Molecular Insights into Protease Inhibition: Design and Interactions of NS2B-NS3 Protease Inhibitors

Dissertation

zur Erlangung des Grades

„Doktor der Naturwissenschaften (Dr. rer. nat.)“

im Promotionsfach Pharmazeutische und Medizinische Chemie

am Fachbereich Chemie, Pharmazie, Geographie und Geowissenschaften

der Johannes Gutenberg-Universität Mainz

Hannah Maus, Master of Science

geb. in Wiesbaden

Submitted to the Faculty of Chemistry, Pharmacy, Geography, and Geoscience.

Dean: [redacted]

Name of the 1st reviewer: [redacted]

Name of the 2nd reviewer: [redacted]

Date of the doctoral examination: 11.09.2023

D77 (Dissertation Universität Mainz)

Declaration of Authorship

I, Hannah Maus, declare that this thesis entitled, “Molecular Insights into Protease Inhibition: Design and Interactions of NS2B-NS3 Protease Inhibitors”, and the work presented in it are my own. I hereby declare that:

- I did this work entirely while in candidature for a research degree at this university.
- I have clearly stated which parts of this thesis have already been submitted for another degree or qualification at this university or other institutions.
- I always attributed clearly where I have consulted the published work of others.
- I always gave the source when I quoted from other works. Excluding those quotations, this thesis is entirely my work.
- I have acknowledged all primary sources of help.
- I have made clear exactly what was done by others and what I have contributed myself.

Place, Date: ______________________ Signature: ______________________

“But if you never try you’ll never know”

− Coldplay, Fix You −

Acknowledgments

I am profoundly grateful and wish to express my sincerest thanks to my doctoral supervisor, [redacted]. Her support and encouragement empowered me to navigate my academic journey freely and prioritize my research effectively. Her mentorship has been invaluable, playing a pivotal role in successfully completing this thesis; for that, I will always be immensely thankful.

I extend my heartfelt appreciation to [redacted] from the Department of Chemistry at Johannes Gutenberg University Mainz. Not only did he take on the responsible task of preparing a second expert assessment of this doctoral thesis, but he also welcomed me into the world of smFRET experiments with open arms. The trust he placed in me and the generous provision of resources and facilities significantly contributed to advancing my work. I am genuinely grateful for the invaluable insights he shared on structuring scientific publications.

I owe [redacted] a debt of gratitude for the delightful collaboration we shared. During moments of uncertainty, his motivating words kept me going, and our discussions on proteases and inhibitors enriched my understanding. Working together could not have been more fulfilling, and I feel fortunate to have had such an exceptional cooperation partner.

A special note of thanks goes to [redacted] for his boundless enthusiasm and thought-provoking ideas. Our collaboration was a source of inspiration, and I am deeply appreciative of his invaluable advice and guidance throughout various situations.

To [redacted], my thanks for introducing me to the smFRET methodology. Your patience and generous sharing of knowledge laid the foundation for numerous projects, and I am grateful for your support at every step.

I extend my appreciation to [redacted] from the Institute of Organic Chemistry and Macromolecular Chemistry (IOMC) at Friedrich Schiller University in Jena, as well as her group. Their hospitality and the enriching experience in Jena were truly exceptional. I am particularly thankful to [redacted] for his tireless support with the ^{19}F-NMR measurements.

To all my collaboration partners, I am indebted for their contributions to various projects. From the *Biozentrum* at Johannes Gutenberg-University, I thank [redacted] and [redacted] for their guidance in CD-spectroscopy and nanoDSF. [redacted] of the Federal University of Alagoas Institute of Chemistry and Biotechnology Maceió, thank you for your valuable collaboration on flaviviral proteases. [redacted] from the Center of Infectiology at the University of Heidelberg, your support in the allosteric inhibitors of DENV and ZIKV NS2B-NS3 proteases project was instrumental. I also want to thank [redacted] from the Institute of Pharmaceutical Chemistry at the Philipps-University Marburg and [redacted] from the Institute of Pharmaceutical and Molecular Biotechnology at the University of Heidelberg for

competitive inhibitors for the smFRET experiments. [redacted] from the Institute of Molecular Biology at Johannes Gutenberg-University, thank you for your efforts in the PAL project. The expertise of all of you significantly enhanced the outcome of my work, and I am genuinely grateful for your collaborative efforts.

A heartfelt acknowledgment goes to all my colleagues from the [redacted] Group. Special mention to my lab mates [redacted], [redacted], [redacted], [redacted], and [redacted]. The camaraderie and productive conversations in the lab were indispensable to my growth as a researcher. A heartfelt thank you to [redacted] for her unwavering support and the ease with which I could discuss scientific and personal matters. Your friendship has been a source of strength.

Gratitude is also due to [redacted] for our collaboration on the flaviviral proteases project and to [redacted] for collaborating on the NBD project. The harmonious teamwork enriched my research experience.

To the students I had the privilege of supervising during my doctoral studies – [redacted] during her master thesis, [redacted], [redacted], and [redacted] during their bachelor theses, [redacted] during her research module, and [redacted] during her internship – thank you for your dedication and hard work. Your contributions were invaluable, and I am incredibly proud of your achievements.

Lastly, my heartfelt gratitude goes to my family and friends. Their unwavering support, encouragement, and belief in me have been invaluable. Without their love and motivation, this thesis would not have been possible.

To each and every one of you, THANK YOU from the bottom of my heart! You made this journey unforgettable and a whole lot of fun!

Table of Contents

List of Abbreviations

[S]	substrate concentration
^{2}DPro	double mutant (S79*C-S158C) of the DENV protease
^{5}ZiPro	quintuple mutant (S84*C-C80S-C143S-S160C -C178S) of the ZIKV protease
A	acceptor
ABC	ATB-binding cassette
abs	absorption
ADE	antibody-dependent enhancement
AMC	7-amino-4-methyl coumarin
APD	avalanche photodiode/ detector
Asp	aspartate
C	capsid protein
c	concentration
CS	conformational selection
D	donor
DENV	dengue virus
DSF	differential scanning fluorometry
E	envelope protein
E_{FRET}	Förster resonance energy transfer efficiency
em	emission
ER	endoplasmatic reticulum
F	fluorescence intensity
F_A	acceptor intensity
FBDD	fragment-based drug design
FCCS	fluorescence cross-correlation spectroscopy
FCS	fluorescence correlation spectroscopy
F_D	donor intensity
FDA	Food and Drug Administration
FRET	Förster resomamce energy transfer
GBS	Guillain-Barré syndrome
$G_{D,A}$	autocorrelation function of the donor (D) or acceptor (A)
G_{DA}	cross-correlation function
gZiPro	glycine-linked ZIKV protease
His	histidine
HPLC	high-pressure liquid chromatography
I_a	allosteric inhibitor

IC_{50}	the concentration of inhibitor needed to reduce the reaction rate by half
I_c^{DENV}	competitive inhibitor of the DENV protease
I_c^{ZIKV}	competitive inhibitor of the ZIKV protease
IF	induced fit
ISC	intersystem crossing
K_d	dissociation constant/ binding affinity constant
k_{FRET}	rate of energy transfer
K_i	inhibition constant
K_i^{app}	apparent K_i
K_M	Michaelis Menten constant
MBP	maltose binding protein
MMP	matrix metalloprotease
N	number of particles
nanoDSF	label-free DSF
NBD	7-nitrobenz-2-oxa-1,3-diazol-4-amine
NMR	nuclear magnetic resonance
NS	non-structural proteins
$NS2B^{cf}$	NS2B cofactor
$NS3^{pro}$	NS3 protease
NTD	neglected tropical disease
PDB	RASC Protein Data Bank
PG	photoreactive group
pNA	*p*-nitroanilide
prM/M	membrane (precursor)
R	distance between donor and acceptor
R_0	Förster radius
RNA	ribonucleic acid
S_0	ground singlet state
S_1	excited singlet state
SAR	structure affinity relationships
SBD	substrate binding domain
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	serine
smFRET	single-molecule Förster resonance energy transfer
T	thermodynamic temperature
T_1	triplet state

TIRF	total internal reflection fluorescence
T_m	melting temperature
TSA	thermal shift assay
uPA	urokinase plasminogen activator
v	reaction rate/ turnover rate
V_{eff}	observation volume
WNV	West Nile virus
ZIKV	Zika virus
ΔT_m	thermal shift
τ_{D0}	donor fluorescence lifetime without acceptor
τ_D	diffusion time
τ_{DA}	lifetime of the donor in the presence of the acceptor

Abstract

The NS2B-NS3 protease of Zika (ZIKV) and dengue (DENV) viruses, which, in addition to host proteases, is responsible for the cleavage of the polyprotein translated from the RNA genome, is essential for viral replication. Inhibition of this protease reduces infectivity, making the protease an interesting drug target. This work investigated protein-ligand interactions of flaviviral proteases from various aspects. The results can be assigned to two different projects.

Project 1. In the first project, competitive and allosteric inhibitors were developed for the ZIKV and DENV proteases. The allosteric inhibitors represent an evolution of previously published allosteric benzo[*d*]thiazole-based inhibitors. Three different approaches were pursued: (i) truncated inhibitors with improved ligand efficiency (LE), (ii) Y-shaped inhibitors based on predictions of molecular docking, and (iii) scaffold hopping and replacement of the benzo[*d*]thiazole backbone. This resulted in allosteric inhibitors with improved properties. The competitive inhibitors were designed using a fragment-based drug design approach. The mechanism of inhibition and selectivity of the most promising compounds were demonstrated using fluorometric assays. Furthermore, the underlying fluorometric assays were optimized to make the development of potential protease inhibitors more reliable in the future. For this purpose, substrates based on 7-nitrobenz-2-oxa-1,3-diazol-4-amine (NBD) were developed and characterized. Compared to previously used fluorophores (7-amino-4-methylcoumarin and Förster resonance energy transfer (FRET) based substrates), excitation and detection wavelengths of NBD are at higher wavelengths, minimizing typical assay interferences and thus reducing the occurrence of false positive results.

Project 2. In this project, the influence of ligands on the conformation of proteases was investigated. Dye-labeled DENV and ZIKV protease cysteine mutants were studied using FRET at the single-molecule level (smFRET). Through smFRET experiments with proteases diffusing freely in solution, it was shown that the conformational equilibrium of the protease, between the so-called open and closed conformation, is shifted by competitive inhibitors in favor of the closed and by allosteric inhibitors in favor of the open conformation. Other biophysical methods (DSF, nanoDSF, and ^{19}F-NMR) also demonstrated opposing effects of allosteric and competitive inhibitors on the ZIKV protease. By studying immobilized proteases using smFRET, the two proteases were shown to have different binding mechanisms with competitive inhibitors. While the ZIKV protease follows an induced fit model, the interaction with the DENV protease occurs by conformational selection. Additionally, photoaffinity labeling was used to investigate the location of the allosteric binding pocket in the DENV protease. For this purpose, trifunctional inhibitors with an affinity moiety, a photoreactive group, and a ligation handle, which allowed the detection of successful photoaffinity labeling in gel, were designed and synthesized. The detected interaction of a tool compound with G121 confirmed the previously proposed allosteric binding pocket.

Kurzdarstellung

Die NS2B-NS3 Proteasen von Zika (ZIKV) und Dengue (DENV) Viren, sind neben Wirtsproteasen für die Spaltung der aus den RNA-Genomen translatierten Polyproteinen verantwortlich und daher für die virale Replikation essenziell. Im Rahmen dieser Arbeit wurden Protein-Ligand Wechselwirkungen der flaviviralen Proteasen untersucht. Die Ergebnisse lassen sich zwei unterschiedlichen Projekten zuordnen.

Projekt 1. Im ersten Projekt wurden kompetitive und allostere Inhibitoren für die ZIKV- und die DENV-Protease entwickelt. Basierend auf zuvor veröffentlichten allosteren Benzo[*d*]thiazol-Inhibitoren wurden drei unterschiedliche Ansätze verfolgt: (i) verkürzte Inhibitoren mit verbesserter Ligand Effizienz (LE), (ii) *Y-shaped* Inhibitoren basierend auf Vorhersagen des molekularen Dockings, und (iii) *Scaffold Hopping* und Austausch des Benzo[*d*]thiazol-Grundgerüsts. Dies resultierte in allosteren Inhibitoren mit verbesserten Eigenschaften. Die kompetitiven Inhibitoren wurden mit Hilfe eines Fragment-basierten Wirkstoffdesign Ansatzes entwickelt. Der Inhibitionsmechanismus sowie die Selektivität der vielversprechendsten Verbindungen wurden in fluorometrischen Assays nachgewiesen. Um die Entwicklung potenzieller Proteaseinhibitoren in Zukunft zuverlässiger zu machen, wurden die zugrundeliegenden fluorometrischen Assays optimiert. Hierfür wurden Substrate basierend auf 7-Nitrobenz-2-oxa-1,3-diazol-4-amin (NBD) entwickelt und charakterisiert. Verglichen mit den zuvor verwendeten Fluorophoren liegen Anregungs- und Detektionswellenlänge des NBDs bei höheren Wellenlängen, was typische Assay-Interferenzen minimiert und folglich das Auftreten falsch-positiver Ergebnisse verringert.

Projekt 2. In diesem Projekt wurde der Einfluss von Liganden auf die Konformation der Proteasen untersucht. Farbstoffmarkierte Cystein Mutanten der DENV und der ZIKV Protease wurden mit Hilfe von FRET auf Einzelmolekül-Ebene (smFRET) untersucht. Durch Experimente in Lösung konnte gezeigt werden, dass das konformative Gleichgewicht der Proteasen durch kompetitive Inhibitoren zugunsten der geschlossenen und durch allostere Inhibitoren zugunsten der offenen Konformation verschoben wird. Durch weitere Methoden (DSF, nanoDSF und ^{19}F-NMR) konnten ebenfalls gegensätzliche Effekte allosterer und kompetitiver Inhibitoren auf die ZIKV-Protease nachgewiesen werden. Die Untersuchung immobilisierter Proteasen zeigte, dass die beiden Proteasen unterschiedliche Bindungsmechanismen mit kompetitiven Inhibitoren aufweisen. Während die ZIKV-Protease einem *Induced Fit* Modell folgt, erfolgt die Wechselwirkung mit der DENV-Protease durch *Conformational Selection*. Mit Hilfe von *Photoaffinity Labeling* wurde zudem die Lage der allosteren Bindetasche in der DENV Protease untersucht. Hierfür wurden Inhibitoren mit einer Affinitätseinheit, einer photoreaktiven Gruppe und einer Ligationsmöglichkeit zum Nachweis der erfolgreichen Photoaffinitätsmarkierung im Gel entworfen und synthetisiert. Die nachgewiesene Wechselwirkung einer Verbindung mit G121 bestätigte die zuvor vorgeschlagene allostere Bindetasche.

1. Introduction

1.1 Proteases in Medicinal Chemistry

Proteins are essential macromolecules that play a fundamental role in all living organisms.[1–3] They are composed of amino acids arranged in a specific sequence, and their three-dimensional structure determines their function.[2] Proteins are of utmost interest to the scientific community because they are involved in various biological processes, including signal transduction, metabolism, and gene regulation.[4–16]

Proteases, a specific class of proteins, catalyze the hydrolysis of peptide bonds within proteins and peptides. They are responsible for the controlled breakdown of proteins and have crucial roles in many biological processes.[17] Proteases regulate cellular pathways, such as protein turnover, cell cycle progression, and apoptosis.[17–25] Additionally, they are involved in crucial physiological functions, such as digestion, blood clotting, and immune response.[17,26–28] Given their central role in protein homeostasis, it is unsurprising that the dysregulation of proteases has been implicated in numerous pathologies, including cancer, neurodegenerative disorders, cardiovascular diseases, and infectious diseases.[29,30] The significance of proteases in disease pathogenesis has spurred intense research efforts to develop therapeutic interventions targeting these enzymes. Medicinal chemistry approaches aim to design molecules that selectively inhibit specific proteases, thereby modulating their activity and restoring normal cellular functions.[30,31]

In the field of antiviral drug discovery, proteases have gained particular attention due to their critical role in viral replication. Flaviviruses, a family of ribonucleic acid (RNA) viruses, rely on proteases for processing viral polyproteins into functional proteins required for viral replication. Targeting flaviviral proteases presents an attractive strategy to combat diseases caused by flaviviruses, such as dengue fever, Zika virus, and yellow fever.[32,33]

1.2 Flaviviral Proteases

1.2.1 Spread, Transmission, and Diseases

In recent decades, there has been a significant increase in the global spread of neglected tropical diseases (NTDs). Incident cases increased from 33.7 million in 1990 to 57.99 million in 2019.[34] Mosquitoes play a substantial role in transmitting these diseases, posing a significant risk to human health. Two prominent examples of mosquito-borne diseases are Zika and dengue fever, caused by the Zika virus (ZIKV) and dengue virus (DENV), respectively. The main vectors responsible for transmitting these viruses are mosquitoes of the *Aedes aegypti* and *Aedes albopictus* species.

Factors such as global warming, increased globalization, and international travel have contributed to the spread of these mosquito vectors from tropical and subtropical regions to more temperate countries.[32] This has expanded the geographical distribution of these diseases (Figure 1A).[35]

Approximately half of the world's population is at risk of DENV infection. This corresponds to about 3.97 billion vulnerable people in around 128 countries.[36] About 400 million people are estimated to be infected with DENV annually. Around a quarter of these infections manifest clinically.[37] In the first half of 2023, 2,162,214 cases and 974 dengue deaths were reported worldwide, including cases reported in Europe (Spain and France, see Figure 1B).[38] The regions most affected by DENV are Southeast Asia, South and Central America, Africa, and Australia (Figure 1A). However, about 70% of the burden is concentrated mainly in Asia.[39] DENV is transmitted to humans by infected female mosquitoes that have acquired the virus from an infected person.[40,41] However, according to the World Health Organization (WHO), isolated cases of non-vector-borne transmission have also been reported.[37,42] A few cases of DENV transmission from mother to fetus during pregnancy and through blood transfusion or organ donation have been documented.[43,44]

After the discovery of ZIKV in 1947, evidence of human infection was found for the first time in Nigeria in 1953, where the virus infected three individuals. Since then, there have been three significant epidemics on the Yap Islands (2007), in French Polynesia (2013), as well as in Brazil and other American countries (2015–2016), resulting in over one million infections.[45–48] More than 86 countries have reported mosquito-borne ZIKV infections.[49,50] Known areas of infection are shown in Figure 1A, with ZIKV being particularly widespread in South America (Brazil) and North America.[51] ZIKV is primarily transmitted to humans by infected mosquitoes. However, ZIKV can also be transmitted from mother to fetus during pregnancy (congenital ZIKV virus), through organ transplantation, sexual contact, or blood transfusions from person to person.[52,53]

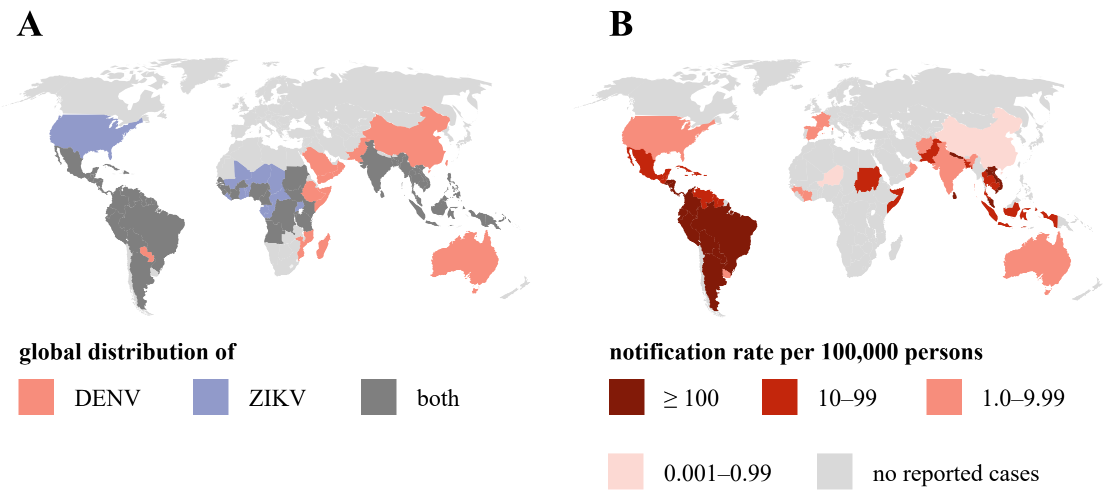


Figure 1. Distribution of infection. ZIKV (blue) and DENV (red) global distribution by country. Affected countries are colored according to data from Collins and Metz (2017).[35] Countries that are affected by both are colored dark grey. **B** Twelve-month DENV notification rate per 100,000 population (June 2022 – May 2023).[38]

The symptoms of dengue fever can range from mild flu-like symptoms such as fever, nausea, rash, and pain, which are associated with dengue fever, to severe life-threatening symptoms such as bleeding, shock, and cardiovascular collapse, which are characteristic of dengue shock syndrome.[54] Based on different surface proteins, DENVs are divided into four serotypes (DENV1–DENV4) associated with varying clinical manifestations of the disease.[55] Studies have shown that after the initial infection with a heterologous DENV serotype, a patient is protected from reinfection or high viral load for a certain period, and there is lifelong immunity to that specific DENV serotype. However, antibody-dependent enhancement (ADE) increases the risk of developing severe dengue after subsequent infection by different serotypes.[56]

ZIKV has an estimated incubation period of three to fourteen days, with most infected individuals not showing any symptoms of illness.[57,58] The most frequently reported symptoms of a viral infection include fever, rash, muscle pain, malaise, and headache, typically lasting between two and seven days.[48,57] If a pregnant woman transmits the virus to her fetus, it can result in microcephaly (abnormal smallness of the head) or other abnormalities in the newborn.[59,60] Pregnancy complications such as preterm birth or fetal loss can also be potential consequences of a ZIKV infection.[61] The infection can also trigger Guillain-Barré syndrome (GBS), which primarily affects adults and older children.[62,63] GBS is an autoimmune disease of the nervous system in which the patient's immune system attacks the nerve cells. Possible symptoms of the syndrome include muscle weakness in the arms and legs, breathing difficulties, or paralysis. The mortality rate associated with GBS is low, but affected individuals can show symptoms for a period ranging from a few weeks to several months.[64,65]

Currently, no specific treatment is available for DENV and ZIKV infections, consequently symptomatic therapies and vector control measures are the primary approaches.[66] While a tetravalent vaccine called

Dengvaxia® by Sanofi-Pasteur has been approved by the Food and Drug Administration (FDA) for DENV, it is only recommended for individuals who have previously recovered from DENV infection.[67] Dengvaxia® demonstrates varying efficacy against different serotypes and has notable side effects in seronegative patients.[68,69] Identifying a fifth serotype in Malaysia in 2007 further complicated vaccine development efforts.[70] However, clinical trials are currently underway for several vaccines targeting DENV and ZIKV.[71–73]

1.2.2 Role of the NS2B-NS3 Protease in the Replication Cycle

DENV and ZIKV are small enveloped flaviviruses with a diameter of approximately 50 nm harboring a single-stranded (+) sense RNA genome comprising about 11,000 base pairs.[45,74,75] The binding of the flavivirus to the host cell and its uptake into the cell through endocytosis is mediated by various receptors specific for the viral envelope proteins (E). The low pH in the endosome induces the fusion of the viral envelope with cellular membranes, releasing the RNA genome into the cytoplasm.[76] After the release, ribosomes translate the single-stranded RNA into the viral precursor polyprotein. This polyprotein consists of approximately 3,000 amino acids, which form the three structural proteins, capsid (C), envelope protein (E), and membrane precursor (prM), as well as the seven non-structural proteins (NS) NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. Co- and post-translational processing by both host proteases and the viral protease (NS2B-NS3 complex) leads to the formation of the ten functional proteins.[77–79] The RNA-dependent RNA polymerase NS5 synthesizes a complementary negative RNA strand, which serves as a template for replicating (+) sense RNA. Assembly of new, initially non-infectious virus particles takes place in the lumen of the rough endoplasmic reticulum (ER). In the trans-Golgi network, the immature virions undergo cleavage of the prM into M, forming mature virus particles that exit the host cell via exocytosis (Figure 2).[77,80–83]

The viral trypsin-like serine protease (NS3^{pro}), which is involved in the cleavage of the viral polyprotein, is essential for virus replication and is formed by the first 180 N-terminal amino acids of the bifunctional NS3 protein. [32,84] The protease cleaves the polyprotein preferentially after two basic amino acid residues, resulting in six cleavage sites within the polyprotein.[79,85,86] The protease's catalytic activity and proper folding are influenced by its interaction with the associated NS2B cofactor (NS2B^{cf}).[32] The cofactor contains three hydrophobic membrane-associated domains that bind the protease on the cytoplasmic side of the ER membrane. The interaction between NS3^{pro} and the NS2B^{cf} occurs through a hydrophilic region of 47 amino acids, spanning residues 49 to 95.[32,79,84,87] Rothan *et al.* showed that inhibiting the NS2B-NS3 DENV2 protease reduces dengue virus infection by approximately 80%, highlighting it as an interesting drug target.[88]

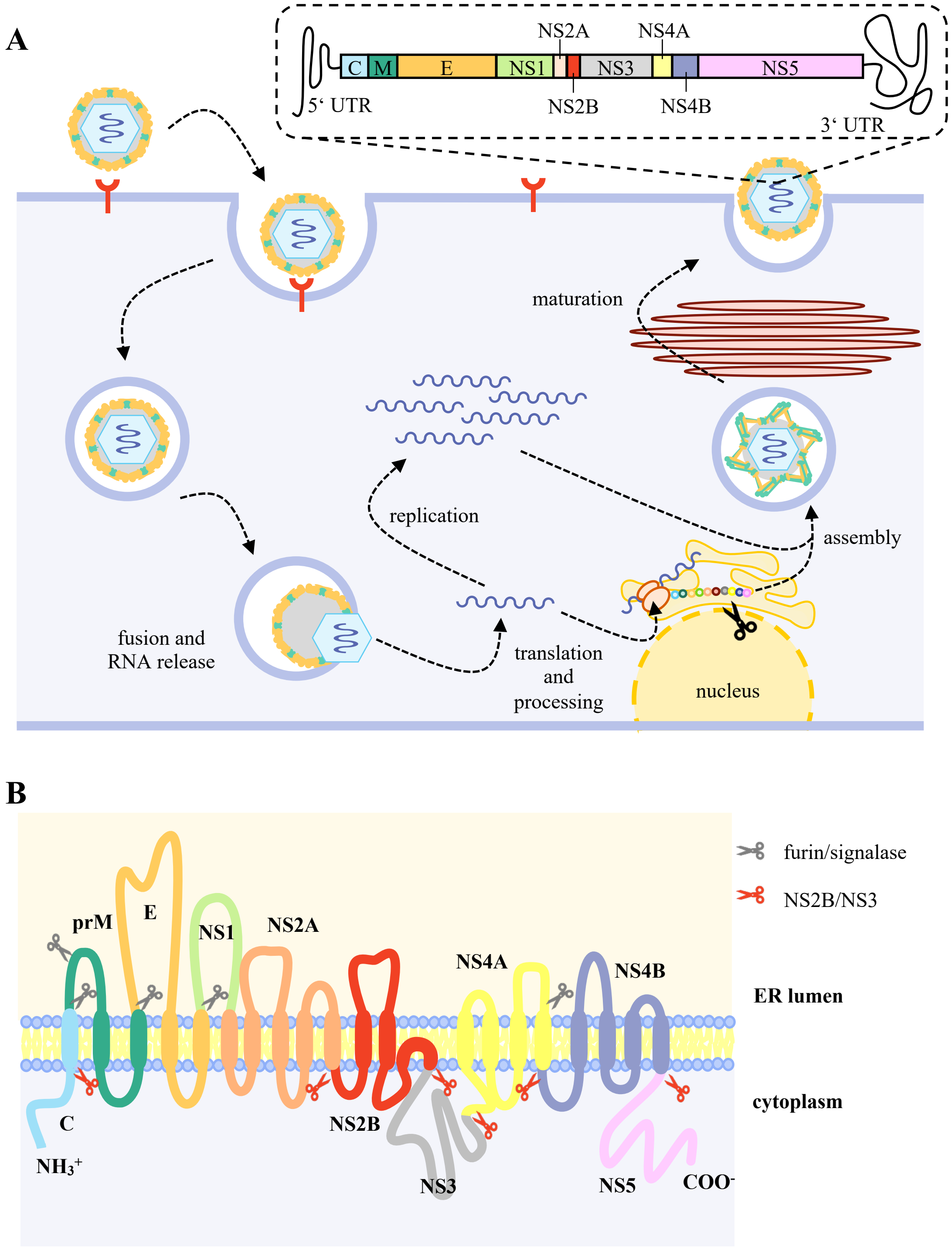


Figure 2. Replication cycle and polyprotein organization of flaviviruses. Figure designed with Microsoft PowerPoint based on Boldescu *et al.*[77] and Guzman *et al.*[80] **A** Replication cycle of flaviviruses. Host cell receptors mediate the uptake of the virus by endocytosis. This is followed by fusion and RNA release, replication of the RNA and its translation and processing, assembly of the virion, and maturation into mature virions. The organization of the single-stranded genome is shown in the box in the upper right corner. **B** Sequential and structural organization of the flaviviral polyprotein at the endoplasmatic reticulum membrane with the cleavage sites of the host (furin and signalase) and viral proteases are indicated.

1.2.3 Binding Site and Catalytic Reaction

According to the MEROPS database, the NS2B-NS3 proteases of DENV and ZIKV belong to clan PA, subclan PA(S), and family S7, based on their evolutionary relationships and catalytic triad composition. The DENV NS2B-NS3 protease's identifier is S07.001, while the ZIKV NS2B-NS3 has the identifier S07.003.[89,90] The PA clan consists of endopeptidases similar to chymotrypsin A, which possess a catalytic triad arranged in the order of histidine (His), aspartate (Asp), and a nucleophilic serine (Ser) or cysteine.[26,90] The catalytic triad of the NS2B-NS3 protease is in the cleft between two β-barrels and consists of the three amino acids His51, Asp75, and Ser135 (Figure 3).[33,91] The substrate binding site of both DENV and ZIKV proteases shows a preference for basic residues (Arg or Lys) in the P1 and P2 position and a short side-chain amino acid (Gly, Ser or Ala) at P1' site.[33,92,93] The proteases exhibit shallow grooves on the protein surface, forming the S3−S1' pockets but only flat S1 pockets.[94–96] The S3 and S4 pockets of the proteases are less selective, accommodating Gly, Ala, and Lys in natural substrates.[97]

The binding mode of the substrate Gly-Lys-Arg was revealed in a crystal structure containing the auto-processed C-terminus of NS2B in the binding pocket (RCSB Protein Data Bank (PDB)[98]-ID: 5GJ4[99], Figure 3). In this structure, the positively charged P1 arginine residue is stabilized by the negatively charged Asp129 and the backbone oxygen of Tyr130 in the S1 pocket. Ser81* and Asp83* side chains of NS2B recognize the negative charge of the P2 lysine residue. The interaction between the P3 glycine carboxylic oxygen with the oxygen of Gly151 and the nitrogen of Gly153 in the S3 site is unspecific. It is worth noting that the substrate in this crystal structure is still connected to NS2B, and this truncated construct may result in distorted interactions between the P3 Gly and P4 Thr residues and the NS3^{pro} domain.[99]

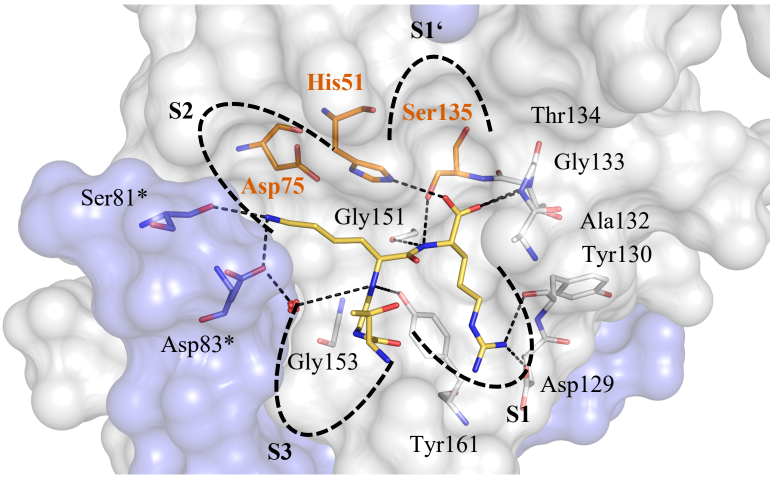


Figure 3. The substrate binding site of ZIKV NS2B-NS3 protease in the closed conformation after linker autocatalysis (PDB-ID: 5GJ4)[99]. NS3 is represented as a white surface, while NS2B is colored blue. The catalytic triad amino acids are depicted as orange sticks. Schechter and Berger's classification is used to indicate the binding sub-pockets.[100] The natural substrate TGKR is visualized with yellow sticks, and polar interactions are represented by black dashed lines. Only the amino acids involved in polar interactions are displayed as sticks to enhance clarity. NS2B residues are marked with an asterisk. The figure was generated using PyMOL and Microsoft PowerPoint.

The proteolytic activity of the proteases results from a nucleophilic attack of Ser135 on the cleaving amide bond, whose nucleophilicity is enhanced by the adjacent His51. Asp75 can initially accept and later donate a proton from histidine, compensating for the positive charge generated on the histidine residue. After the catalytic serine attacks the peptide bond, a tetrahedral transition state is formed, which, upon release of the C-terminal cleavage product, forms the acyl-enzyme intermediate. The nucleophilic water attack releases the N-terminal cleavage product, and the enzyme's active site becomes available for a new catalytic cycle (Figure 4).[90,101,102]

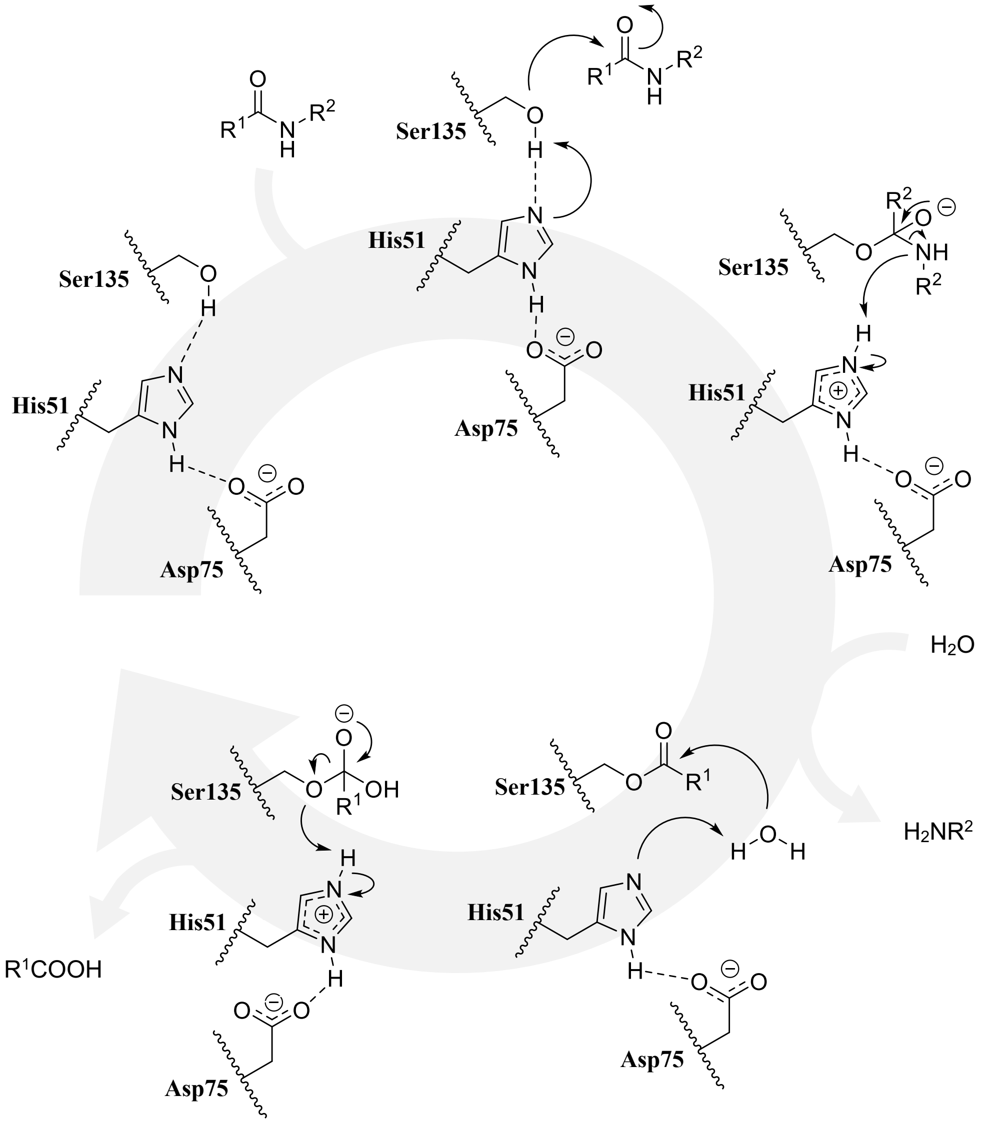


Figure 4. Mechanism of proteolysis of peptides by trypsin-like serine proteases.[101] Numbering of the amino acids of the catalytic center according to 2fom.[94] Figure created with ChemDraw and Microsoft PowerPoint.

In addition to the active site, the NS2B-NS3 protease possesses a proposed allosteric binding site (Figure 5). Ligands can interact either with the allosteric binding site or the active site. The allosteric binding site has been identified by different groups using various methods.[75,103–105] In 2012, Noble *et al.*

described an allosteric binding site formed when the NS2B cofactor is positioned around the NS3 protein. Residues from both NS2B (V78* and M84*) and the NS3 subunit (K73, K74, E88, W89, T118, T120, T122, I123, G124, V147, N152, A164, I165, A166, N167, and T168) form the binding site.[84] This finding was also confirmed by Yildiz *et al.*[103] In 2017, Brecher *et al.* described the allosteric binding site in the open conformation of the protease involving only the NS3 subunit (K73, K74, E88, W89, T118, T120, T122, I123, G124, V147, N152, A164, I165, A166, N167, and T168).[106] This definition of the allosteric binding site differs only slightly from that of Nobel *et al.* resulting in an allosteric pocket localized around Asn152.[84,106] Site-directed mutagenesis of the DENV protease and subsequent blockage of the proposed binding pocket with thio-reactive compounds resulted in decrease in affinity of non-competitive inhibitors to the protease or loss of the protease's activity.[107,108] These experiments also suggest an interaction of the non-competitive inhibitors with the described allosteric binding pocket.[108]

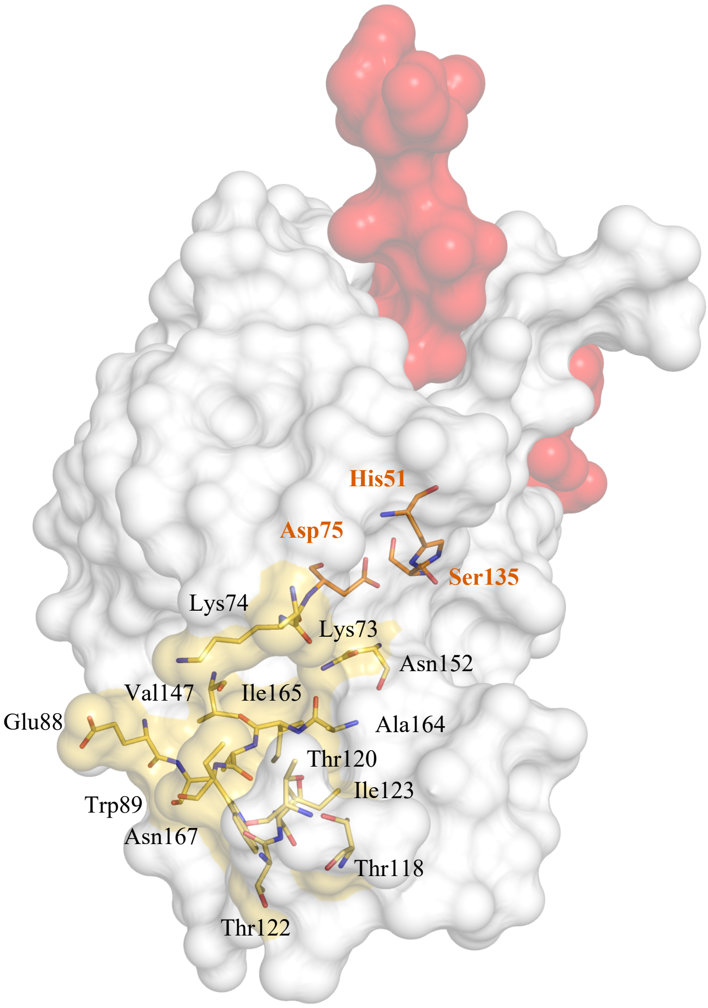


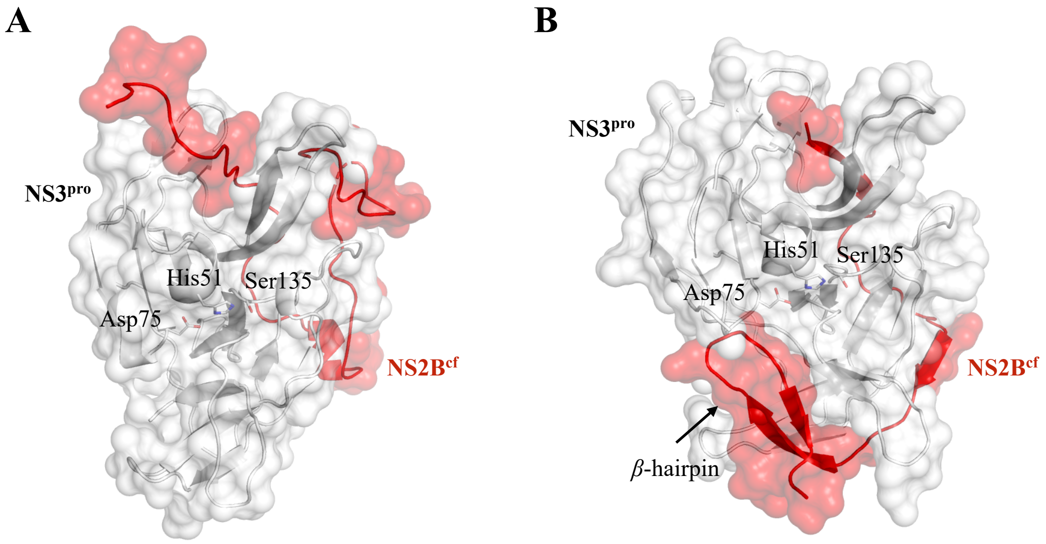
Figure 5. Location of the allosteric binding pocket of DENV NS2B-NS3 protease in the open conformation (PDB-ID: 2fom).[94] Amino acids forming the allosteric binding pocket are shown as yellow sticks.[106] The surface of the allosteric pocket is colored yellow. Amino acids of the catalytic triad are shown as orange sticks. The surface of NS3 is shown in white, and that of NS2B in red. The figure was created using PyMOL.

