M-H<sup>+</sup>]- (calculated for C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>4</sub>S: M = 302.04)

Purity: 100%

### 6.3.3. 2-((5,6-Dihydroxybenzo[d]thiazol-2-yl)carbamoyl)phenyl acetate [12]



Melting point: 265.3 °C

$^1\text{H-NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 12.03 (s, 1H, -NH), 9.31 (s, 2H, Ar-OH), 8.00 (dd,  $J = 7.9, 1.7$  Hz, 1H, C16-H), 7.46 (t,  $J = 7.8$  Hz, 1H, C18-H), 7.25 (s, 1H C6-H), 7.07 (s, 1H, C3-H), 6.98 (q, 2H, C17/C18-H).

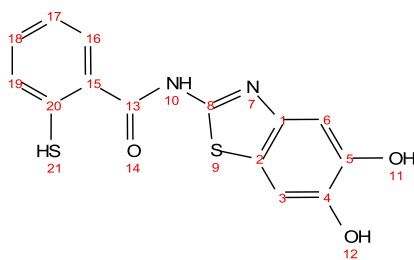
$^{13}\text{C-NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 167.55 (C13), 159.38 (C8), 150.62 (C20), 146.58 (C5), 144.55 (C1), 143.62 (C4), 134.83 (C18), 130.82 (C16), 119.83 (C2), 117.46 (C15/C17/C19), 110.63 (C3), 105.10 (C6).

FT-IR:  $\nu$  [ $\text{cm}^{-1}$ ] = 3447, 3219, 2478, 1826, 1650, 1602, 1536, 1480, 1446, 1355, 1302, 1222, 1161, 1119, 1033, 910, 859, 840, 789, 755, 718, 685.

ESI-MS: 303.0 [ $\text{M-H}^+$ ]- (calculated for  $\text{C}_{14}\text{H}_{10}\text{N}_2\text{O}_4\text{S}$ :  $M = 302.04$ )

Purity: 70 %

### 6.3.4. *N*-(5,6-Dihydroxybenzo[*d*]thiazol-2-yl)-2-mercaptobenzamide [13]



Melting point: 326.2 °C

<sup>1</sup>H-NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 12.82 (s, 1H, -NH), 7.81 (s, 1H, C3-H), 7.64 (s, 1H, C6-H), 7.57 – 7.51 (m, 2H, C17/C18), 7.49 – 7.42 (m, 2H, C16/C19).

<sup>13</sup>C-NMR (75 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 169.43 (C13), 159.66 (C8), 145.70 (C5), 144.74 (C1), 143.11 (C4), 131.14 (C20), 130.67 (C18/C19), 129.71 (C15/C2), 129.13 (C16), 126.41 (C17), 109.67 (C3), 104.18 (C6).

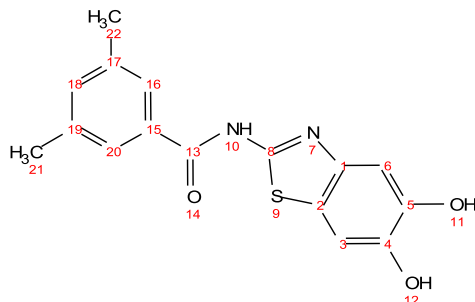
FT-IR: ν [cm<sup>-1</sup>] = 3111, 2524, 2328, 2224, 1664, 1562, 1534, 1496, 1288, 1237, 1054, 899, 742.

ESI-MS: 319.0 [M-H<sup>+</sup>]- (calculated for C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>3</sub>S<sub>2</sub>: M = 318.01)

Purity: 41%

### 6.3.5. *N*-(5,6-Dihydroxybenzo[*d*]thiazol-2-yl)-3,5-dimethylbenzamide

[14]



Melting point: 214.2 °C

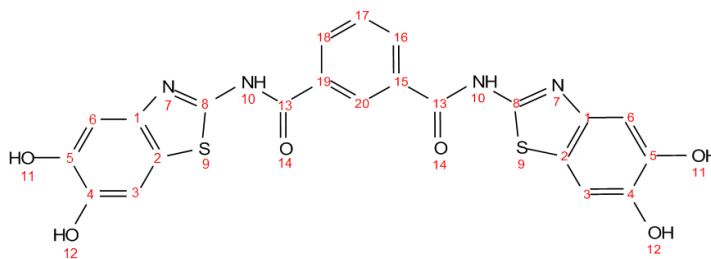
<sup>1</sup>H-NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 12.06 (s, 1H, -NH), 9.28 (s, 2H, Ar-OH), 7.73 (s, 1H, C16-H), 7.34 (s, 1H, C20-H), 7.27 (s, 1H, C18-H), 7.25 (s, 1H, C6-H), 7.12 (s, 1H, C3-H), 2.36 (s, 6H).

<sup>13</sup>C-NMR (75 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 169.75 (C13), 162.02 (C8), 147.53 (C5/C1), 145.32 (C4), 139.80 (C17/C19/C15), 136.25 (C18), 126.25 (C16/C20/C2), 109.04 (C3), 105.36 (C6), 21.17 (C21).

FT-IR: ν [cm<sup>-1</sup>] = 3146, 2906, 1681, 1601, 1561, 1532, 1455, 1381, 1347, 1207, 1163, 1104, 1044, 861, 805, 750, 684.

ESI-MS: 315.1 [M-H<sup>+</sup>]-(calculated for C<sub>16</sub>H<sub>14</sub>N<sub>2</sub>O<sub>3</sub>S: M = 314.07)

Purity: 100%

**6.3.6. N1,N3-Bis(5,6-dihydroxybenzo[d]thiazol-2-yl)isophthalamide [15]**

Melting point: 294.2 °C

<sup>1</sup>H-NMR (300 MHz, DMSO-d<sub>6</sub>) δ [ppm]: 12.42 (s, 1H, -NH), 8.82 (d, J = 1.8 Hz, 1H, C20), 8.31 (dd, J = 7.8, 1.7 Hz, 2H, C16/C18-H), 7.73 (t, J = 7.8 Hz, 1H, C17-H), 7.28 (s, 2H, C3-H), 7.15 (s, 2H, C6-H).

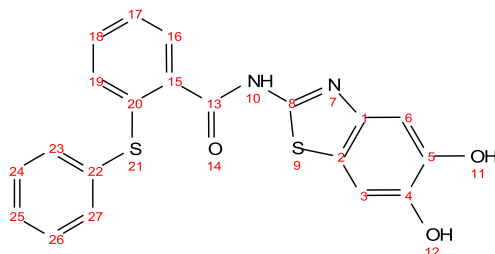
<sup>13</sup>C-NMR (75 MHz, DMSO-d<sub>6</sub>) δ [ppm]: 166.69 (C13), 159.18 (C8), 145.69 (C5), 144.74 (C1), 143.11 (C4), 132.88 (C15), 131.20 (C16), 129.41 (C20), 128.81 (C17), 126.41 (C2), 109.97 (C3), 103.18 (C6).

FT-IR: ν [cm<sup>-1</sup>] = 3311, 3152, 1681, 1656, 1553, 1481, 1451, 1367, 1279, 1212, 1143, 1040, 841, 787, 748, 704.

ESI-MS: 495.5 [M-H<sup>+</sup>]- (calculated for C<sub>22</sub>H<sub>14</sub>N<sub>4</sub>O<sub>6</sub>S<sub>2</sub>: M = 494.04)

Purity: 100%

**6.3.7. *N*-(5,6-Dihydroxybenzo[*d*]thiazol-2-yl)-2-(phenylthio)benzamide**  
**[16]**



Melting point: 287.2 °C

$^1\text{H-NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 7.73 (d,  $J = 7.6$  Hz, 1H, C16-H), 7.47 – 7.35 (m, 7H, Ar-H), 7.25 (s, 1H, C3-H), 7.13 (s, 1H, C6-H), 7.09 (d,  $J = 7.8$  Hz, 1H, C19-H).

$^{13}\text{C-NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 166.88 (C13), 158.66 (C8), 145.69 (C5), 144.74 (C1), 143.11 (C4), 136.72 (C20), 134.77 (C15), 134.07 (C22), 132.28 (C18), 131.02 (C16), 130.70 (C23/C27), 130.13 (C17), 129.26 (C19), 129.24 (C24/C26), 126.41 (C2), 126.13 (C25), 109.97 (C3), 103.18 (C6).

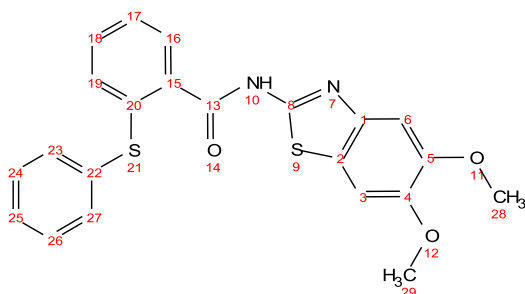
FT-IR:  $\nu$  [ $\text{cm}^{-1}$ ] = 3293, 3202, 3049, 2221, 2038, 1664, 1586, 1543, 1470, 1439, 1360, 1287, 1223, 1209, 1146, 1089, 1045, 1031, 904, 864, 837, 788, 751, 730, 707, 687.

ESI-MS: 395.1 [ $\text{M-H}^+$ ]- (calculated for  $\text{C}_{20}\text{H}_{14}\text{N}_2\text{O}_3\text{S}_2$ :  $M = 394.04$ )

Purity: 100%



***N*-(5,6-Dimethoxybenzo[*d*]thiazol-2-yl)-2-(phenylthio)benzamide [Pro-16]**



Melting point: 328.8 °C

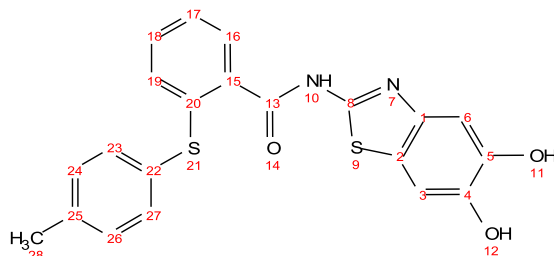
<sup>1</sup>H-NMR (300 MHz, DMSO-d<sub>6</sub>) δ [ppm]: 12.76 (s, 1H, -NH), 7.75 (dd, J = 7.6, 1.6 Hz, 1H, C16-H), 7.60 (s, 1H, C6-H), 7.51-7.42 (m, 3H, C27/C23/C18-H), 7.41-7.32 (m, 4H, C24/C26/C25-H), 7.33 (s, 1H, C3-H) 7.11 (dd, J = 7.9, 1.2 Hz, 1H, C19-H), 3.83 (s, 6H, C29/C28-H).