1.2.4 Structure and Conformations of NS2B-NS3 Proteases

The literature describes at least two different conformations of the NS2B-NS3 proteases' tertiary structure, which mainly differ in their respective orientation of NS2B^{cf} (Figure 6).[84,94,99,106] In the absence of a substrate or ligand at the active site, the C-terminal part of the cofactor assumes an "open" conformation and is loosely bound to NS3 (Figure 6A). When a ligand binds to the active site of the NS2B-NS3 protease, the C-terminus of the cofactor wraps around the NS3 core, forming a "closed" conformation (Figure 6B).[99,109–111] This conformational change of NS2B is crucial for the activity and function of NS3. Due to the occurrence of the closed conformation associated with ligands of the active binding site, the closed conformation is referred to as the catalytically active conformation.[32] In solution, the two conformations are in equilibrium, whereby their relative abundance is significantly influenced by factors such as the expression construct, buffer pH, and ionic strength.[106,112–114] When complexed with competitive inhibitors or substrates, only co-crystals of the closed conformation have been obtained so far (Figure 6B and D).[84,99,115–117] The open conformation was observed without ligands.[94,118] A set of crystal structures of complexes of allosteric inhibitors with the DENV protease in the open conformation have been published (PDB-ID: 6MO0, 6MO1, and 6MO2).[119] However, within the scientific community, the reliability of the crystal structures is discussed. Probably not an allosteric inhibitor but rather, the C-terminus of NS3 is placed in the allosteric pocket.[120–122] Consequently, these structures will not be discussed further in this thesis.

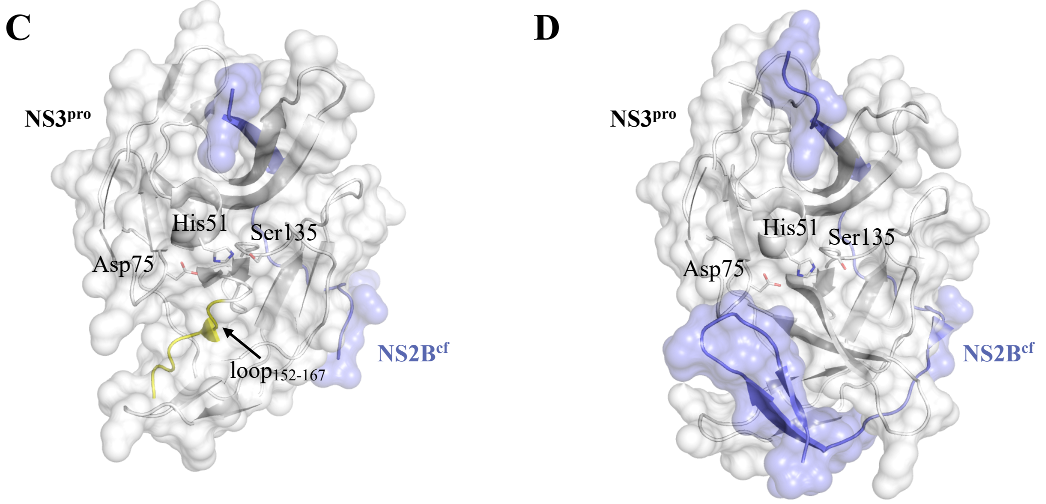
Despite their high similarity, the open conformation of the ZIKV and DENV proteases exhibit a different positioning of the loop$_{152\text{-}167}$, which is involved in substrate recognition in the active site (Figure 6C). The loop bends at Gly151 in the ZIKV protease and extends in the opposite direction compared to the closed conformation. The stabilization of the loop in this position is achieved through hydrophobic interactions.[123] Additionally, the C-terminus of the NS2B^{cf} in the ZIKV protease exhibits increased flexibility, which is explained by Roy *et al.* through sequence variations in positions 91 to 96, resulting in a reduced interaction between the NS2B^{cf} and the NS3^{pro}.[115,123,124]

Brecher *et al.* published a conformational switch assay based on split luciferase complementation to identify potential allosteric inhibitors and demonstrated conformational changes in NS2B. They showed that competitive inhibitors stabilize the closed conformation, while allosteric inhibitors suppress this stabilization. It was therefore speculated that allosteric inhibitors bind to the open conformation or induce a conformational change from the closed to the open conformation.[106]



DENV NS2B-NS3
open

DENV NS2B-NS3
closed



ZIKV NS2B-NS3
open

ZIKV NS2B-NS3
closed

Figure 6. Crystallographically determined conformations of NS2B-NS3 proteases. **A** Structure of the open conformation of DENV2 NS2B-NS3 (PDB-ID: 2fom).[94] **B** Structure of the closed conformation of DENV3 NS2B-NS3 (PDB-ID: 3u1i).[84] **C** Structure of the open conformation of ZIKV NS2B-NS3 (PDB-ID: 5gxj).[118] **D** Structure of the closed conformation of ZIKV NS2B-NS3 (PDB-ID: 5lc0).[115] The cartoon and the transparent surface of NS3 are colored white. NS2B is colored red (DENV) or blue (ZIKV). Amino acids of the catalytic triad are shown as sticks. The loop$_{152\text{-}167}$ is highlighted in yellow.

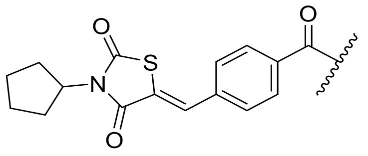
1.2.5 Inhibitors of NS2B-NS3 Proteases

The NS2B-NS3 protease is involved in processing the viral polyprotein and is essential for replicating DENV or ZIKV. This makes the protease an interesting drug target. Both low-molecular-weight non-peptidomimetic inhibitors and substrate-like peptidomimetic inhibitors are described in the literature.[121]

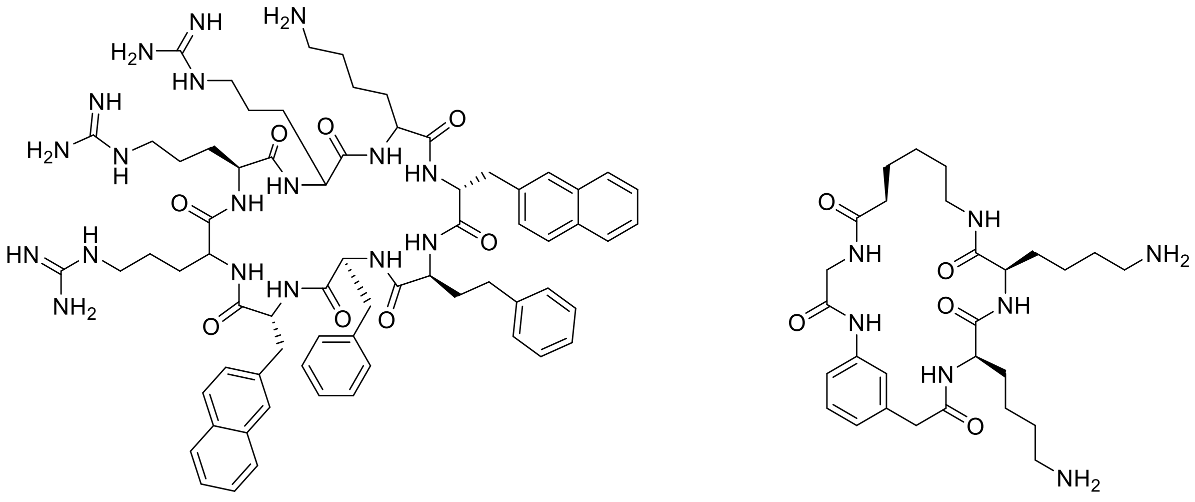
Competitive Inhibitors. Due to the flat nature of the substrate binding site in flaviviral proteases, targeting it with non-peptidomimetic small molecules is challenging. Consequently, extensive research is dedicated to developing protease inhibitors that mimic substrate analogs.[77] However, due to their positive charge, effective inhibitors resembling the preferred dibasic substrates exhibit limited bioavailability *in vivo*.[33,125,126] Among the known protease inhibitors, aprotinin, a macromolecular endogenous inhibitor, inhibits the NS2B-NS3 protease with an IC_{50} value of 65 nM (DENV2). However, other known protease inhibitors such as benzamidine, leupeptin, or pepstatin A show no activity against the DENV protease.[32,127]

Yin *et al.* investigated tetrapeptides with varying C-terminal electrophilic headgroups, also termed “warheads” (Table 1). Aldehydes, trifluoromethyl ketones, or boronic acids were used, which covalently and reversibly bind to the catalytically active serine. The best inhibitors of this series showed K_i values between 0.043 and 5.8 µM (Table 1).[128,129] Based on the peptide aldehyde inhibitors published by Yin *et al.*, Behnam *et al.* synthesized tripeptides with the structure Cap-Arg-Lys-X-NH_2 and were able to generate a non-covalent competitive inhibitor with a K_i value of 2.1 µM by introducing a C-terminal phenyl glycine (Table 1). Thiazolidinedione as a cap structure enhanced the inhibitor’s potency, resulting in a K_i of 0.4 µM (Table 1).[130]

Table 1. Competitive inhibitors, which address the DENV NS2B-NS3 protease.[129,130]

	Cap	**peptide sequence**	**warhead**	**K_i [µM]**
Yildiz *et al.*	Bz	Nle-Lys-Arg-Arg	aldehyde	5.8
	Bz	Nle-Lys-Arg-Arg	trifluormethyl	0.85
	Bz	Nle-Lys-Arg-Arg	boronic acid	0.043
Behnam *et al.*	Bz	Arg-Lys-phenyl glycine	-	2.1
		Arg-Lys-phenyl glycine	-	0.4

Takagi *et al.* published a study on the high-throughput synthesis of cyclic peptides, which are generally considered a privileged class of molecules as they can target so-called undruggable targets such as protein-protein interactions.[131–133] Of 33 synthesized cyclic peptides, 17 showed significant inhibition against the NS2B-NS3 DENV2 protease with IC_{50} values ranging from 0.95 to 15.1 µM.[133] The most potent inhibitor is shown in Figure 7. Braun *et al.* also investigated various macrocyclic inhibitors that resemble the substrate and identified an inhibitor with a K_i value of 142 nM against the DENV4 protease and a K_i value of 1.6 nM for the ZIKV protease (Figure 7).[134]



IC_{50} (DENV2) = 1.1 µM
EC_{50} (DENV2) = 2.0 µM
Takagi *et al.*

K_i (ZIKV) = 1.6 ± 0.3 nM
K_i (DENV4) = 142 ± 10 nM
Braun *et al.*

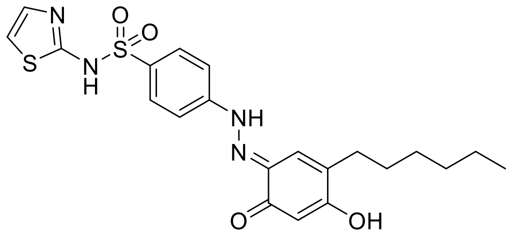
Figure 7. Cyclic inhibitors addressing the ZIKV and DENV NS2B-NS3 proteases published by Takagi *et al.*[133] and Braun *et al.*[134]

Allosteric Inhibitors. Although the flaviviral protease is difficult to inhibit with small molecules, progress has been made in recent years in identifying some non-peptidic allosteric inhibitors. However, the chemical structure of these inhibitors is highly diverse, making structural classification impractical.[32,127] Through high-throughput and virtual screenings, guanidine, cyclohexenyl chalcone, phenol, phthalazine, and benzothiazole derivatives have been identified as allosteric inhibitors of flaviviral proteases (Figure 8).[32,135]

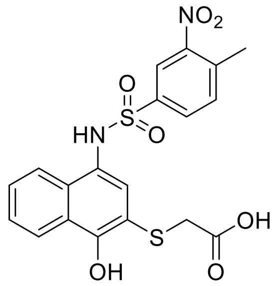
Wu *et al.* synthesized a series of 2-amino benzo[*d*]thiazole derivatives, which selectively and non-competitively inhibit the DENV2 and DENV3 proteases. The most potent inhibitor displayed IC_{50} values of 4.2 µM for DENV2 and 0.99 µM for DENV3, respectively.[135] Millies *et al.* further explored these derivatives and identified new inhibitors with increased solubility. The best inhibitor of this series exhibited an IC_{50} value of 10.0 µM for the DENV2 protease and 0.32 µM for the ZIKV protease.[108]

Pambudi *et al.* conducted a virtual screening of over 660,000 compounds. They discovered a potential allosteric inhibitor, which inhibited the replication of all four DENV serotypes with EC_{50} values below 2.5 µM in a cell-based antiviral activity assay.[105]

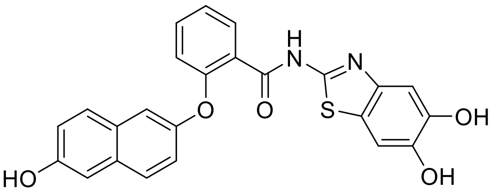
Additionally, Shiryaev *et al.*, in their virtual screening of approximately 275,000 compounds, identified an allosteric inhibitor with a thiazole sulfone amide scaffold, which showed an IC_{50} value of 2.04 µM against the DENV2 protease.[136]



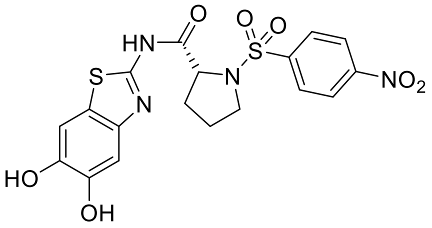
IC_{50} (DENV2) = 2.04 μM
EC_{50} (DENV2) = 59.49 μM
Shiryaev *et al.*



EC_{50} (DENV1) = 0.97 ± 0.42 μM
EC_{50} (DENV2) = 0.98 ± 0.39 μM
EC_{50} (DENV3) = 2.43 ± 0.63 μM
EC_{50} (DENV4) = 0.74 ± 0.48 μM
Pambudi *et al.*



IC_{50} (DENV2) = 4.2 ± 0.44 μM
IC_{50} (DENV3) = 0.99 ± 0.10 μM
Wu *et al.*



IC_{50} (ZIKV) = 0.32 ± 0.05 μM
IC_{50} (DENV2) = 10.00 ± 0.98 μM
Millies *et al.*

Figure 8. Allosteric inhibitors addressing the ZIKV and DENV NS2B-NS3 proteases published by Shiryaev *et al.*[136], Pambudi *et al.*[105], Wu *et al.*[135] and Millies *et al.*[108]

1.3 Investigating Protein-Ligand Interactions

In the field of drug design, understanding the interactions between ligands and proteins is of paramount importance. Ligand-protein interactions are critical in governing various aspects of drug development, including binding affinity, conformational dynamics, binding mechanisms, and interaction location. By investigating these fundamental aspects, researchers can gain crucial insights into the molecular basis of drug action and design novel therapeutics with improved efficacy and reduced side effects.[137–142]

Binding affinity, which refers to the strength and stability of the complex formed between a ligand and its target protein, is a crucial determinant of the potential efficacy of a drug candidate. Accurate determination of binding affinity provides valuable information for identifying and optimizing compounds with high binding affinities.[138–143]

Proteins are dynamic entities that can adopt different conformations to perform their biological functions. Ligand binding can induce conformational changes in proteins, impacting their stability, activity, and interactions with other molecules. Understanding the conformational dynamics resulting from ligand binding is crucial. Investigating the binding mechanism provides insights into how a ligand interacts with its target protein, including the binding mode, specific residues involved, and the interactions formed. Understanding the binding mechanism improves binding specificity and selectivity, reducing the potential for off-target effects, and enables the design of ligands that selectively stabilize specific protein conformations.[137,139,140,144,145]

1.3.1 Binding Affinity and Binding Mode

Fluorometric Assay. Fluorometric assays have become widely used in medicinal chemistry to analyze protease-inhibiting drug candidates.[146,147] In 1973, the introduction of fluorogenic substrates based on Förster resonance energy transfer (FRET) revolutionized protease activity measurement. These substrates comprise a quencher (Q) and a fluorophore (Figure 9A).[148] Due to their superior detection sensitivity, they replaced the previously used colorimetric *p*-nitroanilide (pNA) substrates.[149] In 1976, the fluorescent coumarin derivative 7-amino-4-methyl coumarin (AMC) and its corresponding amides were introduced as a new fluorogenic protease substrate (Figure 9B). This innovation reduced assay interferences compared to FRET-based and colorimetric pNA substrates.[150]

Fluorometric protease assays operate on a general principle: the target enzyme cleaves an amide bond in the substrate, causing a change in the fluorescence properties of the fluorogenic reporter. This can involve the separation of the quencher from the fluorophore (in the case of FRET substrates, Figure 9A) or the cleavage of an internally quenched fluorogenic residue from the substrate (in the case of AMC substrates, Figure 9B).[151] Subsequently, the fluorophore is excited using a specific excitation wavelength within a cuvette fluorometer or a microplate reader, and the resulting fluorescence is measured at the corresponding emission wavelength. The increasing fluorescence intensity over time measures the enzymatic activity (Figure 9C). Introducing an inhibitor reduces the enzyme's reaction rate (v), allowing for evaluating inhibitory activity using various assessment methods.[143,152–155]

Inhibitory Potency. Two primary measurements are commonly used to assess the inhibitory potency of compounds: the inhibition constant (K_i) and the IC_{50} value. K_i represents the equilibrium constant for the dissociation of the inhibitor-bound enzyme complex. At the same time, the IC_{50} value quantifies the concentration of inhibitor needed to reduce the reaction rate of an enzyme-catalyzed reaction by half (Figure 9D). While K_i remains constant for a particular compound with an enzyme, IC_{50} is a relative value influenced by the substrate and its concentration used in the assay.[156] This relationship is mathematically described by the Cheng-Prusoff equation (Equation 1) for competitive inhibitors, which bind to free enzymes and inhibit the enzyme-catalyzed reaction.[157]

$$\mathrm{IC}_{50} = K_{\mathrm{i}}\left(1 + \frac{[\mathrm{S}]}{K_{\mathrm{M}}}\right)$$

Equation 1

The mathematical relationship reveals that the IC_{50} value of a competitive inhibitor is influenced by the K_i value, substrate concentration ([S]), and the substrate's Michaelis Menten constant (K_M) in the assay. Additionally, the relationship between IC_{50} and K_i varies based on the inhibitor's mode of inhibition. In the case of competitive inhibition, the IC_{50} value is always higher than the K_i value. In the case of non-competitive inhibition (also known as simple mixed or allosteric inhibition), the IC_{50} values remain constant regardless of the substrate and its concentration. The IC_{50} value is equal to the K_i value, irrespective of the substrate concentration used in the assay.[156]

Dixon Plot. Dixon introduced a graphical method in 1953 to determine K_i values using limited data.[158] Plotting the inverted initial turnover rate (1/v) against the inhibitor concentration at two different substrate concentrations results in intersecting lines on the left of the vertical axis for competitive inhibitors (Figure 9E), with the intersection indicating $-K_i$. In non-competitive inhibition, the lines meet at a point on the baseline representing $-K$i (Figure 9F).[156,158,159] The Dixon plot consequently enables the determination of K_i from the turnover rate at different substrate and inhibitor concentrations and distinguishes the inhibition type based on the intersection positions.

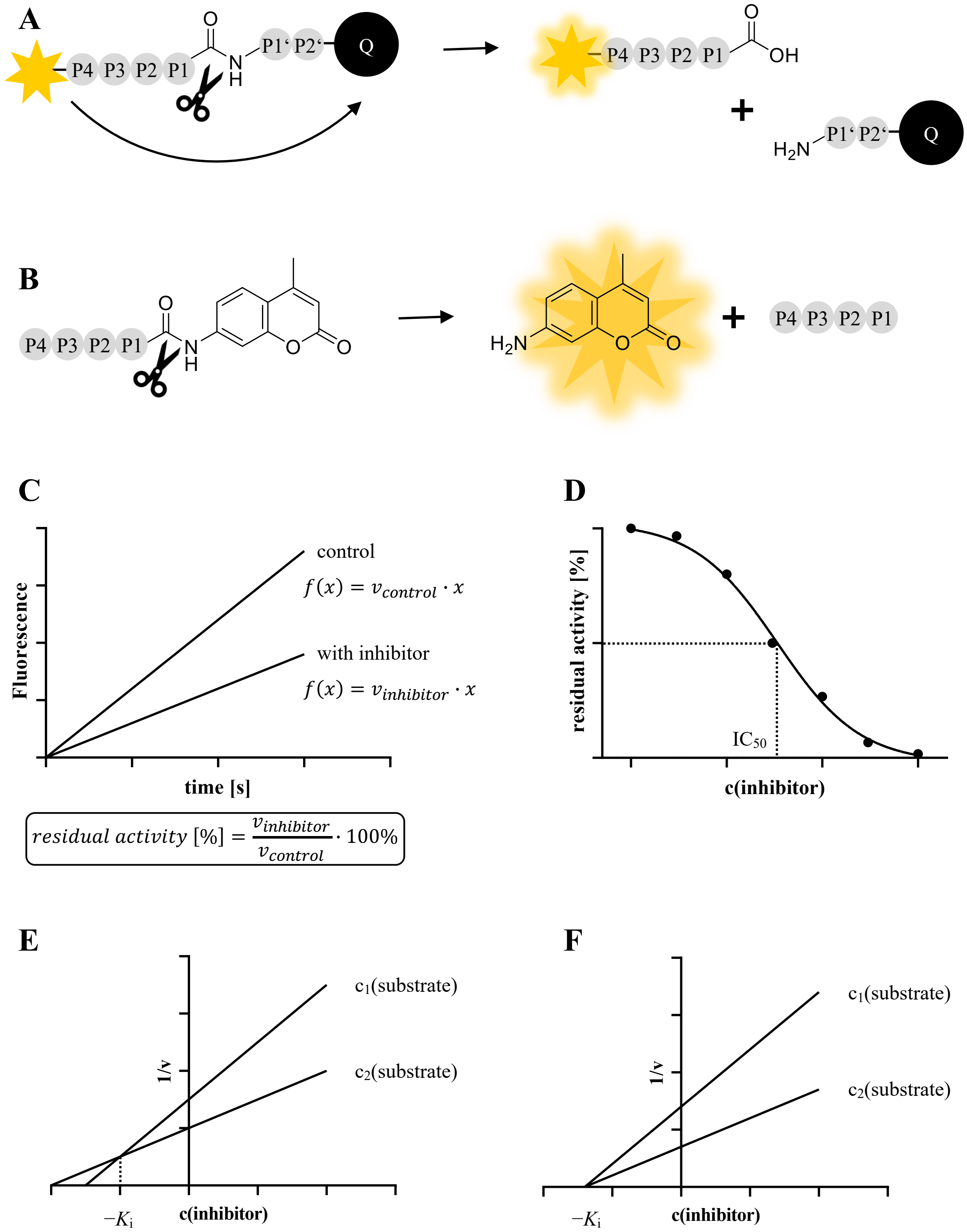


Figure 9. Fluorometric assay. **A** Schematic representation of fluorometric assays based on FRET. Proteolytic cleavage of the peptide allows fluorescence measurement initially suppressed by quencher Q. **B** Schematic representation of fluorometric assays with AMC substrates. The proteolytic release of 7-amino-4-methylcoumarin (AMC) allows its fluorescence to be detected. **C** Typical fluorescence-time curves of a fluorometric assay. The slope of the curve (v) is proportional to the turnover rate of the substrate. Adding an inhibitor causes a lower conversion rate, visible by a slower increasing fluorescence per time. **D** Plot of residual activity versus inhibitor concentration results in a sigmoidal curve. The inflection point of the curve corresponds to the IC_{50} of the inhibitor. **E** Dixon plot for a competitive inhibitor. The straight lines obtained for different substrate concentrations (c_1 and c_2) intersect above the x-axis. Their intersection point corresponds to $-K_i$. **F** Dixon plot for an allosteric inhibitor. The straight lines obtained for different substrate concentrations (c_1 and c_2) meet on the x-axis. Their intersection point corresponds to $-K_i$.

1.3.2 Thermal Stability of Proteins

Thermal shift assays (TSA) enable the investigation of how ligand binding affects the stability of a protein.[160] TSA can be used to determine protein-protein, protein-DNA, protein-lipid, and protein-small molecule interactions.[161–163]

Differential Scanning Fluorometry (DSF). Differential scanning fluorometry (DSF) was introduced in 1997 and has become a valuable tool in drug discovery research, particularly for studying protein interactions and their affinity with other molecules.[164,165] A protein is incubated with a fluorescence dye like SYPRO Orange, which binds to the protein's exposed hydrophobic regions. This binding induces a change in the dye's fluorescence properties.[160] Due to its favorable properties, SYPRO Orange is the dye of choice for DSF. Its high signal-to-noise ratios enhance measurement accuracy, and its excitation wavelength of 500 nm reduces the influence of small molecules on the dye's fluorescence.[160,165]

The protein-dye mixture is heated at a controlled rate, and the resulting fluorescence is analyzed (Figure 10A). According to Bowling *et al.*, proteins exist in a thermodynamic equilibrium where they can be folded or unfolded.[166] When the energy (temperature) of the surrounding environment increases, proteins transition toward the unfolded state. This transition can be quantified to determine a protein's melting temperature (T_m). T_m is defined as the temperature at which 50% of a protein sample is folded, and 50% is unfolded.[167] As the protein unfolds under heating, its hydrophobic regions become exposed, allowing the dye to bind to these regions (Figure 10A). The change in fluorescence properties of the dye can be analyzed.[165] Measured changes in fluorescence intensity as a function of temperature result in a sigmoidal curve. The curve's inflection point corresponds to the midpoint of the protein's unfolding transition, representing T_m.[160]

DSF can be used to determine protein-ligand interactions and assess optimal conditions for producing and storing specific proteins by examining their melting behavior. An advantage of this method is the minimal amount of protein required.[160,165,168] Protein stability can be influenced by pH, ionic strength, and specific ions or molecules. Stabilization of a protein reduces the Gibbs free energy of the complex by forming new molecular interactions, such as hydrogen bonds and van der Waals interactions, or by inducing conformational changes in the protein. The decrease in Gibbs free energy results in enhanced thermal stability and an increased T_m of the protein. By comparing the T_m of a protein under different environmental conditions or in the presence and absence of ligands, the thermal shift (ΔT_m) can be determined (Figure 10B).[169] This shift serves as an indicator of complex formation and/or thermal stabilization. It is important to note that ΔT_m reflects changes in binding interactions and conformational rearrangements within the protein. However, since the thermal stability profile is obtained across a temperature range, it becomes challenging to derive an accurate RT dissociation constant ($K_d = e^{-\frac{\Delta G}{kT}}$; where k is Boltzmann's constant, and T is the thermodynamic temperature) directly from ΔT_m. Focusing solely on T_m might overlook other valuable systemic and thermodynamic information concerning protein stability.[165]

NanoDSF. NanoDSF is a modified label-free DSF method used to determine the thermal stability of proteins.[170,171] It requires a measurement device capable of capturing a broad spectrum of light intensity and light scattering from a sample. This device can measure protein unfolding and aggregation.[172,173] Protein unfolding is assessed through the intrinsic tryptophan and/or tyrosine fluorescence, which changes during thermal unfolding.[170] The fluorescence intensity and maximum of tryptophan strongly rely on its immediate environment.[174] Consequently, the ratio of fluorescence intensities at 330 nm and 350 nm is a suitable indicator for detecting protein structure alterations (Figure 10C).[173,175] The temperature at which protein aggregation occurs after unfolding can be determined through backscattering light.[173] Like DSF, nanoDSF generates a sigmoidal curve of measured fluorescence ratios plotted against temperature, whose inflection point represents T_m (Figure 10D). Alternatively, T_m can be determined from the maximum of the first derivative of these fluorescence ratios as a function of temperature (Figure 10E).[165] Unlike DSF, nanoDSF does not rely on a dye. By omitting the dye, the protein remains closer to its native state, avoiding unwanted effects such as dye autofluorescence or interference with other molecules.

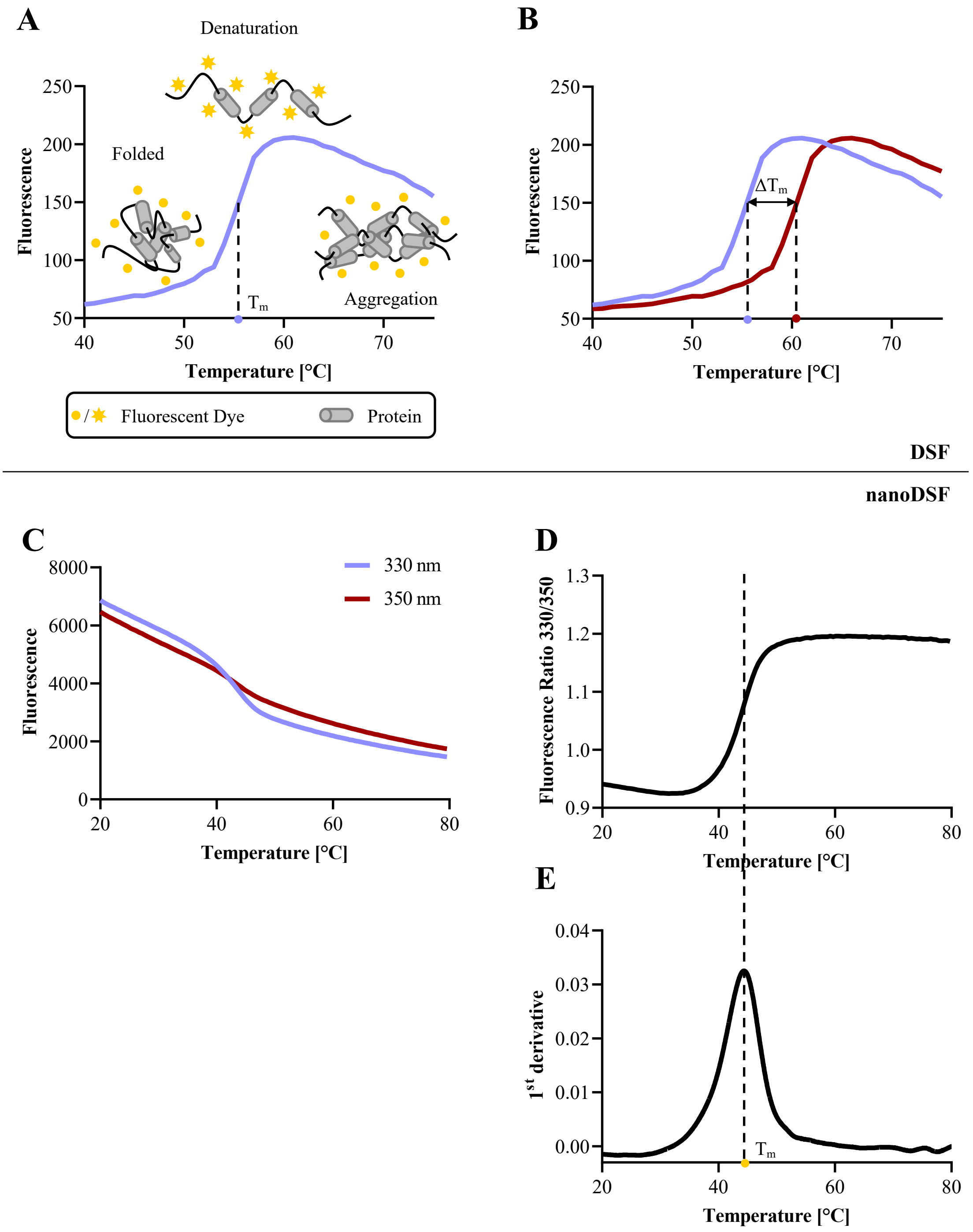


Figure 10. Schematic representation of thermal shift assays. Figure generated with Microsoft PowerPoint according to Gao *et al.*[165] **A** Typical thermal denaturation profile of a protein sample. Fluorescence emission changes with the temperature. The sigmoidal curve demonstrates the cooperative unfolding behavior of the protein when exposed to SYPRO Orange dye (yellow dots/stars), where the dye is bound to the native protein (grey). The curve's peak indicates that all the proteins have unfolded into linear peptides or the protein's hydrophobic core is exposed to SYPRO Orange. Following the peak, the fluorescence decreases due to various mechanisms resulting in less dye bound to the protein. The transition curve's midpoint corresponds to the protein's melting temperature (T_m). **B** DSF curves under two different conditions: in the absence of a ligand (shown in blue) and in the presence of a ligand (shown in red). The difference in melting temperatures between these two conditions is denoted as ΔT_m. **C** Intrinsic fluorescence of tryptophan is measured at both 330 and 350 nm wavelengths and plotted versus temperature during unfolding. **D** Fluorescence ratio at 330 nm and 350 nm (F330/350) of tryptophan plotted against temperature. **E** The melting temperature is calculated by the first derivative of the F330/350 plots.

1.3.3 Conformational Dynamics of Proteins

Studying protein dynamics is crucial for understanding their function and regulation, with implications for protein engineering and drug discovery.[176] Although significant progress has been made in determining high-resolution protein structures [177–179], these models provide a static view of proteins.[176] In reality, proteins are dynamic and constantly fluctuate between multiple conformational states, influenced by environmental conditions and interactions with other molecules.[180,181] Two models, namely induced fit (IF)[182] and conformational selection (CS), describe the interaction between proteins and ligands.[183–186] In the IF model, the protein's bound conformation is formed only after interaction with the ligand. In contrast, the CS model suggests that the protein's pre-existing conformation, sampled without a ligand, is selected for binding. The key distinction lies in whether the conformational change occurs before or after ligand binding (Figure 11).[184,186] Notably, these two models represent limiting cases, and mixed mechanisms are also discussed.[187,188]

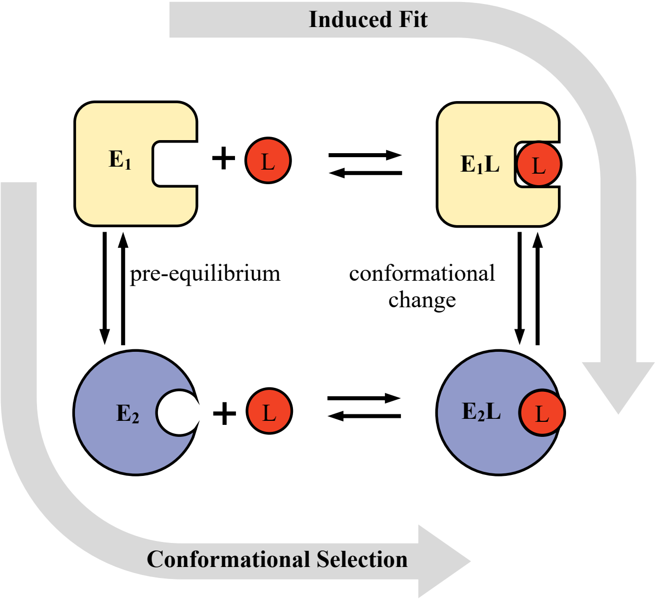


Figure 11. The thermodynamic cycle for molecular recognition involves induced fit (IF) or conformational selection (CS). In CS, the binding competent conformation (E_2) is pre-existing in the solution before the ligand (L) is added, while in IF, the conformational change is induced by ligand binding. Figure created with Microsoft PowerPoint according to Boehr *et al.*[186]

Various experimental techniques have investigated protein structure and dynamics, including X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR), and cryogenic electron microscopy.[177–179] However, these ensemble-based methods are limited in characterizing conformational heterogeneity and real-time conformational changes.[176] Single-molecule techniques, such as single-molecule Förster resonance energy transfer (smFRET) spectroscopy, have emerged as ideal tools for directly measuring distances within biological molecules and observing their dynamic modulation over time. These techniques provide valuable insights into molecular function, motion, and dynamics.[189,190]

Fluorescence and Förster Resonance Energy Transfer (FRET). Förster Resonance Energy Transfer (FRET) involves two molecules, namely donor (D) and acceptor (A), whose absorption spectrum overlaps with the emission spectrum of the donor (Figure 12A). When the donor is excited, energy is transferred non-radiatively to the acceptor, resulting in acceptor fluorescence emission. This process can be explained using a Jablonski diagram (Figure 12B and C).[174]

Upon absorption of a photon close to the resonance energy, a molecule in the ground singlet state (S_0) is promoted to the lowest excited singlet state (S_1). Fast relaxation brings the molecule to the lowest S_1 sublevel in picoseconds. The emission of a photon can occur within nanoseconds, returning the molecule to the ground state. Alternatively, collisional quenching or internal conversion can bring the molecule back to the ground state through non-radiative processes. Additionally, intersystem crossing (ISC) to the first excited triplet state (T_1) can occur in organic molecules. The T_1 state can be relaxed through photon emission (phosphorescence) or non-radiative relaxation.[191]

When an acceptor molecule is near the donor, the fluorescence emission of the donor can be significantly quenched through energy transfer from the donor to the acceptor via Coulomb interaction. Consequently, the acceptor and donor molecules exhibit fluorescent emission following the above mentioned principles (Figure 12C).[191]

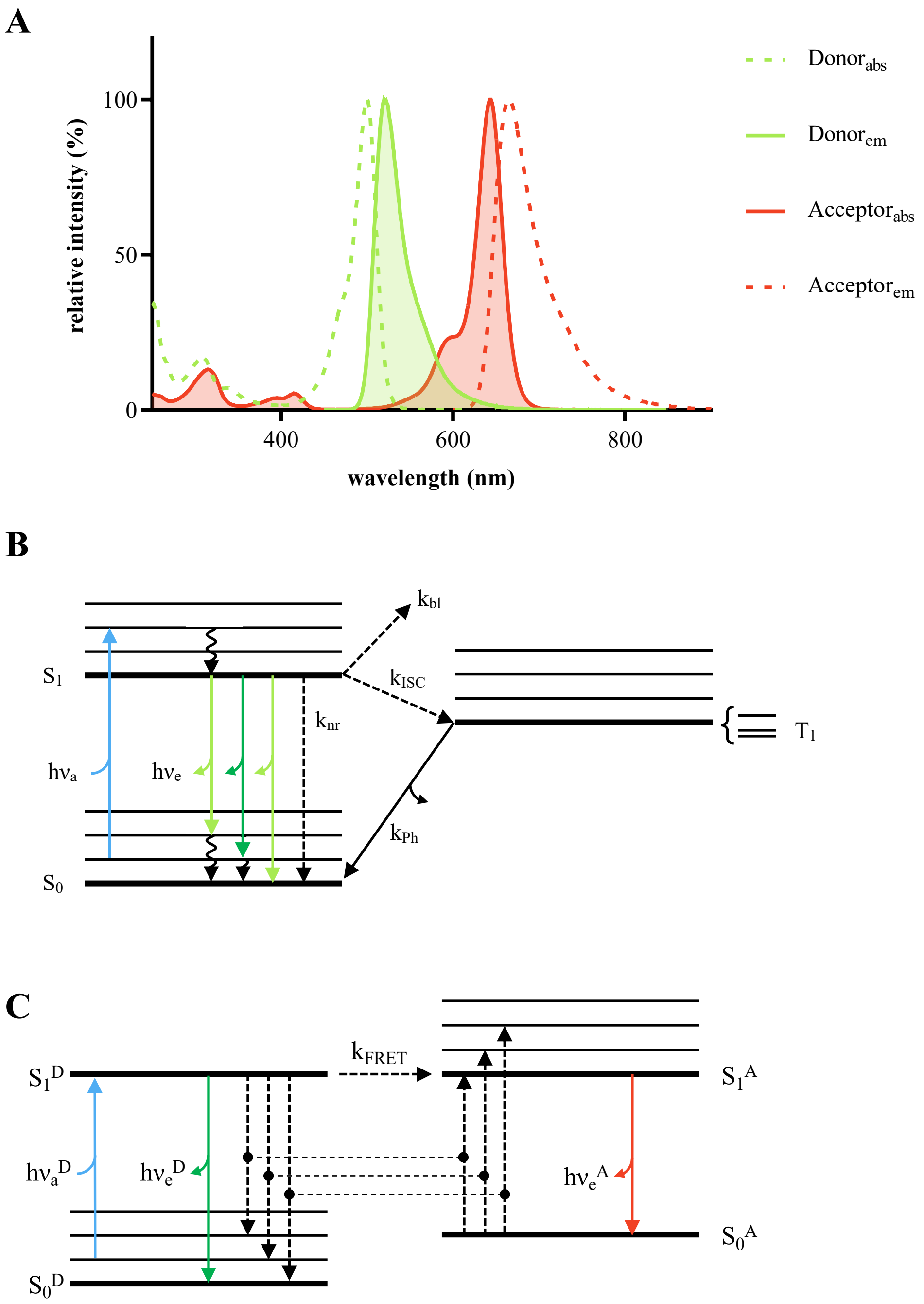


Figure 12. Principle of fluorescence and FRET. **A** Absorption (abs) and emission (em) spectra of a donor (green) and an acceptor (red) dye, exemplarily shown for ATTO 488 (donor) and ATTO 643 (acceptor). The overlapping emission spectrum of the donor and absorption spectrum of the acceptor are shown with solid lines. **B** Jablonski diagram for fluorescence. When a molecule absorbs a photon with energy close to the resonance energy ($h\nu_a$), it transitions from ground singlet state S_0 to the lowest excited singlet state S_1. The molecule can emit a photon ($h\nu_e$), returning it to the ground state. Alternatively, collisional quenching or internal conversion can bring the molecule back to the ground state without photon emission (k_{nr}). In organic dye molecules, an additional process called intersystem crossing (ISC) can transfer the molecule to the first excited triplet state T_1. The molecule can relax to the ground state through photon emission (phosphorescence) or non-radiative relaxation. **C** Jablonski diagram for FRET. FRET involves a donor (D) and an acceptor (A) molecule. Exciting the acceptor to its excited state follows the process described in part **B**. When an acceptor molecule is near the donor, the fluorescence emission of the donor can be significantly quenched. This quenching occurs through energy transfer from the donor to the acceptor via Coulomb interaction, with a rate denoted as k_{FRET}. The acceptor and donor molecules exhibit fluorescent emission following the principles outlined in part **B**. Figure created with Microsoft PowerPoint according to Michalet *et al.*[191]

Parameters of FRET. In the presence of an acceptor molecule close (a few nanometers) to the donor molecule, the donor fluorescence emission can be quenched by energy transfer to the acceptor through Coulomb interaction. The rate of energy transfer, k_{FRET}, depends on the distance between the donor and acceptor (R) and follows an inverse sixth power relationship (Equation 2) where τ_{D0} is the donor fluorescence lifetime in the absence of an acceptor, and R_0 is the dye-pair specific Förster radius.[189,191]

$$k_{FRET} = \frac{1}{\tau_{D0}}\left(\frac{R_0}{R}\right)^6 \qquad \textbf{Equation 2}$$

FRET efficiency (E_{FRET}) depends on the sixth power of the distance ratio (Equation 3), where R is the inter-dye distance, and R_0 is the Förster radius at which $E_{FRET} = 0.5$ (Figure 13A).[189,191,192] For commonly used fluorophores, R_0 is typically around 50–100 Å and depends on the refractive index n, the orientation factor of the donor and the acceptor κ^2, the quantum yield of the donor Φ_D, and the overlap integral *J* between the emission of the donor and the absorption of the acceptor (Equation 4).[174,193]

$$E_{FRET} = \frac{1}{1+\left(\frac{R}{R_0}\right)^6} \qquad \textbf{Equation 3}$$

$$R_0(\text{in Å}) = 0.21 \cdot (n^{-4} \cdot \Phi_D \cdot \kappa^2 \cdot J)^{\frac{1}{6}} \qquad \textbf{Equation 4}$$

E_{FRET} can be determined using the acceptor intensity (F_A) to total emission intensity ($F_A + \gamma F_D$) ratio (Equation 5), with γ incorporating the donor and acceptor quantum yields and detection efficiencies of both channels. Alternatively, the decreased lifetime of the donor molecule in the presence of the acceptor (τ_{DA}) compared to the donor's lifetime in the absence of the acceptor (τ_{D0}) can be used to determine E_{FRET} (Equation 5).[191,194]

$$E_{FRET} = \frac{F_A}{F_A + \gamma F_D} = 1 - \frac{\tau_{DA}}{\tau_{D0}} \qquad \textbf{Equation 5}$$

The strong dependence of E_{FRET} on R has led to FRET being commonly referred to as a "spectroscopic ruler", enabling real-time observation of conformational dynamics in single molecules through monitoring E_{FRET} changes (Figure 13B).[192,195]

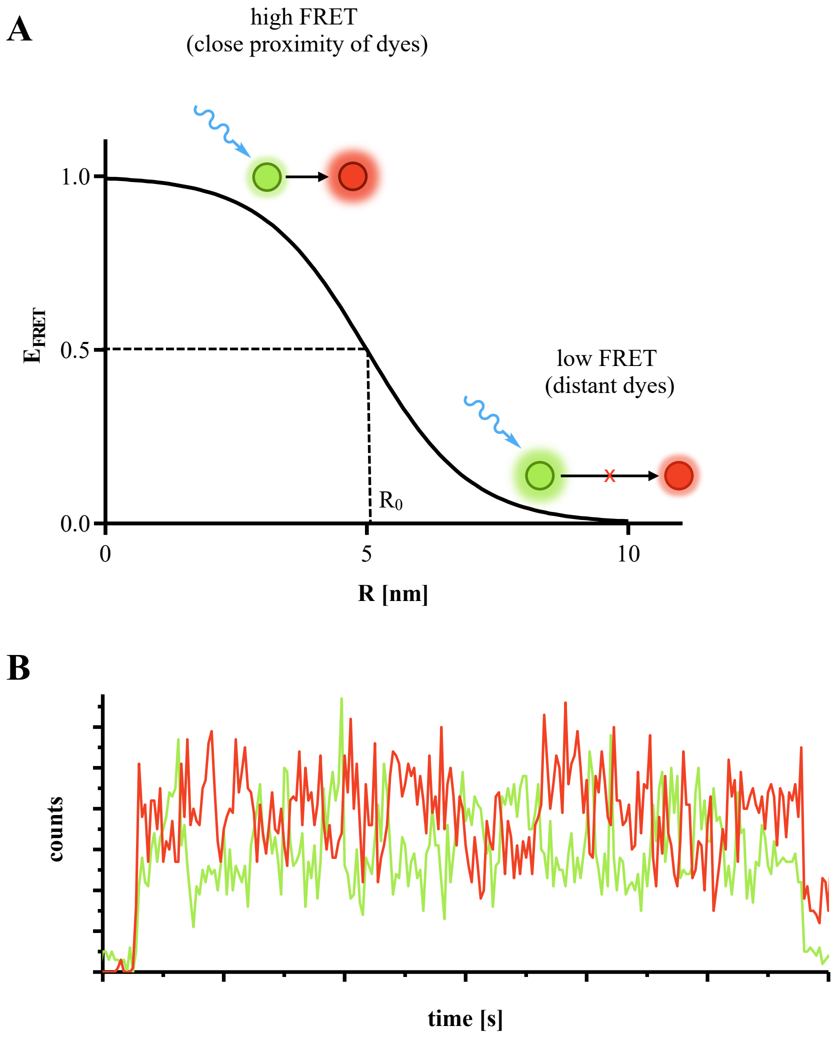


Figure 13. smFRET description. **A** The FRET efficiency (E_{FRET}) can be described as a function of the inter-dye distance (R) for a given Förster radius (R_0). When the donor dye is directly excited, it can either fluoresce or transfer energy to the acceptor dye, depending on its proximity. At $R = R_0$, E_{FRET} is 0.5, indicating an equal fluorescence and energy transfer probability. However, E_{FRET} is greater than 0.5 at smaller distances, favoring energy transfer, while E_{FRET} is less than 0.5 at larger distances, preferring fluorescence. The curve depicts this relationship, illustrating how E_{FRET} changes with varying inter-dye distances. Notably, the E_{FRET} values adjacent to R_0 exhibit a linear relationship. **B** An illustration of a two-color smFRET dataset. The data consists of recorded donor and acceptor fluorophores intensities on an immobilized dye-pair labeled protein. The anti-correlated behavior of the donor and acceptor signals signifies that the fluctuations in intensity arise from energy transfer processes taking place between the fluorophores. Figure created with Microsoft PowerPoint according to Roy *et al.*[196]

SmFRET methods can be divided into two categories based on the time scale of the dynamics studied: slow conformational dynamics (10 ms–s) observed on surface-immobilized molecules and fast conformational dynamics (ns–ms) studied on molecules diffusing in solution.[176] Freely diffusing single molecules offer more straightforward FRET measurements and provide insights into population distributions of inter-dye distances.[196–200] However, the ability to monitor individual molecules over extended periods by immobilization offers dynamic information ranging from milliseconds to minutes.[196]

Experimental Setup. Single-molecule fluorescence spectroscopy is typically conducted using confocal microscopy or total internal reflection fluorescence microscopy (TIRF), requiring high numerical aperture objectives to focus excitation light and minimize background noise. Despite using high-efficiency filters and detectors, such setups' collection efficiency remains below 10% (Figure 14).[191]

In the confocal geometry, a collimated laser beam is directed into the back focal plane of an objective lens, focusing the excitation light to a diffraction-limited volume within the sample (Figures 14A and B). The emitted fluorescence from molecules within this volume is collected by the same objective and transmitted through dichroic mirrors, lenses, and color filters to one or multiple point detectors (avalanche photodiode, APD) (Figure 14E). A crucial element of this setup is the presence of a pinhole in the detection path, carefully sized to allow only light originating from the excitation volume region to reach the detectors. Light from above or below the focal plane is not focused on the pinhole and does not reach the detector. This enables to reject out-of-focus light.[174,197] Signals from freely diffusing molecules manifest as bursts of varying sizes and durations, typically within a few milliseconds (Figure 14A). For immobilized molecules, localization using a scanning device is necessary before recording can begin (Figure 14B). The resulting time traces exhibit one or more fluctuating intensity levels until the molecule eventually undergoes photobleaching after a few seconds.[191,201,202] The ability to perform optical sectioning at any plane within the sample, rather than solely at interfaces with varying refractive indices, makes confocal microscopy more versatile than TIRF. Nevertheless, there are additional dissimilarities that, in certain specific applications, may promote the preference for TIRF.[203]

In TIRF, a laser beam is carefully shaped to reach the interface between the glass and buffer at a critical angle θ. This configuration generates an evanescent wave (with a decay length of a few hundred nanometers) in the sample, selectively exciting fluorescence only from molecules near the surface, resulting in minimal background noise (Figures 14C and D). TIRF can be achieved by illuminating through the objective (Figure 14C) or coupling the laser through a prism (Figure 14D). A wide-field detector (camera) is employed, enabling the simultaneous recording of multiple single-molecule signals, albeit with potentially lower time resolution than point detectors (Figure 14F). Individual intensity trajectories can be extracted from recorded videos, providing similar information to that obtained using the confocal geometry. However, simultaneous recording of fluorescence live times is currently only possible to a limited extent. Wide-field setups, particularly TIRF, are primarily used for surface studies involving molecules bound to surfaces or embedded in membranes.[191,196]

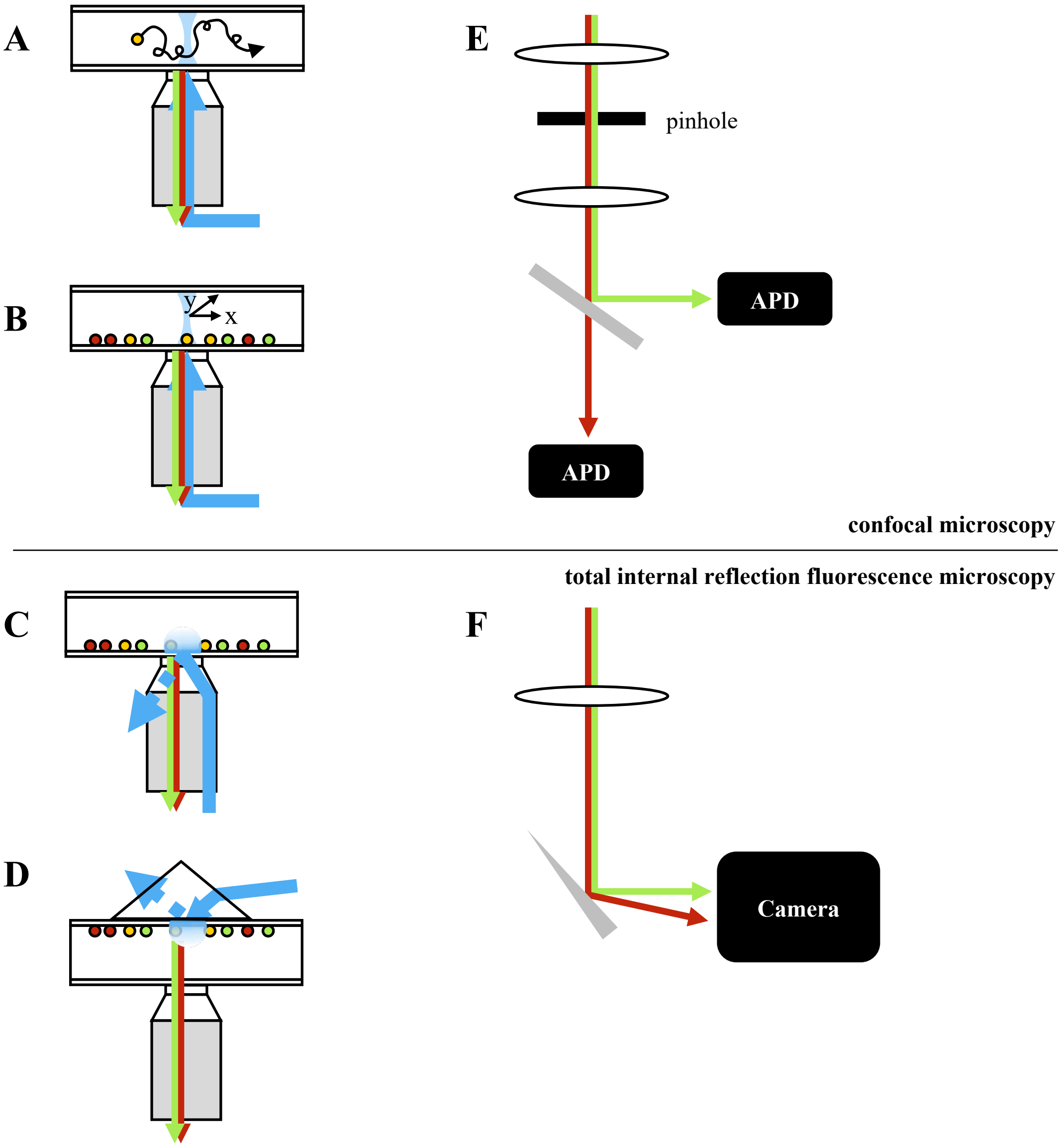


Figure 14. Experimental setup for smFRET. In confocal microscopy (**A, B**), a laser beam focused into a diffraction-limited volume within the sample. The same objective lens collects fluorescence emitted by molecules within this volume. The setup allows the investigation of freely diffusing molecules (**A**) or immobile molecules (**B**). The latter must be initially localized using a scanning device denoted by two arrows (x and y axes) before data acquisition can begin. In the total internal reflection fluorescence (TIRF) geometry (**C, D**), a laser beam is carefully shaped to reach the interface between the glass and buffer at a critical angle. This configuration generates an evanescent wave (depicted by a semicircle) within the sample, with a few hundred nanometers decay length. The evanescent wave selectively excites the fluorescence of molecules near the surface. TIRF illumination can be achieved by directing the light through the objective lens (**C**) or coupling the laser through a prism (**D**). **E** In confocal microscopy, the emitted fluorescence signals are transmitted through dichroic mirrors, lenses, and color filters before reaching one or multiple point detectors (APD). **F** In TIRF, a wide-field detector, such as a camera, is employed. This detector enables the simultaneous recording of multiple single-molecule signals in parallel. Figure created with Microsoft PowerPoint according to Michalet *et al.*[191]

Evaluation Methods. Conformational dynamics studies aim to address three key questions: (i) characterizing the various conformational states adopted by a molecule, (ii) quantifying the rates at which transitions occur between these states, and (iii) identifying transition pathways, correlations, and memory effects, or variations in kinetic constants over time. Fluorescence spectroscopic techniques offer a method to explore these questions by monitoring the temporal changes in fluorescence observables such as spectrum, lifetime, polarization anisotropy, and FRET efficiency. These observables serve as indirect indicators of the underlying conformational dynamics of the molecule.[191]

Fluorescence Correlation Spectroscopy (FCS). Conformational changes occurring at a rapid time scale often cannot be effectively captured by a simple analysis of the observable data (lifetime, E_{FRET}, etc.), necessitating time correlation techniques such as fluorescence correlation spectroscopy (FCS) or related methods.[204–206] In FCS (Figures 15A, C and E), the autocorrelation function G of the fluorescence intensity F(t) (Figure 15A) is calculated for different time lags τ with $\langle F(t)\rangle = \frac{1}{T}\int_0^T F(t)dt = \langle F(t+\tau)\rangle$ for a stationary process (Equation 6).[191,194,207,208]

$$G(\tau) = \frac{\langle F(t)\cdot F(t+\tau)\rangle}{\langle F(t)\rangle^2} - 1 = \frac{\langle \delta F(t)\cdot \delta F(t+\tau)\rangle}{\langle F(t)\rangle^2} \qquad \textbf{Equation 6}$$

To increase the relative fluctuations $\delta F(t) = \frac{F(t)-\langle F(t)\rangle}{\langle F(t)\rangle}$, it is crucial to minimize the number of molecules present within the observation volume V_{eff}. A simple three-dimensional Gaussian profile is a commonly used and reasonably accurate approximation for describing V_{eff}. When considering the absence of other fluorescence-modulating processes, the primary source of fluctuations arises from the diffusion of particles into and out of the illuminated area V_{eff}. Under these conditions, the autocorrelation function can be analytically solved (Equation 7).

$$G(\tau) = \frac{1}{V_{eff}\langle c\rangle}\cdot\frac{1}{1+\frac{\tau}{\tau_D}}\cdot\frac{1}{\sqrt{1+\frac{\tau}{\omega^2\tau_D}}} = \frac{1}{N}\cdot\frac{1}{1+\frac{\tau}{\tau_D}}\cdot\frac{1}{\sqrt{1+\frac{\tau}{\omega^2\tau_D}}} = \frac{1}{N}\cdot G_{Diff}(\tau) \qquad \textbf{Equation 7}$$

Axial and radial radii r_{xy} and r_z characterize V_{eff}. ω expresses the ratio of these radii. The average number of particles in the observation volume is N, with ⟨c⟩ being the average concentration of particles in the observation volume. The diffusion time of the particle, τ_D, represents the time molecules spend on average in the observation volume. [194,207–209]

Fluorescence autocorrelation spectroscopy is a valuable technique for investigating molecular concentration, mobility, transport, and diffusion (Figures 15C and E). By calculating the autocorrelation curve (Figure 15C), FCS enables the determination of diffusion coefficients, providing insights into the diffusing particle’s size or the surrounding medium’s viscosity. However, when it comes to binding analysis, the utility of fluorescence autocorrelation is limited to scenarios where the binding event significantly and selectively reduces the diffusion of the labeled species. Alternatively, both binding partners can be labeled, and cross-correlation analysis (FCCS) can be employed (Figures 15B, D, and F).[210]

Fluorescence Cross-Correlation Spectroscopy (FCCS). When multiple observables are measured simultaneously, cross-correlation analysis can be employed to explore the time scales of correlated variations between these quantities. This analysis is particularly relevant when examining the intensity fluctuations of two distinct fluorescence channels (i and j) in FRET experiments (Figure 15B). Cross-correlation analysis investigates the presence of simultaneous fluctuations in both channels (F_i and F_j), indicating the simultaneous movement of two fluorophores. This can occur when a molecule labeled with both fluorophores or a complex formed by two singly labeled molecules enters and exits the observation volume. The normalized cross-correlation (Figure 15D) can be calculated by Equation 8.[174,194,207,211]

$$G_{ij}(\tau) = \frac{\langle \delta F_i(t) \cdot \delta F_j(t+\tau) \rangle}{\langle F_i(t) \rangle \cdot \langle F_j(t+\tau) \rangle}$$ **Equation 8**

Without FRET and under the assumption of ideal conditions, where both channels have identical effective volume elements (V_{eff}), and entirely separable emission spectra, the correlation curves can be obtained by Equation 9 (autocorrelation) and 10 (cross-correlation). Herein $G_{Diff,i,j,}$ and $G_{Diff,ij}$ are the motion-related parts of the correlation functions (Equation 7), $c_{i,j}$ are the concentrations for the single-labeled species i and j, respectively, and c_{ij} is the concentration of the double-labeled species.[194,207]

Autocorrelation:

$$G_{i,j}(\tau) = \frac{\left(\langle c_{i,j} \rangle G_{Diff,i,j}(\tau) + \langle c_{ij} \rangle G_{Diff,ij}(\tau) \right)}{V_{eff}\left(\langle c_{i,j} \rangle + \langle c_{ij} \rangle \right)^2}$$ **Equation 9**

Cross-correlation:

$$G_{ij}(\tau) = \frac{\langle c_{ij} \rangle G_{Diff,ij}(\tau)}{V_{eff}\left(\langle c_i \rangle + \langle c_{ij} \rangle \right)\left(\langle c_j \rangle + \langle c_{ij} \rangle \right)}$$ **Equation 10**

In FCCS, when combined with the autocorrelation amplitudes, the cross-correlation amplitude offers valuable insights into binding processes, enzyme kinetics, or dynamic colocalization within small, diffusing entities (Figure 15D and F).[210]

FCCS combined with FRET. The combination of FRET with FCS is attractive to study a donor (D) and acceptor (A) labeled system that undergoes conformational fluctuations. In the excitation volume, the donor is excited, and FRET occurs between the dye-pair, resulting in a burst of fluorescence photos originating from both the donor and acceptor.[212] Two types of information can be derived from photo burst analysis: (i) the time average of all photon events in the burst and (ii) the time-resolved study for intraburts kinetics.[194,213] Simultaneously extracted information can be absolute arrival time (intensity information), energy (spectral information), and fluorescence lifetime of the molecules.[214] Auto- and cross-correlation functions are given by Equation 11 (autocorrelation) and 12 (cross-correlation), where $\tau_{0,i}$ is the relaxation time resulting in a non-fluorescent state, S_i describes the ratio between a state of high E_{FRET} and a state of low E_{FRET} and their efficiency values, and τ_r is the relaxation time between these states. In the autocorrelation function, A_i depends on the average fraction of molecules in a non-fluorescent state, e.g. triplett state.[194,215] Equations 11 and 12 notably consider only doubly labeled species. The equations are more complicated for experiments in which both doubly and singly labeled species are observed.

Autocorrelation:
$$G_{A,D} = \frac{G_{Diff,A,D}(\tau)}{V_{eff}\langle c \rangle} \cdot \left(1 + A_{A,D} \cdot e^{-\frac{\tau}{\tau_0}}\right) \cdot \left(1 + S_{A,D} \cdot e^{-\frac{\tau}{\tau_r}}\right)$$
Equation 11

Cross-correlation:
$$G_{AD}(\tau) = \frac{G_{Diff,AD}(\tau)}{V_{eff}\langle c_{AD} \rangle} \cdot \left(1 - S_{AD} \cdot e^{-\frac{\tau}{\tau_r}}\right)$$
Equation 12

The equilibrium between the state with high E_{FRET} and the state with low E_{FRET} results in an anti-correlation factor between donor and acceptor fluorescence intensities.[194] Notably, the cross-correlation G_{AD} encompasses a diffusion term and, in the presence of fluctuating FRET dynamics, an additional rise term, which accounts for the exchange rates between different FRET states. Consequently, fluctuating FRET dynamics delay the decay of the cross-correlation G_{AD} compared to pure diffusion. Therefore, a faster decay in the cross-correlation suggests reduced FRET dynamics or the stabilization of a specific conformation.[216]

Time-correlated single-photon counting (TCSPC). For time-correlated single-photon counting, the sample is excited with a light pulse, and the time between the excitation pulse and the observed photon is stored in a histogram. Notably, the conditions must be adjusted to detect considerably less than one photon per laser pulse (typically one photon per 100 excitation pulses).[174] Three tags allow the unique identification of each detected photon: (i) the channel number gives information for the spectral region and/or polarization of the photon, (ii) the micro time corresponds to the time elapsed between the excitation pulse and the photon detection, and (iii) the macro time corresponds to the number of excitation pulses that have occurred from the start of the experiment until the detection of the photon.[212] Time-resolved measurements of donor fluorescence decays are helpful for accurate FRET measurements offering three main advantages. (i) The FRET efficiency can be determined without instrumental calibrations using the lifetimes of the excited states of the donor reference sample and the donor-acceptor sample. (ii) The joint decay analysis of the donor-only reference and the FRET sample allows to relate the FRET rate constant directly to the inter-dye distance of the acceptor and donor. (iii) The histogram of the donor-acceptor sample contains information about the heterogeneity of the FRET sample so that a distance distribution can be directly resolved.[217] TCSPC analysis is usually limited to confocal microscopy.[174,212] However, in recent years, efforts have been made to develop detectors for wide-field geometries that allow TCSPC analysis.[218,219]

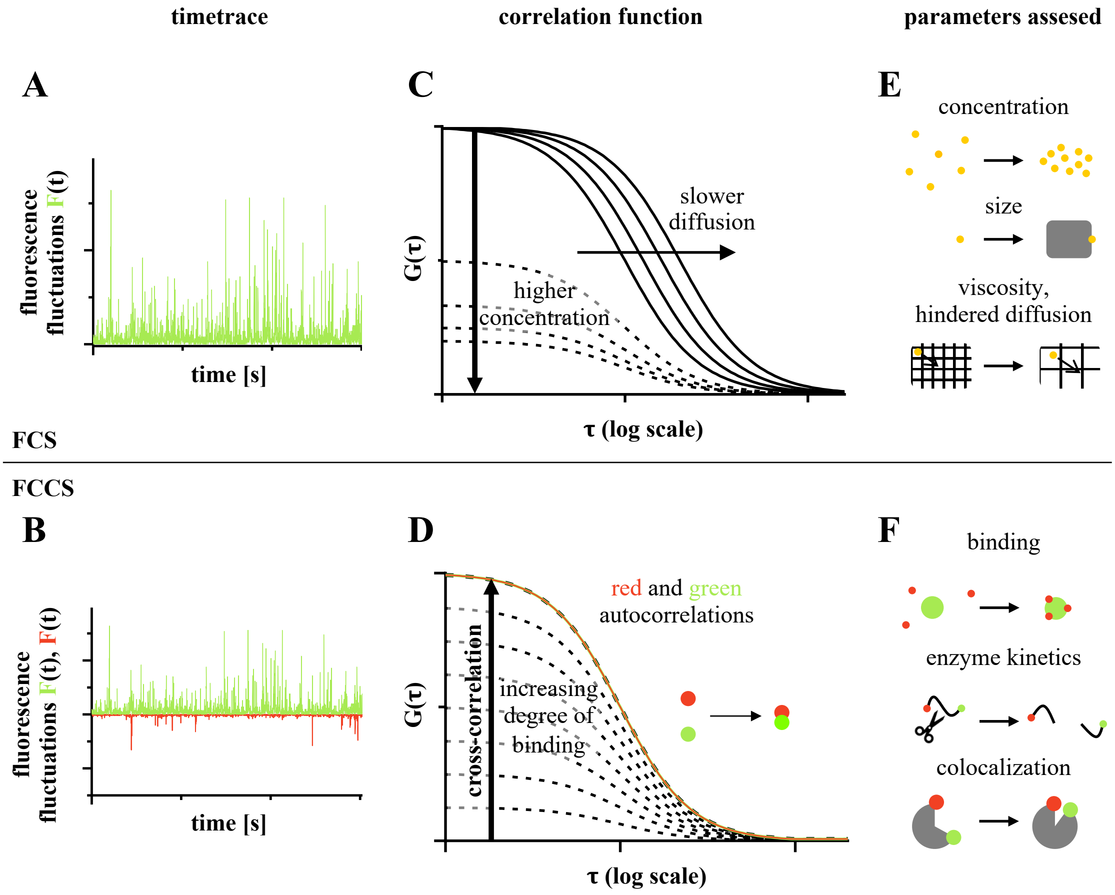


Figure 15. Evaluation of smFRET experiments. **A, B** Time traces collected from one (**A**) or two (**B**) fluorescent dyes for single-color FCS and dual-color FCCS, respectively. **C** Autocorrelation function G(τ). The characteristic decay time of G shows the mobility of the particles (solid lines). The inverse amplitude of G(0), is proportional to the particle concentration (dashed lines). **D** Autocorrelation functions (red and green) and cross-correlation functions between the red and the green traces (black dashed lines). The relative amplitude of the cross-correlation curve is a measure of the degree of binding or colocalization. **E, F** Parameters assessed by FCS (**E**) and FCCS (**F**). Figure created with Microsoft PowerPoint according to Bacia *et al.*[210,211]

Literature Described Experiments. SmFRET experiments have previously been used to study the conformational dynamics of diverse protein classes. Several investigations have focused on ATP-binding cassette (ABC) transporters, membrane protein complexes that utilize ATP energy to transport molecules across the cell membrane.[220–223] Gouridis *et al.* investigated these transporters' substrate binding domain (SBD) dynamics by smFRET of immobilized dye-pair labeled SBD. They showed that 95% of the molecules are in an open conformation without substrates, while the presence of substrates favored the closed conformation.[220] Another extensively studied protein using smFRET is Hsp90, a homodimeric protein involved in the final stages of folding various client proteins, utilizing ATP as a fuel source. It was shown by smFRET experiments that Hsp90 dimers adopt an open V-shaped conformation and undergo a significant conformational change during their active cycle to form a closed conformation.[176,224–228]

SmFRET can help to unravel the interaction mechanism (CS or IF) between a protein and a ligand.[229] Exemplarily, Kim *et al.* utilized two- and three-color smFRET techniques to uncover the highly debated mechanism[186,230,231] underlying the molecular recognition of the maltose binding protein (MBP).[232] Using a FRET pair (Cy3, Cy5) to label the protein, they observed a dynamic equilibrium between an open and a closed state for individual MBP molecules. By introducing a fluorescently labeled ligand (Cy7), they could capture, in real-time, the specific state(s) that the ligand recognized and interacted with selectively for the first time. The authors proposed MBP recognition of maltose to operate *via* an extended induced fit model, in which the protein exists in a conformational equilibrium where the ligand binding induces a structural transition.[232]

1.3.4 Location of Ligand Binding

Photoaffinity Labeling. Singh *et al.* introduced the concept of photoaffinity labeling (PAL) in the 1960s, utilizing photoreactive groups (PG).[233] After acylation, they integrated diazo groups into the enzyme chymotrypsin, forming intramolecular cross-links during photolysis.[144,233–235] This technique has become crucial in studying drug targets and protein-ligand interactions. PAL enables the analysis and characterization of various interactions, including protein-protein, protein-nucleic acid, protein-cofactor, and protein-ligand interactions.[234–236] It facilitates the elucidation of protein structures, identification of unknown ligand targets, and discovery of new binding sites in proteins. Moreover, PAL allows targeted investigations of different binding sites' structures and ligands' binding behavior. In PAL, photoreactive groups covalently attached to the ligands react upon light stimulation, forming a reactive intermediate that rapidly establishes covalent bonds with the nearest reacting partner.[144,234]

Photoreactive Probes. An ideal photoaffinity probe should have a similar structure to the original ligand and exhibit comparable affinity and activity levels. It should be activated at a wavelength that does not harm biological molecules but generates highly reactive intermediates. Phenyldiazirines, benzophenones, and aryl azides are well-known examples of classical photoreactive groups (Figure 16A).[234,236,237]

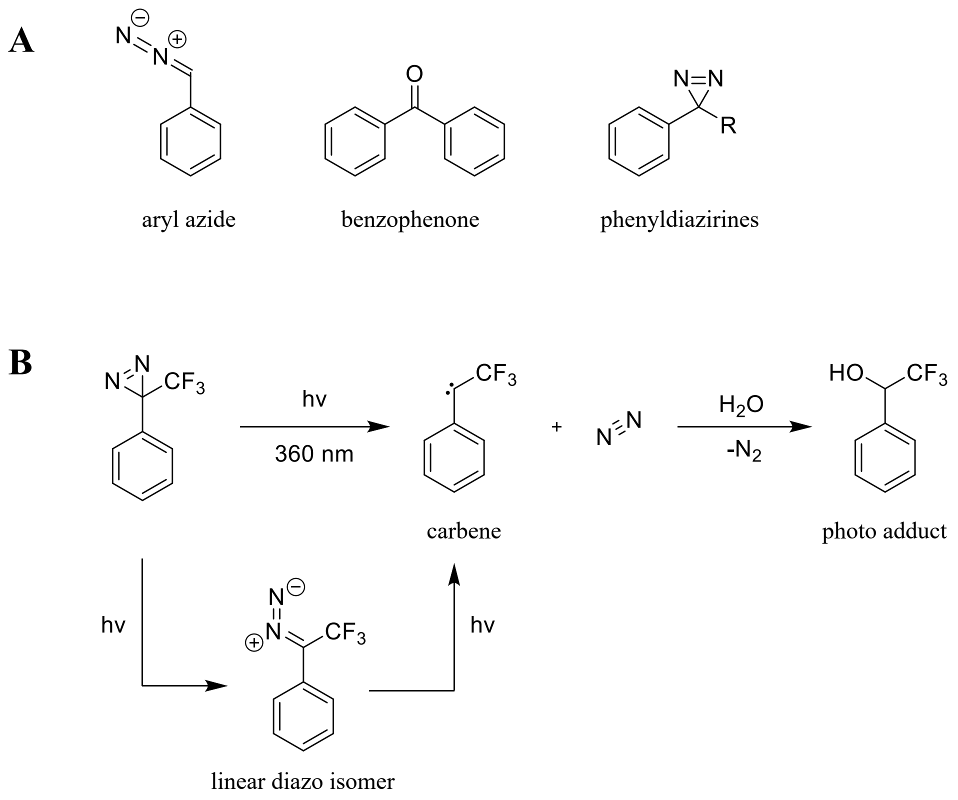


Figure 16. Photoreactive groups **A** Aryl azides, benzophenones, and phenyldiazirines are typically incorporated into photoaffinity probes. **B** Reaction of trifluoromethylphenyldiazirine with water upon irradiation at 360 nm.

To create a photoaffinity probe, a molecule must have two mandatory functionalities, and an optional functionality can be added during the construction of the compound. One mandatory requirement is an affinity unit responsible for reversible binding to the target molecule, which can be based on a known inhibitor of the target protein. Another required component is a photoreactive group that forms a permanent covalent bond with the target molecule (protein) upon irradiation (Figure 17A and B). Additionally, a reporter tag can be incorporated to detect the covalent complex using imaging methods or to enrich labeled proteins through pull-down methods.[234,238–240] Reporter tags can be fluorescence dyes, radioisotopes, or biotin, which allows the photoaffinity labeled complex to be enriched. In addition to the direct introduction of the reporter tag (one-step mechanism, Figure 17A), a ligation handle, in the form of a functional group suitable for click reaction, can also be incorporated into the photoaffinity probe, allowing the reporter group to be attached following photoaffinity labeling of the target molecule (two-step mechanism, Figure 17A).[241] The ligation handle must be unreactive towards functional groups present in the sample, making it bioorthogonal. Azides and alkynes are the most commonly used ligation handles.[240] However, norbornenes (for inverse electron-demand diels-alder reaction)[242] or alkyl thiols, allowing click reaction with maleimides[243–245] (only for cysteine-free target proteins) are viable alternatives. To design an effective photoaffinity probe, it is crucial to consider the structure-activity relationship (SAR) and identify suitable modifiable sites in the molecule where the necessary functionalities can be introduced.[234]

Figure 16B illustrates the principle of a photoreactive group using trifluoromethylphenyldiazirine and its reaction with water upon irradiation at a wavelength of 360 nm. During irradiation, the nitrogen bond in phenyldiazirine is cleaved, generating a highly reactive carbene. This carbene reacts with the nearest reacting partner, such as water, but ideally, the protein of interest, forming a covalent bond and liberating nitrogen. As a side product, a relatively stable linear diazo isomer can be formed, which reacts further to produce the carbene upon continued irradiation.[234,235] The products of photoaffinity labeling can be assessed and characterized using various methods such as LC-MS/MS analysis, SDS-PAGE, or HPLC.[145,234,235,246,247] If no reporter group is present, detecting the covalent bond is usually achieved by observing a shift in the mass peak during LC-MS analysis.[239,240]

Detection Methods. Identifying photo adducts formed by a peptide-based photoprobe interacting with a receptor can yield information at three levels (Figure 17B). At the protein level, intact protein analysis can be achieved through LC-MS analysis or SDS-PAGE. These techniques provide information about the binding protein and the stoichiometry of ligand-target interactions. By incorporating a reporter group into the ligand, the protein's location in the gel can be compared to the protein band detected in a fluorescence scan after Coomassie staining. If the positions of the bands coincide, it indicates successful photoaffinity labeling. To identify the specific binding site of the ligand, protein digestion followed by sequencing can be employed. Typically, enzymatic digestion using trypsin is performed, breaking the protein into smaller peptide fragments. The binding site of the ligand can be detected by analyzing these peptide fragments at the oligopeptide level using LC-MS/MS. Finally, the photo adduct can be analyzed at the amino acid level by sequencing the labeled protein fragment and using MS/MS.[234,235,248]

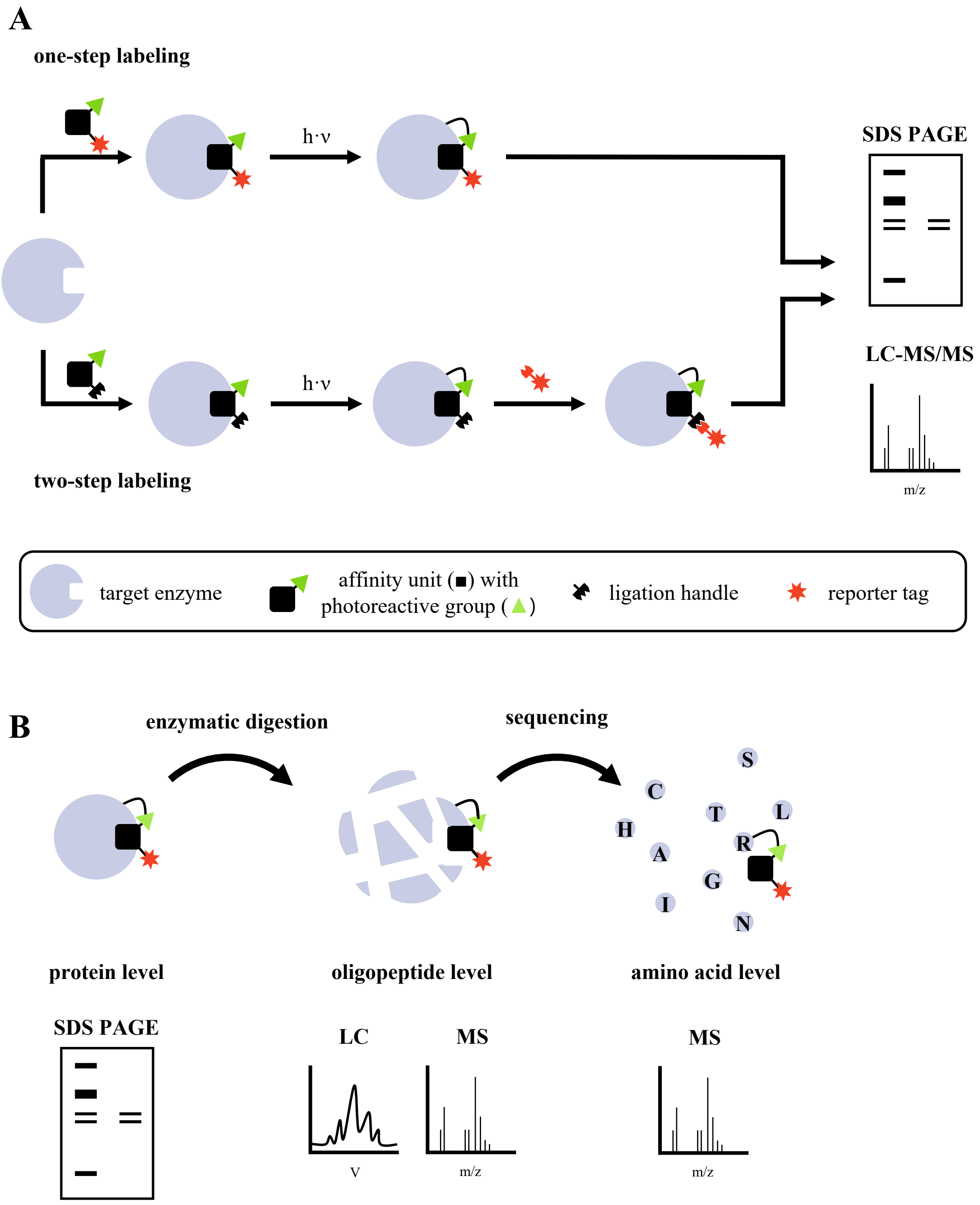


Figure 17. Principle of photoaffinity labeling and detection of photo adducts. **A** One-step and two-step photoaffinity labeling of a target protein. After incubation of the photoaffinity probe consisting of an affinity unit with a photoreactive group and a reporter tag (1-step labeling) or a ligation handle (2-step labeling), the covalent bond between the target enzyme and the photoreactive probe is formed upon irradiation (h·v). **B** Detection of photo adducts can be carried out at three levels: protein level *via* SDS-PAGE gel, oligopeptide level after enzymatic digestion *via* LC-MS/MS, and amino acid level after sequencing *via* MS/MS.