<sup>13</sup>C-NMR (75 MHz, DMSO-d<sub>6</sub>) δ [ppm]: 162.76 (C13), 156.52 (C8), 157.13 (C5), 149.42 (C1), 147.56 (C4), 143.32 (C20), 142.78 (C15), 134.91 (C22), 134.00 (C18), 133.08 (C23,C27) 131.96 (C16), 130.74 (C17), 129.82 (C24,C26), 129.21 (C19), 128.59 (C2), 126.86 (C25), 106.34 (C3), 104.03 (C6), 56.54 (C28), 56.25 (C29).

ESI-MS: 423.2 [M-H<sup>+</sup>]-(berechnet: M = 422.08)

FT-IR: ν [cm<sup>-1</sup>] = 1660, 1544, 1486, 1435, 1286, 1051, 900, 786, 737.

**6.3.8. *N*-(5,6-Dihydroxybenzo[*d*]thiazol-2-yl)-2-(*p*-tolylthio)benzamide**  
**[17]**



Melting point: 258.4 °C

<sup>1</sup>H-NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ [ppm]: 12.57 (s, 1H, -NH), 9.17 (s, 2H, -OH) 7.72 (d, *J* = 7.6 Hz, 1H, C16-H), 7.44 – 7.19 (m, 7H), 7.13 (s, 1H, C6-H), 6.98 (d, *J* = 7.8 Hz, 1H, C3-H), 2.32 (s, 3H, -CH<sub>3</sub>).

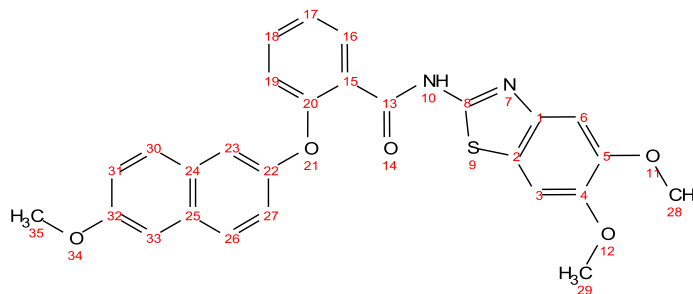
<sup>13</sup>C-NMR (75 MHz, DMSO-*d*<sub>6</sub>) δ [ppm:] 173.32 (C13), 168.42 (C8), 146.10 (C5), 145.98 (C1), 145.73 (C4), 144.50 (C16), 138.89 (C15), 134.16 (C25), 134.23 (C19), 133.91 (C20), 132.7 (C26), 131.89 (C23), 131.88 (C26), 130.93 (C24), 130.08 (C22), 129.66 (C18), 129.21 (C17), 126.17 (C2), 95.88 (C3), 93.02 (C6), 21.19 (C28).

FT-IR: ν [cm<sup>-1</sup>] = 3264, 3055, 1966, 1688, 1587, 1547, 1469, 1361, 1288, 1228, 1210, 1145, 1096, 1032, 905, 862, 838, 788, 751, 731, 715, 681.

ESI-MS: 409.1 [M-H<sup>+</sup>]-(calculated for C<sub>21</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>S<sub>2</sub>: M = 408.06)

Purity: 99%

**6.3.9. *N*-(5,6-Dimethoxybenzo[*d*]thiazol-2-yl)-2-((6-methoxynaphthalen-2-yl)oxy)benzamide [Pro-1]**



Melting point: 201.8 °C

$R_f = 0.77$  (Cy/EtOAc = 1:1)

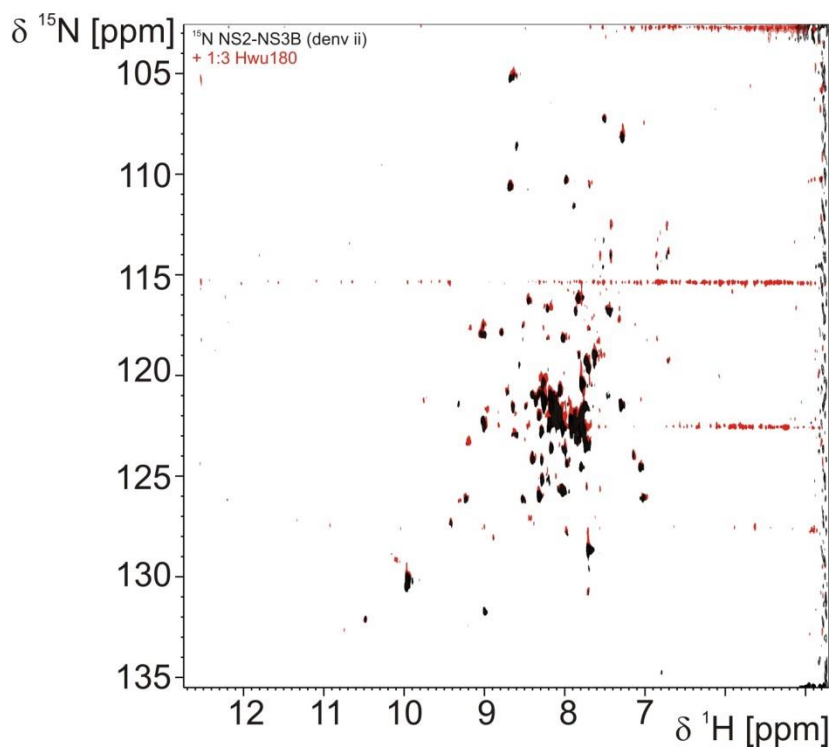
$^1\text{H-NMR}$  (300 MHz, DMSO- $d_6$ )  $\delta$  [ppm]: 12.37 (s, 1H, -NH), 7.89 (d,  $J = 8.9$  Hz, 1H, C16-H), 7.81 (d, 1H, C26-H) 7.77 (m, 1H, C30-H), 7.55 (s, 1H, C18-H), 7.52 (dd,  $J = 8.0, 2.0$  Hz, 2H, C19/C33-H), 7.36 (m, 2H, C23/C17-H), 7.30 (s, 1H, C6-H), 7.26 (s, 1H, C3-H), 7.17 (dd,  $J = 9.0, 2.5$  Hz, 1H, C31-H), 6.95 (dd,  $J = 8.3, 1.0$  Hz, 1H, C27-H), 3.93 (s, 3H, -OCH<sub>3</sub>) 3.84 (s, 6H, -OCH<sub>3</sub>).

$^{13}\text{C-NMR}$  (75 MHz, DMSO- $d_6$ )  $\delta$  [ppm] = 164.89 (-CONH), 157.28 (C8, C32), 155.44 (C20, C22), 152.05 (C5), 149.39 (C1), 147.53 (C4), 142.95 (C18), 133.45 (C25), 131.92 (C16), 130.72 (C24), 129.31 (C30), 125.61 (C26), 123.52 (C2), 121.01 (C17), 119.81 (C15), 118.38 (C27), 116.05 (C19), 106.56 (C31), 104.08 (C23, C33), 56.39 (2x -OCH<sub>3</sub>), 55.91 (-OCH<sub>3</sub>).

ESI-MS: 487.3 [M-H<sup>+</sup>]- (berechnet: M = 486.12)

FT-IR:  $\nu$  [cm<sup>-1</sup>] = 3349, 3066, 2994, 2938, 2842, 1665, 1601, 1532, 1473, 1444, 1291, 1217, 1154, 1024, 987, 939, 883, 715.

## 6.4. NMR spectrum of the DENV 2 NS2B-NS3 linked protease



**Figure 6.1:** NMR spectrum of the DENV 2 NS2B-NS3 linked protease. Black signals represent the DENV 2 NS2-NS3 protease without inhibitor and the red signals the protease after addition of 1 (protease/inhibitor, 1:3).

### 6.5. Homology model of ZIKV NS2B-NS3



Figure 6.2: Structure alignment of the ZIKV NS2B-NS3 protease homology model with the template 2M9P resulting in a RMSD value of 0.880.

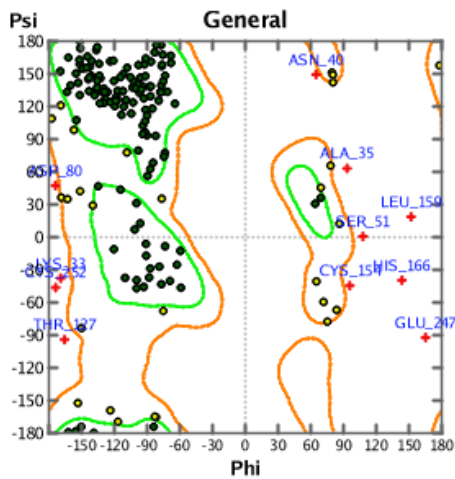


Figure 6.3: Phi-Psi Plot of the ZIKV NS2B-NS3 protease homology model.