2 Projects and Objectives

The results presented in this doctoral thesis can be assigned to two projects centering around flaviviral NS2B-NS3 proteases.

2.1 Project 1: Inhibitors for Flaviviral NS2B-NS3 Proteases

The first project will focus on developing and evaluating potential inhibitors for the flaviviral NS2B-NS3 proteases. Both allosteric and competitive inhibitors can address the target proteins. Potential inhibitors will be designed, synthesized, and investigated as part of this project. The starting point for allosteric inhibitors are previously published benzo[*d*]thiazole compounds[108,135] (Figure 8). Systematic variations of the lead structures are expected to generate new potential allosteric inhibitors with high affinity and improved properties. In the second part of this project, a fragment-based drug design approach will develop competitive inhibitors targeting the ZIKV and DENV NS2B-NS3 protease. Structure affinity relationships (SAR) will be derived from both series of inhibitors.

The inhibition type will be investigated for both the allosteric and competitive inhibitors. For this purpose, besides the Dixon analysis (see Chapter 1.3.1), the examination of inhibitory activity before and after the blockage of the allosteric binding pocket can be utilized. In addition to a high affinity for the target protein, inhibitors should exhibit good selectivity against other proteases, which will be examined for the most promising compounds.

The basis of all these investigations is a fluorometric assay, a frequently employed tool in medicinal chemistry. Therefore, ensuring the assay's reliability with minimal interferences is paramount. Hence, in this project, the optimization of the fluorometric assay will also be pursued alongside the development of potential inhibitors.

The following publications are part of the results obtained in this project:

(1) **Maus, Hannah**‡, Fabian Barthels‡, Stefan J. Hammerschmidt, Katja Kopp, Benedikt Millies, Andrea Gellert, Alessia Ruggieri, Tanja Schirmeister. "SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases." *Bioorganic & Medicinal Chemistry* **2021**, 47 (2021): 116392, doi: 10.1016/j.bmc.2021.116392.

Own contributions: inhibitor synthesis (compounds 34g–l and 36i–n; compounds 34a–f, 36a–f and 36m as part of my master thesis), inhibitor resynthesis (compounds 23c and 25b), DENV and ZIKV NS2B-NS3 inhibition assays (compounds 34g–l and 36i–n; compounds 34a–f, 36a–f and 36m as part of my master thesis), maleimide blockage assay, inhibitor stability assays, protease selectivity assays, and writing of the original draft & editing of the manuscript.

Contribution from other authors: protein expression & purification, inhibitor synthesis (compounds 8a–f, 12a/b, 17, 20a/b, 23a–d, 25a–f, 27, 28), DENV and ZIKV NS2B-NS3 inhibition assays (compounds 8a–f, 12a/b, 17, 20a/b, 23a–d, 25a–f, 27, 28), molecular modeling, cell viability assay, ZIKV, and DENV replication assay, and writing parts of the original draft & editing of the manuscript.

(2) Gabriel Gomes Vilela‡, Wadja Feitosa dos Santos Silva‡, Vitoria de Melo Batista, Leandro Rocha Silva, **Hannah Maus**, Stefan J. Hammerschmidt, Clara Andrezza Crisóstomo Bezerra Costa, Orlando Francisco da Silva Moura, Johnnatan Duarte de Freitas, Grazielle Lobo Coelho, Júlia de Andrade Brandão, Leticia Anderson, Ênio José Bassi, João Xavier de Araújo-Júnior, Tanja Schirmeister, Edeildo Ferreira da Silva-Júnior. "Fragment-based design of α-cyanoacrylates and α-cyanoacrylamides targeting Dengue and Zika NS2B/NS3 proteases." *New Journal of Chemistry* **2022**, 46 (42): 20322–20346, doi: 10.1039/D2NJ01983C.

Own contributions: DENV and ZIKV inhibition assays, protease selectivity assays, and manuscript editing.

Contribution from other authors: molecular docking, inhibitor synthesis, protein expression & purification, molecular dynamics simulations, MM/PBSA calculations, cytotoxicity assays, evaluation of *in vitro* antiviral activity, quantification of viral load, and writing of the original draft & editing of the manuscript.

(3) **Hannah Maus**‡, Patrick Müller‡, Mergim Meta, Sabrina N. Hoba, Stefan J. Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister, Fabian Barthels. "Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates" *Chemistry – A European Journal* **2023**, *accepted*, doi: 10.1002/chem.202301855.

Own contributions: kinetic characterization of substrates, protease inhibition assays, investigations on assay interferences, writing of the first draft (kinetic characterization of NBD-based substrates, autocleavage of a DENV and ZIKV NS2B/NS3 protease substrate, harnessing NBD autocleavage for FRET substrate replacement, applications for protease inhibitor investigation, mitigation of typical protease assay interferences, conclusion, fluorometric assays) & editing of the manuscript.

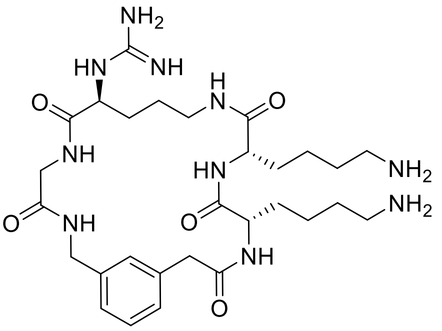
Contribution from other authors: substrate synthesis, inhibitor synthesis, protein expression & purification, substrate stability assays, molecular docking, and writing parts of the first draft (introduction, synthesis, smartphone-based fluorometer for usage in classroom applications, chemistry) & editing of the manuscript.

2.2 Project 2: NS2B-NS3 Protease-Ligand Interactions

This project will examine the interactions of allosteric and competitive inhibitors with flaviviral proteases. Aspects considered in this context include the impact of ligands on the conformation of ZIKV and DENV protease, the underlying binding mechanism, and the position of the allosteric inhibitors' binding.

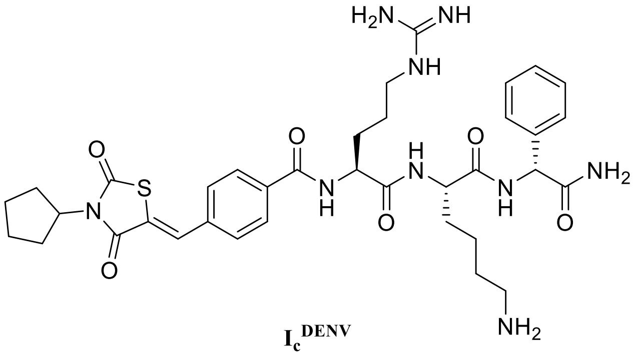
To investigate real-time protein conformational changes, smFRET is a particularly suitable method. In this project, the influence of different ligands, including an allosteric inhibitor (I_a) from project 1[249], as well as literature-known competitive inhibitors for the DENV (I_c^{DENV})[130] and the ZIKV (I_c^{ZIKV})[134] NS2B-NS3 protease, on the proteases' conformation will be investigated (Figure 18) using this method. Additionally, the binding mechanism of the ZIKV and DENV proteases with competitive inhibitors will be examined by determining the kinetic rates of conformational changes through smFRET.[187,250]

Furthermore, this project will investigate the allosteric inhibitor's interaction location. While the binding site of competitive inhibitors with their target protein is known, the localization of the binding pocket for allosteric inhibitors is only speculative. One approach for investigating the location of the interaction between the ligand and the target protein is PAL, which will be employed in this work to gain information about the site of the allosteric inhibitor's binding.



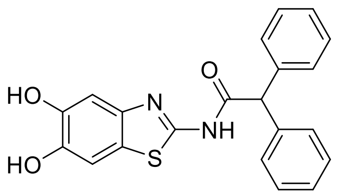
$\mathbf{I_c^{ZIKV}}$

K_i (Braun *et al.*) = 1.6 nM



$\mathbf{I_c^{DENV}}$

K_i (Behnam *et al.*) = 0.4 µM



$\mathbf{I_a}$

IC_{50} (ZIKV NS2B-NS3) = 0.95 µM

IC_{50} (DENV NS2B-NS3) = 11.12 µM

Figure 18. Model ligands for investigating conformational dynamics of the ZIKV and the DENV NS2B-NS3 protease. The competitive inhibitor for the ZIKV NS2B-NS3 protease (I_c^{ZIKV}) was published by Braun *et al.*[116] The competitive inhibitor for the DENV NS2B-NS3 protease (I_c^{DENV}) was published by Behnam *et al.*[130] My group published the allosteric inhibitor (I_a), part of project 1 of this doctoral thesis.[249]

The following publications and submitted manuscripts are part of the results obtained in this project:

(4) Christian Götz, Gerald Hinze, Andrea Gellert, **Hannah Maus**, Franziska von Hammerstein, Stefan J. Hammerschmidt, Luca M. Lauth, Ute A. Hellmich, Tanja Schirmeister, Thomas Basché. "Conformational dynamics of the dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies." *The Journal of Physical Chemistry B* **2021**, 125 (25): 6837–6846, doi: 10.1021/acs.jpcb.1c01797.

Own contributions: protein labeling of S79*C-S158C with ATTO 488 and Cy5 (together with Andrea Gellert), protein expression and purification (as part of my internship), fluorometric enzyme assay (activity of dye-labeled S79*C-S158C and control experiments not described in the publication), and manuscript editing.

Contributions from other authors: protein expression & purification, protein labeling, fluorometric enzyme assay, smFRET data curation, data analysis, and writing of the first draft & editing of the manuscript.

(5) **Hannah Maus**, Gerald Hinze, Stefan J. Hammerschmidt, Thomas Basché, Tanja Schirmeister. "A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease." *Protein Science* **2023**, 32 (1): e4526, doi: 10.1002/pro.4526.

Own contributions: protein labeling, fluorometric enzyme assays, smFRET data curation, and writing of the first draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, smFRET data analysis, kinetic modeling, and manuscript editing.

(6) **Hannah Maus**[‡], Stefan J. Hammerschmidt[‡] Gerald Hinze, Fabian Barthels, Victor H. Pérez Carrillo, Ute A. Hellmich, Thomas Basché, Tanja Schirmeister. "The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR." *European Journal of Medicinal Chemistry* **2023**, 258 (February): 115573, doi: 10.1016/j.ejmech.2023.115573.

Own contributions: protein labeling for smFRET experiments, smFRET data curation, DSF experiments, nanoDSF experiments, CD spectroscopy experiments, fluorometric enzyme assays, protein labeling for ^{19}F-NMR experiments (together with Stefan J. Hammerschmidt), ^{19}F-NMR data curation (together with Stefan J. Hammerschmidt and Victor H. Pérez), and writing of the original draft (introduction, results, discussion, conclusion, material and methods) & editing of the manuscript.

Contributions of others: protein expression & purification, site-directed mutagenesis, smFRET data analysis, protein labeling for ^{19}F-NMR experiments, ^{19}F-NMR data curation, ^{19}F-NMR data analysis, and writing parts of the original draft (results of ^{19}F-NMR, material and methods) & editing of the manuscript.

(7) **Hannah Maus**, Gerald Hinze, Stefan J. Hammerschmidt, Tanja Schirmeister, Thomas Basché. "Conformational Selection and Induced Fit: The Behavior of Two Homologous Proteases" *Angewandte Chemie* **2023**, *submitted.*

Own contributions: protein labeling, smFRET data curation, and writing of the original draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, smFRET data analysis, and manuscript editing.

(8) **Hannah Maus**, Andrea Gellert, Olivia Englert, Jia-Xuan Chen, Tanja Schirmeister, Fabian Barthels. "Designing Photo Affinity Tool Compounds for the Investigation of the DENV NS2B-NS3 Protease Allosteric Binding Pocket" *RSC Medicinal Chemistry* **2023**, *submitted.*

Own contributions: tool compound synthesis, fluorometric enzyme assay, PAL experiments, SDS-PAGE gel analysis, molecular docking experiments, and writing of the original draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, mass spectrometry experiments, and manuscript editing.

3 List of Publications and Manuscripts

3.1 Publications and Manuscripts as Part of this Doctoral Thesis

3.1.1 Project 1: Inhibitors for Flaviviral NS2B-NS3 Proteases

(1) **Maus, Hannah**‡, Fabian Barthels‡, Stefan J. Hammerschmidt, Katja Kopp, Benedikt Millies, Andrea Gellert, Alessia Ruggieri, Tanja Schirmeister. “SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases.” *Bioorganic & Medicinal Chemistry* **2021**, 47 (2021): 116392, doi: 10.1016/j.bmc.2021.116392.

(2) Gabriel Gomes Vilela‡, Wadja Feitosa dos Santos Silva‡, Vitoria de Melo Batista, Leandro Rocha Silva, **Hannah Maus**, Stefan J. Hammerschmidt, Clara Andrezza Crisóstomo Bezerra Costa, Orlando Francisco da Silva Moura, Johnnatan Duarte de Freitas, Grazielle Lobo Coelho, Júlia de Andrade Brandão, Leticia Anderson, Ênio José Bassi, João Xavier de Araújo-Júnior, Tanja Schirmeister, Edeildo Ferreira da Silva-Júnior. “Fragment-based design of α-cyanoacrylates and α-cyanoacrylamides targeting Dengue and Zika NS2B/NS3 proteases.” *New Journal of Chemistry* **2022**, 46 (42): 20322–20346, doi: 10.1039/D2NJ01983C.

(3) **Hannah Maus**‡, Patrick Müller‡, Mergim Meta, Sabrina N. Hoba, Stefan J. Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister, Fabian Barthels. “Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates” *Chemistry – A European Journal* **2023**, *accepted*, doi: 10.1002/chem.202301855.

3.1.2 Project 2: NS2B-NS3 Protease-Ligand Interactions

(4) Christian Götz, Gerald Hinze, Andrea Gellert, **Hannah Maus**, Franziska von Hammerstein, Stefan J. Hammerschmidt, Luca M. Lauth, Ute A. Hellmich, Tanja Schirmeister, Thomas Basché. “Conformational dynamics of the dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies.” *The Journal of Physical Chemistry B* **2021**, 125 (25): 6837–6846, doi: 10.1021/acs.jpcb.1c01797.

(5) **Hannah Maus**, Gerald Hinze, Stefan J. Hammerschmidt, Thomas Basché, Tanja Schirmeister. “A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease.” *Protein Science* **2023**, 32 (1): e4526, doi: 10.1002/pro.4526.

(6) **Hannah Maus**‡, Stefan J. Hammerschmidt‡ Gerald Hinze, Fabian Barthels, Victor H. Pérez Carrillo, Ute A. Hellmich, Thomas Basché, Tanja Schirmeister. “The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and

^{19}F-NMR." *European Journal of Medicinal Chemistry* **2023**, 258 (February): 115573, doi: 10.1016/j.ejmech.2023.115573.

(7) **Hannah Maus**, Gerald Hinze, Stefan J. Hammerschmidt, Tanja Schirmeister, Thomas Basché. "Conformational Selection and Induced Fit: The Behavior of Two Homologous Proteases" *Angewandte Chemie* **2023**, *submitted.*

(8) **Hannah Maus**, Andrea Gellert, Olivia Englert, Jia-Xuan Chen, Tanja Schirmeister, Fabian Barthels. "Designing Photo Affinity Tool Compounds for the Investigation of the DENV NS2B-NS3 Protease Allosteric Binding Pocket" *RSC Medicinal Chemistry* **2023**, *submitted.*

3.2 Publications Beyond this Doctoral Thesis

3.2.1 Research Articles

(9) Stefan J. Hammerschmidt‡, Hannah Maus‡, Annabelle C. Weldert‡, Michael Gütschow, Christian Kersten. "Improving binding entropy by higher ligand symmetry? – A case study with human matriptase" *RSC Medicinal Chemistry* **2023**, 5 (14): 969–982, doi: 10.1039/D3MD00125C.

(10) Sumit Kumar, Neha Sharma, Willyenne M. Dantas, Jessica C. Frutuoso do Nascimento, **Hannah Maus**, Ronaldo Nascimento de Oliveira, Unnat Pandit, Agam P. Singh, Tanja Schirmeister, Puja Panwar Hazari, Lindomar Pena, Brijesh Rathi. "A potent candidate against Zika virus infection: Synthesis, bioactivity, radiolabeling, and biodistribution studies." *New Journal of Chemistry* **2022**, 46 (39): 18764–18775, doi: 10.1039/D2NJ02482A.

(11) Felix Schalk‡, Janis Fricke‡, Soohyun Um, Benjamin H. Conlon, **Hannah Maus**, Nils Jäger, Thorsten Heinzel, Tanja Schirmeister, Michael Poulsen, Christine Beemelmanns. "GNPS-guided discovery of xylacremolide C and D, evaluation of their putative biosynthetic origin and bioactivity studies of xylacremolide A and B." *RSC advances* **2021**, 11 (31): 18748–18756, doi: 10.1039/D1RA00997D.

(12) Armin Welker, Christian Kersten, Chrsitin Müller, Ramakanth Madhugiri, Collin Zimmer, Patrick Müller, Robert A. Zimmermann, Stefan J. Hammerschmidt, Hannah Mus, John Ziebuhr, Christoph Sotriffer, Tanja Schirmeister. "Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2" *ChemMedChem* **2021**, 16 (2): 340–354, doi: 10.1002/cmdc.202000548.

Front Cover: Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2 (2/2021). *ChemMedChem* **2021** 16 (2): 301–301, doi: 10.1002/cmdc.202000987.

3.2.2 Review Articles

(13) Patrick Müller[‡], **Hannah Maus**[‡], Stefan J. Hammerschmidt[‡], Philip M. Knaff, Volker Mailänder, Tanja Schirmeister, Christian Kersten. "Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy" *Current Medicinal Chemistry* **2022**, 29 (4): 635–665, doi: 10.2174/0929867328666210526111318.

(14) Érica Erlanny da Silva Rodrigues, **Hannah Maus**, Stefan J. Hammerschmidt, Alessia Ruggieri, Elane Conceição dos Santos, Ênio José Bassi, Leticia Anderson, Pedro Gregório Vieira Aquino, João Xavier de Araújo-Júnior, Fenju Wei, Xinyong Liu, Peng Zhan, Tanja Schirmeister, Edeildo Ferreira da Silva-Júnior. "The Medicinal Chemistry of Zika Virus." *Human Viruses: Diseases, Treatments and Vaccines: The New Insights.* Cham: Springer International Publishing **2021**, 233–295, doi: 10.1007/978-3-030-71165-8_13.

4 Project 1: Inhibitors for Flaviviral NS2B-NS3 Proteases

4.1 SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B-NS3 Proteases.

4.1.1 Context, Project Summary, and Own Contributions

Developing substrate-analogous inhibitors is a reasonable approach to achieving protease inhibition.[251,252] However, the active site's shallow and charged nature hindered inhibitor development, as their high affinity binding properties did not translate well in cell-based assays due to poor membrane permeability.[86,253,254] Recent research efforts have made progress in overcoming these challenges, exemplarily by the design of allosteric inhibitors.[103,255,256] An allosteric binding pocket around Ala125 was initially identified using surface cysteine mutagenesis and cysteine reactive probes.[103,108] Docking studies of previously published benzo[*d*]thiazole inhibitors suggested that the addressed binding site is near Asn152. These inhibitors demonstrated non-competitive behavior and were confirmed through single cysteine mutagenesis studies to bind to the allosteric site effectively.[108,135]

We synthesized and tested new inhibitors based on these lead structures to explore SAR by systematically varying the amino acid linker, sulfonamide/amide connector, and heteroaromatic system (Figure 19). Based on docking and molecular dynamics studies of the *R*- and *S*- enantiomers predicted to bind with their tosyl moieties in opposing subsites, we designed Y-shaped inhibitors to combine both binding modes into one molecule. Additionally, we explored truncated inhibitors with only the 5,6-dihydroxybenzo[*d*]thiazole fragment attached to different aromatic residues. Furthermore, we evaluated a scaffold hopping strategy by replacing the benzo[*d*]thiazole core heterocycle with different (hetero)aromatic structures.

Our work optimized allosteric inhibitors for the DENV2 and ZIKV NS2B/NS3 proteases and discovered new lead structures with good inhibitory properties. Y-shaped and truncated inhibitors demonstrated promising results, showing inhibition in the low micromolar range for the DENV2 protease and increased ligand efficacy. Inversion of the amide bond increased the compounds' stability under assay conditions. We also explored the off-target selectivity of our inhibitors against various serine and cysteine proteases, which yielded excellent results. In a cell-based assay, selected compounds exhibited antiviral potential by attenuating DENV2 and ZIKV replication. Two particularly promising compounds, *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)-4-iodobenzamide and 5,6-dihydroxy-*N*-phenylbenzofuran-2-carboxamide, demonstrated good starting points for further NS2B-NS3 inhibitor development with ligand efficiencies above 0.4. The best Y-shaped inhibitor showed IC_{50} values of 0.95 µM for ZIKV and 11.12 µM for DENV protease, respectively (Figure 19).

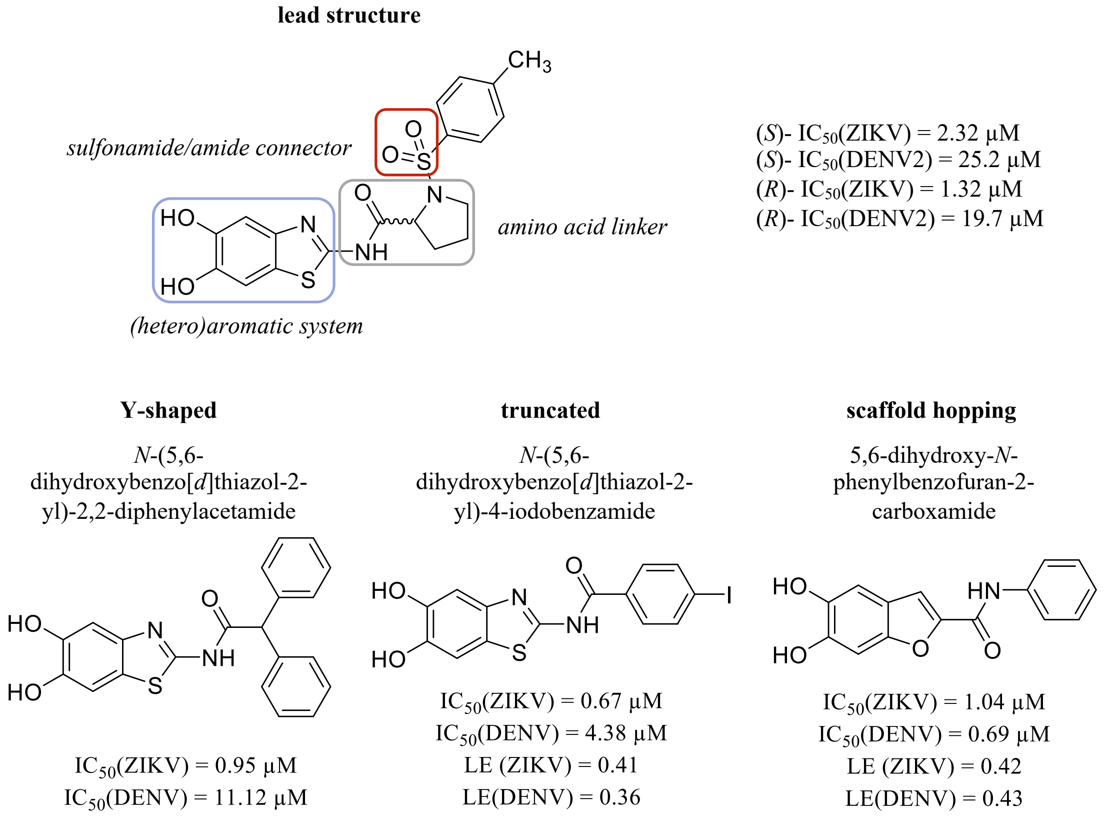


Figure 19. Most promising inhibitors of the article “SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B-NS3 Proteases.“[249] based on the lead structure published by Millies *et al.*[108]

Own contributions: inhibitor synthesis (compounds 34g–l and 36i–n; compounds 34a–f, 36a–f and 36m as part of my master thesis), inhibitor resynthesis (compounds 23c and 25b), DENV and ZIKV NS2B-NS3 inhibition assays (compounds 34g–l and 36i–n; compounds 34a–f, 36a–f and 36m as part of my master thesis), maleimide blockage assay, inhibitor stability assays, protease selectivity assays, and writing of the original draft & editing of the manuscript.

Contribution from other authors: protein expression & purification, inhibitor synthesis (compounds 8a–f, 12a/b, 17, 20a/b, 23a–d, 25a–f, 27, 28), DENV and ZIKV NS2B-NS3 inhibition assays (compounds 8a–f, 12a/b, 17, 20a/b, 23a–d, 25a–f, 27, 28), molecular modeling, cell viability assay, ZIKV, and DENV replication assay, and writing parts of the original draft & editing of the manuscript.

This work has been published in Bioorganic & Medicinal Chemistry (impact factor 3.46).

Article reprinted with permission from *Bioorganic & Medicinal Chemistry* **2021**, 47 (2021): 116392 “SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases.” © 2021 Elsevier B.V. (Netherlands).

The appended Supporting Information represents an abridged version. The full version can be accessed online at doi: 10.1016/j.bmc.2021.116392.

4.1.2 Publication

The following publication quoted within “” from page 45 to page 72 is the same as the manuscript cited on page 44.

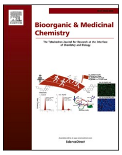
“

Bioorg. Med. Chem. 47 (2021) 116392

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc



SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases

Hannah Maus [a,1], Fabian Barthels [a,1], Stefan Josef Hammerschmidt [a], Katja Kopp [b], Benedikt Millies [a], Andrea Gellert [a], Alessia Ruggieri [b], Tanja Schirmeister [a,*]

[a] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany*
[b] *Department of Infectious Diseases, Molecular Virology, Center for Integrative Infectious Disease Research, University of Heidelberg, Im Neuenheimer Feld 344, 69120 Heidelberg, Germany*

ARTICLE INFO

Keywords:
SAR
Flaviviruses
Allosteric inhibition
NS2B/NS3 protease

ABSTRACT

In recent years, dengue virus (DENV) and Zika virus (ZIKV), both mosquito-borne members of the *Flaviviridae* family, have emerged as intercontinental health issues since their vectors have spread from their tropical origins to temperate climate zones due to climate change and increasing globalization. DENV and ZIKV are positive-sense, single-stranded RNA viruses, whose genomes consist of three structural (capsid, membrane precursor, envelope) and seven non-structural (NS) proteins, all of which are initially expressed as a single precursor polyprotein. For virus maturation, the polyprotein processing is accomplished by host proteases and the viral NS2B/NS3 protease complex, whose inhibitors have been shown to be effective antiviral agents with loss of viral pathogenicity. In this work, we elucidate new structure–activity relationships of benzo[*d*]thiazole-based allosteric NS2B/NS3 inhibitors. We developed a new series of Y–shaped inhibitors, which, with its larger hydrophobic contact surface, should bind to previously unaddressed regions of the allosteric NS2B/NS3 binding pocket. By scaffold-hopping, we varied the benzo[*d*]thiazole core and identified benzofuran as a new lead scaffold shifting the selectivity of initially ZIKV-targeting inhibitors to higher activities towards the DENV protease. In addition, we were able to increase the ligand efficiency from 0.27 to 0.41 by subsequent inhibitor truncation and identified *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)-4-iodobenzamide as a novel sub-micromolar NS2B/NS3 inhibitor. Utilizing cell-based assays, we could prove the antiviral activity *in cellulo*. Overall, we report new series of sub-micromolar allosteric DENV and ZIKV inhibitors with good efficacy profile in terms of cytotoxicity and protease inhibition selectivity.

1. Introduction

1.1. The global burden

The mosquito-borne dengue virus (DENV) and Zika virus (ZIKV) are major health concerns in the tropical and sub-tropical regions all over the world, putting about half of the world's population at risk. While DENV is endemic in many countries, leading to 100–400 million infections annually, ZIKV is more commonly known for epidemic outbreaks.[1–3] In 2015, Brazil experienced a ZIKV outbreak with about 200,000 reported cases solely in 2016.[4] Both viruses are transmitted by the mosquitoes *Aedes aegypti* and *Aedes albopictus*. In the last decades, these vectors began to spread considerably in temperate climate zones due to increasing globalization and climate change.[5] The appearance of *Aedes albopictus* on the northern side of the Alps recently raised the awareness of this health issue in middle-European countries.[6,7] The first autochthone DENV infections in Europe were reported in 2010 from south France and Croatia.[8,9] In 2012, the first endemic outbreak on the island of Madeira led to > 2,000 documented cases.[10] Now, several European countries report autochthone infections on an annual basis. 2019, the first autochthone ZIKV infections in Europe were reported from the southern French department of Var. Besides the vectorial transmission route, ZIKV was also found to be contracted via unprotected sexual intercourse.[11] This is of particular interest since its emergence in Southern America was correlated with a striking increase of neurological disorders and microcephaly in neonates.[12] Although ZIKV

---

[*] Corresponding author.
E-mail address: schirmei@uni-mainz.de (T. Schirmeister).
[1] Authors contributed equally.

https://doi.org/10.1016/j.bmc.2021.116392
Received 27 May 2021; Received in revised form 3 August 2021; Accepted 30 August 2021

infections tend to be asymptomatic or accompanied by self-limiting mild symptoms (fever, rash, conjunctivitis), some patients develop severe neurological conditions such as the Guillain-Barré syndrome.[13–15] Typical DENV infections show the very unpleasant symptoms of dengue fever. It is characterized by high fever and flu-like symptoms that are usually self-limiting. Some patients, however, develop severe dengue syndrome. Respiratory distress, fluid accumulation, plasma leaking, severe bleedings, and organ impairments can lead to a potentially fatal outcome. Although there is a vaccination against DENV available, the antibody-dependent enhancement of recurring DENV infections increases the risk for severe courses if a vaccinated patient gets infected by another DENV serotype. Hence, the approval of this vaccination is limited to people who have already recovered from dengue infection.[16–18] To date, there is no approved causative treatment for patients, suffering DENV or ZIKV infections, making it inevitable to pursue further research on small molecule drugs against these diseases.[19]

1.2. *Functions of NS2B/NS3 in viral replication*

Amongst others, the five serotypes of dengue virus (DENV1–5) and ZIKV belong to the family of *Flaviviridae*, genus flavivirus.[20,21] Their whole genome is a single-stranded (+)-sense RNA, which is translated into a single precursor polyprotein, consisting of the structural proteins capsid, membrane, and envelope, that are components of the virion, and the seven non-structural (NS) proteins involved in viral replication and maturation. The functional proteins are released after processing by the host proteases furin and signalase as well as the viral NS2B/NS3 protease.[22–26] On top of this essential role, the NS2B/NS3 protease was also found to modulate the human immune system by cleaving the stimulator of interferon genes (STING).[27] Based on these functions, the NS2B/NS3 protease is considered a promising target for developing specific anti-flaviviral therapeutics.[28–31]

1.3. *Structure of NS2B/NS3*

The NS3 protein consists of two domains. The *N*-terminal serine protease domain ($NS3_{pro}$) and the *C*-terminal helicase ($NS3_{hel}$). For its proteolytic activity, $NS3_{pro}$ requires NS2B as its cofactor which is composed of a transmembrane region, anchoring the protease to the endoplasmic reticulum, and its *C*-terminal cofactor region ($NS2B_{cf}$), essential for proper folding of the $NS3_{pro}$ catalytic domain.[32,33] Crystallography of the flaviviral $NS2B_{cf}$-$NS3_{pro}$ revealed the existence of at least two distinct conformations.[28,34,35] The catalytically active conformation is known as the *closed* conformation with $NS2B_{cf}$ wrapped around the protease domain forming a β-turn with its *C*-terminus contributing to the formation of the S2- and S3-binding pockets.[35,36] Besides that, an inactive *open* conformation was identified, in which $NS2B_{cf}$ is loosely bound to $NS3_{pro}$.[35] In solution, an equilibrium between these conformations could be observed, whereas an increased proportion of the closed conformation was achieved by the addition of substrate or active site-directed inhibitors via conformational selection or induced fit mechanisms.[28,33,86] An allosteric binding pocket around Ala125 was first identified with surface cysteine mutagenesis and cysteine reactive probes.[37] It was proposed, that binding to this site may rearrange the 120 s loop, thus locking the protein in its open conformation.[38]

1.4. *Obstacles in inhibitor development*

To achieve protease inhibition, developing substrate analogous inhibitors is a reasonable approach.[39,40] In the case of the flaviviral NS2B/NS3 protease, however, such inhibitors may feature poor pharmacological properties due to the substrate preferences for dibasic residues in *P*1 and *P*2 amino acids (AA). Inhibitor development is also hindered by the shallow and charged nature of the active site. Therefore, it is comprehensible that, although some substrate competitive inhibitors display high-affinity binding properties, their efficacy in cell-based assays stayed far behind expectations due to poor membrane permeability.[34,41,42] Some recent research efforts, however, did overcome these issues.[30,43,44]

In our group, we focused on inhibitor design by addressing the allosteric site. Previous docking studies of our inhibitors suggested the allosteric binding site to be located close to Asn152. Since Asn152 was found to be addressable from both the active and the allosteric site, this residue was referred to as a “molecular switch” between the open and closed conformation.[37,45] The allosteric binding mode of our inhibitors was perfectly demonstrated by their non-competitive behavior as well as by single cysteine mutagenesis studies.[37,46]

Based on these previously reported inhibitors including the lead structures **1a,b**, **2**, and **3a,b** (Figure 1), we synthesized and tested new series of inhibitors to explore further structure–activity relationships (SAR).[37] A second objective was to optimize the inhibitors in terms of their ligand efficacy (LE), generating an improved starting point for following drug development.

2. Results and discussion

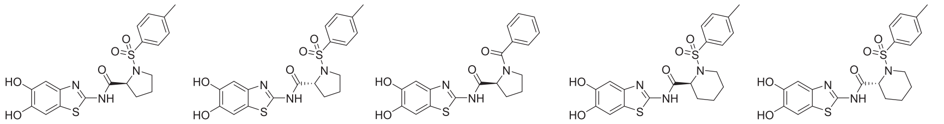
The individual moieties of the lead structures **1a,b**, **2**, and **3a,b**, namely AA linker, sulfonamide/amide connector, and heteroaromatic system, were systematically varied to investigate SAR.

Since the exchange of the (*R*/*S*)-proline moiety in **1a,b** to pipecolic acid in **3a,b** led to an improved inhibition of the NS2B/NS3 protease, further AA exchanges were explored in this work. Previously, the exchange of the sulfonamide moiety in **1a** to an amide linker in compound **2** resulted in modest changes of inhibitory potency so that both linking groups were used here depending on the synthetic context. A second approach was based on docking and molecular dynamics studies of the enantiomers **1a** and **1b** which are predicted to bind with their tosyl moieties in opposing subsites (Figure 2). Consequently, Y-shaped inhibitors were designed to combine both binding modes into one molecule. A third series of compounds was designed to increase the ligand efficacy by truncating previous full-scale inhibitors to a smaller scaffold consisting of only the 5,6-dihydroxybenzo[*d*]thiazole fragment, which was previously identified as the important pharmacophore, attached to different aromatic residues. Last, a scaffold hopping strategy was evaluated, for which the benzo[*d*]thiazole itself was exchanged to different (hetero)aromatic structures.

2.1. *Chemistry*

For the first SAR strategy, inhibitors based on different AA were synthesized according to a previously published route.[37] 2-Amino-5,6-dimethoxybenzo[*d*]thiazole **5** was coupled with different Boc-protected AA **4a–f** mediated by 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium tetrafluoroborate (TBTU), followed by BBr_3-deprotection of the Boc- and methoxy-groups to yield the corresponding hydrobromide salts **7a–f**. The sulfonamides **8a–f** were obtained by reaction of the deprotected amines with *p*-tosyl chloride (Scheme 1**A**).

For synthetic stability reasons, the tyrosine derivative **12a**, the aspartic acid derivative **12b** (Scheme 1**B**), and the β-alanine derivative **17** (Scheme 1**C**) were synthesized with an amide linker instead of a sulfonamide linking group. As shown in previous work, the exchange between sulfonamides and amides in these structures did not significantly influence the IC_{50} values of the inhibitors.[37] For compounds **12a, b**, the Boc- and benzyl-protected AA were reacted with **5** to form the precursor amides **9a,b**. The Boc group was removed with HCl in dioxane to give **10a,b**, followed by benzoylation to yield the dimethoxy derivates **11a,b**. Deprotection with BBr_3 gave the final compound **12a**. For the deprotection of **11b**, BBr_3-deprotection led to unwanted side reactions, hence, EtSH/$AlCl_3$ was used instead for deprotection of the methoxy and benzyl groups to obtain **12b**. The β-alanine derivative **17** was obtained



1a
IC_{50}(ZIKV) = 2.32 µM
IC_{50}(DENV) = 25.2 µM
K_D(ZIKV) = 0.24 µM

1b
IC_{50}(ZIKV) = 1.32 µM
IC_{50}(DENV) = 19.7 µM
K_D(ZIKV) = 0.49 µM

2
IC_{50} (ZIKV) = 1.55 µM
Inhibition (DENV) = 28% (20 µM)

3a
IC_{50}(ZIKV) = 0.51 µM
IC_{50}(DENV) = 13 µM

3b
IC_{50}(ZIKV) = 0.58 µM
IC_{50}(DENV) = 29 µM

Fig. 1. Structures of previously reported benzo[*d*]thiazole-based inhibitors **1a,b**, **2** and **3a,b**.[37]

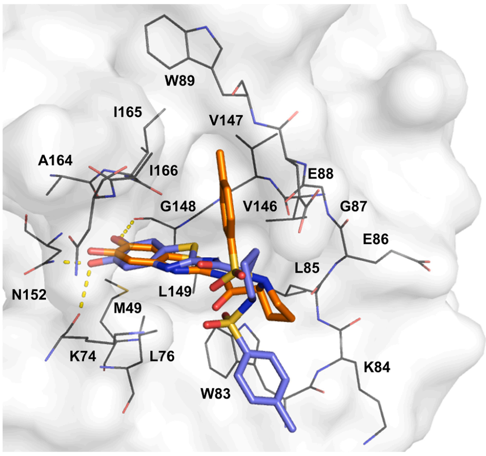


Fig. 2. Predicted binding modes of the previously investigated enantiomeric proline derivates **1a** (orange) and **1b** (purple) in the DENV2 allosteric binding site (pdb: 2FOM).[35,37] The catechol fragment shows conserved interactions with Lys74, Gly148, and Asn152 for both (*R*)- and (*S*)-enantiomers, while the aromatic *para*-tosyl moieties interact with distinct hydrophobic subsites. The design of inhibitors with an enlarged interaction surface targeting both subsites may lead to the generation of more potent inhibitors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

by TBTU-mediated coupling of **5** to Boc-protected *β*-alanine. After Boc-deprotection by HCl in dioxane and benzoylation, the dimethoxy derivative **16** was obtained. Subsequently, BBr_3-deprotection gave the final product **17** (Scheme 1**C**).

Initial Y-shaped inhibitors designed to bind to the two subsites addressed by **1a** and **1b** were synthesized based on the aspartic acid compound **12b**. Hence, asparagine-derivates **20a,b** were synthesized starting from **11b** by hydrolysis of the benzyl ester to yield the carboxylic acid **18**. Coupling to *N,N*-diethyl amine and aniline led to the dimethoxy precursors **19a,b**, followed by deprotection to yield the final compounds **20a,b** (Scheme 2**A**). Furthermore, Y–shaped inhibitors **23a–d** were obtained by the synthesis of different carboxylic acids **21a–d**.[47–49] Subsequent coupling with **5** and deprotection with BBr_3 yielded the final compounds **23a–d** (Scheme 2**B**).