## 6.6. Mass spectrometry of WT ZIKV NS2B-NS3 protease

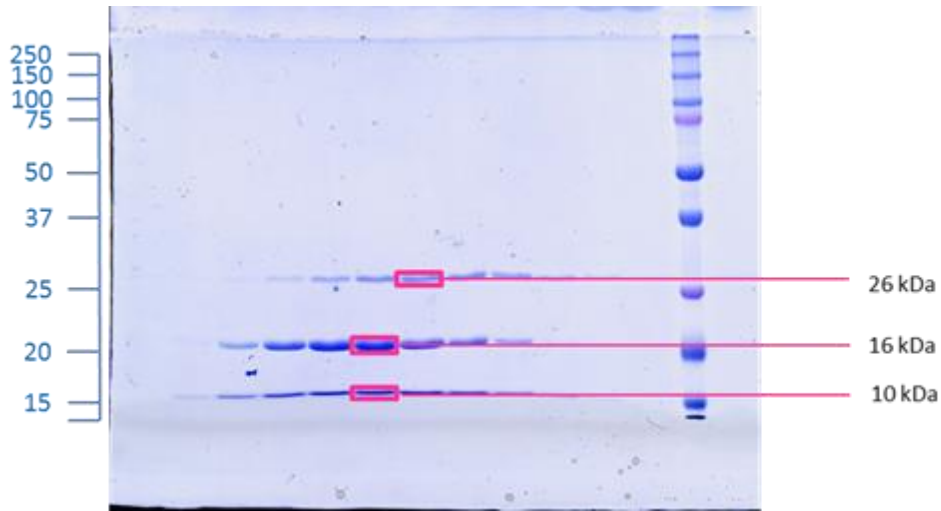


Figure 6.4: SDS-page (12%) from 28.02.2016. Protein band extraction for mass spectrometry analysis of WT ZIKV NS2B-NS3 protease experiencing autocatalytic cleavage highlighted in magenta.

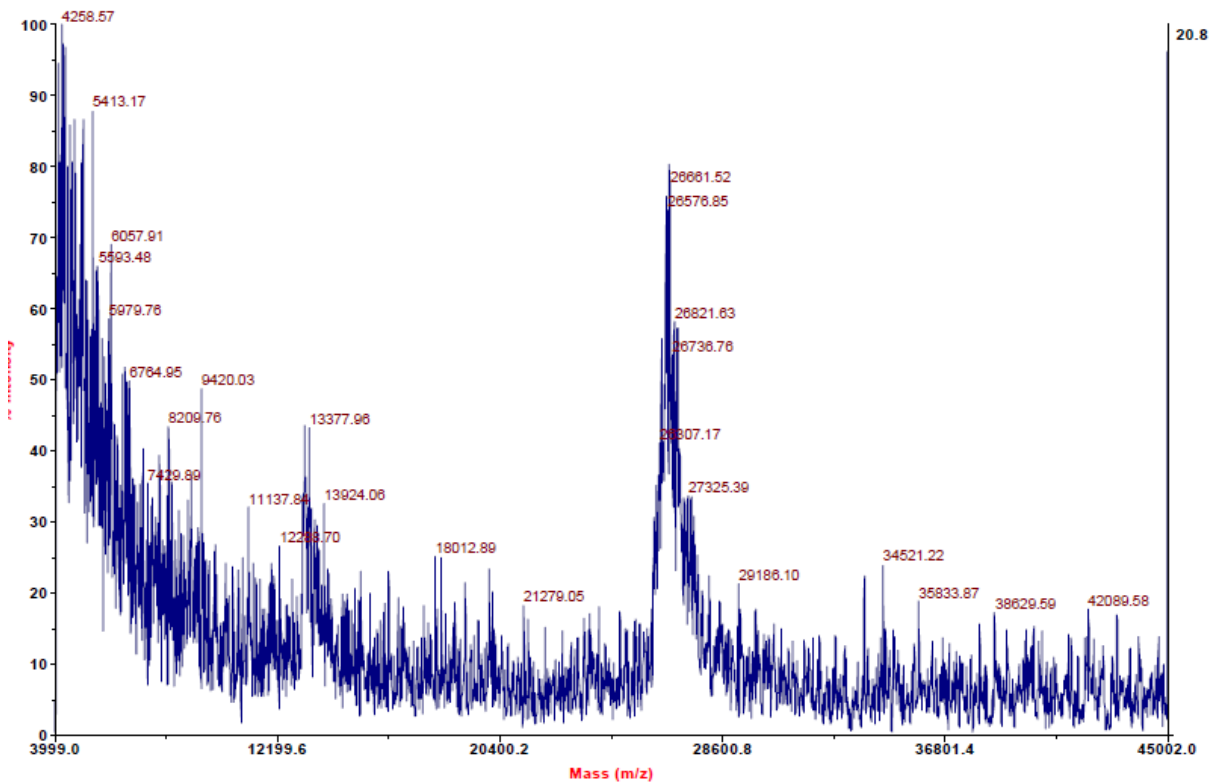


Figure 6.5: Mass spectrum of the ~ 26 kDa protein band (see Figure 6.4).

## Appendix

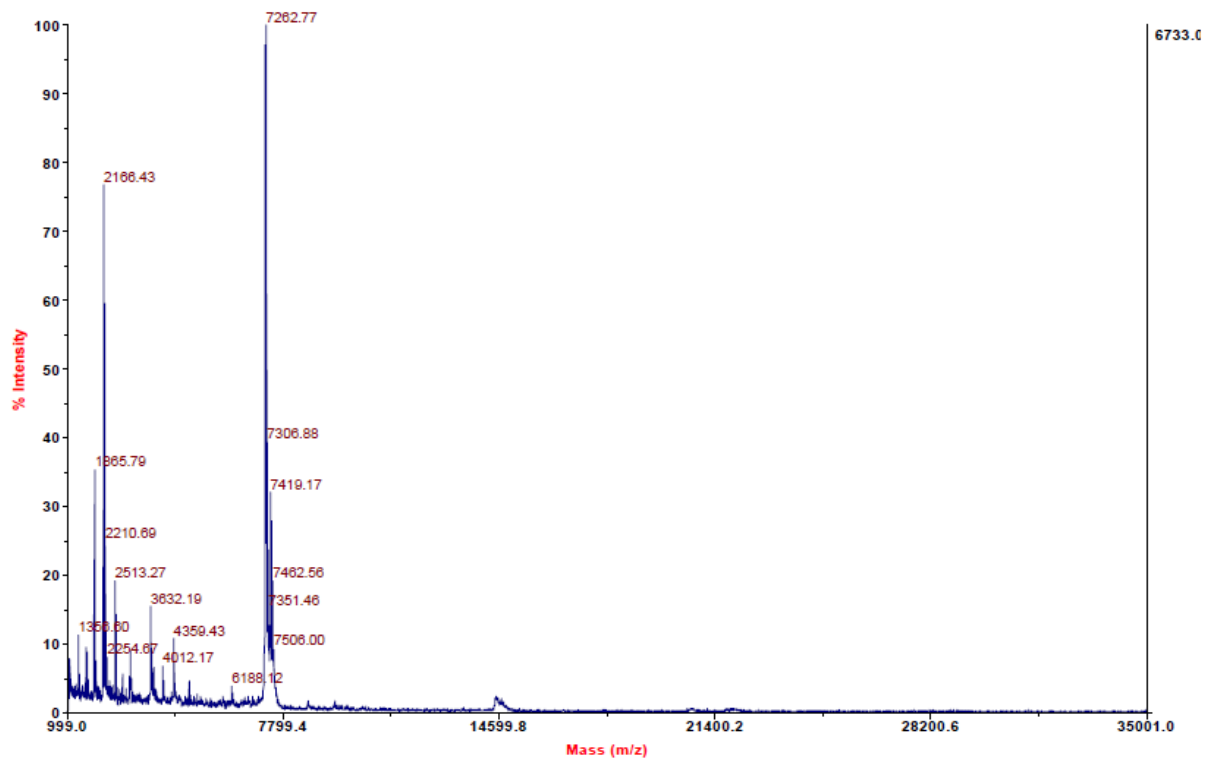


Figure 6.6: Mass spectrum of the ~ 10 kDa protein band (see Figure 6.4).

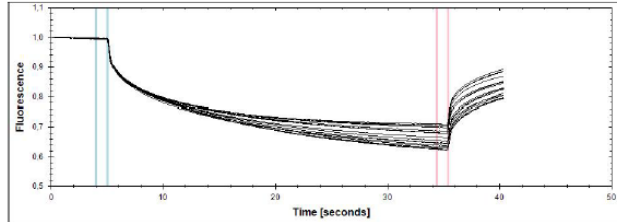
## 6.7. MST measurements

### 6.7.1. Linked ZIKV NS2B-NS3 (R95A) protease with 1

#### Report

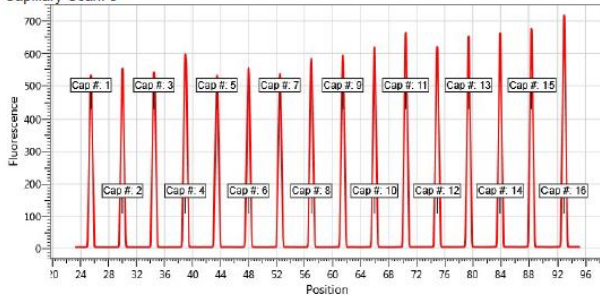
Ntp-Filename: 20161115 ZIKV R65A (150nM), HWu 180 (625uM)  
 Experiment Name: Experiment 15.11.2016 11:26:34  
 MST Power: 40% , LED Power: 60%

Normalized Fluorescence Timetrace

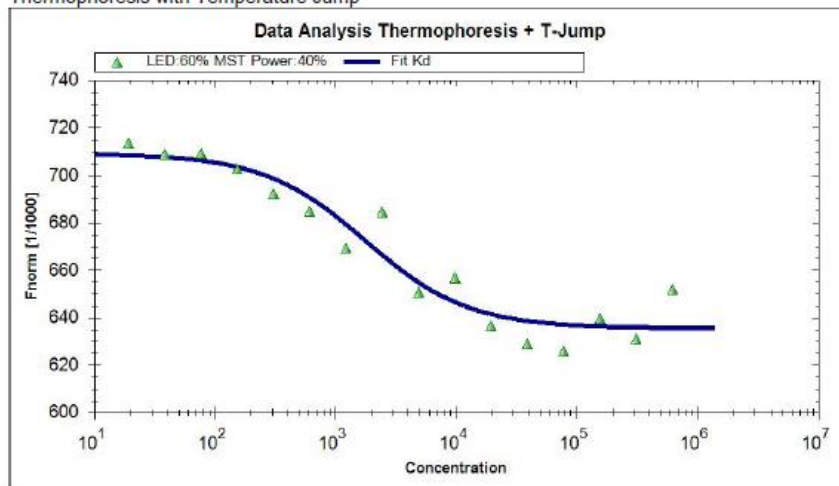


hot region: HotStart=4 HotLength=1  
 cold region: ColdStart=34.4 ColdLength=1

Capillary-Scan: 0



Thermophoresis with Temperature Jump



Fitting for Kd Formular

Fitted Parameter	Fitted Value
Dissociation Constant	1700+/-231
Fluo.Conc	150
Bound	635.59
Unbound	709.22
Amplitude	73.63