To increase the ligand efficiency of parent compounds **1a,b**, truncated inhibitors, consisting of only a benzo[*d*]thiazole fragment and substituted benzoic acids, were synthesized. By amide coupling between the benzo[*d*]thiazoles **5** or **26** and the corresponding benzoic acids with TBTU or the respective benzoyl chlorides, the dimethoxy derivates **24a–f** were obtained (Scheme 3**A**). The 6-nitrobenzo[*d*]thiazole derivates **27** and **28** were synthesized accordingly (Scheme 3**B**). Methoxy derivatives **24a–f** were deprotected by BBr_3, to obtain the final compounds **25a–f** with free hydroxy groups (Scheme 3**A**).

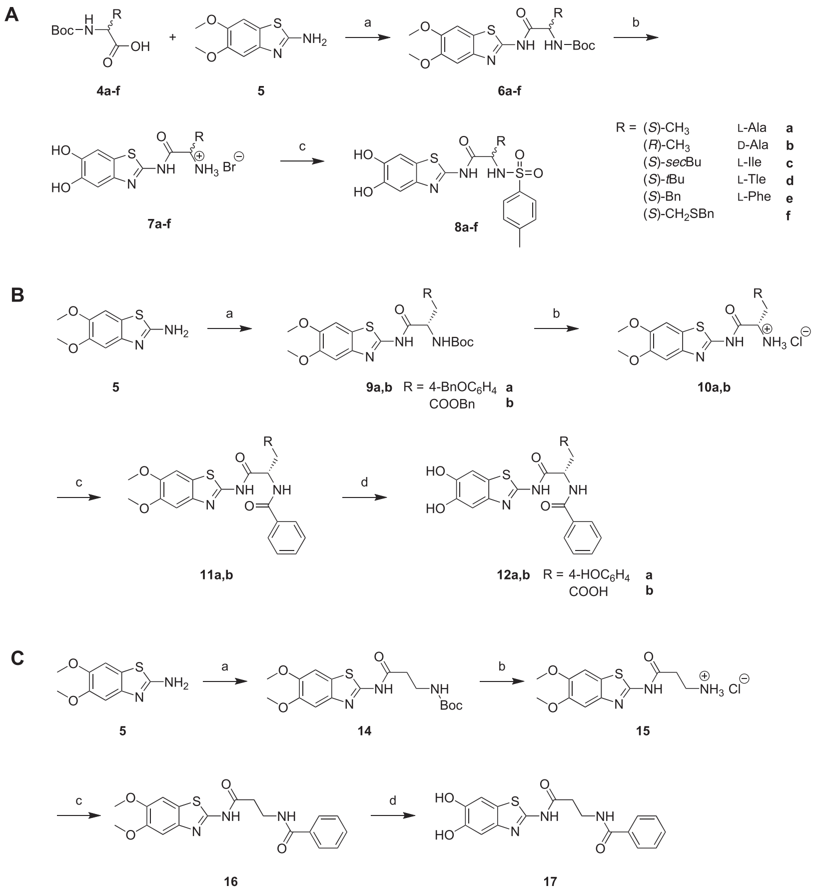
In order to vary and optimize the benzo[*d*]thiazole core fragment, a scaffold hopping strategy was evaluated for the synthesis of different heterocyclic systems. Since various heteroaromatics were synthetically more accessible as carboxylic acid derivatives than as amines, the amide bond was inverted compared to the previously described compounds. The corresponding (hetero)aromatic carboxylic acids **32a–m** were synthesized according to literature procedures.[50–59] In this way, a variety of different amides **33a–k** and **35a–n** could be synthesized by TBTU-mediated coupling of the (hetero)aromatic carboxylic acids with either aniline, *p*-iodoaniline (Scheme 4**B**), prolyl anilide **31a** or pipecolyl anilide **31b** (Scheme 4**C**). The latter were synthesized from Boc-protected proline or pipecolic acid **29a,b**, and aniline via condensation catalyzed by triethylamine and ethyl chloroformate (Scheme 4**A**).[60] Subsequent deprotection of the dimethoxy compounds **33a–k** and **35a–n** by BBr_3 yielded the catechols **34a–k** and **36a–n** (Scheme 4**B,C**). Inhibitor **34l** was obtained by reducing the nitro group of compound **34j** using tin chloride.

2.2. Structure–activity relationships

To evaluate the inhibition potency of all final compounds, these were evaluated by means of fluorometric enzyme assays with recombinantly expressed DENV and ZIKV NS2B/NS3 protease and Boc-Gly-Arg-Arg-AMC as substrate.[61,62] Starting from parent compounds **1a,b**, **2**, and **3a,b**, which previously showed sub-micromolar IC_{50} values on the ZIKV protease, inhibitors containing a variety of AA were synthesized and tested for inhibition against DENV2 and ZIKV NS2B/NS3 proteases (Table 1).

The most potent amino acid-based inhibitors were found to have IC_{50} values in the low micromolar range, however, none of the inhibitors showed improved inhibition compared to the previously published parent compounds. Inhibition selectivity towards the ZIKV protease was consistent for all tested inhibitors **8a–17**, as they caused a maximum of 63% inhibition at an inhibitor concentration of 20 µM for the homologous DENV2 protease. A significant SAR was observed for the alanine derivatives **8a,b**. The (*S*)-derivative **8a** showed inhibition in the low micromolar range (6.48 µM), while the (*R*)-enantiomer **8b** inhibited ZIKV protease only to 26% at 20 µM. The other compounds of this series, which were based on hydrophobic (*S*)-AA (isoleucine **8c**, *tert*-leucine **8d**, and phenylalanine **8e**), showed similar IC_{50} values ranging from 3.66 to 5.74 µM. Compound **8f** that was based on an *S*-benzyl-modified cysteine residue, displayed the lowest IC_{50} of this compound series (2.13 µM). Moderate polar decoration, like the introduction of a hydroxyl group, yielded the tyrosine-based inhibitor **12a** and did not significantly change inhibition potency compared to the phenylalanine-based derivate **8e** (4.59 µM).

However, this trend was reversed for more polar substituents. The aspartic acid-based inhibitor **12b** showed only 45% inhibition of the ZIKV protease at 20 µM. Therefore, we concluded that the allosteric binding pocket may tolerate hydrophobic residues of different sizes, but anionic substituents potentially abolish the inhibition due to non-beneficial interactions with the pocket-determining non-polar (Trp83, Leu85, V147) and acidic residues (Glu86, Glu88). Lengthening of the AA



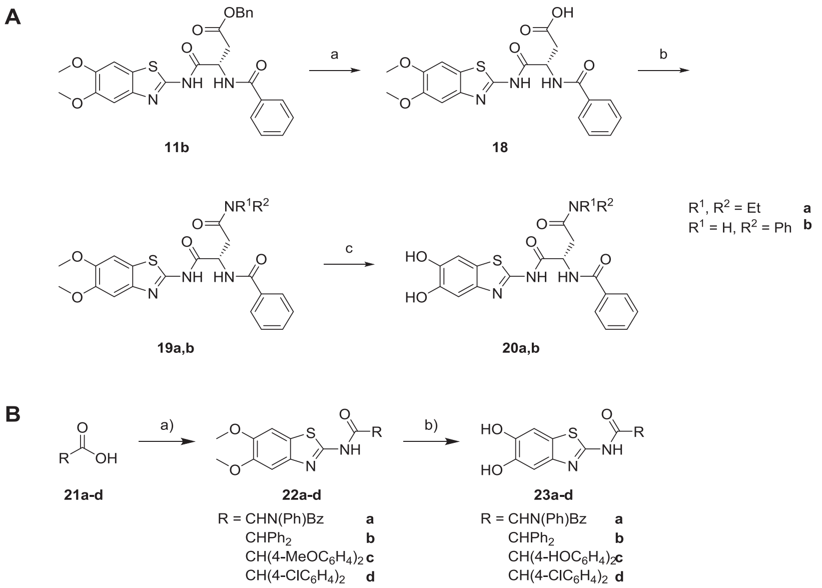
Scheme 1. Synthesis of the AA-based compounds **8a–f**, **12a,b**, and **17**. (**A**) Synthesis of sulfonamides **8a–f**. Reagents and conditions: a) TBTU, HOBt, DIPEA, DMF, 2 d, 34–82%. b) BBr_3, DCM, –78 °C to r.t., 16 h, 66–99%. c) *p*-TsCl, pyridine, 16 h, 32–70%. (**B**) Synthesis of the tyrosine- (**12a**) and aspartate-based inhibitor **12b**. Reagents and conditions: a) TBTU, HOBt, DIPEA, DMF, 2 d, 95–97%. b) HCl, dioxane, r.t., 30 min, 87–95%. c) PhCOCl, NEt_3, DCM, 16 h, 71–76%. d) BBr_3, DCM, –78 °C to r.t., 16 h, 77% or $AlCl_3$, EtSH, DCM, 0 °C to r.t., 16 h, 87%. (**C**) Synthesis of the *β*-alanine derivative **17**. Reagents and conditions: a) Boc-*β*-Ala-OH, TBTU, HOBt, DIPEA, DMF, 3 d, 78%. b) HCl, dioxane, r.t., 30 min, 70%. c) PhCOCl, NEt_3, DCM, 16 h, 40%. d) BBr_3, DCM, –78 °C to r.t., 16 h, 87%.

linker by one methylene unit from (*S*)-alanine **8a** to *β*-alanine **17** resulted in a strong decrease of inhibition, demonstrating the importance of the distance between the benzo[*d*]thiazole unit and the aromatic group for binding into the allosteric binding pocket.

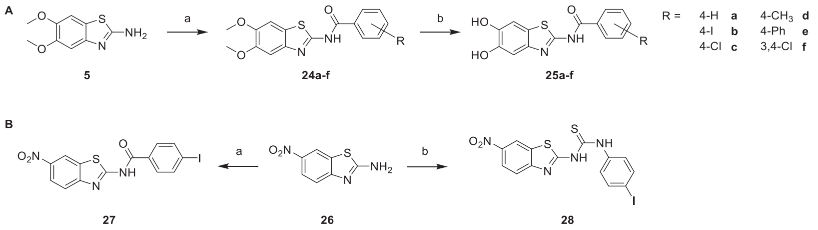
To increase the affinity of the benzo[*d*]thiazole inhibitors, a second strategy was applied to optimize the hydrophobic interaction surface of the inhibitor while maintaining the core benzo[*d*]thiazole pharmacophore. Previous docking and molecular dynamics studies of **1a** and **1b** proposed that *p*-tosyl moieties bind to opposing subsites corresponding to the orientation of the proline stereoisomers (Figure 2). Thus, Y-shaped inhibitors were designed to address both subsites and to increase the overall inhibition potency.[37] Due to the observation that mainly hydrophobic AA lead to potent inhibition, we decided to functionalize the aspartic acid-based inhibitor **12b** at the carboxyl group by hydrophobic substituents. In addition to the resulting asparagine-based inhibitors (**20a,b**), diphenyl-derived Y-shaped compounds (**23a–d**) were also designed and evaluated for their inhibition (Table 2).

The asparagine-based inhibitor **20a**, which includes an *N,N*-diethyl amide substitution, showed very low inhibition at 20 μM (13%). On the other hand, the corresponding aniline substituted derivate **20b** showed a significantly improved IC_{50} value of 5.48 μM for the ZIKV protease and 9.95 μM for the DENV2 protease. From these results, the hydrophobic yet structurally circumscribed binding region of the allosteric binding pocket becomes apparent. Increasing the inhibition strength seems to depend on the hydrophobic contact area as well as on the relative orientation of the Y-shape arms, and thus, should be an optimizable parameter.

Aromatic structures appear to be favored in this process, hence, four additional structures **23a–d** with diphenyl-derived Y-shapes were investigated. Inhibitor **23a**, incorporating a rigid *N*-phenyl peptoid-structure, led to a slight decrease of affinity (2.07 μM) compared to the proline amide derivative **2** (1.55 μM). The 2,2-diphenylacetic acid derivate **23b** showed the lowest IC_{50} value of this series (0.95 μM), albeit with no increased inhibition for the DENV2 protease (11.12 μM).



Scheme 2. Synthesis of Y-shaped inhibitors. (**A**) Synthesis of asparagine-based Y-shaped inhibitors **20a,b**. Reagents and conditions: a) LiOH, THF, H_2O, 1 h, 98%. b) TBTU, HOBt, DIPEA, DMF, 2 d, 35–65%. c) BBr_3, DCM, –78 °C to r.t., 16 h, 44–85%. (**B**) Synthesis of Y-shaped inhibitors **23a–d**. Reagents and conditions: a) **5**, TBTU, HOBt, DIPEA, DMF, 2 d, 29–68%. b) BBr_3, DCM, –78 °C to r.t., 16 h, 50–85%.

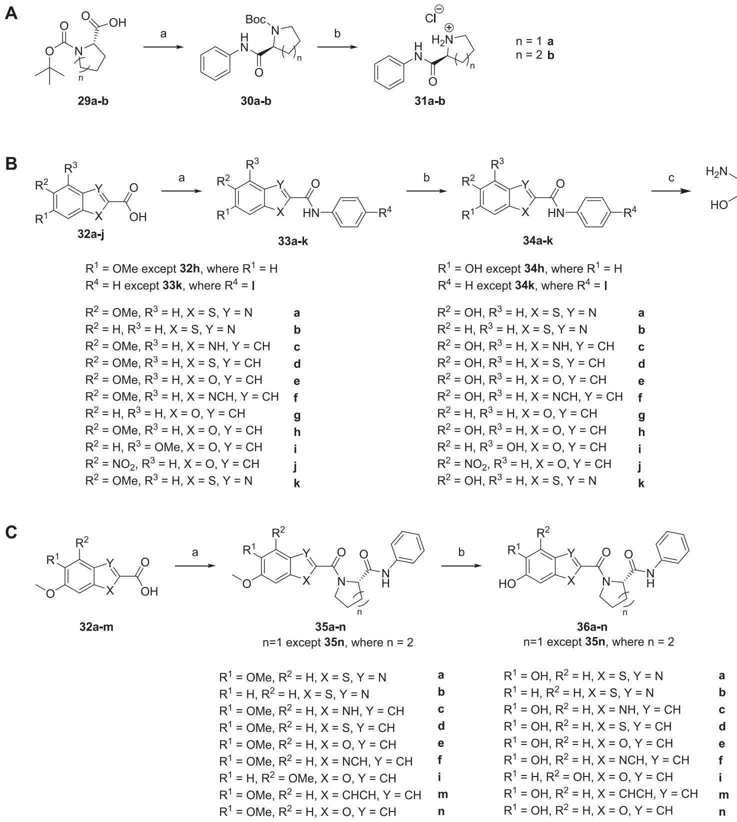


Scheme 3. Synthesis of truncated inhibitors. (**A**) Synthesis of catechol derivatives **25a–f**. Reagents and conditions: a) R-PhCOOH, TBTU, HOBt, DIPEA, DMF, 2 d, 24–99% or PhCOCl, NEt_3, DCM, 16 h, 72%. b) BBr_3, DCM, –78 °C to r.t., 16 h, 29–99%. (**B**) Synthesis of the 6-nitrobenzo[*d*]thiazole derivative **27** and **28**. Reagents and conditions: a) 4-iodobenzoic acid, TBTU, DMF, 2 d, 74%. b) 4-iodophenyl isothiocyanate, DBU, DCM, r.t., 7 d, 68%.

However, the improvement in inhibitory strength is not yet optimal related to the increase in molecular weight, as the ligand efficiency between parent compound **1a** and Y-shape inhibitors changed only marginally from 0.27 to 0.21–0.31 (**23a–d**). To improve the inhibition potency, the hypothesis was pursued to rationally address the polar residues Glu86/Glu88 of the allosteric binding pocket. For this purpose, similar to tyrosine derivative **12a**, two hydroxyl groups were introduced as *para*-decoration substituents. However, the resulting tetrahydroxy derivate **23c** led to a significant decrease in inhibitory potency (54%). Calculation of the protonation state revealed that one of the two additional hydroxyl groups forms a negatively charged phenolate under assay conditions, possibly leading to repulsion of Glu86/Glu88 and low inhibition as for aspartic acid derivative **12b**. In contrast, the dichloro derivative **23d** did not show increased activity compared with the undecorated analog **23b**, so we suggest that 2,2-diphenylacetic acid derivatives cannot be further improved by substitution. With the derivatives **20b** and **23a,b** we were able to show that the Y-shape strategy is successful and is a promising idea in future structure-based inhibitor designs.

To increase the ligand efficiency at the same or smaller molecular size, the synthesis and evaluation of truncated inhibitors was performed as a third SAR strategy. In this series, truncated inhibitors consisting only of a benzo[*d*]thiazole scaffold and a substituted benzoic acid were evaluated for their inhibition potency (Table 3). Root compound **25a** *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl) benzamide showed an IC_{50} of 9.19 μM for the ZIKV protease and 26.95 μM for the DENV protease. Thus, the starting point of this SAR series was an inhibitor about 2–3 times weaker than the parent compounds **1a,b**. Subsequentially, the *para*-substituent of the benzamide was varied to modulate the structure.

Exemplary, the introduction of a halogen atom in the *para*-position, was accompanied by an improved IC_{50} for DENV2 protease as demonstrated for iodine **25b** (4.38 μM) or chlorine substituents **25c** (5.97 μM). Other derivates bearing a *para*-methyl group (**25d**), a *para*-phenyl substituent (**25e**), or two chlorine atoms in *meta* and *para* positions (**25f**), inhibited the DENV2 protease, with IC_{50} value of 4.86–7.80 μM, in the similar range. For ZIKV, however, the iodine derivate **25b** displayed a



Scheme 4. Synthesis of inhibitors with different (hetero)aromatic moieties. (**A**) Reagents and conditions: a) aniline, ethyl chloroformate, triethylamine, THF, 0 °C to 70 °C, 20 h, 89–90%. b) HCl in dioxane, r.t., 15 h, quantitative. (**B**) Reagents and conditions: a) aniline, TBTU, HOBt, DIPEA, EtOAc, 0 °C to r.t., 3 d, 66%–quantitative. b) BBr_3, DCM, –78 °C to r.t., 16 h, 20%–quantitative. c) $SnCl_2$, HCl, r.t. to 70 °C, 16 h, 56%. (**C**) Reagents and conditions: a) **31a** or **31b**, TBTU, HOBt, DIPEA, EtOAc, 0 °C to r.t., 3 d, 55%–quantitative. b) BBr_3, DCM, –78 °C to r.t., 16 h, 13–96%.

Table 1
Inhibition of the DENV2 and ZIKV NS2B/NS3 proteases by amino acid-based inhibitors.[a]

Compound	Included AA	ZIKV	DENV2
		IC_{50} [μM]/%	IC_{50} [μM]/%
8a	(*S*)-Ala	6.48 ± 1.41	39%
8b	(*R*)-Ala	26%	14%
8c	(*S*)-Ile	5.35 ± 0.33	33%
8d	(*S*)-*tert*-Leu	5.74 ± 0.79	44%
8e	(*S*)-Phe	3.66 ± 0.49	33%
8f	(*S*)-Cys(SBn)	2.13 ± 0.11	57%
12a	(*S*)-Tyr	4.59 ± 0.50	63%
12b	(*S*)-Asp	45%	32%
17	*β*-Ala	38%	23%

[a] IC_{50} values are indicated by mean ± standard deviation from at least three independent measurements. Percentage inhibition was determined at 20 μM inhibitor concentration.

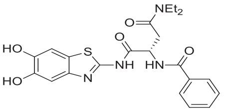
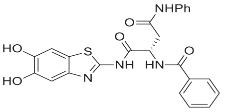
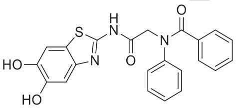
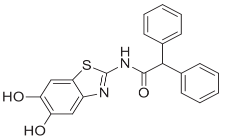
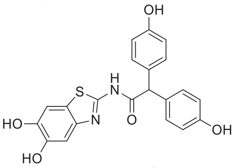
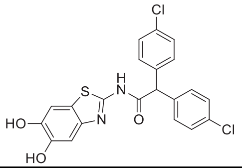
significantly higher potency (0.67 μM) than the unsubstituted derivate **25a** (9.19 μM) which makes **25b** the best inhibitor of this series. Exemplarily, the substrate conversion plot in the presence of inhibitor **25b** is showing sub-micromolar inhibition (Figure 3).

For the other substituted benzamides, no clear structural trend was recognized for ZIKV protease inhibition. Wherein the *para*-chlorine (**25c**, 6.66 μM) and *para*-phenyl (**25e**, 1.22 μM) derivatives showed improved inhibition, the inhibition was decreased for the *para*-methyl (**25d**, 15.40 μM) and the 3,4-dichloro derivative (**25f**, 39%) compared to the root compound **25a**. To follow up the best inhibitor **25b**, two additional *para*-iodine substituted inhibitors were synthesized. Derivate **27** harbors a nitro-benzothiazole fragment instead of the catechol-benzo[*d*]thiazole motif. **27** inhibits both proteases at lower potencies (49% for ZIKV and 36% for DENV2). Compared to compound **27**, compound **28** incorporates a thiourea linker instead of an amide linker. The replacement of the linker is accompanied by a significant increase in inhibition (IC_{50} = 6.51 μM for ZIKV and 11.21 μM for DENV2).

In summary, by shortening the inhibitor scaffold, the ligand efficacy of the compounds was increased from 0.27 resp. 0.28 for the parent compounds **1a,b** to 0.41 for the iodine-substituted inhibitor **25b**. In contrast, the ligand efficacy for the best Y-shape inhibitor **23b** could be increased to only 0.31, illustrating that optimizing the root scaffold is more effective than increasing the hydrophobic contact area of the inhibitor.

As the fourth SAR optimization strategy, the core benzo[*d*]thiazole

Table 2
Inhibition of the DENV2 and ZIKV NS2B/NS3 proteases by inhibitors with a Y-shaped structure.[a]

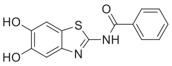
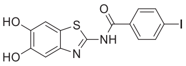
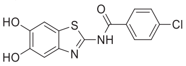
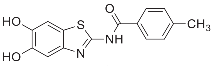
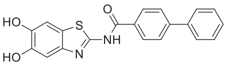
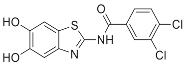
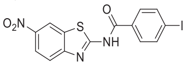
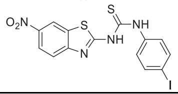
Compound	Structure	ZIKV	DENV2
		IC_{50} [μM]/%	IC_{50} [μM]/%
20a		13%	16%
20b		5.48 ± 0.35	9.95 ± 0.34
23a		2.07 ± 0.16	56%
23b		0.95 ± 0.13	11.12 ± 0.49
23c		54%	45%
23d		0.96 ± 0.12	58%

[a] IC_{50} values are indicated by mean ± standard deviation from at least three independent measurements. Percentage inhibition was determined at 20 μM inhibitor concentration.

itself was varied by scaffold hopping alterations, leading to a variety of inhibitors with (hetero)cyclic structures that were evaluated for their inhibition towards the NS2B/NS3 proteases (Table 4). Since various heteroaromatics were synthetically more accessible as carboxylic acid derivatives than as amines, the amide bond was inverted compared to the previously described compounds. Therefore, an amide-inverted prototype scaffold **36a**, analogously to parent compound **2**, was first synthesized and tested. This root compound **36a**, based on a benzo[*d*]thiazole, showed a similar IC_{50} value (4.00 μM) on the ZIKV protease compared to the inhibitor with inverted amide bonds **2** (1.55 μM, Figure 1). Furthermore, the inhibition of DENV2 protease could even be significantly increased by the inversion of the amide bond (**36a** IC_{50} = 11.59 μM, **2** 28% at 20 μM).

Subsequently, two series of heteroaromatic scaffold hops were evaluated: 1) those (**36a–n**) containing a proline linker and derived from parent compound **2** and 2) those (**34a–l**) derived from the truncated benzamide derivatives **25a** and lacking a proline linker. An overview of all derivatives can be found in Table 4 and Scheme 4.

Table 3
Inhibition of the DENV2 and ZIKV NS2B/NS3 proteases by differently substituted truncated inhibitors.[a]

Compound	Structure	ZIKV	DENV2
		IC_{50} [μM]/%	IC_{50} [μM]/%
25a		9.19 ± 0.33	26.95 ± 1.61
25b		0.67 ± 0.32	4.38 ± 0.38
25c		6.66 ± 0.41	5.97 ± 0.50
25d		15.40 ± 4.70	6.39 ± 0.20
25e		1.22 ± 0.13	7.80 ± 0.89
25f		39%	4.86 ± 0.33
27		49%	36%
28		6.51 ± 0.44	11.21 ± 0.87

[a] IC_{50} values are indicated by mean ± standard deviation from at least three independent measurements. Percentage inhibition was determined at 20 μM inhibitor concentration.

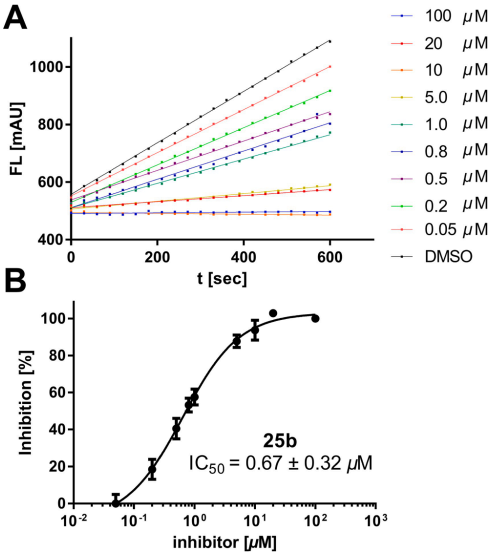
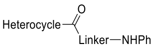
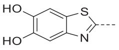
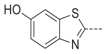
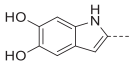
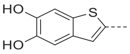
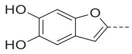
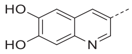
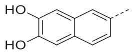
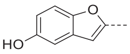
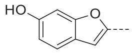
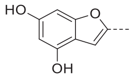
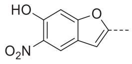
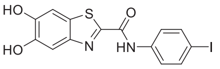
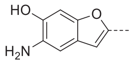


Fig. 3. (**A**) Fluorometric assay with top inhibitor **25b** showing enzyme inhibition from linear substrate conversion plots. (**B**) Determination of IC_{50} values [μM] for compound **25b** by regression to the four-parameter Hill equation.

Table 4

Inhibition of the DENV2 and ZIKV NS2B/NS3 proteases by inhibitors with different (hetero)aromatic scaffolds.[a]

				
Compound	(Hetero)cycle	Linker	ZIKV	DENV2
			IC_{50} [μM]/%	IC_{50} [μM]/ %
36a		(*S*)-proline	4.00 ± 0.34	11.59 ± 0.90
34a		none	9.19 ± 0.33	26.95 ± 1.60
36b		(*S*)-proline	19% at 20 μM	47% at 20 μM
34b		none	25% at 20 μM	20% at 20 μM
36c		(*S*)-proline	44% at 20 μM	36% at 20 μM
34c		none	3.44 ± 1.21	6.90 ± 2.30
36d		(*S*)-proline	3.43 ± 0.14	4.01 ± 0.47
34d		none	1.59 ± 0.08	2.73 ± 0.90
36e		(*S*)-proline	2.43 ± 0.28	6.43 ± 0.76
34e		none	1.04 ± 0.26	0.69 ± 0.16
36n		(*S*)-pipecolic acid	2.90 ± 0.04	4.73 ± 0.28
36f		(*S*)-proline	44% at 20 μM	23% at 20 μM
34f		none	6.38 ± 0.50	3.63 ± 0.21
36m		(*S*)-proline	42% at 20 μM	9% at 20 μM
34g		none	n. i.	n. i.
34h		none	n. i.	n. i.
36i		(*S*)-proline	20.13 ± 3.48	27.69 ± 2.82
34i		none	62% at 100 μM	51% at 100 μM
34j		none	48% at 200 μM	19% at 200 μM
34k			1.30 ± 0.08	2.48 ± 0.16
34l		none	52% at 200 μM	34% at 200 μM

[a] IC_{50} values are indicated by mean ± standard deviation from at least three independent measurements. Percentage inhibition was determined at 20 μM, 100 μM, 200 μM inhibitor concentration. n. i.: no inhibition at an inhibitor concentration of 200 μM

Truncated benzo[*d*]thiazole derivate **34a** (without proline linker) exhibited comparable activities compared to the inverted-amide derivate **25a**, thus, it was proven for both **36a** and **34a** that the orientation of the amide bonds does not influence the NS2B/NS3 inhibition. In the series of heterocycles containing only one heteroatom **34c–f** and **36c–f**, the benzofuran derivate **34e** showed with IC_{50} values of 0.69 μM for DENV2 and 1.04 μM for ZIKV the best inhibition profile of the whole series. Remarkably, the benzofuran derivative **34e** showed an abrogation of the previously observed inhibition selectivity, which was manifested by the fact that benzo[*d*]thiazole inhibitors inhibited ZIKV protease by a factor of 2–10 better than the DENV2 protease.

Besides this observation, the benzo[*d*]thiazole and benzofuran derivates **34b**, **36b**, **34g**, and **34h** contained only one single 5′- or 6′-hydroxyl group, and consequently, these derivates displayed reduced inhibition compared to their corresponding catechol analogs **34a**, **36a**, and **34e**. This confirms the previously shown trend that both hydroxyl groups are necessary for effective inhibition potency.[46] For benzofurans **34i** and **36i**, the hydroxyl groups were arranged in *meta*-positions instead of the conserved *ortho*-configuration (**34e** and **36e**) resulting in loss of inhibition potency with IC_{50} values for **36i** of 20.13 μM for ZIKV and 27.69 μM for DENV2. For the other modifications, the exchange of a hydroxyl group for nitro or amine substituents also resulted in a partial loss of inhibitory activity as can be seen for compounds **34j** and **34l**. Quinoline **36f** as a core scaffold led to significantly decreased inhibition, as well as the naphthalene derivative **36m**. Based on these results, it can be hypothesized that one heteroatom is necessary for the core scaffold and that scaffolds containing five-membered rings are superior to scaffolds with six-membered rings such as quinolines or naphthalenes. Similar to the transformation from **1** to **3**, the replacement of (*S*)-proline by pipecolic acid resulted in a slight improvement of the DENV2 IC_{50} values of proline derivate **36e** (6.43 μM) compared to pipecolyl derivate **36n** (4.73 μM). In analogy to the improvement of inhibition of the DENV2 protease upon the inversion of the amide bond (**25a** (26.95 μM) vs. **34a** (11.59 μM)), amide inversion of the iodinated top inhibitor **25b** (4.38 μM) also resulted in an improvement of DENV2 inhibition (**34k**, 2.48 μM). In summary, besides the original benzo[*d*]thiazole scaffold, benzo[*b*]thiophenes (**34d** and **36d**) and benzofurans (**34e** and **36e**) were identified as promising scaffolds for the synthesis of allosteric DENV2 and ZIKV inhibitors.

To validate our hypothesis of allosteric inhibition, IC_{50} values were determined for the top inhibitor **25b** at various substrate concentrations (50–200 μM). Evaluation by Dixon plot analysis revealed that the inhibition is not substrate competitive, supporting our hypothesis of an allosteric binding (**SI** Figure 5**A**). The apparent K_i was determined to be 7.67 μM and thus lies in the same range as the determined IC_{50} for **25b**. In a previous work, we showed by single-site cysteine mutagenesis and subsequent maleimide blockage that the lead structure **1b** binds into an allosteric binding pocket.[37] Here, we have also applied this assay to the top inhibitors of this work (**23b**, **25b**, and **34e**). Thus, the IC_{50} values of compounds **23b**, **25b**, and **34e** were compared for the DENV2 wild-type protease, the T122C mutant protease, and the T122C mutant protease after incubation with *N*-benzylmaleimide (BMI, 250 nM). For all three inhibitors, significantly higher IC_{50} values were measured for the BMI pre-incubated protease in comparison to the native T122C protease (**23b**: 5.14 μM/50.80 μM, **25b**: 1.60 μM/25.12 μM, **34e**: 1.22 μM/2.22 μM, **SI** Figure 5**B**) suggesting that the inhibitors in this work also bind near the allosteric Thr122 residue and that this binding interferes with the sterically demanding BMI label.

2.3. *Off-target selectivity*

To investigate how the benzo[*d*]thiazole and benzofuran inhibitors affect the activity of other serine and cysteine proteases, a selectivity panel was constructed from a matrix of seven NS2B/NS3 inhibitors and seven different proteases. A representative set of compounds for each SAR optimization series was tested for their selectivity against various in-house serine (trypsin, thrombin, α-chymotrypsin, and urokinase plasminogen activator (uPA)) and cysteine proteases (rhodesain, *Staphylococcus aureus* sortase A, SARS-CoV-2 M^{pro}). The serine protease NS2B/NS3 is structurally most related to trypsin, but the latter does not have an NS2B cofactor and thus does not feature an allosteric binding pocket. Neither trypsin nor any of the other proteases showed appreciable inhibition at an inhibitor concentration of 20 μM. The results of the inhibition data are summarized in Table 5.

Table 5
Protease inhibition selectivity of representative compounds towards various serine and cysteine proteases.[a]

Compound	Residual Enzyme Activity in [%] at 20 μM						
	Trypsin	Thrombin	Chymotrypsin	uPA	SARS-CoV-2 M^{pro}	Rhodesain	Sortase A
1b	96 ± 4	97 ± 3	100 ± 7	100 ± 0	94 ± 1	93 ± 6	83 ± 4
23b	85 ± 6	95 ± 2	96 ± 2	62 ± 4	77 ± 6	86 ± 0	73 ± 0
25a	80 ± 6	75 ± 0.3	77 ± 1	100 ± 15	90 ± 1	91 ± 6	57 ± 1
25b	94 ± 5	65 ± 3	79 ± 7	100 ± 10	89 ± 4	77 ± 4	74 ± 1
34e	88 ± 2	81 ± 2	78 ± 7	82 ± 4	99 ± 2	93 ± 2	81 ± 11
36e	96 ± 1	97 ± 3	58 ± 2	91 ± 5	52 ± 6	93 ± 4	64 ± 0
36n	95 ± 3	93 ± 1	77 ± 4	94 ± 8	57 ± 10	90 ± 3	74 ± 2

[a] Residual enzyme activities at 20 μM of corresponding compound are indicated by mean ± standard deviation from at least three independent measurements. The residual enzyme activity is expressed relative to the DMSO control in [%].

2.4. *Inhibitor stability*

Due to the known susceptibility of phenolic compounds to oxidation, selected compounds (**23b**, **25a**, **25b**, **34a**, **34e**, **36e**) were tested for their chemical stability in the assay buffer (Tris pH 9.0) and under cell culture conditions (HEPES pH 7.4). For this purpose, the conditions of the respective assay were simulated, and the stability was determined by HPLC/MS after 15 min, 90 min, and 20 h of incubation time. The stability was interpreted as the integral of the corresponding compound in the 210 nm UV trace as a proportion of the total integral.

Except for compound **25a** (recovery of the compound after 15 min, 90 min, 20 h: 67%, 61%, 0%), all substances that were active in the fluorometric enzyme assay showed sufficient stability in the assay buffer for the duration of the assay (Tris pH 9.0, 83–100%). Under cell-assay conditions (HEPES pH 7.4), the stability of the compounds was even significantly increased due to the lower pH value (95–100%). The inversion of the amide bond increased the stability of the oxidation-labile dihydroxy group (compare **25a** with **34a**: 95%, 94%, 88% after 15 min, 90 min, 20 h in Tris buffer), which means that the new benzofuran derivatives are not only more potent than the previous derivatives, but also have higher stability.[63]

2.5. *Docking studies*

Until now, no crystal structure of the flaviviral NS2B/NS3 protease in complex with a benzo[*d*]thiazole inhibitor was solved. However, previously a specific allosteric binding site for non-competitive benzo[*d*]thiazole inhibitors of the NS2B/NS3 protease had been proposed by biochemical and computational investigations.[37,38,46,64,65] This prominent cavity is located on the back side of the protein as viewed from the active site (Ser135, Asp175, and His51) and is formed by the AAs Met49, Lys74, Leu76, Trp83, Lys84, Leu85, Glu86, Gly87, Glu88, Trp89, Val146, Vak147, Gly148, Leu149, Asn152, Ala164, Ile165, and Ile166.

The allosteric binding site was shown to be existing for both the open and closed conformations of the DENV2 and ZIKV NS2B/NS3.[32,38] However, the most stable predicted binding modes of benzo[*d*]thiazole ligands were previously found for the open conformation of the DENV2 protease (PDB code: 2FOM).[35,37] Hence, molecular docking studies of a comprehensive set of inhibitors were performed to the 2FOM-protein receptor by FlexX docking within the LeadIT worksuite.[66]

The key interactions between the NS2B/NS3 protein and the respective ligands are exemplarily shown for a representative top inhibitor from each investigated SAR series **8e**, **23b**, **25b**, and **34e** (Figure 4). The docking of the non-covalently bound ligands resulted in conformations that aligned well with the docking poses obtained in our previous study.[37] When docked into the allosteric site the benzo[*d*]thiazole moiety was inserted into the lipophilic sub-pocket generated by the residues Leu76, Gly148, and Ile165. Here, the catechol motif of all four ligands forms a hydrogen-bonding network to the residues Asn152, Lys74, and Gly148.

In their docking pose, the two Y-shaped inhibitors **8e** and **23b** show beneficial interactions with the two spatially separated hydrophobic subsites (Figure 4 **A,B**). One arm of each inhibitor interacts hydrophobically with the subsite determining residues Trp89/Ala166 resp. Val146/147 analogously to the overlay of the two proline-derived parent compounds (Figure 2).

However, compared to the linear truncated inhibitor **25b**, the catechol motif in the Y-shape inhibitors is slightly twisted because the sterically demanding Y-shape substituent exerts a geometric constraint on the benzothiazole moiety (Figure 4**C**). This could be the reason why there was only a slight increase in inhibitory potency with the Y-shaped structures despite the increased hydrophobic contact area: The selected Y-shape geometry does not yet optimally fit the binding pocket. Of note, the connector moiety (AA resp. sulfonamide) showed only few interactions with the protein (Lys74, Trp84, and Asn167) indicating that the type of connector is of minor importance if the hydrophobic moieties can find one of the hydrophobic subsites.

The truncated inhibitor **25b** incorporates an aromatic iodine substituent oriented toward the acidic residues Glu86/Glu88. A favorable carboxylate-halogen interaction could determine the reason for the high inhibitory potency of this compound despite its small size.[67] The inhibitors **8e**, **23b**, and **25b** have higher inhibitory potency for the ZIKV

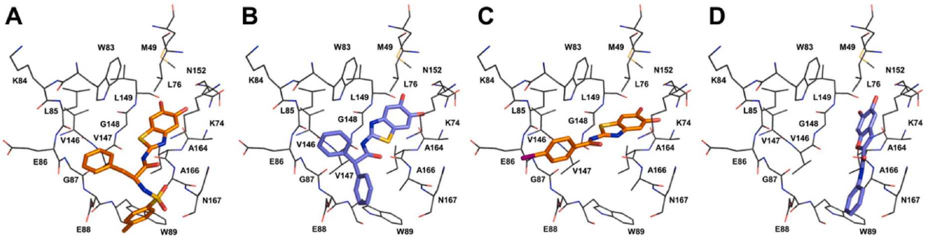


Fig. 4. Docking poses of each top inhibitor from the respective SAR series in complex with the DENV2 NS2B/NS3 protease (PDB: 2FOM) highlighting the proposed interaction features upon allosteric binding. (**A**) Docking pose for phenylalanine-derived inhibitor **8e**. (**B**) Docking pose for Y-shaped inhibitor **23b**. (**C**) Docking pose for truncated inhibitor **25b**. (**D**) Docking pose for benzofuran-derived inhibitor **34e**.

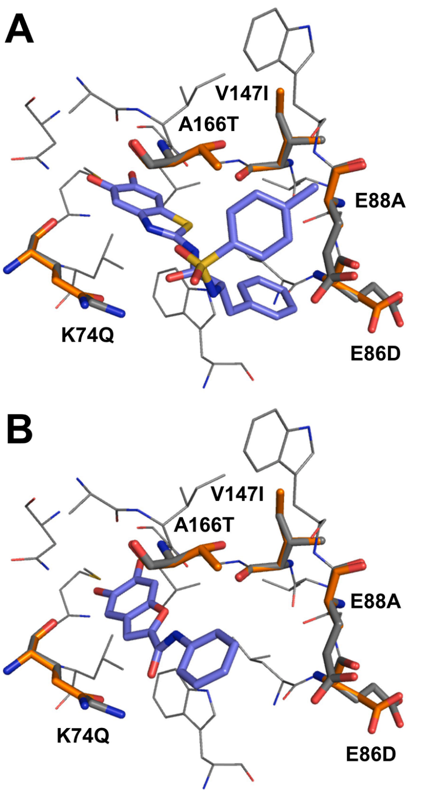


Fig. 5. Residue differences between DENV2 (grey) and ZIKV (orange) NS2B/NS3 within the allosteric binding site explaining the protease inhibitor selectivity by five differences in AA composition (K74Q, G87A, E88A, V147I, and A166T). (**A**) Docking pose for phenylalanine-derived inhibitor **8e**. (**B**) Docking pose for benzofuran-derived inhibitor **34e**. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

protease than for the DENV2 homolog, while this trend is reversed for the benzofuran derivative **34e**, which preferentially inhibits the DENV2 NS2B/NS3. The differences in inhibitory potencies may be explained by the binding mode of the inhibitors. The benzofuran motif has a slightly tilted binding mode compared to the benzothiazole fragment, which allows better interaction of the amide carbonyl atom with Lys74 in the DENV2 protease (Figure 4**D**). This residue is replaced in the ZIKV protease by a K74Q exchange, by which the positive charge at the pocket entry is lost and might explain the inverse selectivity of the benzofuran inhibitors for the DENV2 protease (Figure 5**B**).