Kd Formula (law of mass action)

$$f(c) = \text{unbound} + (\text{bound} - \text{unbound}) / 2 * (\text{FluoConc} + c + \text{Kd} - \sqrt{(\text{FluoConc} + c + \text{Kd})^2 - 4 * \text{FluoConc} * c})$$

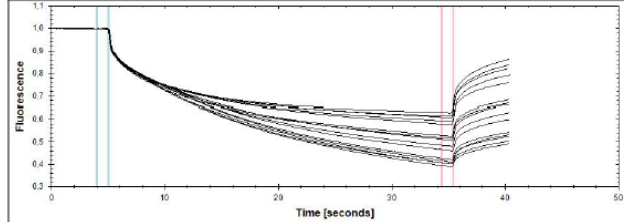


## 6.7.2. Linked ZIKV NS2B-NS3 (S135A) protease with 1

### Report

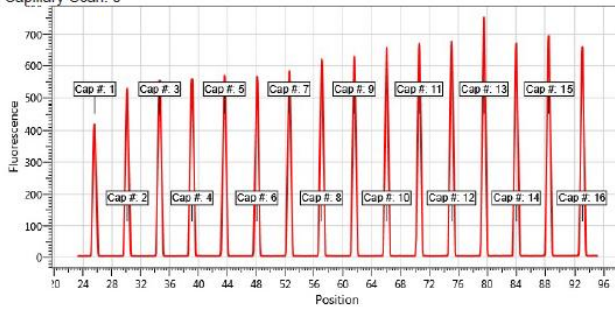
Ntp-Filename: ZIKV HWu180  
 Experiment Name: ZIKV S209A 150nM, HWu180 500uM 19.09.2017 17:24:58  
 MST Power: 60% , LED Power: 50%

Normalized Fluorescence Timetrace

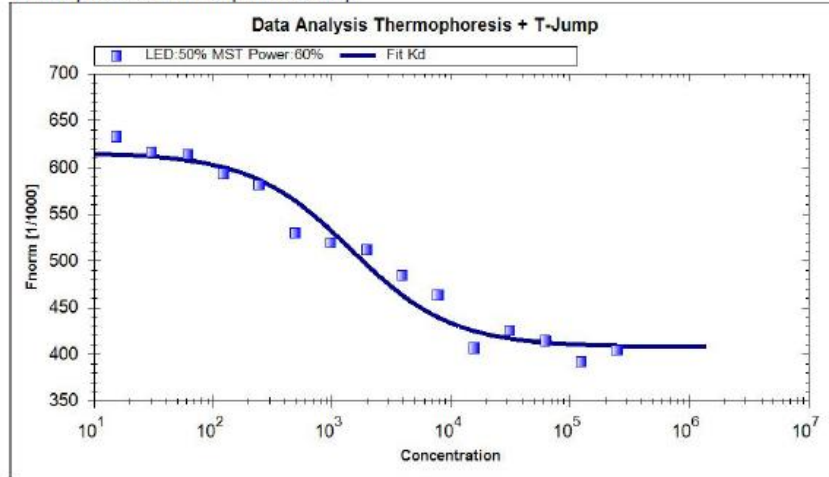


hot region: HotStart=4 HotLength=1  
 cold region: ColdStart=34,4 ColdLength=1

Capillary-Scan: 0



Thermophoresis with Temperature Jump



Fitting for Kd Formular

Fitted Parameter	Fitted Value
Dissociation Constant	1370+/-119
Fluo.Conc	150
Bound	408.21
Unbound	615.74
Amplitude	207.53

Kd Formula (law of mass action)

$$f(c) = \text{unbound} + (\text{bound} - \text{unbound}) / 2 * (\text{FluoConc} + c + Kd - \sqrt{(\text{FluoConc} + c + Kd)^2 - 4 * \text{FluoConc} * c})$$

## 6.8. Crystallisation of the linked ZIKV NS2B-NS3 (R95A) protease

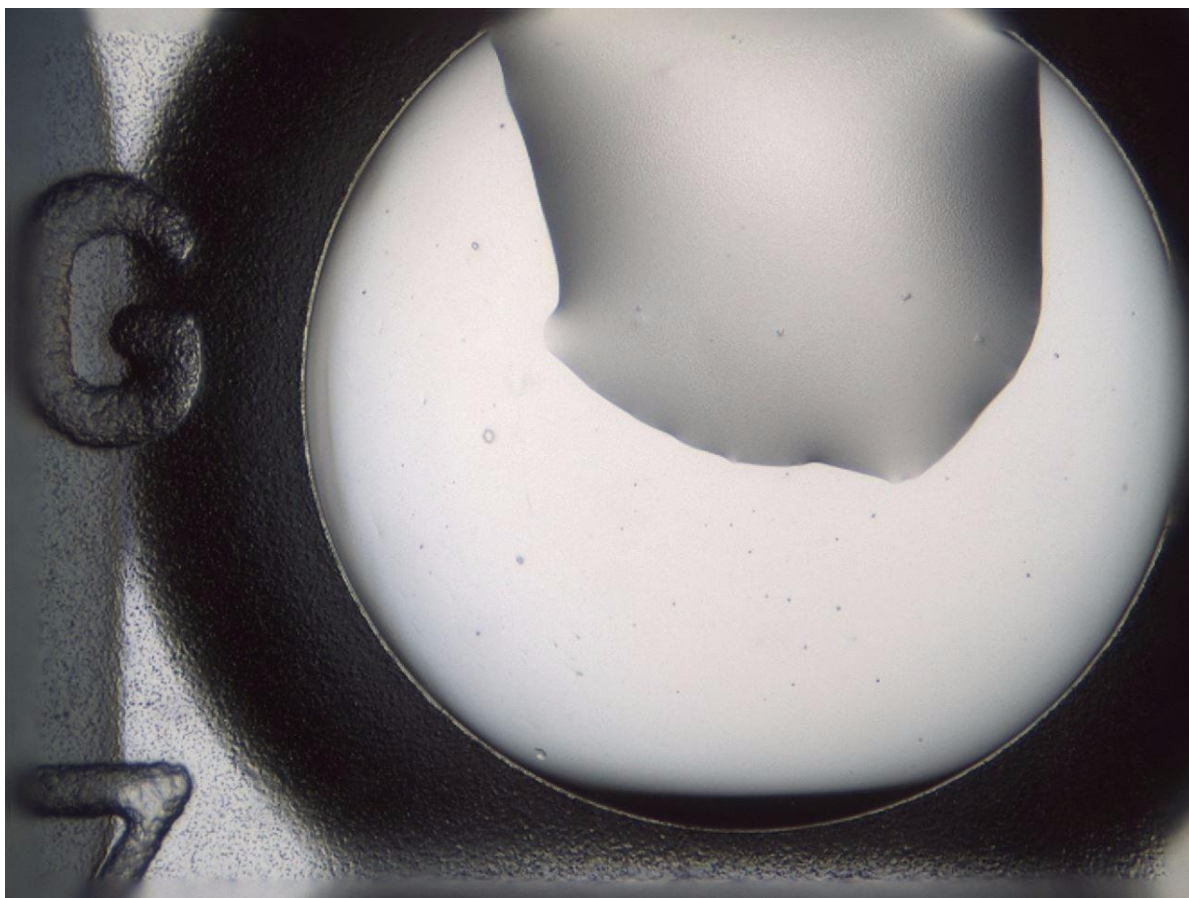


Figure 6.7: Picture was taken from the documentation system Formulatrix Rock Imager™, MarXtal University Marburg. Conditions: 35% PEG 200, 100 mM MES, pH 9.