In a sequence-based alignment between DENV2 (pdb: 2FOM) and ZIKV (pdb: 5GXJ) proteases, five core residue differences of the allosteric binding pocket were identified (K74Q, G87A, E88A, V147I, and A166T), which might explain the protease selectivity. The superior inhibitory potency of large and hydrophobic Y-shaped structures for the ZIKV protease might be explained by an $\mathrm{Glu88_{DENV2}}$ to $\mathrm{Ala88_{ZIKV}}$ exchange, which reduces the polarity of the respective DENV2 subsite and further enhances the space for hydrophobic substituents in the ZIKV protease as shown for the phenylalanine derivate **8e** (Figure 5**A**).

2.6. *Antiviral activity*

To determine the antiviral activity of NS2B/NS3 inhibitors in a cellular context, we studied the suppression of viral replication by a luciferase replication assay. Briefly, Huh7 cells were transfected by electroporation of DENV2 (strain New Guinea C, NGC) or ZIKV (strain H/PF/2013) subgenomic luciferase reporter replicon RNAs, expressing the luciferase gene and all NS genes to monitor viral replication.[68,69] After 4 h, 10 μM of the respective inhibitor was added. As a quantity for the determination of viral replication, luciferase activity was measured at 24, 48, 72, and 96 h after electroporation. A depiction of the DENV2/ZIKV RNA replication time curves for active compounds can be found in **SI** Figure 1.

Prior to viral replication analyses, the effect of inhibitors on mammalian cell integrity, potential cytotoxicity was determined by luminescent cell viability assay at varying compound concentrations. While most of the compounds had no measurable effects on Huh7 cell viability at concentrations between 1.25 and 10 μM, only the dimethoxy derivate **24b** showed cytotoxic effect by a reduction of substrate turnover at concentrations ≥ 5 μM (**SI** Figure 2). Therefore, the measured decrease in RNA replication of **24b** (Figure 6**A**) is probably not solely due to the antiviral properties of this compound but is caused by its cytotoxicity, which is why this compound was not used for further investigations.

The effect on DENV2 replication was tested for the top inhibitors of

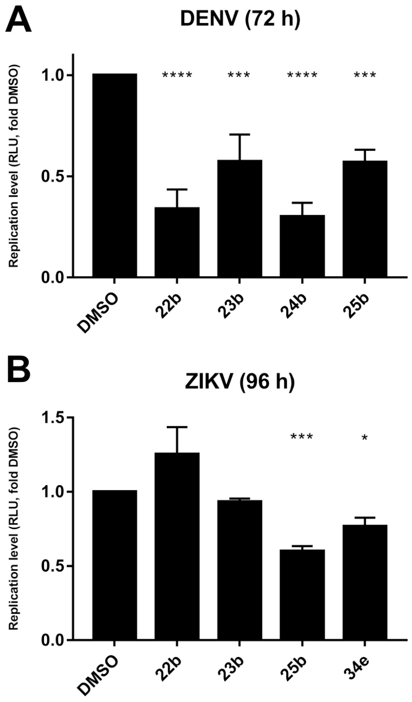


Fig. 6. ZIKV and DENV2 RNA replication quantified by luciferase activity assays. (**A**) Inhibition of DENV2 NGC replication by the most potent compounds (10 μM) 72 h after electroporation. (**B**) Inhibition of ZIKV HPF replication by the most potent compounds (10 μM) 96 h after electroporation. Shown are means ± standard deviation (n = 3). ****, $P < 0.0001$; ***, $P < 0.001$; *, $P < 0.05$. Statistical significance was calculated by using one-way ANOVA.

each SAR optimization series (Tables 1–4): Truncated *para*-iodine derivate **25b**, Y-shaped 2,2-diphenylacetic acid inhibitor **23b**, truncated benzofuran inhibitor **34e**, proline-containing benzofuran inhibitor **36e**, and pipecolic acid-containing benzofuran inhibitor **36n**. In addition, the corresponding dimethoxy derivates **24b**, **22b**, **33e**, **35e**, and **35n** were also tested for their effect on DENV2 replication since it was shown in our previous work that the dimethoxy compounds can act as prodrugs. After enzymatic metabolization, the corresponding dihydroxy derivatives are released.[37] At a concentration of 10 μM, both the *para*-iodine derivate **25b** and the Y-shaped 2,2-diphenylacetic acid inhibitor **23b** showed a significant decrease in RNA copies of DENV2 replication after 72 h (Figure 6A). Furthermore, the attenuation of viral replication was even more pronounced for the corresponding dimethoxy derivates **24b** and **22b**, which agrees with our previous study. The increased activity can probably be attributed to a higher membrane permeability of these prodrugs. However, the other benzofuran-consisting inhibitors tested for DENV2 inhibition showed no significant reduction of viral replication. The reason for this could not be conclusively elucidated here, as these had even shown a selectivity shift towards better inhibition of DENV2 protease in the fluorometric enzyme assay (Table 4).

To further investigate the DENV2 insusceptibility to the benzofuran-based inhibitors, sequence analysis was performed to identify possible amino acid differences between the recombinantly expressed NS2B/NS3 protease and DENV2 NGC protease (**SI** Figure 3). The protein construct used in the fluorometric assay in fact showed two point mutations near the allosteric pocket (T120A, N167Q). To study the activity and inhibition of the native cellular protease sequence in the context of the fluorometric assay, a double mutant of our NS2B/NS3 protease was recombinantly expressed by site-directed mutagenesis and tested for both protease activity compared to the wildtype and inhibition by selected compounds. The activity of the DENV2 protease T120A/N167Q-mutant appears to be slightly decreased compared with the wild type (**SI** Figure 4, wildtype: $k_{cat} = 0.709 \pm 0.039$ mAU s^{-1}; T120A/N167Q: $k_{cat} = 0.447 \pm 0.003$ mAU s^{-1}). Inhibition by benzofuran inhibitor **34e** and *para*-iodine derivative **25b** was evaluated for the T120A/N167Q-mutant by determination of the IC_{50} values (**34e**: 0.73 ± 0.015 μM, **25b**: 8.01 ± 1.25 μM). Since neither the activities nor the inhibition rates did severely differ, it can be assumed that the two mutations have only minimal influence on the inhibition of the compounds, and insusceptibility to benzofurans is most likely due to cellular causes (metabolism, permeability, etc.).

Besides DENV2 replication inhibition, ZIKV H/PF/2013 reporter virus replication was analyzed by a luciferase assay for the benzofuran inhibitor **34e**, the *para*-iodine derivate **25b**, the 2,2-diphenylacetic acid inhibitor **23b**, and its corresponding dimethoxy derivate **22b**, which showed efficient reduction of DENV2 replication. Luciferase activity was significantly reduced by 10 μM of benzofuran inhibitor **34e** and *para*-iodine derivate **25b** after 96 h. Compounds **22b** and **23b**, on the other hand, showed no significant effect on ZIKV replication. That our inhibitors have a generally lower efficiency to inhibit ZIKV replication compared to DENV2 inhibition could be since ZIKV replication factories are additionally surrounded by intermediate filaments in cage-like structures. This could lead to stronger protection of the ZIKV protease from inhibitors compared to DENV2.[70]

3. Conclusions

In this work, we have optimized allosteric inhibitors for the DENV2 and ZIKV NS2B/NS3 proteases. Hereby, we found new lead structures with good inhibitory properties. By exchanging the individual moieties of the lead structures **1a,b**, different series of inhibitors were synthesized and investigated regarding their inhibitory effects on the NS2B/NS3 proteases of DENV2 and ZIKV. Replacing the amino acid linker with (*R/S*)-alanine (**8a,b**), isoleucine (**8c**), *tert*-leucine (**8d**), phenylalanine (**8e**), tyrosine (**12a**), and aspartic acid (**12b**) did not significantly improve the inhibitory effect. Nevertheless, the ZIKV protease was inhibited by the compounds **8a–f** and **12a** in the low micromolar range ($IC_{50} = 2.13–6.48$ μM). In the series of newly designed Y-shaped inhibitors (Table 2), 2,2-diphenylacetic acid derivate **23b** was the most promising compound with a sub-micromolar IC_{50} for ZIKV (0.95 μM). Y-shaped compounds **23b** and **20b** as well as AA-based compounds with hydrophobic side chains **8c–f** demonstrated with IC_{50} values in the low micromolar range that addressing both binding subsites with Y-shaped inhibitors is a promising strategy for further improvement. By truncating the inhibitor scaffold (Table 3), we were able for the first time to achieve inhibition in the low micromolar range for the DENV2 protease and simultaneously increased the ligand efficacy to obtain new starting points for further DENV2 drug discovery. The iodine-substituted inhibitor **25b** was the most promising compound of this series with an IC_{50} of 4.38 μM for DENV2 and a sub-micromolar IC_{50} of 0.67 μM for ZIKV, respectively. By exchanging the benzo[*d*]thiazole core heterocycle, we could show that other heteroaromatic systems can serve as scaffolds besides benzo[*d*]thiazoles. In this SAR series, the benzofuran derivate **34e** showed the best inhibitory properties (IC_{50}(DENV2) = 0.69 μM, IC_{50}(ZIKV) = 1.04 μM). In contrast to the benzo[*d*]thiazole-based compounds, **34e** inhibits both the ZIKV and the DENV2 NS2B/NS3 proteases in the same order of magnitude. Screening of selected compounds against various serine and cysteine proteases demonstrated excellent off-target selectivity of our inhibitors. The most potent inhibitors and their respective methoxy prodrugs were administered in a cell-based assay for their antiviral potential to interfere with DENV2 and ZIKV replication. The compounds **22b**, **23b**, and **25b** highlighted their antiviral potential by significantly attenuating DENV2 replication, whereas **25b** and **34e** showed reduced ZIKV replication. In summary, we identified two promising compounds that resemble suitable starting points for upcoming drug development. First, *N*-(5,6-dihydroxybenzo[*d*]thiazol-2-yl)-4-iodobenzamide (**25b**) and second, 5,6-dihydroxy-*N*-phenylbenzofuran-2-carboxamide (**34e**), both of them containing only 20 and 21 heavy atoms, respectively, and with ligand efficiencies above 0.4 are good starting points for further NS2B/NS3 inhibitor development.

4. Experimental section

4.1. Expression constructs

A pET15b expression vector, harboring a thrombin-cleavable *N*-terminal hexahistidine-tagged DENV2 NS2B cofactor domain covalently linked to the NS3 protease domain *via* a $Gly_4SerGly_4$-linker (GenBank ID: AY037116.1) with two point mutations in the NS3 region (I30A and L31A for better solubility, herein referred to as WT), was kindly provided by the group of W. Diederich (University of Marburg, Germany).[71] The pET11a vector, containing the sequence of the French Polynesia ZIKV strain (GenBank ID: KJ776791.1) protease was used as described previously.[37,61,72] To abolish autocatalytic cleavage of the covalently linked construct, a point mutation (R95A) was introduced. The construct contains a *N*-terminal hexahistidine tag provided with a tobacco etch virus (TEV) protease cleavage site. The DENV2 T120A-N167Q mutant for validation of antiviral activity measurements and the DENV2 T122C mutant suitable for maleimide coupling based binding site verification were generated using the Kapa HiFi PCR kit (Kapa Biosystems) and the following primers:

(i) T120A: 5′–GACCAAT**GCG**GGTACCATTGGCG–3′ and 5′–GGTACC**CGC**ATTGGTCTTAAACAGGC–3′,
(ii) N167Q: 5′–CGATTGCC**CAG**ACGGAAAAGTCCATTGAAGATAACC–3′ and 5′–CTTTTCCGT**CTG**GGCAATCGCGGACACGTAG–3′,
(iii) T122C: 5′-CAATACGGGT**TGC**ATTGGCGCGGTTAGCCTGG–3′ and 5′–CGCCAAT**GCA**ACCCGTATTGGTCTTAAACAGGCCC–3′.

4.2. *Protein preparation*

The DENV2 and ZIKV proteases were expressed in *E. coli* BL21-Gold (DE3) (Agilent Technologies) cells according to the literature.[61] Briefly, cells were grown in LB media with the presence of 100 µg/mL ampicillin at 37 °C. Overexpression was induced by the addition of 1 mM IPTG at an OD_{600} nm of ~0.8 and incubated at 20 °C for 16 h. Cells were harvested by centrifugation, 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, 2 mM imidazole, 0.1% (v/v) $Triton_{X\text{-}100}$, RNase, lysozyme) and lysed by sonication (Sonopuls, Bandelin). Cell debris was removed by centrifugation and the supernatant was subjected to affinity chromatography using a HisTrap HP column (GE Healthcare, Chicago, USA). Histidine-rich proteins and other impurities were removed by subsequent increasing amounts of elution buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 250 mM imidazole). Eluted fractions were further purified by size exclusion chromatography (*SEC*) using a HiLoad 16/600 Superdex 75 pg (GE Healthcare, Chicago, USA) and eluted in *SEC* buffer (20 mM Tris-HCl pH 8, 150 mM NaCl). Prior to storage at –80 °C, proteins were concentrated, and flash-frozen with liquid nitrogen.

4.3. *Molecular Modeling*

Docking calculations were performed applying the FlexX algorithm as implemented in LeadIT.[73] The non-covalent mode was started from the program GUI with advanced settings. A standardized ligand preparation protocol was executed with MOE2020 prior docking. Ligand minimization was performed using the MMFF94x force field.[74] 3D coordinates of minimized structures were exported from MOE as single libraries in SDF file format to be imported to LeadIT. Modeling the putative binding mode of inhibitors to the allosteric pocket of DENV2 NS2B/NS3 protease was done using the crystal structure in the open conformation thereof (pdb entry: 2FOM). The binding site was defined as 11 Å shell around the core residue K73 including 24 receptor defining residues in the NS2B-protein chain: K73, K74, L76, W83, L85, E86, G87, E88, W89, T118, T120, I123, V146, V147, G148, L149, Y150, G151, N152, V154, A164, I165, A166, and N167. Chemical receptor properties were kept in FlexX default settings: No alternate protonation states from ligand preparation; Ligand binding driven by entropy and enthalpy (hybrid approach); Full score contribution threshold = 0.3; No score contribution threshold = 0.7; Consider *R/S* stereo mode during docking; Maximum allowed overlap volume = 2.9 Å^3; Clash factor = 0.6; Consider hydrogens in internal clash tests; Maximum number of solutions per iteration = 500; Maximum number of solutions per fragmentation = 500, Number of poses to keep = Top 10. A torsion subgraph was defined to force the thiazole-attached amide bond to be restricted in a planar constitution with the aromatic system as determined by QM-calculations.[37]

4.4. *Fluorometric assays*

The determination of the inhibitory activity of the compounds against the proteases was performed with an assay based on the fluorogenic substrates or FRET-based substrates. The inhibitors and the substrate were prepared as stock solutions in DMSO. The fluorescence was measured in white flat-bottom 96-well microtiter plates from Greiner Bio-One using a Tecan Infinite F2000 PRO plate reader. Measurements were performed in at least three independent experiments. In each well a total volume of 200 µL was used, consisting of 180 µL buffer, 5 µL enzyme solution, 10 µL inhibitor in DMSO or pure DMSO as control, and 5 µL solution of the corresponding substrate. Initial screenings were performed at inhibitor concentrations of 20 µM. IC_{50} values were determined with dilution series between 0.01 µM and 100 µM. The fluorescence was measured every 30 s for 10 min at 25 °C with the corresponding excitation and emission wavelengths. IC_{50} values were calculated with GraFit (Version 6.0.12; Erithacus Software Limited, East Grinstead, West Sussex, UK)[75] by fitting the remaining enzymatic activity to the four-parameter IC_{50} equation

$$Y = \frac{Y_{max} - Y_{min}}{1 + \left(\frac{[I]}{IC_{50}}\right)^{s}} + Y_{min}$$

with Y [ΔF/min] as the substrate hydrolysis rate, Y_{max} as the maximum value of the dose–response curve, measured at inhibitor concentrations of [I] = 0 µM, Y_{min} as the minimum value, obtained at high inhibitor concentrations, and s as the Hill coefficient.[76]

Buffers and Substrates. Rhodesain (50 mM Na-acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 10 µM Z-Phe-Arg-AMC),[77] SARS-CoV-2 M^{pro} (20 mM Tris pH 7.5, 0,1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 µM Dabcyl-KTSAVLQ|SGFRKME-Edans),[78] DENV2 and ZIKV NS2B/NS3 (50 mM Tris pH 9.0, 20% (v/v) Glycerol, 1 mM Chaps, 100 µM Boc-Gly-Arg-Arg-AMC),[62] urokinase plasminogen activator (50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 µM Z-Gly-Gly-Arg-AMC),[79] *Staphylococcus aureus* sortase A (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$,0.5 mM Gly_4, 25 µM Abz-LPETG-Dap(dnp)–OH),[80] thrombin (50 mM Tris pH 8.0, 100 mM NaCl, 5 mM $CaCl_2$, 0.01% Tween-20, 200 µM Z-Gly-Gly-Arg-AMC), trypsin (50 mM Tris pH 8.0, 100 mM NaCl, 5 mM EDTA, 40 µM Z-Phe-Arg-AMC), and α-chymotrypsin (50 mM Tris pH 8.0, 100 mM NaCl, 5 mM EDTA, 75 µM Succ-Leu-Tyr-AMC).[81]

4.5. *Antiviral activity and cytotoxicity*

4.5.1. *Cell culture*

Huh7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 2 mM L–glutamine, 1x non-essential AA, 100 U/mL penicillin, 100 µg/mL streptomycin (all from GIBCO, Life Technologies), and 10% fetal calf serum (Capricorn).

4.5.2. *Plasmids*

Plasmids encoding the wildtype ZIKV H/PF/2013 sub-genomic replicon expressing a *Renilla* luciferase gene (pFK-sgR2A-H/PF/2013) as well as the replication-defective mutant (pFK-sgR2A-H/PF/2013-GAA) were previously described.[69] Plasmid encoding the DENV2 strain New Guinea C (NGC) sub-genomic replicon expressing a *Firefly* luciferase gene (pDVWSK601ΔCprME-LucUbi) were also described previously.[68] Plasmids were kindly provided by R. Bartenschlager (Heidelberg).

4.5.3. ***In vitro** transcription*

10 µg plasmid were linearized by restriction using *Xho*I and *Xba*I for ZIKV-based and DENV2-based plasmids, respectively. Linearized plasmid was purified using the Nucleospin Extract II kit (Macherey-Nagel). *In vitro* transcription reactions were carried out in 100 µL final volume containing the linearized plasmid DNA in transcription buffer containing 80 mM HEPES pH 7.5, 12 mM $MgCl_2$, 2 mM spermidine, 40 mM dithiothreitol (DTT), 3.125 mM of ATP, CTP, and UTP, 1.5625 mM GTP (all from Roche), 100 U RNAsin ribonuclease inhibitor (Promega), and 1 mM cap structure analog. For DENV2 transcripts, m7G(5′)ppp(5′)G RNA cap structure analog (New England Biolabs) was used. For ZIKV transcripts, 3′-O-Me-m7G(5′)ppp(5′)G RNA cap structure analog (New England Biolabs). Reactions were supplemented with 80 U T7 RNA polymerase (Promega). After incubation for 3 h at 37 °C, 80 U T7 RNA polymerase was added and transcription was carried out for another 3 h. Transcription was terminated by the addition of 10 U RQ1 RNase-free DNase (Promega) and incubation at 37 °C for 30 min. RNA was purified using phenol:chloroform:isoamyl alcohol (25:24:1) (Applichem), precipitated with isopropanol, and the RNA pellet was dissolved in RNase-free water. RNA integrity was determined by using agarose gel electrophoresis and concentration was determined by measuring optical density at 260 nm.

4.5.4. Replication analysis and luciferase assays

Quantification of luciferase activity was used to determine ZIKV and DENV2 RNA replication as described previously.[68,69] In brief, single-cell suspensions of Huh7 cells were prepared by trypsinization and washed once with phosphate-buffered saline. Cells were resuspended at a concentration of 1×10^7 cells per mL in Cytomix containing 2 mM ATP and 5 mM glutathione. 10 µg of *in vitro* transcribed RNA was mixed with 400 µL of the cell suspension ($4x10^6$ cells) and transfected by electroporation using a Gene Pulser system (Bio-Rad) in a cuvette with a gap width of 0.4 cm (Bio-Rad) at 975 µF and 270 V.[82] Cells were resuspended in 15 mL culture medium. Cells were seeded in duplicate wells in a 12-well plate: 500 µL for the time points 4 h and 24 h and 1 mL for the time points 48 h, 72 h, and 96 h. 4 h after electroporation, cells were treated with 10 µM compound or an equivalent volume of DMSO, as vehicle control, in culture medium supplemented with 15 mM HEPES (Gibco, Life technologies). Compound- or DMSO-containing culture medium was replenished after 48 h. Cells were lysed at 4, 24, 48, 72, and 96 h after electroporation by addition of 250 µL luciferase lysis buffer (0.1% (v/v) Triton$_{X\text{-}100}$, 25 mM glycylglycine, 15 mM $MgSO_4$, 15 mM K_3PO_4 pH 7.8, 4 mM EGTA, 10% (v/v) glycerol, and 1 mM DTT). For detection of *Renilla* luciferase activity, 100 µL lysate was mixed with 200 µL luciferase assay buffer (25 mM glycylglycine, 15 mM $MgSO_4$, 15 mM K_3PO_4 pH 7.8, and 4 mM EGTA) supplemented with 14 or 28 nM coelanterazine (P.J.K). For detection of firefly luciferase activity, 100 µL lysate was mixed with 350 µL luciferase assay buffer freshly supplemented with 1 mM DTT and 2 mM ATP, and ᴅ-luciferin substrate (200 µM ᴅ-Luciferin, P.J.K.) in 25 mM glycylglycine. All measurements were done in duplicates by using a tube luminometer (Berthold Technologies). Replication efficiency was determined relative to the 4 h values, which reflect input transcript levels.

4.5.5. Cell viability assay

To determine the impact of compound treatment on cell viability, Huh7 cells were seeded at a density of $4x10^3$ cells per well in white-walled 96-well plates (Greiner Bio-One) and one day after seeding cells were treated with 1.25, 2.5, 5, 10, 20, and 40 µM compound or equivalent volumes of DMSO for 96 h. Cell viability was measured using the CellTiter-Glo Luminescent Cell Viability Assay (Promega) following manufacturer instructions. Cell viability was determined by measurement with a plate luminometer (Berthold Technologies) and values normalized to untreated cells.

4.5.6. Statistical significance

Statistical analysis was performed by using the GraphPad Prism.[83] Statistical significance was calculated by using one-way ANOVA for single time points and by using 2way ANOVA for replication curves. ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

4.6. Chemistry

All reagents and solvents were of analytical grade quality and purchased from Sigma-Aldrich, Carbolution, BLDpharmatech, or FisherScientific. Chemicals were used without further purification. Solvents were purified by distillation and desiccated by standard methods if necessary. ^{1}H and ^{13}C spectra were recorded on a *Bruker* Fourier 300 using DMSO-d_6, $CDCl_3$, MeOD-d_4 as solvent. Chemical shifts δ are given in parts per million (ppm) using residual proton peaks of the solvent as internal standard. HPLC and mass spectra were obtained by LC-MS consisting of a 1100 series HPLC system from *Agilent* with an *Agilent* Poroshell 120 EC-C_{18} 150 × 2.10 mm, 4 μm column. Detection wavelength was 254 nm. The purities of the inhibitors were higher than 95% in all cases. The molecular mass was detected by an *Agilent* 1100 series LC/MSD Trap with electron spray ionization (ESI) in positive mode. Column chromatography was performed with silica gel (0.06–0.02 mm) obtained from *Carl Roth*. All reactions were monitored by thin-layer chromatography using *Macherey-Nagel* ALUGRAM Xtra SIL G/UV254 silica gel 60 plates for detection at 254 nm. Melting points (uncorrected) were determined in open capillaries using a Schorpp MPM-H3 melting point device. Specific rotations $[\alpha]_D$[20] were measured on a P3000 polarimeter from Krüss and are reported in $cm^3\ g^{-1}\ dm^{-1}$.

***2-Amino-5,6-dimethoxybenzo*[*d*] thiazole (5)**: was prepared according to the literature.[84]

(Hetero)aromatic carboxylic acids (32a–l): were prepared according to literature procedures.

*5,6-Dimethoxybenzo[d]thiazole-2-carboxylic acid **(32a)***:[52] 400 mg, 1.65 mmol, yield: 89%. mp: 140–142 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.73 (s, 1H), 7.67 (s, 1H), 3.87 (s, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 161.6, 150.3, 149.9, 147.4, 129.4, 105.5, 103.4, 103.2, 56.0, 55.8 ppm.

*6-Methoxybenzo[d]thiazole-2-carboxylic acid **(32b)***:[52] 360 mg, 1.72 mmol, yield: 82%. mp: 120–122 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 9.17 (s, 1H), 8.06 (d, J = 9.0 Hz, 1H), 7.76 (d, J = 2.5 Hz, 1H), 7.22 (dd, J = 9.0, 2.5 Hz, 1H), 3.86 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 157.6, 153.5, 147.6, 135.2, 123.5, 115.8, 104.8, 55.8 ppm.

*5,6-Dimethoxy-1H-indole-2-carboxylic acid **(32c)***:[53,54] 330 mg, 1.49 mmol, yield: 99%. mp: decomp. > 205 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.45 (s, 1H), 7.07 (s, 1H), 6.96 (d, J = 2.1 Hz, 1H), 6.89 (s, 1H), 3.77 (s, 3H), 3.74 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 162.7, 149.2, 145.6, 132.5, 126.6, 119.9, 107.7, 102.7, 94.5, 55.7, 55.5 ppm.

*5,6-Dimethoxybenzo[b]thiophene-2-carboxylic acid **(32d)***:[55] 320 mg, 1.34 mmol, yield: 94%. mp: 256–258 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.94 (s, 1H), 7.55 (s, 1H), 7.45 (s, 1H), 3.84 (s, 3H), 3.81 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 163.7, 150.4, 148.5, 135.2, 132.3, 132.0, 130.1, 106.3, 104.2, 55.9, 55.6 ppm.

*5,6-Dimethoxybenzofuran-2-carboxylic acid **(32e)***:[56,57] 220 mg, 1.00 mmol, yield: 99%. mp: decomp. > 250 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.52 (s, 1H), 7.31 (s, 1H), 7.21 (s, 1H), 3.83 (s, 3H), 3.79 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 160.0, 150.7, 150.5, 147.2, 145.0, 118.6, 114.1, 103.2, 95.6, 56.0, 55.9 ppm.

*6,7-Dimethoxyquinoline-3-carboxylic acid **(32f)***:[58] 390 mg, 1.67 mmol, yield: 99%. mp: 265–267 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 9.10 (s, 1H), 8.72 (s, 1H), 7.52 (s, 1H), 7.43 (s, 1H), 3.96 (s, 3H), 3.92 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 166.9, 153.9, 149.9, 147.7, 146.6, 136.1, 122.4, 122.3, 107.5, 106.7, 55.9, 55.8 ppm.

*5-Methoxybenzofuran-2-carboxylic acid **(32g)***: 160 mg, 0.83 mmol, yield: 91%. mp: 217–219 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.61–7.59 (m, 1H), 7.58 (s, 1H), 7.26 (d, J = 2.6 Hz, 2H), 7.08 (dd, J = 9.0, 2.6 Hz, 2H), 3.79 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 160.1, 156.1, 150.1, 146.8, 127.5, 117.2, 113.7, 112.8, 104.3, 55.7 ppm.

*6-Methoxybenzofuran-2-carboxylic acid **(32h)***: 190 mg, 0.99 mmol, yield: 87%. mp: 211–213 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.64 (d, J = 8.7 Hz, 1H), 7.58 (d, J = 1.0 Hz, 1H), 7.28 (dd, J = 2.3, 1.0 Hz, 1H), 6.97 (dd, J = 8.7, 2.3 Hz, 1H), 3.83 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 160.1, 160.1, 156.5, 145.3, 123.4, 120.0, 113.9, 113.8, 95.9, 55.8 ppm.

*4,6-Dimethoxybenzofuran-2-carboxylic acid **(32i)***: 530 mg, 2,39 mmol, yield: 99%. mp: 235–237 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.45 (d, J = 1.0 Hz, 1H), 6.85 (dd, J = 1.9, 1.0 Hz, 1H), 6.44 (d, J = 1.9 Hz, 1H), 3.87 (s, 3H), 3.81 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 161.6, 160.0, 157.2, 154.4, 144.2, 111.1, 111.0, 95.3, 88.5, 56.0, 55.9 ppm.

*6-Methoxy-5-nitrobenzofuran-2-carboxylic acid **(32j)***: 590 mg, 2.5 mmol, yield: 75%. decomp. > 250 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 8.34 (s, 1H), 7.69 (dd, J = 9.1, 1.0 Hz, 2H), 3.98 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 159.5, 157.2, 152.5, 147.7, 138.0, 120.0, 119.0, 113.9, 97.7, 57.3 ppm.

*6,7-Dimethoxy-2-naphthoic acid **(32m)***:[50,51] 500 mg, 2.17 mmol, yield: 99%. mp: 245–247 °C. ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 12.80 (s, 1H), 8.44 (s, 1H), 7.80 (s, 2H), 7.49 (s, 1H), 7.36 (s, 1H), 3.91 (s, 3H), 3.89 (s, 3H) ppm. ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 167.8, 151.1, 149.8, 131.5, 128.9, 128.0, 126.4, 126.0, 123.5, 107.7, 106.3, 55.6 ppm.

***tert*-Butyl (*S*)-2-(phenylcarbamoyl)pyrrolidine-1-carboxylate (30a)**: was prepared according to the literature.[60] 1.80 g, 6.20 mmol, yield: 89%. mp: 187–189 °C. $R_f = 0.37$ (CH/EtOAc 2:1). $[\alpha]_D^{20} = -80$ ($c = 5$ in MeCN). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 9.96$ (s, 1H), 7.59 (d, $J = 7.6$ Hz, 2H), 7.30 (t, $J = 7.6$ Hz, 2H), 7.03 (t, $J = 7.6$ Hz, 1H), 4.27–4.17 (m, 1H), 3.46–3.29 (m, 2H), 2.28–2.09 (m, 1H), 1.97–1.72 (m, 3H), 1.39 (s, 3H), 1.27 (s, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 171.5$, 153.2, 139.1, 128.7, 123.2, 119.2, 119.0, 78.5, 60.4, 46.6, 31.0, 28.2, 28.0, 23.4 ppm.

***tert*-Butyl (*S*)-2-(phenylcarbamoyl)piperidine-1-carboxylate (30b)**: was prepared according to the literature.[60] 355 mg, 1.47 mmol, yield: 90%. mp: 183–185 °C. $R_f = 0.66$ (CH/EtOAc 2:1). $[\alpha]_D^{20} = -131$ ($c = 10$ in CHCl$_3$). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 9.89$ (s, 1H), 7.58 (d, $J = 7.7$ Hz, 2H), 7.29 (t, $J = 7.7$ Hz, 2H), 7.03 (t, $J = 7.7$ Hz, 1H), 4.65 (s, 1H), 3.94–3.69 (m, 1H), 3.34 (s, 2H), 2.17–1.99 (m, 1H), 1.80–1.50 (m, 4H), 1.36 (s, 11*H*) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.8$, 155.3, 139.0, 128.7, 123.2, 119.3, 78.9, 28.0, 27.5, 24.2, 19.6 ppm.

(*S*)-*S*-Benzyl-*N*-(*tert*-butoxycarbonyl)cysteine (4f): was prepared according to the literature.[85] 1100 mg, 3.31 mmol, yield: 81%. mp: 76 °C. $R_f = 0.48$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -44$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.35$–7.27 (m, 5H), 4.18–4.05 (m, 1H), 3.75 (s, 2H), 2.89–2.57 (m, 2H), 1.39 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 172.5$, 155.3, 138.3, 128.8, 128.3, 126.8, 78.2, 53.2, 35.2, 32.4, 28.2 ppm.

***N*-Benzoyl-*N*-phenylglycine (21a)**: was prepared according to the literature.[47] 826 mg, 3.24 mmol, yield: 81%. mp: 160 °C. $R_f = 0.43$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.77$ (s, 1H), 7.32–7.09 (m, 10*H*), 4.49 (s, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.6$, 169.8, 143.9, 135.8, 129.2, 128.8, 128.5, 128.1, 127.6, 126.8, 52.1 ppm.

2,2-Diphenylacetic acid (21b): was prepared according to the literature.[48] 1060 mg, 5.00 mmol, yield: 99%. mp: 146 °C. $R_f = 0.80$ (CH/EtOAc). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.71$ (s, 1H), 7.32 (s, 8H), 7.28–7.19 (m, 2H), 5.06 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 173.3$, 139.5, 128.4, 128.3, 126.8, 56.2 ppm.

2,2-Bis-(4-methoxyphenyl)acetic acid (21c): was prepared according to the literature.[49] 568 mg, 2.09 mmol, yield: 95%. mp: 107 °C. $R_f = 0.64$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.48$ (s, 1H), 7.24–7.17 (m, 4H), 6.90–6.83 (m, 4H), 4.90 (s, 1H), 3.71 (s, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 173.9$, 158.0, 131.9, 129.4, 113.7, 55.0, 54.6 ppm.

1-(4-Iodophenyl)-3-(6-Nitrobenzo [*d*] thiazol-2-yl)thiourea (28): A mixture of 2-amino-6-nitrobenzo[*d*]thiazole (195 mg, 1.0 mmol, 1.1 eq.), 4–iodophenyl isothiocyanate (234 mg, 0.9 mmol, 1.0 eq.), and DBU (179 µL, 1.2 mmol, 1.33 eq.) in DCM (5 mL) was stirred at room temperature for 7 d. The mixture was evaporated and conc. HCl (5 mL) was added. The crude product was filtrated and washed with HCl (3x 10 mL) and pentane (3x 10 mL) to yield **28** (280 mg, 0.61 mmol, yield: 68%). mp: 231 °C. $R_f = 0.20$ (EtOAc). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 11.22$ (s, 1H), 8.87 (d, $J = 2.4$ Hz, 1H), 8.26 (dd, $J = 8.8$, 2.4 Hz, 1H), 7.74–7.63 (m, 3H), 7.54 (d, $J = 8.8$ Hz, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 166.8$, 164.5, 153.5, 151.5, 143.0, 142.6, 138.9, 138.2, 137.7, 137.3, 132.2, 125.1, 122.6, 120.9, 119.6, 119.2 ppm. MS (ESI) m/z [M + H$^+$] calcd for $C_{14}H_9IN_4O_2S_2$ 456.92 (100%), 457.92 (15.1%), 458.92 (9.0%), found 456.89 (100%), 457.87 (16.3%), 458.88 (8.0%); purity (HPLC) = 96%.

General procedure for the amide bond formation mediated by TBTU

The respective carboxylic acid (1.0 eq.), TBTU (1.1 eq.), and HOBt were dissolved in dry DMF and DIPEA (4.0 eq.) was added while cooling with an ice-water bath. The corresponding amine (1.0 eq.) was added after 30 min and the solution was stirred for 2 d at r.t. Water and EtOAc were added and the aqueous phase was extracted three times with EtOAc and the combined organic phases were washed with 2 M HCl, saturated NaHCO$_3$-solution, and saturated NaCl-solution three times each. The organic phase was dried with MgSO$_4$ and the solvent was evaporated under reduced pressure.

(*S*)-2-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)propionamide (6a): 864 mg, 2.27 mmol, yield: 48%. mp: 93–100 °C. $R_f = 0.37$ (CH/EtOAc 2:1). $[\alpha]_D^{20} = -39$ ($c = 10$ in EtOAc). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.31$ (s, 1H), 7.23 (s, 1H), 4.01–3.81 (m, 7H), 1.56–1.34 (m, 12H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 171.8$, 157.3, 155.8, 149.7, 147.9, 141.3, 123.3, 103.0, 102.9, 80.9, 56.5, 56.3, 50.6, 28.4, 18.1 ppm.

(*R*)-2-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)propionamide (6b): 561 mg, 1.47 mmol, yield: 77%. mp: 83–85 °C. $R_f = 0.37$ (CH/EtOAc 2:1). $[\alpha]_D^{20} = +43$ ($c = 10$ in EtOAc). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.31$ (s, 1H), 7.23 (s, 1H), 4.01–3.81 (m, 7H), 1.56–1.34 (m, 12H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 171.8$, 157.3, 155.8, 149.7, 147.9, 141.3, 123.3, 103.0, 102.9, 80.9, 56.5, 56.3, 50.6, 28.4, 18.1 ppm.

(*S*),(*S*)-2-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-3-methylpentanoicamide (6c): 308 mg, 0.82 mmol, yield: 99%. mp = 99–103 °C. $R_f = 0.47$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -14$ ($c = 10$ in EtOAc). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.39$ (s, 1H), 7.17 (s, 1H), 4.40–4.25 (m, 1H), 3.89 (s, 3H), 3.89 (s, 3H), 1.83–1.68 (m, 1H), 1.46–1.27 (m, 10*H*), 1.05–0.83 (m, 1H), 0.78–0.58 (m, 6H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 171.5$, 157.9, 156.0, 149.9, 148.0, 140.5, 122.8, 102.9, 102.7, 80.4, 59.4, 56.5, 56.4, 37.6, 28.4, 24.7, 15.5, 11.1 ppm.

(*S*)-2-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)-3,3-dimethylbutyramide (6d): 1500 mg, 3.55 mmol, yield: 82%. mp = 185–195 °C. $R_f = 0.51$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +9$ ($c = 10$ in DCM). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.54$ (s, 1H), 7.29 (s, 1H), 4.22 (d, $J = 8.6$ Hz, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 1.37 (s, 9H), 0.96 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.9$, 170.4, 155.9, 148.9, 147.0, 142.5, 122.9, 103.7, 103.5, 78.4, 62.1, 56.0, 55.7, 34.0, 28.2, 26.5 ppm.

(*S*)-2-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)-3-phenylpropionamide (6e): 370 mg, 0.81 mmol, yield: 72%. mp: 96–98 °C. $R_f = 0.70$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +24$ ($c = 10$ in EtOAc). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.32$ (s, 1H), 7.24 (s, 1H), 7.21–7.14 (m, 3H), 7.09–6.98 (m, 2H), 4.87–4.60 (m, 1H), 3.96 (s, 3H), 3.90 (s, 3H), 3.26–2.99 (m, 2H), 1.39 (s, 9H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 170.9$, 157.8, 155.6, 150.0, 148.1, 139.8, 135.7, 129.2, 128.8, 127.3, 122.6, 102.8, 102.6, 77.4, 56.5, 56.4, 56.1, 38.3, 28.3 ppm.

***tert*-Butyl (*S*)-(3-(benzylthio)-1-((5,6-dimethoxybenzo [*d*] thiazol-2-yl)amino)-1-oxopropan-2-yl)carbamate (6f)**: 305 mg, 0.60 mmol, yield: 34%. mp: 99 °C. $R_f = 0.51$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -61$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.46$ (s, 1H), 7.56 (s, 1H), 7.37–7.17 (m, 6H), 4.64–4.42 (m, 1H), 3.93–3.72 (m, 8H), 2.87–2.58 (m, 2H), 1.40 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta =$ 170.3, 155.2, 154.0, 148.9, 147.0, 142.5, 138.2, 128.9, 128.3, 126.9, 122.9, 103.7, 103.5, 78.5, 55.9, 55.74, 54.3, 35.0, 32.6, 28.1 ppm.

(*S*)-3-(4-(Benzyloxy)phenyl)-2-(*tert*-butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)propionamide (9a): 1039 mg, 1.84 mmol, yield: 97%. mp: 86–88 °C. $R_f = 0.59$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -38$ ($c = 10$ in DCM). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.46$–7.23 (m, 7H), 6.87 (d, $J = 8.3$ Hz, 2H), 6.75 (d, $J = 8.3$ Hz, 2H), 4.95 (s, 2H), 4.81–4.63 (m, 1H), 3.96 (s, 3H), 3.91 (s, 3H), 3.18–2.86 (m, 2H), 1.40 (s, 9H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 170.9$, 157.9, 157.3, 155.6, 149.7, 147.8, 142.0, 137.0, 130.2, 128.6, 128.0, 127.9, 127.6, 123.5, 115.0, 103.4, 102.8, 70.0, 69.9, 56.5, 56.3, 38.7, 37.6, 28.4 ppm.