## 6.9. Phylogenetic trees

### 6.9.1. Evolution of flaviviral polyprotein

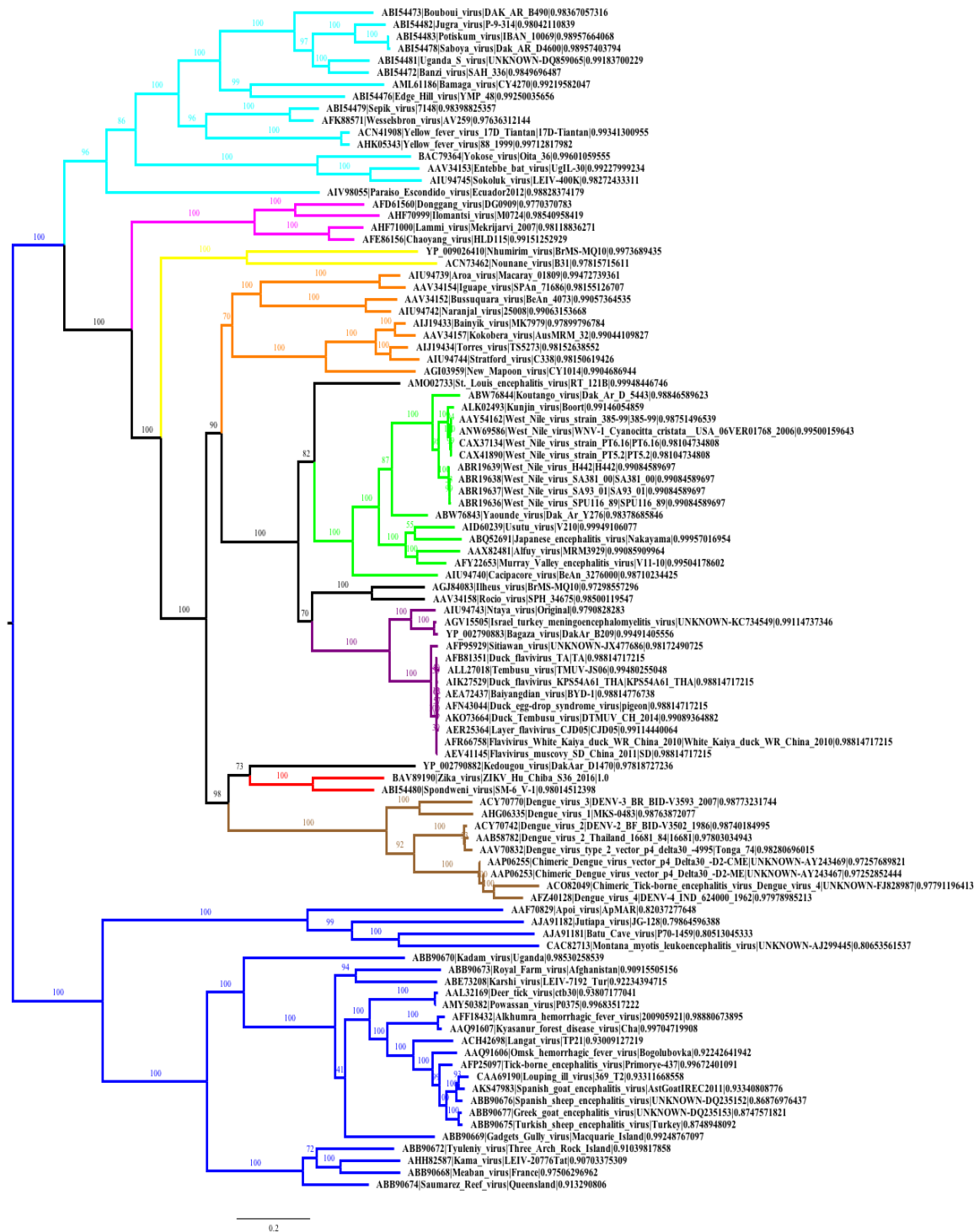


Figure 6.8: The evolutionary history of *flaviviruses* based on the entire polyprotein.

## 6.9.2. Evolution of flaviviral NS2B-NS3 protease

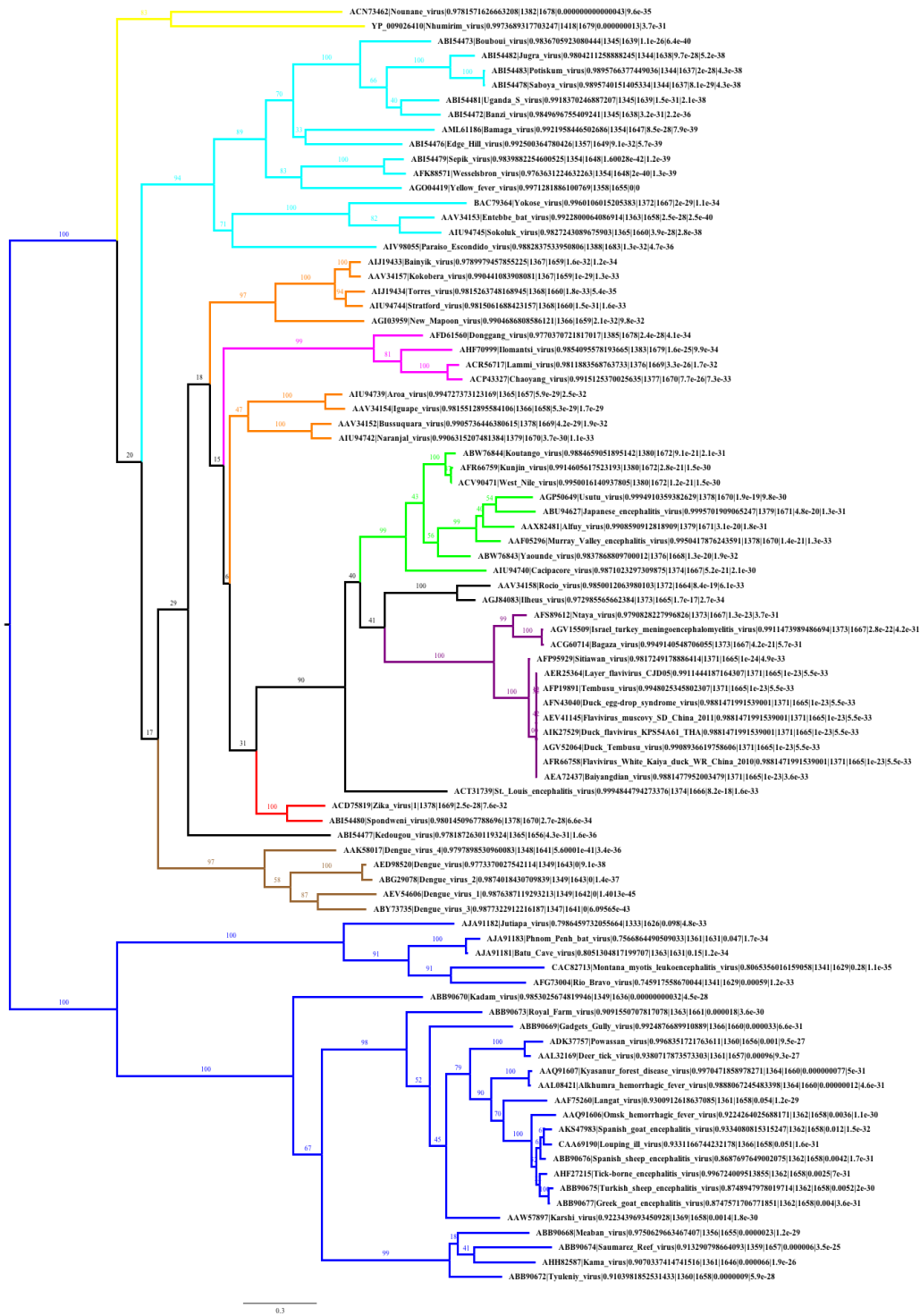


Figure 6.9: The evolutionary history of *flaviviruses* based on the NS2B-NS3 protease sequence.

### 6.9.3. Alignment of additionally selected flaviviral constructs

TBEV	-----GSSHHHHHHGSENLYFQGSQLVAEWSGCVEWHPELVNEGGEVSLRVRODAMGNFH	55
YFV	-----GSSHHHHHHGSENLYFQGSGLLELKLGEVSWEEEAIEISGSSARYDVALSEQGEFK	55
ZIKV	-----GSSHHHHHHGSENLYFQGSMDYIERAGDITWEKDAEVTGNSPRLDVALDESDFDS	55
SWV	-----GSSHHHHHHGSENLYFQGSMDYIEKVCDISWDKAEITGTSRPLDVALDDSGDFDS	55
WNV	-----GSSHHHHHHGSENLYFQGSMDMIERTADISWESDAEITGSSERVDVRLDDGDNFQ	55
UTV	-----GSSHHHHHHGSENLYFQGSDLWLERAADITWESDAITGTSQRLDVKLDDDGDFH	55
JEV	-----GSSHHHHHHGSENLYFQGSMDWLERAADISWEMDAITGSSRRLDVKLDDDGDFH	55
DENV2	MGSSHHHHHSSGLVPRGSHMLAADLELERAADVRWEEQAEISGSSPILSITISEDGSMS	60
DENV4	-----GSSHHHHHHGSENLYFQGSDLLEKAAVQWDEMADITGSSPIIEVQKQDEDSGFS	55
DENV1	-----GSSHHHHHHGSENLYFQGSDLLEKAAEVSWEEEAIEHSGASHNILVEVQDDGTMK	55
DENV3	-----GSSHHHHHHGSENLYFQGSDLTVEKAAADVTEEEAEQTGVSHNLMITVDDGTMK	55
	* * . . : : : : : : * . : . * :	
TBEV	LTELEKEERMGGGGSGGGSDLVFSGQGRERGRDRPFVVKDGVYRIFSPGLFWGQNVQVW	115
YFV	LLSEKVPWDGGGGSGGGSDLVLDIPTPKIIEECEHLEDGIYGIQSTFL-GASQRGV	114
ZIKV	LVEDDGPMPRGGGGSGGGSG-ALWDVPAPKEVKKG-ETTDGVYRVMTRRL-LGSTQVGV	112
SWV	LIQDDGPPTRGGGGSGGGSG-AMWDIPSPREVKKG-ETTAGVYRIMTRKLL-LGSTQVGA	112
WNV	LMNDPGAPWKGGGGGSGGGGG-VLWDTPSPKEYKKG-DTTTGYYRIMTRGLL-GSYQAGA	112
UTV	LINDPGVPWKGGGGSGGGGG-VFWDTPSPRTYPKG-DTSPGVYRIMTRRL-GAYQIGV	112
JEV	LIDDPGVPWKGGGGGSGGGGG-VFWDTPSPKPCSKG-DTTTGYYRIMARGIL-GTYQAGV	112
DENV2	IKNEEEEQTLGGGGSGGGGAG-VLWDVPSPPVVGKA-ELEDGAYRIKQKGA-GYSQIGA	117
DENV4	IRDVEETNMIGGGSGGGGG-ALWDVPSPAATKKA-ALSEGYYRIMQRGLF-GKTQVGV	112
DENV1	IKNEERDDTLGGGGSGGGGG-VLWDTPSPPEVERA-VLDDGIYRIMQRGLL-GRSQVGV	112
DENV3	IKDDETENILGGGGSGGGGG-VLWDVPSPPETQKA-ELEEGYYRIKQQGIF-GKTQVGV	112
	: . * * * * * . . : . . . * * : * * *	
TBEV	GYGSKVLHTMWHVTRGAALSIDDAVAGPYWADVRELVVVCYGGAWSLEEKWKG-ETVQVH	174
YFV	GVAQGGVFHTMWHVTRGAFLVRNGKLI PSWASVKEDLVAYGGSWKLEGRWDGEEVQLI	174
ZIKV	GVMQEGVFHTMWHVTRKGSALRSGEGRLDPIYWGDKQLVSYCGPWKLDAAWDGHSEVQLL	172
SWV	GVMHEGVFHTMWHVTRKGSALRSGEGRLDPIYWGDKQLVSYCGPWKLDGKWDGVSEVQLI	172
WNV	GVMVEGVFHTLWHTTKGAALMSGEGRLDPIYWGSVKEDRLCYGGPWKLDKHWNGQVQMI	172
UTV	GVMYEGVLHTLWHTTRGAAIRSGEGRLDPIYWGCVKEDRIYGGPWKLDKRWGLDDVQLI	172
JEV	GVMYENVFHTLWHTTRGAAIRSGEGRLDPIYWGSVKEDRIAYGGPWRFRDKWNGTDDVQVI	172
DENV2	GVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQVL	177
DENV4	GIHMEGVFHTMWHVTRGAVLMSGEGKLTPIYWGSVKEDRIAYGGPWRFRDKWNGTDDVQVL	172
DENV1	GVFQENVFHTMWHVTRGAVLMYQGRLEPSWASVKKDLISYGGGWRLQGSWNTGEEVQVI	172
DENV3	GVQKEGVFHTMWHVTRGAVLTHNGKRLEPNWASVKKDLISYGGGWRLSAQWQKGEVQVI	172
	* . . : * * * * * : : * * . * * * : * . . * * :	
TBEV	AFPPGKAHEVHQCPGELIDTGRKLGAIPIIDLKGTSGSPILNAQGVVVGLYGNGLKT-	233
YFV	AAVPGKNVNVNQTKPSLFKVRNNGEIGAVALDYPSGTSGPSIVNRNGEVIGLYGNGLLVG	234
ZIKV	AVPPGERARNIQTLPGIFKTK-DGDIGAVALDYPSGTSGPSILDKCGRVIGLYGNGVVVK	231
SWV	AVAPGERARNVQTKPGVFKTT-DGEIGALALDFPSSGPSIIDKNHVIIGLYGNGVVVK	231
WNV	VVEPGKNVNVNQTKPGVFKTP-EGEIGAVTLDFFPTGTSGPSIVDKNGDVIIGLYGNVIMP	231
UTV	VVAPGKAAVNIQTKPGIFKTP-QGEIGAVSLDYPSGTSGPSILDKNGDVIIGLYGNVVLG	231
JEV	VVEPGKAAVNIQTKPGVFRTP-FGEVGAVALDYPRGTSGPSILDSNGDIIGLYGNVELG	231
DENV2	ALEPGKNPRAVQTKPGLFKTN-TGTIGAVSLDFSPGTSGPSIVDKKGVVVGLYGNVVT	236
DENV4	AIEPGKNPKHVQTKPGLFKTL-TGEIGAVTLDFFKPGTSGPSIINRKGKVIIGLYGNVVT	231
DENV1	AVEPGKNPKNVQTAGTFKTP-EGEIGAIALDFKPGTSGPSIVNREGKVIIGLYGNVVT	231
DENV3	AVEPGKNPKNFQTMPGIFQTT-TGEIGAIALDFKPGTSGPSIINREGKVVVGLYGNVVT	231
	. * * : * * . : : * * : * * * * * : * : * * * * :	
TBEV	NETYVSSIAQGEAEKSRPNLPQAVVGTGWTSK	265
YFV	DNSFVSAISQTEVKEEGKEELQEIPT--MLKK	264
ZIKV	NGSYVSAITQGRREEETPVECFE-PS--MLKK	260
SWV	SGSYVSAIMQTEKMEEPAVDCFE-ED--MLRK	260
WNV	NGSYISAIVQGERMDEPI PAGFE-PE--MLRK	260
UTV	NGSYVSAIVQGEREEPLPDAYN-AD--MLRK	260
JEV	DGSYVSAIVQGERQEEVPPEAYT-PN--MLRK	260
DENV2	SGAYVSAIANTEKS-IEDNPEIE-DD--IFRK	264
DENV4	SGDYVSAITQAERI-GEPTYEVD-ED--IFRK	259
DENV1	SGTYVSAIAQAKASQEGPLPEIE-DE--VFRK	260
DENV3	NGGYVSGIAQTNAEPDGPTEPEL-EE--MFKK	260
	. : * * * : . *	