Benzyl (*S*)-3-((*tert*-butoxycarbonyl)amino)-4-((5,6-dimethoxybenzo [*d*] thiazol-2-yl)amino)-4-oxobutanoate (9b): 604 mg, 1.17 mmol, yield: 95%. mp = 73–75 °C. $R_f = 0.36$ (CH/EtOAc 1:2). $[\alpha]_D^{20} = -17$ ($c = 10$ in DCM). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.35$–7.30 (m, 5H), 7.29 (s, 1H), 7.22 (s, 1H), 4.91–4,69 (m, 1H), 5.14 (d, $J = 2.2$ Hz, 2H) 3.94 (s, 3H), 3.93 (s, 3H), 3.20 (dd, $J = 17.5$, 4.6 Hz, 1H), 2.87 (dd, $J = 17.5$, 5.6 Hz, 1H), 1.47 (s, 9H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta =$

171.3, 169.1, 155.9, 155.6, 149.4, 147.7, 142.4, 135.1, 128.6, 128.5, 128.3, 123.8, 103.5, 102.7, 81.5, 67.2, 56.4, 56.2, 51.3, 35.7, 28.3 ppm.

3-(*tert*-Butoxycarbonylamino)-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)propionamide (14): 1.58 g, 4.10 mmol, yield: 78%. mp: 130–135 °C. R$_f$ = 0.23 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.53$ (s, 1H), 7.28 (s, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.30–3.19 (m, 2H), 2.63 (t, $J = 6.8$ Hz, 2H), 1.36 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.0$, 156.3, 155.5, 148.9, 146.9, 142.5, 122.8, 103.7, 103.5, 77.7, 55.9, 55.7, 36.0, 35.6, 28.2 ppm.

(*S*)-2-benzamido-*N^1*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)-*N^4*, *N^4*-diethylsuccinamide (19a): 200 mg, 0.41 mmol, yield: 35%. mp: 121–123 °C. R$_f$ = 0.32 (CH/EtOAc 1:1). $[\alpha]_D{}^{20} = +2$ ($c = 10$ in DCM). ^{1}H NMR: (300 MHz, CDCl$_3$) $\delta = 7.90$ (d, $J = 7.2$ Hz, 1H), 7.59–7.39 (m, 3H), 7.29 (s, 1H), 7.19 (s, 1H), 5.28 (td, $J = 6.8$, 3.0 Hz, 1H), 3.93 (s, 3H), 3.92 (s, 3H), 3.54–3.24 (m, 5H), 2.84–2.63 (m, 1H), 1.28–1.19 (m, 6H) ppm. ^{13}C NMR: (75 MHz, CDCl$_3$) $\delta = 170.3$, 169.6, 167.4, 156.4, 149.5, 147.7, 141.6, 133.0, 132.2, 128.7, 127.4, 123.2, 103.2, 102.6, 56.4, 56.1, 50.6, 42.6, 40.8, 34.4, 14.1, 13.0 ppm.

(*S*)-2-Benzamido-*N*^{1}-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)-*N*^{4}-phenylsuccinamide (19b): 88 mg, 0.18 mmol, yield: 65%. mp: 205–208 °C. R$_f$ = 0.17 (CH/EtOAc 1:1). $[\alpha]_D{}^{20} = +3$ ($c = 10$ in DCM). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.93$ (d, $J = 7.5$ Hz, 2H), 7.65 (d, $J = 8.0$ Hz, 2H), 7.57–7.42 (m, 4H), 7.35–7.23 (m, 3H), 7.05 (t, $J = 7.4$ Hz, 1H), 5.14–5.01 (m, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.22–3.01 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.0$, 169.3, 166.9, 156.7, 149.4, 147.4, 143.0, 139.5, 134.3, 131.9, 129.1, 128.7, 128.0, 123.8, 123.3, 119.8, 104.2, 104.0, 56.5, 56.2, 51.5, 37.5 ppm.

(2-((5,6-Dimethoxybenzo [*d*] thiazol-2-yl)amino)-2-oxoethyl)-*N*-phenylbenzamide (22a): 300 mg, 0.67 mmol, yield: 52%. mp: 231 °C. R$_f$ = 0.61 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta =$ 12.41 (s, 1H), 7.53 (s, 1H), 7.32 (d, $J = 4.8$ Hz, 1H), 7.31–7.09 (m, 10*H*), 4.80 (s, 2H), 4.01–3.87 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta =$ 169.7, 167.6, 156.1, 148.9, 147.0, 143.8, 142.5, 135.4, 129.8, 128.8, 128.4, 127.8, 127.5, 126.6, 122.8, 103.7, 103.6, 55.9, 55.7, 53.1 ppm.

***N*-(5,6-Dimethoxybenzo [*d*] thiazol-2-yl)-2,2-diphenylacetamide (22b)**: 550 mg, 136 mmol, yield: 68%. mp: 234 °C. R$_f$ = 0.59 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.69$ (s, 1H), 7.54 (s, 1H), 7.40–7.24 (m, 11*H*), 5.40 (s, 1H), 4.04–3.85 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.3$, 156.2, 148.9, 147.0, 142.4, 138.9, 128.5, 128.5, 127.1, 122.8, 103.6, 103.5, 59.7, 55.9, 55.7 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{23}H_{20}N_2O_3S$ 405.12 (100%), 406.12 (24.9%), 407.12 (4.5%), found 405.23 (100%), 406.19 (24.8%), 407.15 (4.0%); purity (HPLC) = 99%.

***N*-(5,6-Dimethoxybenzo [*d*] thiazol-2-yl)-2,2-bis-(4-methoxyphenyl)acetamide (22c)**: 120 mg, 0.25 mmol, yield: 25%. mp: 242 °C. R$_f$ = 0.40 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.52$ (s, 1H), 7.53 (s, 1H), 7.31–7.17 (m, 5H), 6.91 (d, $J = 8.4$ Hz, 4H), 5.25 (s, 1H), 4.00–3.84 (m, 6H), 3.72 (s, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 172.0$, 158.3, 156.3, 148.9, 147.0, 142.5, 131.3, 129.5, 122.8, 113.9, 103.6, 103.5, 55.9, 55.7, 55.1, 54.4 ppm.

***N*-(5,6-Dimethoxybenzo [*d*] thiazol-2-yl)-2,2-bis-(4-chlorophenyl)acetamide (22d)**: 230 mg, 0.48 mmol, yield: 54%. mp: 251 °C. R$_f$ = 0.48 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta =$ 7.54 (s, 1H), 7.46–7.27 (m, 9H), 5.47 (s, 1H), 3.81 (d, $J = 4.4$ Hz, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.8$, 156.0, 148.9, 147.1, 142.4, 137.5, 132.0, 130.4, 128.6, 122.8, 103.6, 103.5, 56.0, 55.7, 54.5 ppm.

***N*-(5,6-Dimethoxybenzo [*d*] thiazol-2-yl)-4-iodobenzamide (24b)**: 223 mg, 0.51 mmol, yield 51%. mp: 298 °C. R$_f$ = 0.66 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.81$ (s, 1H), 7.95 (d, $J = 8.6$ Hz, 2H), 7.88 (d, $J = 8.6$ Hz, 2H), 7.59 (s, 1H), 7.30 (s, 1H), 4.06–3.82 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 172.5$, 158.7, 149.0, 147.2, 137.5, 137.0, 131.6, 130.1, 129.4, 122.9, 103.7, 100.9, 56.0, 55.6 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{16}H_{13}IN_2O_3S$ 440.97 (100%), 441.97 (17.3%), 442.97 (4.5%), found 441.05 (100%), 442.05 (16.3%), 443.00 (4.3%); purity (HPLC) = 97%.

4-Chloro-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)benzamide (24c): 378 mg, 1.08 mmol, yield: 85%. mp: 254 °C. R$_f$ = 0.30 (DCM/MeOH 49:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 7.89$ (d, $J = 8.3$ Hz, 2H), 7.38 (d, $J = 8.3$ Hz, 2H), 7.34 (s, 1H), 7.07 (s, 1H) 3.90–3.76 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 165.1$, 157.8, 149.5, 147.6, 142.4, 138.0, 131.4, 130.7, 129.2, 123.3, 104.2, 103.6, 56.4, 56.2 ppm.

4-Methyl-*N*-(5,6-dimethoxybenzothiazol-2-yl)benzamide (24d): 502 mg, 1.53 mmol yield: 99%. mp: 254 °C. R$_f$ = 0.30 (DCM/MeOH 49:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 8.02$ (d, $J = 7.8$ Hz, 2H), 7.57 (s, 1H), 7.35 (d, $J = 7.8$ Hz, 2H), 7.30 (s, 1H), 3.82 (s, 6H), 2.38 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 165.8$, 157.8, 149.4, 147.5, 143.5, 142.7, 123.4, 129.7, 129.6, 128.7, 104.2, 103.7, 56.4, 56.2, 21.6 ppm.

***N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)- [1,1́-biphenyl] -4-carboxamide (24e)**: 295 mg, 0.75 mmol, yield: 75%. mp: 273 °C. R$_f$ = 0.70 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.71$ (s, 1H), 8.27–8.19 (m, 2H), 7.87 (d, $J = 8.2$ Hz, 2H), 7.79 (d, $J = 7.6$ Hz, 2H), 7.58–7.48 (m, 4H), 7.32 (s, 1H), 3.88–3.81 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.4$, 157.1, 149.0, 147.1, 138.8, 144.3, 139.1, 129.9, 129.6, 128.9, 128.1, 126.9, 126.7, 126.7, 107.4, 106.4, 98.9, 59.7, 55.9 ppm.

3,4-Dichloro-*N*-(5,6-dimethoxybenzo [*d*] thiazol-2-yl)benzamide (24f): 95 mg, 0.29 mmol, yield: 24%. mp: 150 °C. R$_f$ = 0.30 (DCM/MeOH 49:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 8.43$ (s, 1H), 8.12 (d, $J = 8.5$ Hz, 1H), 7.89 (d, $J = 8.5$ Hz, 1H), 7.64 (s, 1H), 7.36 (s, 1H), 4.00–3.84 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 172.5$, 164.1, 149.5, 147.7, 142.2, 135.8, 133.2, 131.4, 130.7, 129.0, 123.4, 104.2, 103.5, 56.44, 56.2 ppm.

4-Iodo-*N*-(6-nitrobenzo [*d*] thiazol-2-yl) benzamide (27): 171 mg, 0.30 mmol, yield: 74%. mp: 211 °C. R$_f$ = 0.21 (CH/EtOAc 3:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 13.34$ (s, 1H), 9.10 (d, $J = 2.4$ Hz, 1H), 8.31 (dd, $J = 9.0$, 2.4 Hz, 1H), 8.01–7.88 (m, 5H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 177.6$, 165.8, 150.4, 143.1, 138.3, 137.6, 130.2, 121.8, 119.1, 101.7 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{14}H_8IN_3O_3S$ 425.93 (100%), 426.94 (15.1%), 427.93 (4.5%), found 425.88 (100%), 426.91 (14.6%), 427.90 (2.0%); purity (HPLC) = 97%.

5,6-Dimethoxy-*N*-phenylbenzo [*d*] thiazole-2-carboxamide (33a): 180 mg, 0.57 mmol, yield: 90%. mp: 205–207 °C. R$_f$ = 0.76 (CH/EtOAc 1:2). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 10.87$ (s, 1H), 7.89 (d, $J = 7.4$ Hz, 2H), 7.79 (s, 1H), 7.59 (s, 1H), 7.38 (t, $J = 7.4$ Hz, 2H), 7.15 (t, $J = 7.4$ Hz, 1H), 3.90 (s, 3H), 3.89 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 161.8$, 158.3, 149.9, 147.0, 138.0, 129.4, 128.8, 124.3, 120.6, 104.9, 103.6, 56.0, 55.8 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{16}H_{14}N_2O_3S$ 315.1, found 315.2.

6-Methoxy-*N*-phenylbenzo [*d*] thiazole-2-carboxamide (33b): 180 mg, 0.63 mmol, yield: 88%. mp: 163–165 °C. R$_f$ = 0.67 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 10.99$ (s, 1H), 8.08 (d, $J = 9.1$ Hz, 1H), 7.96–7.85 (m, 2H), 7.81 (d, $J = 2.6$ Hz, 1H), 7.44–7.31 (m, 2H), 7.24 (dd, $J = 9.1$, 2.6 Hz, 1H), 7.21–7.09 (m, 1H), 3.88 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 161.8$, 158.7, 158.3, 147.1, 138.3, 138.0, 128.7, 124.8, 124.4, 120.7, 117.4, 104.8, 55.9 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{15}H_{12}N_2O_2S$ 285.1, found 285.1.

5,6-Dimethoxy-*N*-phenyl-1*H*-indole-2-carboxamide (33c): 140 mg, 0.47 mmol, yield: 69%. mp: decomp. > 215 °C. R$_f$ = 0.23 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 11.44$ (s, 1H), 10.02 (s, 1H), 7.84–7.67 (m, 2H), 7.55–7.27 (m, 3H), 7.16–7.02 (m, 2H), 6.93 (s, 1H), 3.79 (s, 3H), 3.78 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta =$ 159.7, 148.8, 145.5, 139.2, 132.0, 129.7, 128.7, 123.3, 120.1, 119.9, 104.1, 102.8, 94.6, 55.8, 55.5 ppm. MS (ESI) *m/z* $[M + H^+]$ calcd for $C_{17}H_{16}N_2O_3$ 297.1, found 297.2.

5,6-Dimethoxy-*N*-phenylbenzo [*b*] thiophene-2-carboxamide (33d): 160 mg, 0.50 mmol, yield: 99%. mp: 204–206 °C. R$_f$ = 0.36 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 10.38$ (s, 1H), 8.19 (s, 1H), 7.75 (d, $J = 7.7$ Hz, 2H), 7.59 (s, 1H), 7.44 (s, 1H), 7.36 (t, $J = 7.7$ Hz, 2H), 7.11 (t, $J = 7.7$ Hz, 1H), 3.86 (s, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 160.5$, 149.9, 148.5, 138.8, 137.5, 134.3, 132.5, 128.8, 125.7, 123.8, 120.4, 106.2, 104.3, 55.9 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{17}H_{15}NO_3S$ 314.1, found 314.2.

5,6-Dimethoxy-*N*-phenylbenzofuran-2-carboxamide (33e): 180 mg, 0.61 mmol, yield: 90%. mp: 187–189 °C. R$_f$ = 0.37 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.33 (s, 1H), 7.81 (d, J = 7.7 Hz, 2H), 7.67 (s, 1H), 7.36 (t, J = 7.7 Hz, 2H), 7.32–7.24 (m, 2H), 7.11 (d, J = 7.7 Hz, 1H), 3.86 (s, 3H), 3.82 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 156.7, 150.3, 149.8, 147.7, 147.2, 138.6, 128.7, 123.8, 120.4, 118.9, 111.1, 103.3, 95.5, 56.0, 55.9 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{17}H_{15}NO_4$ 298.10 (100%), 299.10 (18.4%), 300.11 (1.6%), found 298.18 (100%), 299.18 (16.7%), 300.17 (1.0%); purity (HPLC) = 99%.

6,7-Dimethoxy-*N*-phenylquinoline-3-carboxamide (33f): 140 mg, 0.45 mmol, yield: 70%. mp: 201–203 °C. R$_f$ = 0.53 (DCM/MeOH 19:1). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.48 (s, 1H), 9.17 (d, J = 2.2 Hz, 1H), 8.74 (d, J = 2.2 Hz, 1H), 7.82 (d, J = 7.7 Hz, 2H), 7.49 (s, 1H), 7.47 (s, 1H), 7.38 (t, J = 7.7 Hz, 2H), 7.13 (t, J = 7.7 Hz, 1H), 3.98 (s, 3H), 3.94 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 164.4, 153.6, 150.0, 146.5, 146.1, 139.1, 134.1, 128.7, 125.8, 123.8, 122.1, 120.3, 107.5, 106.4, 55.9, 55.8 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{17}H_{15}NO_4$ 298.1, found 298.2.

5-Methoxy-*N*-phenylbenzofuran-2-carboxamide (33g): 170 mg, 0.64 mmol, yield: 94%. mp: 168–170 °C. R$_f$ = 0.40 (CH/EtOAc 4:1). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 10.48 (s, 1H), 7.87–7.77 (m, 2H), 7.73–7.66 (m, 3H), 7.62 (d, J = 9.1 Hz, 1H), 7.42–7.33 (m, 2H), 7.34–7.27 (m, 2H), 7.18–7.04 (m, 2H), 3.81 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 156.7, 156.1, 149.4, 138.4, 128.7, 127.8, 124.1, 120.5, 116.5, 112.6, 110.8, 104.2, 55.6 ppm.

6-Methoxy-*N*-phenylbenzofuran-2-carboxamide (33h): 165 mg, 0.62 mmol, yield: 91%. mp: 134–136 °C. R$_f$ = 0.4 (CH/EtOAc 4:1). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 10.37 (s, 1H), 7.86–7.76 (m, 2H), 7.74–7.64 (m, 2H), 7.36 (dd, J = 8.6, 7.3 Hz, 2H), 7.30–7.23 (m, 1H), 7.17–7.06 (m, 1H), 6.99 (dd, J = 8.6, 2.3 Hz, 1H), 3.85 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 159.7, 156.7, 155.8, 148.0, 138.5, 128.7, 123.9, 123.2, 120.4, 120.3, 113.4, 110.8, 95.9, 55.7 ppm.

4,6-Dimethoxy-*N*-phenylbenzofuran-2-carboxamide (33i): 160 mg, 0.54 mmol, yield: 79%. mp: 144–146 °C. R$_f$ = 0.25 (CH/EtOAc 4:1). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 10.25 (s, 1H), 7.84–7.75 (m, 2H), 7.76–7.70 (m, 2H), 7.41–7.30 (m, 2H), 7.16–7.05 (m, 1H), 6.89–6.83 (m, 2H), 6.49 (d, J = 1.9 Hz, 1H), 3.91 (s, 3H), 3.84 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 161.2, 156.6, 154.2, 146.6, 138.6, 128.7, 123.8, 120.3, 111.1, 108.1, 95.2, 88.5, 55.9, 55.8, 52.8 ppm.

6-Methoxy-5-nitro-*N*-phenylbenzofuran-2-carboxamide (33j): 760 mg, 2.4 mmol, yield: 97%. mp: 184–186 °C. R$_f$ = 0.44 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.51 (s, 1H), 8.44 (s, 1H), 7.85 (d, J = 0.9 Hz, 1H), 7.81–7.73 (m, 2H), 7.67 (d, J = 0.9 Hz, 1H), 7.43–7.32 (m, 2H), 7.20–7.08 (m, 1H), 4.00 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 157.2, 156.5, 152.7, 150.6, 138.7, 138.5, 129.2, 124.7, 121.0, 120.4, 119.8, 111.3, 98.0, 57.8 ppm.

5,6-Dimethoxy-*N*-(4-iodophenyl)benzo [*d*] thiazole-2-carboxamide (33k): 230 mg, 0.52 mmol, yield: 76%. mp: 181–185 °C. R$_f$ = 0.55 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.96 (s, 1H), 7.78–7.67 (m, 5H), 7.57 (s, 1H), 3.89 (d, J = 2.0 Hz, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 161.4, 158.4, 150.0, 150.0, 146.9, 137.9, 137.4, 129.5, 122.7, 104.9, 103.5, 88.3, 56.0, 55.8 ppm.

(*S*)-1-(5,6-Dimethoxybenzo [*d*] thiazole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35a): 210 mg, 0.51 mmol, yield: 81%. mp: 237–239 °C. R$_f$ = 0.47 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = –10 (c = 5 in DMSO). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.2 (s, 1H), 7.74–7.52 (m, 3H), 7.38–7.20 (m, 3H), 7.11–6.39 (m, 1H), 4.76–4.56 (m, 1H), 4.36–4.23 (m, 1H), 3.88 (s, 3H), 3.82 (s, 3H), 3.79–3.72 (m, 1H), 2.37–1.82 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 171.2, 169.0, 159.0, 149.9, 149.7, 147.5, 139.5, 128.7, 128.5, 123.3, 119.2, 105.5, 103.2, 62.3, 55.9, 55.5, 48.7, 32.1, 21.8 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{21}H_{21}N_3O_4S$ 412.1, found 412.1.

(*S*)-1-(6-Methoxybenzo [*d*] thiazole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35b): 150 mg, 0.39 mmol, yield: 54%. mp: 209–211 °C. R$_f$ = 0.28 (CH/EtOAc 2:1). $[\alpha]_D^{20}$ = –6 (c = 5 in DMSO). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.21 (s, 1H), 8.05 (d, J = 9.0 Hz, 1H), 7.78–7.57 (m, 2H), 7.55–7.48 (m, 1H), 7.34–7.17 (m, 3), 7.14–6.95 (m, 1H), 5.58–5.39 (m, 1H), 4.34–4.17 (m, 1H), 3.84 (s, 3H), 3.78–3.64 (m, 1H), 2.47–1.85 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.9, 169.9, 159.0, 147.6, 147.2, 139.3, 128.7, 125.3, 124.8, 123.3, 119.3, 117.0, 104.3, 62.2, 55.8, 49.6, 32.2, 21.7 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{19}N_3O_3S$ 382.1, found 382.1.

(*S*)-1-(5,6-Dimethoxy-1*H*-indole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35c): 250 mg, 0.63 mmol, yield: 93%. mp: 113–115 °C. R$_f$ = 0.33 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = –62 (c = 5 in DMSO). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 11.26 (s, 1H), 10.10 (s, 1H), 7.71–7.53 (m, 2H), 7.36–7.24 (m, 2H), 7.14–7.00 (m, 2H), 6.96–6.88 (m, 2H), 4.74–4.64 (m, 1H), 4.09–3.85 (m, 2H), 3.40 (s, 6H), 2.31–2.06 (m, 1H), 2.06–1.89 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.8, 160,4, 148.7, 145.3, 139.3, 131.1, 129.0, 128.7, 123.2, 120.2, 119.2, 105.4, 102.8, 94.5, 61.7, 55.7, 55.4, 48.8, 29.1, 25.1 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{22}H_{23}N_3O_4$ 394.2, found 394.2.

(*S*)-1-(5,6-Dimethoxybenzo [*b*] thiophene-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35d): 210 mg, 0.50 mmol, yield: 99%. mp: 95–97 °C. R$_f$ = 0.13 (CH/EtOAc 2:1). $[\alpha]_D^{20}$ = –16 (c = 5 in MeCN). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.10 (s, 1H), 7.90 (s, 1H), 7.62 (d, J = 7.7 Hz, 2H), 7.55 (s, 1H), 7.45 (s, 1H), 7.30 (d, J = 7.7 Hz, 2H), 7.04 (t, J = 7.7 Hz, 1H), 4.70–4.60 (m, 1H), 4.04–3.92 (m, 2H), 3.84 (s, 3H), 3.83 (s, 3H), 2.33–2.22 (m, 1H), 2.18–2.06 (m, 1H), 2.05–1.89 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.4, 161.1, 149.8, 148.3, 139.2, 136.9, 133.5, 132.8, 128.7, 126.8, 123.2, 119.2, 106.3, 103.9, 62.0, 55.8, 55.6, 49.4, 29.2, 25.2 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{22}H_{22}N_2O_4S$ 411.1, found 411.2.

(*S*)-1-(5,6-Dimethoxybenzofuran-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35e): 250 mg, 0.63 mmol, yield: 93%. mp: 184–186 °C. R$_f$ = 0.17 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = –80 (c = 5 in MeCN). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.11 (s, 1H), 7.67–7.55 (m, 2H), 7.44 (s, 1H), 7.36–7.24 (m, 3H), 7.15 (s, 1H), 7.01–6.98 (m, 1H), 4.72–4.62 (m, 1H), 4.05–3.97 (m, 2H), 3.84 (s, 3H), 3.80 (s, 3H), 2.33–2.18 (m, 1H), 2.16–1.83 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.4, 157.6, 150.1, 149.5, 147.8, 147.1, 139.2, 128.7, 123.2, 119.2, 118.6, 112.5, 103.1, 95.5, 61.6, 61.3, 59.8, 56.0, 55.8, 48.6, 48.1, 32.3, 29.0, 25.0, 21.7, 20.7, 14.1 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{22}H_{22}N_2O_5$ 395.15 (100%), 396.16 (23.8%), 397.16 (2.7%), found 395.20 (100%), 396.17 (26.7%), 397.16 (2.4%); purity (HPLC) = 99%.

(*S*)-1-(6,7-Dimethoxyquinoline-3-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35f): 30 mg, 0.57 mmol, yield: 89%. mp: 112–114 °C. R$_f$ = 0.18 (DCM/MeOH 49:1). $[\alpha]_D^{20}$ = –44 (c = 5 in MeCN). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.13 (s, 1H), 8.83 (d, J = 2.1 Hz, 1H), 8.46 (d, J = 2.1 Hz), 7.65 (t, J = 7.8 Hz, 2H), 7.48 (s, 1H), 7.42 (s, 1H), 7.32 (t, J = 7.8 Hz, 2H), 7.05 (t, J = 7.8 Hz, 1H), 4.73–4.62 (m, 1H), 3.96 (s, 3H), 3.92 (s, 3H), 3.74–3.60 (m, 2H), 2.38–2.25 (m, 1H), 2.38–2.04 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.4, 166.7, 153.3, 150.0, 146.3, 145.2, 139.2, 133.2, 128.7, 127.2, 123.2, 122.0, 119.2, 107.4, 106.3, 61.0, 55.8, 54.9, 29.8, 25.2. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{23}H_{23}N_3O_4$ 406.2, found 406.2.

(*S*)-1-(4,6-Dimethoxybenzofuran-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (35i): 260 mg, 0.67 mmol, yield: 98%. mp: 125–128 °C. R$_f$ = 0.43 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = –50 (c = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 10.36–9.96 (m, 1H), 7.66–7.52 (m, 2H), 7.41–7.23 (m, 3H), 7.09–6.99 (m, 1H), 6.88 (s, 1H), 6.50–6.35 (m, 1H), 5.14–4.59 (m, 1H), 4.09–3.94 (m, 1H), 3.93–3.63 (m, 6H), 2.46–1.82 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.3, 161.0, 157.4, 156.2, 154.1, 146.7, 139.1, 128.7, 123.2, 119.3, 119.2, 110.8, 109.2, 95.1, 88.3, 61.6, 61.3, 55.9, 55.7, 48.6, 48.2, 32.3, 29.0, 25.0, 21.7 ppm.

(*S*)-1-(6,7-Dimethoxy-2-naphthoyl)-*N*-phenylpyrrolidine-2-carboxamide (35m): 150 mg, 0.37 mmol, yield: 86%. mp: 173–175 °C. R$_f$

= 0.33 (CH/EtOAc 1:2). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.10 (s, 1H), 8.01 (s, 1H), 7.79 (d, J = 8.3 Hz, 1H), 7.72 (s, 1H), 7.65 (d, J = 8.3 Hz, 1H), 7.52–7.41 (m, 2H), 7.39–7.19 (m, 3H), 7.11–6.99 (m, 1H), 4.70–4.59 (m, 1H), 3.93–3.82 (m, 6H), 3.73–3.55 (m, 2H), 2.35–2.24 (m, 1H), 2.04–1.80 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.7, 168.6, 150.3, 149.8, 139.2, 131.7, 129.6, 128.7, 127.9, 126.1, 125.6, 123.2, 122.7, 119.2, 107.1, 106.3, 60.8, 55.5, 50.2, 29.8, 25.2 ppm. MS (ESI) m/z [M + H^{+}] calcd for $C_{24}H_{24}N_2O_4$ 405.2, found 405.2.

(*S*)-1-(5,6-Dimethoxybenzofuran-2-carbonyl)-*N*-phenylpiperidine-2-carboxamide (35n): 155 mg, 0.38 mmol, yield: 56%. mp: 104–106 °C. R_f = 0.22 (CH/EtOAc 2:1). $[\alpha]_D^{20}$ = –148 (c = 10 in MeCN). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 10.01 (s, 1H), 7.67–7.57 (m, 2H), 7.37–7.26 (m, 3H), 7.21 (s, 1H), 7.12–7.01 (m, 1H), 5.15 (d, J = 3.9 Hz, 1H), 4.36 (d, J = 12.8 Hz, 1H), 3.92–3.70 (m, 6H), 2.23 (d, J = 14.0 Hz, 1H), 1.91–1.63 (m, 3H), 1.61–1.37 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 169.8, 160.6, 149.8, 149.1, 147.5, 147.0, 138.9, 128.7, 123.5, 119.7, 118.4, 103.1, 95.5, 55.9, 55.9, 28.1, 27.8, 24.7, 20.2 ppm. MS (ESI) m/z [M + Na^{+}] calcd for $C_{23}H_{24}N_2O_5$ 431.16 (100%), 432.16 (24.9%), 433.17 (2.7%), found 431.30 (100%), 432.30 (26.0%), 433.26 (2.4%); purity (HPLC) = 99%.

General procedure for the deprotection of Boc-groups by HCl

The Boc-protected amide was placed in a round-bottom flask and dissolved in HCl (4 M in dioxane). After stirring for 30 min at room temperature, crystallization was induced by the addition of pentane and the product was obtained by filtration.

(*S*)-2-Amino-*N*-(5,6-dimethoxybenzo[*d*] thiazol-2-yl)-3-(4-benzoxyphenyl)propionamide hydrochloride (10a): 387 mg, 0.77 mmol, yield: 87%. mp: 215–219 °C. $[\alpha]_D^{20}$ = +46 (c = 10 in DMSO). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.58 (s, 1H), 7.47–7.25 (m, 6H), 7.26–7.11 (m, 2H), 7.01–6.86 (m, 2H), 5.07 (s, 1H), 5.03 (s, 1H), 4.43–4.33 (m, 1H), 3.85–3.76 (m, 6H), 3.24–3.05 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 170.3, 157.6, 155.4, 149.1, 147.3, 142.3, 137.1, 130.7, 128.4, 127.8, 127.7, 127.0, 123.0, 114.8, 103.8, 103.7, 69.2, 56.0, 55.8, 53.3, 35.8 ppm.

(*S*)-Benzyl 3-amino-4-((5,6-dimethoxybenzo[*d*] thiazol-2-yl) amino)-4-oxobutanoate hydrochloride (10b): 414 mg, 0.92 mmol, yield: 95%. mp: 178–180 °C. $[\alpha]_D^{20}$ = +35 (c = 10 in H_2O). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.58 (s, 1H), 7.38–7.23 (m, 6H), 5.13 (s, 2H), 4.55–4.42 (m, 1H), 3.83 (s, 3H), 3.81 (s, 3H), 3.20 (d, J = 6.2 Hz, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO–d_6) δ = 168.8, 167.0, 156.1, 149.1, 147.3, 141.9, 135.6, 128.4, 128.2, 128.1, 122.9, 103.8, 103.4, 66.4, 56.0, 55.8, 49.2, 34.8 ppm.

3-Amino-*N*-(5,6-dimethoxybenzo[*d*] thiazol-2-yl)propionamide hydrochloride (15): 291 mg, 0.92 mmol, yield: 70%. mp: 225–235 °C. ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.54 (s, 1H), 7.29 (s, 1H), 3.81 (s, 3H), 3.79 (s, 3H), 3.18–3.03 (m, 2H), 2.98–2.86 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 169.1, 156.1, 149.0, 147.0, 142.5, 122.8, 103.8, 103.6, 56.0, 55.8, 34.3, 32.5 ppm.

(*S*)-*N*-Phenylpyrrolidine-2-carboxamide hydrochloride (31a): 310 mg, 1.38 mmol, yield: 99%. mp: 237–239 °C. $[\alpha]_D^{25}$ = –36 (S, c = 5 in DMSO). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.02 (s, 1H), 10.19 (s, 1H), 8.66 (s, 1H), 7.65 (d, J = 7.7 Hz, 2H), 7.34 (t, J = 7.7 Hz, 2H), 7.10 (t, J = 7.7 Hz, 1H), 4.47–4.35 (m, 1H), 3.33–3.19 (m, 2H), 2.47–2.36 (m, 1H), 2.04–1.84 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO–d_6) δ = 166.8, 138.3, 128.9, 124.1, 119.5, 59.5, 45.6, 29.8, 23.7 ppm.

(*S*)-*N*-Phenylpiperidine-2-carboxamide hydrochloride (31b): 355 mg, 1.47 mmol, yield: 90%. decomp. > 270 °C. $[\alpha]_D^{20}$ = +3 (c = 3 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.09 (s, 1H), 9.64 (d, J = 9.4 Hz, 1H), 8.83 (d, J = 10.9 Hz, 1H), 7.74–7.56 (m, 2H), 7.42–7.24 (m, 2H), 7.16–6.96 (m, 1H), 4.11–3.89 (m, 1H), 3.31–3.17 (m, 1H), 2.97–2.82 (m, 1H), 2.34–2.22 (m, 1H), 1.86–1.42 (m, 5H) ppm. ^{13}C NMR: (75 MHz, DMSO–d_6) δ = 167.3, 138.4, 128.8, 123.9, 119.4, 57.6, 43.3, 27.1, 21.7, 21.2 ppm.

4.7. General procedure for the formation of amides by acid chlorides

The respective amine or the corresponding hydrochloride (1.0 eq.) was dissolved in dichloromethane and triethylamine (1.1 eq., in the case of the hydrochlorides 2.1 eq.) was added, followed by the addition of the acyl chloride (1.1 eq.) while cooling the reaction mixture with an ice-water bath. The cooling bath was removed, and the mixture was stirred for 16 h. After removal of the solvent under reduced pressure, the residue was purified by silica column chromatography.

(*S*)-*N*-(3-(4-(benzyloxy)phenyl)-1-((5,6-dimethoxybenzo[*d*] thiazol-2-yl)amino)-1-oxopropan-2-yl)benzamide (11a): 241 mg, 0.42 mmol, yield: 71%. mp: 182–184 °C. R_f = 0.38 (CH/EtOAc 1:1). $[\alpha]_D^{20}$ = –18 (c = 5 in DCM). ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 7.85–7.77 (m, 2H), 7.47–7.24 (m, 10*H*), 7.01 (d, J = 8.6 Hz, 2H), 6.79 (d, J = 8.6 Hz, 2H), 5.37 (t, J = 6.9 Hz, 1H), 4.95 (d, J = 1.2 Hz, 2H), 3.99 (s, 3H), 3.89 (s, 3H), 3.28 (d, J = 6.9 Hz, 2H) ppm. ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 170.6, 167.8, 158.1, 157.4, 149.8, 148.0, 141.6, 137.0, 133.4, 132.2, 130.4, 128.7, 128.7, 128.1, 127.8, 127.6, 127.4, 123.0, 115.2, 103.3, 102.9, 70.0, 56.5, 56.3, 55.4, 37.5 ppm.

(*S*)-Benzyl 3-benzamido-4-((5,6-dimethoxybenzo[*d*] thiazol-2-yl)amino)-4-oxobutanoat (11b): 286 mg, 0.51 mmol, yield: 76%. mp: 191–192 °C. R_f = 0.28 (CH/EtOAc 1:1). $[\alpha]_D^{20}$ = –6 (c = 10 in DCM). ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 7.81 (d, J = 7.2 Hz, 2H), 7.56–7.48 (m, 1H), 7.47–7.38 (m, 2H), 7.37–7.29 (m, 5H), 7.27 (s, 1H), 7.20 (s, 1H), 5.40–5.29 (m, 1H), 5.19 (s, 2H), 3.94 (s, 3H), 3.91 (s, 3H), 3.27 (dd, J = 17.0, 4.5 Hz, 1H), 2.95 (dd, J = 17.0, 5.8 Hz, 1H) ppm. ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 171.9, 168.7, 168.0, 156.1, 149.6, 147.8, 142.2, 135.2, 132.6, 128.9, 128.8, 128.7, 128.5, 128.3, 127.5, 123.7, 103.5, 102.7, 70.0, 67.5, 56.5, 56.2, 50.2, 35.1 ppm.

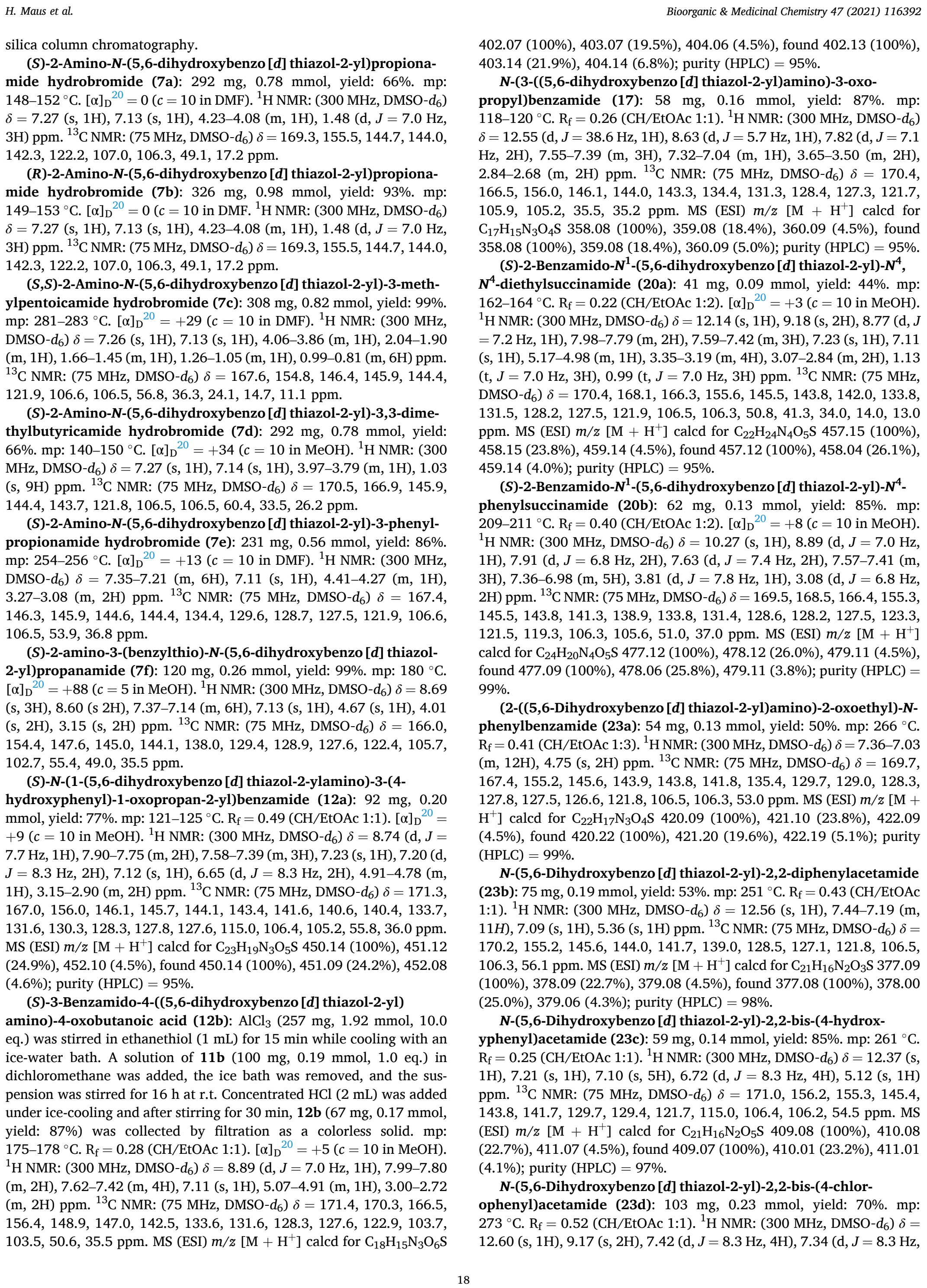
***N*-(3-((5,6-dimethoxybenzo[*d*] thiazol-2-yl)amino)-3-oxopropyl)benzamide (16)**: 120 mg, 0.31 mmol, yield: 40%. mp: 124–127 °C. R_f = 0.28 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.88–7.80 (m, 2H), 7.53 (s, 1H), 7.52–7.39 (m, 3H), 7.28 (s, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.59 (dt, J = 6.6, 5.5 Hz, 2H), 2.79 (t, J = 6.6 Hz, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 170.5, 166.8, 156.8, 149.3, 147.3, 143.0, 134.8, 131.6, 128.7, 127.6, 123.3, 104.1, 103.9, 56.4, 56.2, 35.9, 35.6 ppm.

***N*-(5,6-Dimethoxybenzo[*d*] thiazol-2-yl)-benzamide (24a)**: 326 mg, 1.04 mmol, yield: 72%. mp: 190–195 °C, R_f = 0.39 (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, $CDCl_3$) δ = 8.13–8.01 (m, 2H), 7.60–7.48 (m, 1H), 7.45–7.34 (m, 2H), 7.24 (s, 1H), 6.57 (s, 1H), 3.94 (s, 3H), 3.51 (s, 3H) ppm. ^{13}C NMR: (75 MHz, $CDCl_3$) δ = 165.8, 159.2, 149.1, 147.6, 140.7, 133.1, 131.9, 129.1, 128.4, 122.9, 102.5, 102.5, 56.3, 55.7 ppm.

(*S*)-3-benzamido-4-((5,6-dimethoxybenzo[*d*] thiazol-2-yl) amino)-4-oxobutanoic acid (18): **11b** (2.19 g, 4.22 mmol, 1 eq.) was dissolved in THF and LiOH (1.01 g, 41.29 mmol, 10 eq.), dissolved in water, was added. After stirring for 1 h, the organic solvent was removed under reduced pressure. The aqueous phase was acidified with HCl and extracted with ethyl acetate three times. The combined organic phases were washed with water and brine, dried with Mg_2SO_4 and the solvent was evaporated to yield the **18** (1.77 g, 4.12 mmol, yield: 98%). mp: 210–211 °C. R_f = 0.34 (CH/EtOAc 1:2). $[\alpha]_D^{20}$ = +39 (c = 10 in DMF). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.99–7.84 (m, 2H), 7.64–7.41 (m, 4H), 7.29 (s, 1H), 5.14–4.89 (m, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.01–2.75 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 171.8, 170.7, 166.9, 156.8, 149.4, 147.4, 143.0, 134.1, 132.0, 128.7, 128.0, 123.4, 104.1, 104.0, 56.4, 56.2, 51.1, 35.9 ppm.