Figure 6.10: Sequence alignment of the newly designed flaviviral NS2B-NS3 protease constructs highlighting the active (cyan) and allosteric (magenta) site.



## 7. Bibliography

Ai, J. W., Zhang, Y. and Zhang, W. (2016) ‘Zika virus outbreak: “A perfect storm”’, *Emerging Microbes and Infections*, 5, pp. 2–4. doi: 10.1038/emi.2016.42.

Andreyev, A. Y., Fahy, E., Guan, Z., Kelly, S., Li, X., McDonald, J. G., Milne, S., Myers, D., Park, H., Ryan, A., Thompson, B. M., Wang, E., Zhao, Y., Brown, H. A., Merrill, A. H., Raetz, C. R. H., Russell, D. W., Subramaniam, S. and Dennis, E. A. (2010) ‘Subcellular organelle lipidomics in TLR-4-activated macrophages’, *Journal of Lipid Research*, 51(9), pp. 2785–2797. doi: 10.1194/jlr.M008748.

Baronti, C., Piorkowski, G., Charrel, R. N., Boubis, L., Leparco-goffart, I. and Lamballerie, D. (2014) ‘Complete Coding Sequence of Zika Virus from a French Polynesia Outbreak in 2013’, *Genome Announcements*, 2(3), pp. 2013–2014. doi: 10.1128/genomeA.00500-14.

Barrows, N. J., Campos, R. K., Liao, K., Prasanth, K. R., Soto-acosta, R., Yeh, S., Schott-lerner, G., Pompon, J., Sessions, O. M., Bradrick, S. S. and Garcia-blanco, M. A. (2018) ‘Biochemistry and Molecular Biology of Flaviviruses’, *Chemical Reviews*, 118, pp. 4448–4482. doi: 10.1021/acs.chemrev.7b00719.

Baud, D., Gubler, D. J., Schaub, B., Lanteri, M. C. and Musso, D. (2017) ‘An update on Zika virus infection’, *The Lancet*. Elsevier Ltd, 390(November 4), pp. 2099–2109. doi: 10.1016/S0140-6736(17)31450-2.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) ‘The Protein Data Bank’, *Nucleic Acids Research*, 28(1), pp. 235–242. doi: 10.1093/nar/28.1.235.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., Drake, J. M., Brownstein, J. S., Hoen, A. G., Sankoh, O., Myers, M. F., George, D. B. and Jaenisch, T. (2013) ‘The global distribution and burden of dengue’, *Nature*. Nature Publishing Group, 496, pp. 504–507. doi: 10.1038/nature12060.

Blaney, J. E., Johnson, D. H., Firestone, C., Hanson, C. T., Murphy, B. R. and Whitehead, S. S. (2001) ‘Chemical Mutagenesis of Dengue Virus Type 4 Yields Mutant Viruses Which Are Temperature Sensitive in Vero Cells or Human Liver Cells and Attenuated in Mice’, *Journal of Virology*, 75(20), pp. 9731–9740. doi: 10.1128/JVI.75.20.9731.

*BLAST: Basic Local Alignment Search Tool* (2020). Available at: <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (Accessed: 5 July 2020).



Bonaldo, M. C., Gómez, M. M., Ac, A., Vieira, F., Abreu, S. De, Ferreira-de-brito, A., Miranda, R. M. De, Castro, M. G. De and Lourenço-de-oliveira, R. (2017) 'Genome analysis of yellow fever virus of the ongoing outbreak in Brazil reveals polymorphisms', *Memórias do Instituto Oswaldo Cruz, Rio de Janeiro*, 112(6), pp. 447–451. doi: 10.1590/0074-02760170134.

Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., Moyes, C. L., Farlow, A. W., Scott, T. W. and Hay, S. I. (2012) 'Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus', *PLoS Neglected Tropical Diseases*, 6(8), pp. 1–15. doi: 10.1371/journal.pntd.0001760.

Brecher, M., Li, Z., Liu, B., Zhang, J., Koetzner, C. A., Alifarag, A., Jones, S. A., Lin, Q., Kramer, L. D. and Li, H. (2017) 'A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease', *PLoS Pathogens*, 13(5), pp. 1–29. doi: 10.1371/journal.ppat.1006411.

Capeding, M. R., Tran, N. H., Hadinegoro, S. R. S., Ismail, H. I. H. M., Chotpitayasunondh, T., Chua, M. N., Luong, C. Q., Rusmil, K., Wirawan, D. N., Nallusamy, R., Pitisuttithum, P., Thisyakorn, U., Yoon, I. K., Van Der Vliet, D., Langevin, E., Laot, T., Hutagalung, Y., Frago, C., Boaz, M., *et al.* (2014) 'Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: A phase 3, randomised, observer-masked, placebo-controlled trial', *The Lancet*. Elsevier Ltd, 384(October 11), pp. 1358–1365. doi: 10.1016/S0140-6736(14)61060-6.

Chambers, T. J., Hahn, C. S., Galler, R. and Rice, C. M. (1990) 'Flavivirus genome: organization, expression and replication', *Annual Review of Microbiology*, 44, pp. 649–688.

*Clustal Omega, Multiple Sequence Alignment, EMBL-EBI* (2020). Available at: <https://www.ebi.ac.uk/Tools/msa/clustalo/> (Accessed: 5 July 2020).

*Codon usage* (1996). Available at: <http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/in-vitro-genetics/codon-usage.html> (Accessed: 5 July 2020).

Collins, M. H. and Metz, S. W. (2017) 'Progress and Works in Progress : Update on Flavivirus Vaccine Development', *Clinical Therapeutics*. Elsevier HS Journals, Inc., 39(8), pp. 1519–1536. doi: 10.1016/j.clinthera.2017.07.001.

Crooks, G., Hon, G., Chandonia, J. and Brenner, S. (2004) 'WebLogo: a sequence logo generator', *Genome Research*, 14, pp. 1188–1190. doi: 10.1101/gr.849004.1.

*Dengvaxia, European Medicines Agency* (2020). Available at: <https://www.ema.europa.eu/en/medicines/human/EPAR/dengvaxia> (Accessed: 5 July 2020).



## Bibliography

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E. and Finn, R. D. (2019) 'The Pfam protein families database in 2019', *Nucleic acids research*, 47(D1), pp. D427–D432. doi: 10.1093/nar/gky995.

Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S. P., Yin, Z., Keller, T. H., Vasudevan, S. G. and Hommel, U. (2006) 'Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus', *Nature Structural and Molecular Biology*, 13(4), pp. 372–373. doi: 10.1038/nsmb1073.

Ferreira-De-Lima, V. H. and Lima-Camara, T. N. (2018) 'Natural vertical transmission of dengue virus in *Aedes aegypti* and *Aedes albopictus*: A systematic review', *Parasites and Vectors*. *Parasites & Vectors*, 11(1), pp. 1–8. doi: 10.1186/s13071-018-2643-9.

*First FDA-approved vaccine for the prevention of dengue disease in endemic regions*, FDA (2020). Available at: <https://www.fda.gov/news-events/press-announcements/first-fda-approved-vaccine-prevention-dengue-disease-endemic-regions> (Accessed: 5 July 2020).

Frey, S., Essbauer, S., Zöllner, G., Klempa, B., Dobler, G. and Pfeffer, M. (2014) 'Full genome sequences and preliminary molecular characterization of three tick-borne encephalitis virus strains isolated from ticks and a bank vole in Slovak Republic', *Virus Genes*, 48, pp. 184–188. doi: 10.1007/s11262-013-0985-0.

Fulmali, P., Sapkal, G., Athawale, S., Gore, M., Mishra, A. and Bondre, V. (2011) 'Introduction of Japanese Encephalitis Virus Genotype I, India', *Emerging Infectious Diseases*, 17(2), pp. 319–321. doi: 10.1099/jmm.0.46204-0.

*GenBank Overview* (2020). Available at: <https://www.ncbi.nlm.nih.gov/genbank/> (Accessed: 5 July 2020).

Gibbs, A. C., Steele, R., Liu, G., Tounge, B. A. and Montelione, G. T. (2018) 'Inhibitor Bound Dengue NS2B-NS3pro Reveals Multiple Dynamic Binding Modes', *Biochemistry*, 57(10), pp. 1591–1602. doi: 10.1021/acs.biochem.7b01127.

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison III, C. A. and Smith, H. O. (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nature Methods*, 6(5), pp. 343–345. doi: 10.1038/NMETH.1318.

Gould, E. A. and Solomon, T. (2008) 'Pathogenic flaviviruses', *Lancet*, 371(February 9), pp. 500–509. doi: 10.1590/S0034-72802009000200010.

- Green, T. J. (2019) *Questionable Ligand Density: 6MO0, 6MO1, 6MO2*. Available at: <https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47086.html> (Accessed: 29 June 2020).
- Grinter, R. (2019) *Questionable Ligand Density: 6MO0, 6MO1, 6MO2*. Available at: <https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47072.html> (Accessed: 19 June 2020).
- Halgren, T. A. (1996) 'Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94', *Journal of Computational Chemistry*, 17, pp. 490–519.
- von Hammerstein, F., Lauth, L. M., Hammerschmidt, S., Wagner, A., Schirmeister, T. and Hellmich, U. A. (2019) 'Cis autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs', *FEBS Letters*. doi: 10.1002/1873-3468.13507.
- Hedstrom, L. (2002) 'Serine protease mechanism and specificity', *Chemical Reviews*, 102(12), pp. 4501–4523. doi: 10.1021/cr000033x.
- Hellwinkel, D. and Bohnet, S. (1987) 'Dibenzocycloocten-, Dibenzochalcocin- und Diarenochalconindione', *Chemische Berichte*, 120, pp. 1151–1173.
- Ishikawa, T., Yamanaka, A. and Konishi, E. (2014) 'A review of successful flavivirus vaccines and the problems with those flaviviruses for which vaccines are not yet available', *Vaccine*. Elsevier Ltd, 32(12), pp. 1326–1337. doi: 10.1016/j.vaccine.2014.01.040.
- Jeffery, J. A. L., Yen, N. T., Nam, V. S., Nghia, L. T., Hoffmann, A. A., Kay, B. H. and Ryan, P. A. (2009) 'Characterizing the Aedes aegypti Population in a Vietnamese Village in Preparation for a Wolbachia-Based Mosquito Control Strategy to Eliminate Dengue', 3(11). doi: 10.1371/journal.pntd.0000552.
- Kalayanarooj, S. (2011) 'Clinical Manifestations and Management of Dengue/DHF/DSS', *Tropical Medicine and Health*, 39(4), pp. 83–87. doi: 10.2149/tmh.2011-s10.
- Katzelnick, L. C., Gresh, L., Halloran, M. E., Mercado, J. C., Kuan, G., Gordon, A., Balmaseda, A. and Harris, E. (2017) 'Antibody-dependent enhancement of severe dengue disease in humans', *Science*, 358, pp. 929–932. doi: 10.1126/science.aan6836.
- Kim, Y. M., Gayen, S., Kang, C. B., Joy, J., Huang, Q., Chen, A. S., Wee, J. L. K., Ang, M. J. Y., Lim, H. A., Hung, A. W., Li, R., Noble, C. G., Lee, L. T., Yip, A., Wang, Q. Y., Chia, C. S. B., Hill, J., Shi, P. Y. and Keller, T. H. (2013) 'NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex', *Journal of Biological Chemistry*, 288(18), pp. 12891–12900. doi: 10.1074/jbc.M112.442723.

## Bibliography

Kufareva, I. and Abagyan, R. (2012) 'Homology Modeling Exercise', *Methods in Molecular Biology*, 857, pp. 231–257. doi: 10.1007/978-1-61779-588-6.

Kuiper, B. D., Slater, K., Spellmon, N., Holcomb, J., Medapureddy, P., Muzzarelli, K. M., Yang, Z., Ovadia, R., Amblard, F., Kovari, I. A., Schinazi, R. F. and Kovari, L. C. (2017) 'Increased activity of unlinked Zika virus NS2B / NS3 protease compared to linked Zika virus protease', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 492(4), pp. 668–673. doi: 10.1016/j.bbrc.2017.03.108.

Kuno, G., Chang, G. J., Tsuchiya, K. R., Karabatsos, N. and Cropp, C. B. (1998) 'Phylogeny of the Genus Flavivirus', *Journal of Virology*, 72(1), pp. 73–83.

Lambrechts, L., Fansiri, T., Pongsiri, A., Thaisomboonsuk, B., Klungthong, C., Richardson, J. H., Ponlawat, A., Jarman, R. G. and Scott, T. W. (2012) 'Dengue-1 Virus Clade Replacement in Thailand Associated with Enhanced Mosquito Transmission', *Journal of Virology*, 86(3), pp. 1853–1861. doi: 10.1128/JVI.06458-11.

Le, S. Q. and Gascuel, O. (2008) 'An improved general amino acid replacement matrix', *Molecular Biology and Evolution*, 25(7), pp. 1307–1320. doi: 10.1093/molbev/msn067.

Leatherbarrow, R. J. (2010) 'GraFit Version 7', *Erithacus Software Ltd.*, Horley, UK.

Lei, J., Hansen, G., Nitsche, C., Klein, C. D., Zhang, L. and Hilgenfeld, R. (2016) 'Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor', *Science*, 353(6298), pp. 503–505.

Leysen, P., Clercq, E. D. E. and Neyts, J. (2000) 'Perspectives for the Treatment of Infections with Flaviviridae', *Clinical Microbiology Review*, 13(1), pp. 67–82.

Li, Y., Zhang, Z., Phoo, W. W., Loh, Y. R., Li, R., Yang, H. Y., Jansson, A. E., Hill, J., Keller, T. H., Nacro, K., Luo, D. and Kang, C. B. (2018) 'Structural Insights into the Inhibition of Zika Virus NS2B-NS3 Protease by a Small-Molecule Inhibitor', *Structure*. Elsevier Ltd., 26(4), pp. 555–564. doi: 10.1016/j.str.2018.02.005.

Lindenbach, B. D., Thiel, H.-J. and Rice, C. M. (2007) 'Flaviviridae: The Virus and Their Replication', *Fields Virology*, 5th Editio, pp. 1101–1133.

Mahawaththa, M. C., Pearce, B. J. G., Szabo, M., Graham, B., Klein, C. D., Nitsche, C. and Otting, G. (2017) 'Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease', *Antiviral Research*. Elsevier B.V, 142, pp. 141–147. doi: 10.1016/j.antiviral.2017.03.011.

McOmie, J. F. W., Watts, M. L. and West, D. E. (1968) 'Demethylation of aryl methyl ethers by boron tribromide', *Tetrahedron*, 24, pp. 2289–2292.

*MEROPS, the Peptidase Database* (2020). Available at: <https://www.ebi.ac.uk/merops/> (Accessed: 5 July 2020).

Montalbetti, C. A. G. N., Falque, V., Park, M. and Ox, A. (2005) 'Amide bond formation and peptide coupling', *Tetrahedron*, 61(740), pp. 10827–10852. doi: 10.1016/j.tet.2005.08.031.

Moureau, G., Charrel, R., Holmes, E. C., Grard, G., Gould, E. A. and Lamballerie, X. De (2010) 'Genomics and evolution of Aedes -borne flaviviruses', *Journal of General Virology*, 91, pp. 87–94. doi: 10.1099/vir.0.014506-0.

Mustafa, M. S., Rasotgi, V., Jain, S. and Gupta, V. (2015) 'Discovery of fifth serotype of dengue virus (denv-5): A new public health dilemma in dengue control', *Medical Journal Armed Forces India*, 71(1), pp. 67–70. doi: 10.1016/j.mjafi.2014.09.011.

Natarajan, S. (2010) 'NS3 protease from flavivirus as a target for designing antiviral inhibitors against dengue virus', *Genetics and Molecular Biology*, 33(2), pp. 214–219.

Nikolay, B., Dupressoir, A., Firth, C., Faye, O., Boye, C. S., Diallo, M. and Sall, A. A. (2013) 'Comparative full length genome sequence analysis of usutu virus isolates from Africa', *Virology Journal*. *Virology Journal*, 10(217), pp. 1–9. doi: 10.1186/1743-422X-10-217.

Nitsche, C. (2019) 'Proteases from dengue, West Nile and Zika viruses as drug targets', *Biophysical Reviews*. *Biophysical Reviews*, 11(2), pp. 157–165. doi: 10.1007/s12551-019-00508-3.

Nitsche, C., Holloway, S., Schirmeister, T. and Klein, C. D. (2014) 'Biochemistry and Medicinal Chemistry of the Dengue Virus Protease', *Chemical Reviews*, 114, pp. 11348–11381.

Nitsche, C., Passioura, T., Varava, P., Mahawaththa, M. C., Leuthold, M. M., Klein, C. D., Suga, H. and Otting, G. (2019) 'De Novo Discovery of Nonstandard Macrocyclic Peptides as Noncompetitive Inhibitors of the Zika Virus NS2B-NS3 Protease', *ACS Medicinal Chemistry Letters*. American Chemical Society, 10(2), pp. 168–174. doi: 10.1021/acsmchemlett.8b00535.

Noble, C. G., Seh, C. C., Chao, A. T. and Shi, P. Y. (2011) 'Ligand-Bound Structures of the Dengue Virus Protease Reveal the Active Conformation', *Journal of Virology*, 86(1), pp. 438–446. doi: 10.1128/jvi.06225-11.

## Bibliography

- Oliva, C. F., Jacquet, M., Gilles, J., Lemperiere, G., Maquart, P., Vreysen, M. J. B., Boyer, S. and Quilici, S. (2012) 'The Sterile Insect Technique for Controlling Populations of *Aedes albopictus* (Diptera : Culicidae) on Reunion Island : Mating Vigour of Sterilized Males', *PLoS ONE*, 7(11), pp. 1–8. doi: 10.1371/journal.pone.0049414.
- Ong, S. and Jaal, Z. (2015) 'Investigation of mosquito oviposition pheromone as lethal lure for the control of *Aedes aegypti* (L.) (Diptera : Culicidae)', *Parasites & Vectors*, 8(28), pp. 1–7. doi: 10.1186/s13071-015-0639-2.
- Pan America Health Organization (PAHO) (2014) 'Technical Note on Transgenic Mosquitoes Engineered for *Aedes Aegypti* Control. Washington,.pdf'.
- Phoo, W. W., Li, Y., Zhang, Z., Lee, M. Y., Loh, Y. R., Tan, Y. B., Ng, E. Y., Lescar, J., Kang, C. and Luo, D. (2016) 'Structure of the NS2B-NS3 protease from Zika virus after self-cleavage', *Nature Communications*. Nature Publishing Group, 7(13410), pp. 1–8. doi: 10.1038/ncomms13410.
- Piccirillo, E., Merget, B., Sottriffer, C. A. and Antonia, T. (2016) 'Conformational flexibility of DENV NS2B / NS3pro : from the inhibitor effect to the serotype influence', *Journal of Computer-Aided Molecular Design*. Springer International Publishing, 30, pp. 251–270. doi: 10.1007/s10822-016-9901-8.
- Pybus, O. G., Suchard, M. A., Lemey, P., Bernardin, F. J., Rambaut, A. and Delwart, E. L. (2012) 'Unifying the spatial epidemiology and molecular evolution of emerging epidemics', *Proceedings of the National Academy of Sciences of the United States of America*, 109(37), pp. 15066–15071. doi: 10.1073/pnas.1206598109.
- Rarey, M., Kramer, B., Lengauer, T. and Klebe, G. (1996) 'A fast flexible docking method using an incremental construction algorithm', *Journal of Molecular Biology*, 261(3), pp. 470–489. doi: 10.1006/jmbi.1996.0477.
- Rasmussen, S. A., Jamieson, D. J., Honein, M. A., Ph, D. and Petersen, L. R. (2016) 'Zika Virus and Birth Defects — Reviewing the Evidence for Causality', *The New England Journal of Medicine*, 375(20), pp. 1981–1987.
- Rather, I. A., Parray, H. A., Lone, J. B., Paek, W. K., Lim, J., Bajpai, V. K. and Park, Y.-H. (2017) 'Prevention and Control Strategies to Counter Dengue Virus Infection', *Frontiers in Cellular and Infection Microbiology*, 7(July), pp. 1–8. doi: 10.3389/fcimb.2017.00336.
- Rios, M. (2009) 'Climate change and vector-borne viral diseases potentially', *International Society of Blood Transfusion Science Series*, 4, pp. 87–94.

- Roy, A., Lim, L., Srivastava, S., Lu, Y. and Song, J. (2017) 'Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants', *PLoS ONE*, 12(7), pp. 1–22. doi: 10.1371/journal.pone.0180632.
- Sakkas, H., Bozidis, P., Giannakopoulos, X., Sofikitis, N. and Papadopoulou, C. (2018) 'An Update on Sexual Transmission of Zika Virus', *Pathogens*, 7(3), p. 66. doi: 10.3390/pathogens7030066.
- Sasmono, R. T., Wahid, I., Trimarsanto, H., Yohan, B., Wahyuni, S., Hertanto, M., Yusuf, I., Mubin, H., Ganda, I. J., Latief, R., Bifani, P. J., Shi, P. and Schreiber, M. J. (2015) 'Infection , Genetics and Evolution Genomic analysis and growth characteristic of dengue viruses from Makassar , Indonesia', *Infection, Genetics and Evolution*. Elsevier B.V., 32, pp. 165–177. doi: 10.1016/j.meegid.2015.03.006.
- Schneider, N., Lange, G., Hindle, S., Klein, R. and Rarey, M. (2013) 'A consistent description of HYdrogen bond and DEhydration energies in protein-ligand complexes: Methods behind the HYDE scoring function', *Journal of Computer-Aided Molecular Design*, 27(1), pp. 15–29. doi: 10.1007/s10822-012-9626-2.
- Shiryayev, S. A., Farhy, C., Pinto, A., Huang, C. T., Simonetti, N., Ngono, A. E., Dewing, A., Shresta, S., Pinkerton, A. B., Cieplak, P., Strongin, A. Y. and Terskikh, A. V. (2017) 'Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists', *Antiviral Research*. Elsevier B.V, 143, pp. 218–229. doi: 10.1016/j.antiviral.2017.04.015.
- Slon Campos, J. L., Mongkolsapaya, J. and Sreaton, G. R. (2018) 'The immune response against flaviviruses', *Nature Immunology*. Springer US, 19(11), pp. 1189–1198. doi: 10.1038/s41590-018-0210-3.
- Stamatakis, A. (2006) 'RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models', *Bioinformatics*, 22(21), pp. 2688–2690. doi: 10.1093/bioinformatics/btl446.
- Stamatakis, A., Hoover, P. and Rougemont, J. (2008) 'A rapid bootstrap algorithm for the RAxML web servers', *Systematic Biology*, 57(5), pp. 758–771. doi: 10.1080/10635150802429642.
- Steuer, C., Heinonen, K. H., Kattner, L. and Klein, C. D. (2009) 'Optimization of Assay Conditions fo r Dengue Virus Protease: Effect of Various Polyols and Nonionic Detergents', *Journal of Biomolecular Screening*, 14(9), pp. 1102–1108. doi: 10.1177/1087057109344115.

## Bibliography

Sun, J., Lin, J., Yan, J., Fan, W., Lu, L. and Lv, H. et al. (2011) 'Dengue Virus Serotype 3 Subtype III , Zhejiang Province , China', *Emerging Infectious Diseases*, 17(2), pp. 321–323. doi: 10.1038/nm1144.

Trapani, G., Franco, M., Latrofa, A., Reho, A. and Liso, G. (2001) 'Synthesis , in vitro and in vivo cytotoxicity , and prediction of the intestinal absorption of substituted 2-ethoxycarbonyl-imidazo [ 2 , 1- b ] benzothiazoles', *European Journal of Pharmaceutical Sciences*, 14, pp. 209–216.

Vannice, K. S., Wilder-Smith, A., Barrett, A. D. T., Carrijo, K., Cavaleri, M., de Silva, A., Durbin, A. P., Endy, T., Harris, E., Innis, B. L., Katzelnick, L. C., Smith, P. G., Sun, W., Thomas, S. J. and Hombach, J. (2018) 'Clinical development and regulatory points for consideration for second-generation live attenuated dengue vaccines', *Vaccine*. The Authors, 36(24), pp. 3411–3417. doi: 10.1016/j.vaccine.2018.02.062.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009) 'Jalview Version 2-a multiple sequence alignment editor and analysis workbench', *Bioinformatics*, 25(9), pp. 1189–1191. doi: 10.1093/bioinformatics/btp033.

Wu, H., Bock, S., Snitko, M., Berger, T., Weidner, T., Holloway, S., Kanitz, M., Diederich, W. E., Steuber, H., Walter, C., Hofmann, D., Weißbrich, B., Spannaus, R., Acosta, E. G., Bartenschlager, R., Engels, B., Schirmeister, T. and Bodem, J. (2015) 'Novel dengue virus NS2B/NS3 protease inhibitors', *Antimicrobial Agents and Chemotherapy*, 59(2), pp. 1100–1109. doi: 10.1128/AAC.03543-14.

Yao, Y., Huo, T., Lin, Y. L., Nie, S., Wu, F., Hua, Y., Wu, J., Kneubehl, A. R., Vogt, M. B., Rico-Hesse, R. and Song, Y. (2019) 'Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease', *Journal of the American Chemical Society*. American Chemical Society, 141(17), pp. 6832–6836. doi: 10.1021/jacs.9b02505.

Yildiz, M. (2013) 'Allosteric inhibition of the NS2B-NS3 protease from dengue virus', *ACS Chemical Biology*, 8(12), pp. 2744–2752. doi: 10.1021/cb400612h.Allosteric.

Zhu, L., Yang, J., Li, H., Sun, H., Liu, J. and Wang, J. (2015) 'Conformational change study of dengue virus NS2B-NS3 protease using 19F NMR spectroscopy', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 461, pp. 677–680. doi: 10.1016/j.bbrc.2015.04.090.





## Curriculum vitae

**Franziska v. Hammerstein, Dipl. - Chem.**

**E-mail**

franziska.hammerstein@uni-mainz.de

### University education

---

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]
[REDACTED]	[REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]



**Scientific posters, talks and publications**

---

[Redacted text block containing multiple lines of blacked-out content]