4.8. General procedure for the deprotection of methoxy groups by BBr_3

At –78 °C, the respective aryl ether (1.0 eq.) was dissolved in DCM and BBr_3 (1 M in DCM, 8.0 eq.) was added dropwise. The cooling bath was removed after 30 min and stirring was continued for 16 h at r.t. While cooling with an ice-water bath, DCM (20 mL) was added, followed by 2 mL of MeOH. The product was obtained by filtration or the solvent was evaporated under reduced pressure and the residue was purified by



silica column chromatography.

(*S*)-2-Amino-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)propionamide hydrobromide (7a): 292 mg, 0.78 mmol, yield: 66%. mp: 148–152 °C. $[\alpha]_D^{20} = 0$ ($c = 10$ in DMF). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.27$ (s, 1H), 7.13 (s, 1H), 4.23–4.08 (m, 1H), 1.48 (d, $J = 7.0$ Hz, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.3$, 155.5, 144.7, 144.0, 142.3, 122.2, 107.0, 106.3, 49.1, 17.2 ppm.

(*R*)-2-Amino-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)propionamide hydrobromide (7b): 326 mg, 0.98 mmol, yield: 93%. mp: 149–153 °C. $[\alpha]_D^{20} = 0$ ($c = 10$ in DMF. ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.27$ (s, 1H), 7.13 (s, 1H), 4.23–4.08 (m, 1H), 1.48 (d, $J = 7.0$ Hz, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.3$, 155.5, 144.7, 144.0, 142.3, 122.2, 107.0, 106.3, 49.1, 17.2 ppm.

(*S*,*S*)-2-Amino-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-3-methylpentoicamide hydrobromide (7c): 308 mg, 0.82 mmol, yield: 99%. mp: 281–283 °C. $[\alpha]_D^{20} = +29$ ($c = 10$ in DMF). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.26$ (s, 1H), 7.13 (s, 1H), 4.06–3.86 (m, 1H), 2.04–1.90 (m, 1H), 1.66–1.45 (m, 1H), 1.26–1.05 (m, 1H), 0.99–0.81 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 167.6$, 154.8, 146.4, 145.9, 144.4, 121.9, 106.6, 106.5, 56.8, 36.3, 24.1, 14.7, 11.1 ppm.

(*S*)-2-Amino-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-3,3-dimethylbutyricamide hydrobromide (7d): 292 mg, 0.78 mmol, yield: 66%. mp: 140–150 °C. $[\alpha]_D^{20} = +34$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.27$ (s, 1H), 7.14 (s, 1H), 3.97–3.79 (m, 1H), 1.03 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.5$, 166.9, 145.9, 144.4, 143.7, 121.8, 106.5, 106.5, 60.4, 33.5, 26.2 ppm.

(*S*)-2-Amino-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-3-phenylpropionamide hydrobromide (7e): 231 mg, 0.56 mmol, yield: 86%. mp: 254–256 °C. $[\alpha]_D^{20} = +13$ ($c = 10$ in DMF). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.35$–7.21 (m, 6H), 7.11 (s, 1H), 4.41–4.27 (m, 1H), 3.27–3.08 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 167.4$, 146.3, 145.9, 144.6, 144.4, 134.4, 129.6, 128.7, 127.5, 121.9, 106.6, 106.5, 53.9, 36.8 ppm.

(*S*)-2-amino-3-(benzylthio)-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)propanamide (7f): 120 mg, 0.26 mmol, yield: 99%. mp: 180 °C. $[\alpha]_D^{20} = +88$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 8.69$ (s, 3H), 8.60 (s 2H), 7.37–7.14 (m, 6H), 7.13 (s, 1H), 4.67 (s, 1H), 4.01 (s, 2H), 3.15 (s, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 166.0$, 154.4, 147.6, 145.0, 144.1, 138.0, 129.4, 128.9, 127.6, 122.4, 105.7, 102.7, 55.4, 49.0, 35.5 ppm.

(*S*)-*N*-(1-(5,6-dihydroxybenzo [*d*] thiazol-2-ylamino)-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)benzamide (12a): 92 mg, 0.20 mmol, yield: 77%. mp: 121–125 °C. $R_f = 0.49$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +9$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 8.74$ (d, $J = 7.7$ Hz, 1H), 7.90–7.75 (m, 2H), 7.58–7.39 (m, 3H), 7.23 (s, 1H), 7.20 (d, $J = 8.3$ Hz, 2H), 7.12 (s, 1H), 6.65 (d, $J = 8.3$ Hz, 2H), 4.91–4.78 (m, 1H), 3.15–2.90 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 171.3$, 167.0, 156.0, 146.1, 145.7, 144.1, 143.4, 141.6, 140.6, 140.4, 133.7, 131.6, 130.3, 128.3, 127.8, 127.6, 115.0, 106.4, 105.2, 55.8, 36.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{23}H_{19}N_3O_5S$ 450.14 (100%), 451.12 (24.9%), 452.10 (4.5%), found 450.14 (100%), 451.09 (24.2%), 452.08 (4.6%); purity (HPLC) = 95%.

(*S*)-3-Benzamido-4-((5,6-dihydroxybenzo [*d*] thiazol-2-yl) amino)-4-oxobutanoic acid (12b): $AlCl_3$ (257 mg, 1.92 mmol, 10.0 eq.) was stirred in ethanethiol (1 mL) for 15 min while cooling with an ice-water bath. A solution of **11b** (100 mg, 0.19 mmol, 1.0 eq.) in dichloromethane was added, the ice bath was removed, and the suspension was stirred for 16 h at r.t. Concentrated HCl (2 mL) was added under ice-cooling and after stirring for 30 min, **12b** (67 mg, 0.17 mmol, yield: 87%) was collected by filtration as a colorless solid. mp: 175–178 °C. $R_f = 0.28$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +5$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 8.89$ (d, $J = 7.0$ Hz, 1H), 7.99–7.80 (m, 2H), 7.62–7.42 (m, 4H), 7.11 (s, 1H), 5.07–4.91 (m, 1H), 3.00–2.72 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 171.4$, 170.3, 166.5, 156.4, 148.9, 147.0, 142.5, 133.6, 131.6, 128.3, 127.6, 122.9, 103.7, 103.5, 50.6, 35.5 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{18}H_{15}N_3O_6S$ 402.07 (100%), 403.07 (19.5%), 404.06 (4.5%), found 402.13 (100%), 403.14 (21.9%), 404.14 (6.8%); purity (HPLC) = 95%.

***N*-(3-((5,6-dihydroxybenzo [*d*] thiazol-2-yl)amino)-3-oxopropyl)benzamide (17)**: 58 mg, 0.16 mmol, yield: 87%. mp: 118–120 °C. $R_f = 0.26$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.55$ (d, $J = 38.6$ Hz, 1H), 8.63 (d, $J = 5.7$ Hz, 1H), 7.82 (d, $J = 7.1$ Hz, 2H), 7.55–7.39 (m, 3H), 7.32–7.04 (m, 1H), 3.65–3.50 (m, 2H), 2.84–2.68 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.4$, 166.5, 156.0, 146.1, 144.0, 143.3, 134.4, 131.3, 128.4, 127.3, 121.7, 105.9, 105.2, 35.5, 35.2 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{17}H_{15}N_3O_4S$ 358.08 (100%), 359.08 (18.4%), 360.09 (4.5%), found 358.08 (100%), 359.08 (18.4%), 360.09 (5.0%); purity (HPLC) = 95%.

(*S*)-2-Benzamido-*N*^{1}-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-*N*^{4},*N*^{4}-diethylsuccinamide (20a): 41 mg, 0.09 mmol, yield: 44%. mp: 162–164 °C. $R_f = 0.22$ (CH/EtOAc 1:2). $[\alpha]_D^{20} = +3$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.14$ (s, 1H), 9.18 (s, 2H), 8.77 (d, $J = 7.2$ Hz, 1H), 7.98–7.79 (m, 2H), 7.59–7.42 (m, 3H), 7.23 (s, 1H), 7.11 (s, 1H), 5.17–4.98 (m, 1H), 3.35–3.19 (m, 4H), 3.07–2.84 (m, 2H), 1.13 (t, $J = 7.0$ Hz, 3H), 0.99 (t, $J = 7.0$ Hz, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.4$, 168.1, 166.3, 155.6, 145.5, 143.8, 142.0, 133.8, 131.5, 128.2, 127.5, 121.9, 106.5, 106.3, 50.8, 41.3, 34.0, 14.0, 13.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{22}H_{24}N_4O_5S$ 457.15 (100%), 458.15 (23.8%), 459.14 (4.5%), found 457.12 (100%), 458.04 (26.1%), 459.14 (4.0%); purity (HPLC) = 95%.

(*S*)-2-Benzamido-*N*^{1}-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-*N*^{4}-phenylsuccinamide (20b): 62 mg, 0.13 mmol, yield: 85%. mp: 209–211 °C. $R_f = 0.40$ (CH/EtOAc 1:2). $[\alpha]_D^{20} = +8$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 10.27$ (s, 1H), 8.89 (d, $J = 7.0$ Hz, 1H), 7.91 (d, $J = 6.8$ Hz, 2H), 7.63 (d, $J = 7.4$ Hz, 2H), 7.57–7.41 (m, 3H), 7.36–6.98 (m, 5H), 3.81 (d, $J = 7.8$ Hz, 1H), 3.08 (d, $J = 6.8$ Hz, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.5$, 168.5, 166.4, 155.3, 145.5, 143.8, 141.3, 138.9, 133.8, 131.4, 128.6, 128.2, 127.5, 123.3, 121.5, 119.3, 106.3, 105.6, 51.0, 37.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{24}H_{20}N_4O_5S$ 477.12 (100%), 478.12 (26.0%), 479.11 (4.5%), found 477.09 (100%), 478.06 (25.8%), 479.11 (3.8%); purity (HPLC) = 99%.

(2-((5,6-Dihydroxybenzo [*d*] thiazol-2-yl)amino)-2-oxoethyl)-*N*-phenylbenzamide (23a): 54 mg, 0.13 mmol, yield: 50%. mp: 266 °C. $R_f = 0.41$ (CH/EtOAc 1:3). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 7.36$–7.03 (m, 12H), 4.75 (s, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.7$, 167.4, 155.2, 145.6, 143.9, 143.8, 141.8, 135.4, 129.7, 129.0, 128.3, 127.8, 127.5, 126.6, 121.8, 106.5, 106.3, 53.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{22}H_{17}N_3O_4S$ 420.09 (100%), 421.10 (23.8%), 422.09 (4.5%), found 420.22 (100%), 421.20 (19.6%), 422.19 (5.1%); purity (HPLC) = 99%.

***N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2,2-diphenylacetamide (23b)**: 75 mg, 0.19 mmol, yield: 53%. mp: 251 °C. $R_f = 0.43$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.56$ (s, 1H), 7.44–7.19 (m, 11*H*), 7.09 (s, 1H), 5.36 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 170.2$, 155.2, 145.6, 144.0, 141.7, 139.0, 128.5, 127.1, 121.8, 106.5, 106.3, 56.1 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{21}H_{16}N_2O_3S$ 377.09 (100%), 378.09 (22.7%), 379.08 (4.5%), found 377.08 (100%), 378.00 (25.0%), 379.06 (4.3%); purity (HPLC) = 98%.

***N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2,2-bis-(4-hydroxyphenyl)acetamide (23c)**: 59 mg, 0.14 mmol, yield: 85%. mp: 261 °C. $R_f = 0.25$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.37$ (s, 1H), 7.21 (s, 1H), 7.10 (s, 5H), 6.72 (d, $J = 8.3$ Hz, 4H), 5.12 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 171.0$, 156.2, 155.3, 145.4, 143.8, 141.7, 129.7, 129.4, 121.7, 115.0, 106.4, 106.2, 54.5 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{21}H_{16}N_2O_5S$ 409.08 (100%), 410.08 (22.7%), 411.07 (4.5%), found 409.07 (100%), 410.01 (23.2%), 411.01 (4.1%); purity (HPLC) = 97%.

***N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2,2-bis-(4-chlorophenyl)acetamide (23d)**: 103 mg, 0.23 mmol, yield: 70%. mp: 273 °C. $R_f = 0.52$ (CH/EtOAc 1:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 12.60$ (s, 1H), 9.17 (s, 2H), 7.42 (d, $J = 8.3$ Hz, 4H), 7.34 (d, $J = 8.3$ Hz, 5H), 7.22 (s, 1H), 7.09 (s, 1H), 5.36 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 169.5$, 155.1, 145.6, 144.0, 137.5, 132.0, 130.3, 128.6, 121.8, 106.5, 106.3, 54.6 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{22}H_{14}Cl_2N_2O_3S$ 445.01 (100%), 446.01 (22.7%), 447.01 (63.9%), found 444.98 (100%), 445.98 (27.6%), 446.96 (75.2%); purity (HPLC) = 95%.

***N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)benzamide (25a)**: 105 mg, 0.37 mmol, yield: 99%. mp: 253–260 °C, $R_f = 0.38$ (DCM/MeOH 9:1). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 11.73$ (s, 1H), 8.32 (s, 2H), 8.20–8.04 (m, 2H), 7.68–7.59 (m, 1H), 7.59–7.45 (m, 2H), 7.27 (s, 1H), 7.16 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 165.4$, 156.3, 145.7, 144.1, 141.7, 132.6, 132.3, 128.6, 128.2, 122.0, 106.5, 106.4 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_{10}N_2O_3S$ 287.04 (100%), 288.04 (15.1%), 289.04 (4.5%), found 287.11 (100%), 288.10 (16.3%), 289.09 (4.2%); purity (HPLC) = 99%.

***N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-4-iodobenzamide (25b)**: 40 mg, 0.01 mmol, yield: 29%. mp: 228 °C. $R_f = 0.31$ (CH/EtOAc 1:3). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 8.20$–8.07 (m, 2H), 7.70–7.56 (m, 2H), 7.26 (s, 1H), 7.14 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 168.0$, 145.7, 144.1, 137.4, 131.2, 130.1, 128.7, 121.7, 117.7, 106.4, 105.8 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_9IN_2O_3S$ 412.94 (100%), 413.94 (15.1%), 414.93 (4.5%), found 413.03 (100%), 414.01 (20.0%), 415.00 (3.0%); purity (HPLC) = 99%.

4-Chloro-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)benzamide (25c): 29 mg, 0.09 mmol, yield: 32%. mp: 293 °C. $R_f = 0.36$ (DCM/MeOH 49:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 7.87$ (d, $J = 8.2$ Hz, 2H), 7.37 (d, $J = 8.2$ Hz, 2H), 7.01 (s, 1H). 6.89 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 164.7$, 158.1, 145.7, 144.1, 137.5, 131.2, 130.1, 128.7, 121.7, 106.4, 106.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_9ClN_2O_3S$ 321.00 (100%), 322.01 (15.1%), 323.00 (32.0%), found 321.04 (100%), 322.03 (15.9%), 323.02 (35.9%); purity (HPLC) = 98%.

4-Methyl-*N*-(5,6-dihydroxybenzothiazol-2-yl)benzamide (25d): 135 mg, 0.45 mmol, yield: 99%. mp: 331 °C. $R_f = 0.57$ (DCM/MeOH 19:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 8.00$ (d, $J = 8.0$ Hz, 2H), 7.34 (d, $J = 8.0$ Hz, 2H), 7.26 (s, 1H), 7.14 (s, 1H), 2.38 (s, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 165.3$, 156.6, 145.7, 144.1, 143.0, 129.2, 128.3, 121.8, 106.4, 106.2, 21.1 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{15}H_{12}N_2O_3S$ 301.06 (100%), 302.06 (16.2%), 303.05 (4.5%), found 301.06 (100%), 302.06 (10.0%), 303.05 (2.8%); purity (HPLC) = 99%.

***N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)- [1,1′-biphenyl] -4-carboxamide (25e)**: 163 mg, 0.45 mmol, yield: 94%. mp: 273 °C. $R_f = 0.36$ (CH/EtOAc 1:2). ^{1}H NMR: (600 MHz, DMSO-d_6) $\delta = 8.21$ (d, $J = 8.4$ Hz, 2H), 7.86 (d, $J = 8.4$ Hz, 2H), 7.78 (d, $J = 7.3$ Hz, 2H), 7.51 (t, $J = 7.7$ Hz, 2H), 7.43 (t, $J = 7.3$ Hz, 1H), 7.27 (s, 1H), 7.14 (s, 1H) ppm. ^{13}C NMR: (151 MHz, DMSO-d_6) $\delta = 158.4$, 158.2, 145.7, 144.0, 144.0, 138.9, 129.1, 128.9, 128.4, 127.0, 126.8, 106.4 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{20}H_{14}N_2O_3S$ 363.07 (100%), 364.08 (21.6%), 365.07 (4.5%), found 363.05 (100%), 364.03 (20.7%), 365.03 (4.2%); purity (HPLC) = 99%.

3,4-Dichloro-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)benzamide (25f): 36 mg, 0.10 mmol, yield: 52%. mp: 272 °C. $R_f = 0.53$ (DCM/MeOH 19:1). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 8.41$–8.31 (m, 1H), 8.05 (d, $J = 8.4$ Hz, 1H), 7.81 (d, $J = 8.6$ Hz, 1H), 7.36–7.24 (m, 1H), 7.21–7.09 (m, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 164.0$, 163.4, 156.3, 146.3, 145.8, 144.4, 144.2, 143.5, 141.8, 135.6, 135.3, 133.1, 132.3, 131.5, 130.9, 130.2, 128.4, 122.0, 121.4, 106.6, 105.7, 105.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_8Cl_2N_2O_3S$ 354.96 (100%), 355.97 (15.1%), 356.96 (63.9%), found 355.01 (100%), 356.02 (15.6%), 356.96 (69.9%); purity (HPLC) = 98%.

5,6-Dihydroxy-*N*-phenylbenzo [*d*] thiazole-2-carboxamide (34a): 90 mg, 0.31 mmol, yield: 97%. mp: 297–299 °C. $R_f = 0.47$ (CH/EtOAc 1:2 + 0.1% TFA). ^{1}H NMR: (300 MHz, MeOD-d_4): $\delta = 7.76$ (d, $J = 7.9$ Hz, 2H), 7.51 (s, 1H), 7.41–7.33 (m, 3H), 7.16 (t, $J = 7.9$ Hz, 1H) ppm. ^{13}C NMR: (75 MHz, MeOD-d_4): $\delta = 161.5$, 160.2, 149.0, 148.4, 148.2, 138.8, 131.3, 130.6, 129.9, 125.9, 121.7, 109.3, 106.7 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_{10}N_2O_3S$ 287.04 (100%), 288.04 (15.1%), 289.04 (4.5%), found 287.11 (100%), 288.08 (17.3%), 289.07 (4.0%); purity (HPLC) = 99%.

6-Hydroxy-*N*-phenylbenzo [*d*] thiazole-2-carboxamide (34b): 20 mg, 0.07 mmol, yield: 20%. decomp. > 225 °C. $R_f = 0.62$ (DCM/MeOH 19:1 + 0.1% TFA). ^{1}H NMR: (600 MHz, DMSO-d_6): $\delta = 10.96$ (s, 1H), 10.24 (s, 1H), 8.01 (d, $J = 8.9$ Hz, 1H), 7.89 (d, $J = 7.6$ Hz, 2H), 7.51 (d, $J = 2.4$ Hz, 1H), 7.41–7.35 (m, 2H), 7.17–7.09 (m, 2H). ^{13}C NMR: (151 MHz, DMSO-d_6): $\delta = 160.7$, 158.4, 157.2, 146.2, 138.3, 138.1, 128.8, 125.0, 124.4, 120.7, 117.4, 107.0 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{14}H_{10}N_2O_2S$ 271.05 (100%), 272.05 (15.1%), 273.05 (4.5%), found 271.12 (100%), 272.10 (17.3%), 273.09 (3.8%); purity (HPLC) = 97%.

5,6-Dihydroxy-*N*-phenyl-1*H*-indole-2-carboxamide (34c): 40 mg, 0.15 mmol, yield: 44%. mp: 207–209 °C. $R_f = 0.18$ (DCM/MeOH 19:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 11.11$ (s, 1H), 9.91 (s, 1H), 9.05 (s, 1H), 8.50 (s, 1H), 7.82–7.70 (m, 2H), 7.41–7.28 (m, 2H), 7.17 (s, 1H), 7.11–7.00 (m, 1H), 6.90 (s, 1H), 6.83 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 159.9$, 145.6, 141.9, 139.3, 132.2, 129.1, 128.6, 123.1, 120.0, 119.9, 104.9, 103.7, 97.1 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{15}H_{12}N_2O_3$ 269.08 (100%), 270.09 (16.2%), 271.09 (1.2%), found 269.15 (100%), 270.13 (15.8%), 271.10 (0.8%); purity (HPLC) = 96%.

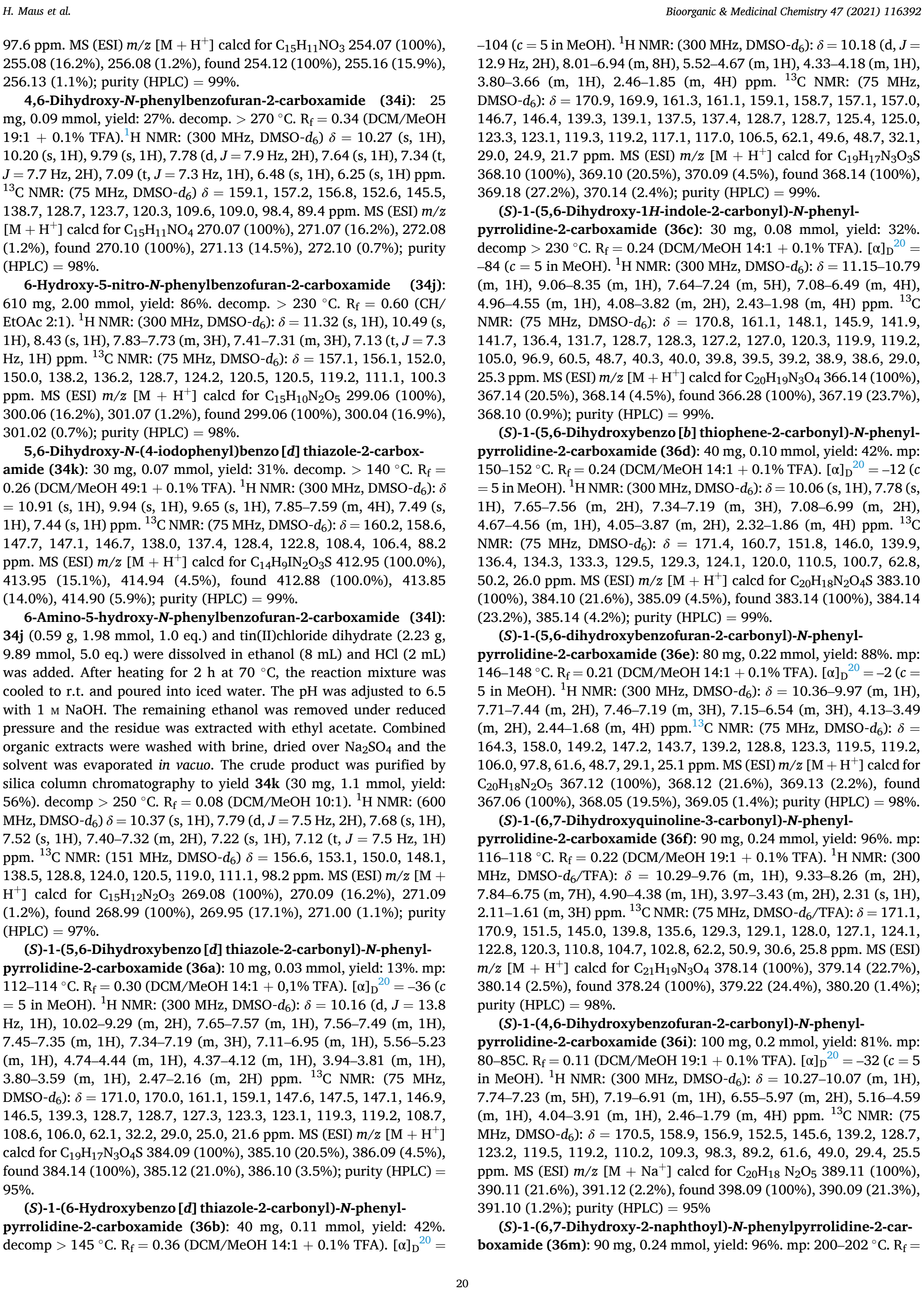
5,6-Dihydroxy-*N*-phenylbenzo [*b*] thiophene-2-carboxamide (34d): 90 mg, 0.32 mmol, yield: 99%. decomp. > 215 °C. $R_f = 0.24$ (DCM/MeOH 14:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 10.25$ (s, 1H), 9.69 (s, 1H), 9.33 (s, 1H), 8.08 (s, 1H), 7.74 (d, $J = 7.7$ Hz, 2H), 7.35 (t, $J = 7.7$ Hz, 2H), 7.30–7.22 (m, 2H), 7.09 (t, $J = 7.7$ Hz, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 160.7$, 147.5, 145.4, 138.9, 136.2, 133.2, 132.2, 128.7, 125.5, 123.7, 120.2, 109.5, 107.2 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{15}H_{11}NO_3S$ 286.05 (100%), 287.05 (16.2%), 288.04 (4.5%), found 286.00 (100%), 286.97 (18.8%), 288.00 (4.3%); purity (HPLC) = 99%.

5,6-Dihydroxy-*N*-phenylbenzofuran-2-carboxamide (34e): 80 mg, 0.30 mmol, yield: 88%. mp: 180–182 °C. $R_f = 0.18$ (DCM/MeOH 14:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6): $\delta = 10.27$–10.14 (m, 1H), 7.86–7.75 (m, 2H), 7.59–7.50 (m, 1H), 7.34 (t, $J = 7.9$ Hz, 2H), 7.16–6.92 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): $\delta = 156.9$, 149.5, 147.4, 147.1, 143.9, 138.7, 128.7, 123.8, 120.4, 118.7, 111.2, 106.1, 97.7 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{15}H_{11}NO_4$ 270.07 (100%), 271.07 (16.2%), 272.08 (1.2%), found 270.14 (100%), 271.12 (17.5%), 272.10 (0.8%); purity (HPLC) = 99%.

6,7-Dihydroxy-*N*-phenylquinoline-3-carboxamide (34f): 30 mg, 0.10 mmol, yield: 29%. mp: 224–226 °C. $R_f = 0.24$ (DCM/MeOH 14:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6/TFA): $\delta = 9.32$ (s, 2H), 8.11–6.74 (m, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6/TFA): $\delta = 161.7$, 157.6, 151.0, 141.6, 139.3, 138.8, 136.2, 129.0, 125.7, 124.6, 124.4, 120.8, 110.5, 102.5 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{16}H_{12}N_2O_3$ 281.08 (100%), 282.09 (17.3%), 283.09 (1.4%), found 281.15 (100%), 282.14 (18.4%), 283.10 (0.9%); purity (HPLC) = 98%.

5-Hydroxy-*N*-phenylbenzofuran-2-carboxamide (34g): 40 mg, 0.16 mmol, yield: 42%. mp: 229–231 °C. $R_f = 0.26$ (CH/EtOAc 4:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 10.43$ (s, 1H), 9.45 (s, 1H), 7.86–7.76 (m, 2H), 7.64–7.58 (m, 2H), 7.51 (d, $J = 8.9$ Hz, 1H), 7.42–7.30 (m, 2H), 7.18–7.11 (m, 1H), 7.08 (d, $J = 2.5$ Hz, 1H), 6.96 (dd, $J = 8.9$, 2.5 Hz, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 156.8$, 154.0, 149.2, 148.8, 138.4, 128.7, 128.0, 124.0, 120.5, 116.5, 112.3, 110.6, 106.4 ppm. MS (ESI) m/z $[M + H^+]$ calcd for $C_{15}H_{11}NO_3$ 254.07 (100%), 255.08 (16.2%), 256.08 (1.2%), found 254.13 (100%), 255.10 (16.0%), 256.09 (1.0%); purity (HPLC) = 99%.

6-Hydroxy-*N*-phenylbenzofuran-2-carboxamide (34h): 40 mg, 0.16 mmol, yield: 42%. mp: 237–239 °C. $R_f = 0.3$ (CH/EtOAc 2:1 + 0.1% TFA) ^{1}H NMR: (300 MHz, DMSO-d_6) $\delta = 10.31$ (s, 1H), 9.99 (s, 1H), 7.80 (d, $J = 7.7$ Hz, 2H), 7.64 (s, 1H), 7.59 (d, $J = 8.5$ Hz, 1H), 7.35 (t, $J = 7.9$ Hz, 2H), 7.11 (t, $J = 7.4$ Hz, 1H), 7.02 (s, 1H), 6.86 (dd, $J = 8.5$, 2.1 Hz, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) $\delta = 157.9$, 156.8, 156.0, 147.4, 138.5, 128.7, 123.9, 123.2, 120.4, 119.2, 113.9, 111.1,



97.6 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{15}H_{11}NO_3$ 254.07 (100%), 255.08 (16.2%), 256.08 (1.2%), found 254.12 (100%), 255.16 (15.9%), 256.13 (1.1%); purity (HPLC) = 99%.

4,6-Dihydroxy-*N*-phenylbenzofuran-2-carboxamide (34i): 25 mg, 0.09 mmol, yield: 27%. decomp. > 270 °C. R_f = 0.34 (DCM/MeOH 19:1 + 0.1% TFA).[1]H NMR: (300 MHz, DMSO-d_6) δ = 10.27 (s, 1H), 10.20 (s, 1H), 9.79 (s, 1H), 7.78 (d, *J* = 7.9 Hz, 2H), 7.64 (s, 1H), 7.34 (t, *J* = 7.7 Hz, 2H), 7.09 (t, *J* = 7.3 Hz, 1H), 6.48 (s, 1H), 6.25 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 159.1, 157.2, 156.8, 152.6, 145.5, 138.7, 128.7, 123.7, 120.3, 109.6, 109.0, 98.4, 89.4 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{15}H_{11}NO_4$ 270.07 (100%), 271.07 (16.2%), 272.08 (1.2%), found 270.10 (100%), 271.13 (14.5%), 272.10 (0.7%); purity (HPLC) = 98%.

6-Hydroxy-5-nitro-*N*-phenylbenzofuran-2-carboxamide (34j): 610 mg, 2.00 mmol, yield: 86%. decomp. > 230 °C. R_f = 0.60 (CH/EtOAc 2:1). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 11.32 (s, 1H), 10.49 (s, 1H), 8.43 (s, 1H), 7.83–7.73 (m, 3H), 7.41–7.31 (m, 3H), 7.13 (t, *J* = 7.3 Hz, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 157.1, 156.1, 152.0, 150.0, 138.2, 136.2, 128.7, 124.2, 120.5, 120.5, 119.2, 111.1, 100.3 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{15}H_{10}N_2O_5$ 299.06 (100%), 300.06 (16.2%), 301.07 (1.2%), found 299.06 (100%), 300.04 (16.9%), 301.02 (0.7%); purity (HPLC) = 98%.

5,6-Dihydroxy-*N*-(4-iodophenyl)benzo[*d*] thiazole-2-carboxamide (34k): 30 mg, 0.07 mmol, yield: 31%. decomp. > 140 °C. R_f = 0.26 (DCM/MeOH 49:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.91 (s, 1H), 9.94 (s, 1H), 9.65 (s, 1H), 7.85–7.59 (m, 4H), 7.49 (s, 1H), 7.44 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 160.2, 158.6, 147.7, 147.1, 146.7, 138.0, 137.4, 128.4, 122.8, 108.4, 106.4, 88.2 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{14}H_9IN_2O_3S$ 412.95 (100.0%), 413.95 (15.1%), 414.94 (4.5%), found 412.88 (100.0%), 413.85 (14.0%), 414.90 (5.9%); purity (HPLC) = 99%.

6-Amino-5-hydroxy-*N*-phenylbenzofuran-2-carboxamide (34l): **34j** (0.59 g, 1.98 mmol, 1.0 eq.) and tin(II)chloride dihydrate (2.23 g, 9.89 mmol, 5.0 eq.) were dissolved in ethanol (8 mL) and HCl (2 mL) was added. After heating for 2 h at 70 °C, the reaction mixture was cooled to r.t. and poured into iced water. The pH was adjusted to 6.5 with 1 M NaOH. The remaining ethanol was removed under reduced pressure and the residue was extracted with ethyl acetate. Combined organic extracts were washed with brine, dried over Na_2SO_4 and the solvent was evaporated *in vacuo*. The crude product was purified by silica column chromatography to yield **34k** (30 mg, 1.1 mmol, yield: 56%). decomp > 250 °C. R_f = 0.08 (DCM/MeOH 10:1). ^{1}H NMR: (600 MHz, DMSO-d_6) δ = 10.37 (s, 1H), 7.79 (d, *J* = 7.5 Hz, 2H), 7.68 (s, 1H), 7.52 (s, 1H), 7.40–7.32 (m, 2H), 7.22 (s, 1H), 7.12 (t, *J* = 7.5 Hz, 1H) ppm. ^{13}C NMR: (151 MHz, DMSO-d_6) δ = 156.6, 153.1, 150.0, 148.1, 138.5, 128.8, 124.0, 120.5, 119.0, 111.1, 98.2 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{15}H_{12}N_2O_3$ 269.08 (100%), 270.09 (16.2%), 271.09 (1.2%), found 268.99 (100%), 269.95 (17.1%), 271.00 (1.1%); purity (HPLC) = 97%.

(*S*)-1-(5,6-Dihydroxybenzo[*d*] thiazole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36a): 10 mg, 0.03 mmol, yield: 13%. mp: 112–114 °C. R_f = 0.30 (DCM/MeOH 14:1 + 0,1% TFA). $[\alpha]_D^{20}$ = –36 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.16 (d, *J* = 13.8 Hz, 1H), 10.02–9.29 (m, 2H), 7.65–7.57 (m, 1H), 7.56–7.49 (m, 1H), 7.45–7.35 (m, 1H), 7.34–7.19 (m, 3H), 7.11–6.95 (m, 1H), 5.56–5.23 (m, 1H), 4.74–4.44 (m, 1H), 4.37–4.12 (m, 1H), 3.94–3.81 (m, 1H), 3.80–3.59 (m, 1H), 2.47–2.16 (m, 2H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 171.0, 170.0, 161.1, 159.1, 147.6, 147.5, 147.1, 146.9, 146.5, 139.3, 128.7, 128.7, 127.3, 123.3, 123.1, 119.3, 119.2, 108.7, 108.6, 106.0, 62.1, 32.2, 29.0, 25.0, 21.6 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{19}H_{17}N_3O_4S$ 384.09 (100%), 385.10 (20.5%), 386.09 (4.5%), found 384.14 (100%), 385.12 (21.0%), 386.10 (3.5%); purity (HPLC) = 95%.

(*S*)-1-(6-Hydroxybenzo[*d*] thiazole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36b): 40 mg, 0.11 mmol, yield: 42%. decomp > 145 °C. R_f = 0.36 (DCM/MeOH 14:1 + 0.1% TFA). $[\alpha]_D^{20}$ = –104 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.18 (d, *J* = 12.9 Hz, 2H), 8.01–6.94 (m, 8H), 5.52–4.67 (m, 1H), 4.33–4.18 (m, 1H), 3.80–3.66 (m, 1H), 2.46–1.85 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.9, 169.9, 161.3, 161.1, 159.1, 158.7, 157.1, 157.0, 146.7, 146.4, 139.3, 139.1, 137.5, 137.4, 128.7, 128.7, 125.4, 125.0, 123.3, 123.1, 119.3, 119.2, 117.1, 117.0, 106.5, 62.1, 49.6, 48.7, 32.1, 29.0, 24.9, 21.7 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{19}H_{17}N_3O_3S$ 368.10 (100%), 369.10 (20.5%), 370.09 (4.5%), found 368.14 (100%), 369.18 (27.2%), 370.14 (2.4%); purity (HPLC) = 99%.

(*S*)-1-(5,6-Dihydroxy-1*H*-indole-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36c): 30 mg, 0.08 mmol, yield: 32%. decomp > 230 °C. R_f = 0.24 (DCM/MeOH 14:1 + 0.1% TFA). $[\alpha]_D^{20}$ = –84 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 11.15–10.79 (m, 1H), 9.06–8.35 (m, 1H), 7.64–7.24 (m, 5H), 7.08–6.49 (m, 4H), 4.96–4.55 (m, 1H), 4.08–3.82 (m, 2H), 2.43–1.98 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.8, 161.1, 148.1, 145.9, 141.9, 141.7, 136.4, 131.7, 128.7, 128.3, 127.2, 127.0, 120.3, 119.9, 119.2, 105.0, 96.9, 60.5, 48.7, 40.3, 40.0, 39.8, 39.5, 39.2, 38.9, 38.6, 29.0, 25.3 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{19}N_3O_4$ 366.14 (100%), 367.14 (20.5%), 368.14 (4.5%), found 366.28 (100%), 367.19 (23.7%), 368.10 (0.9%); purity (HPLC) = 99%.

(*S*)-1-(5,6-Dihydroxybenzo[*b*] thiophene-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36d): 40 mg, 0.10 mmol, yield: 42%. mp: 150–152 °C. R_f = 0.24 (DCM/MeOH 14:1 + 0.1% TFA). $[\alpha]_D^{20}$ = –12 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.06 (s, 1H), 7.78 (s, 1H), 7.65–7.56 (m, 2H), 7.34–7.19 (m, 3H), 7.08–6.99 (m, 2H), 4.67–4.56 (m, 1H), 4.05–3.87 (m, 2H), 2.32–1.86 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 171.4, 160.7, 151.8, 146.0, 139.9, 136.4, 134.3, 133.3, 129.5, 129.3, 124.1, 120.0, 110.5, 100.7, 62.8, 50.2, 26.0 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{18}N_2O_4S$ 383.10 (100%), 384.10 (21.6%), 385.09 (4.5%), found 383.14 (100%), 384.14 (23.2%), 385.14 (4.2%); purity (HPLC) = 99%.

(*S*)-1-(5,6-dihydroxybenzofuran-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36e): 80 mg, 0.22 mmol, yield: 88%. mp: 146–148 °C. R_f = 0.21 (DCM/MeOH 14:1 + 0.1% TFA). $[\alpha]_D^{20}$ = –2 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.36–9.97 (m, 1H), 7.71–7.44 (m, 2H), 7.46–7.19 (m, 3H), 7.15–6.54 (m, 3H), 4.13–3.49 (m, 2H), 2.44–1.68 (m, 4H) ppm.[13]C NMR: (75 MHz, DMSO-d_6): δ = 164.3, 158.0, 149.2, 147.2, 143.7, 139.2, 128.8, 123.3, 119.5, 119.2, 106.0, 97.8, 61.6, 48.7, 29.1, 25.1 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{18}N_2O_5$ 367.12 (100%), 368.12 (21.6%), 369.13 (2.2%), found 367.06 (100%), 368.05 (19.5%), 369.05 (1.4%); purity (HPLC) = 98%.

(*S*)-1-(6,7-Dihydroxyquinoline-3-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36f): 90 mg, 0.24 mmol, yield: 96%. mp: 116–118 °C. R_f = 0.22 (DCM/MeOH 19:1 + 0.1% TFA). ^{1}H NMR: (300 MHz, DMSO-d_6/TFA): δ = 10.29–9.76 (m, 1H), 9.33–8.26 (m, 2H), 7.84–6.75 (m, 7H), 4.90–4.38 (m, 1H), 3.97–3.43 (m, 2H), 2.31 (s, 1H), 2.11–1.61 (m, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6/TFA): δ = 171.1, 170.9, 151.5, 145.0, 139.8, 135.6, 129.3, 129.1, 128.0, 127.1, 124.1, 122.8, 120.3, 110.8, 104.7, 102.8, 62.2, 50.9, 30.6, 25.8 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{21}H_{19}N_3O_4$ 378.14 (100%), 379.14 (22.7%), 380.14 (2.5%), found 378.24 (100%), 379.22 (24.4%), 380.20 (1.4%); purity (HPLC) = 98%.

(*S*)-1-(4,6-Dihydroxybenzofuran-2-carbonyl)-*N*-phenylpyrrolidine-2-carboxamide (36i): 100 mg, 0.2 mmol, yield: 81%. mp: 80–85C. R_f = 0.11 (DCM/MeOH 19:1 + 0.1% TFA). $[\alpha]_D^{20}$ = –32 (*c* = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.27–10.07 (m, 1H), 7.74–7.23 (m, 5H), 7.19–6.91 (m, 1H), 6.55–5.97 (m, 2H), 5.16–4.59 (m, 1H), 4.04–3.91 (m, 1H), 2.46–1.79 (m, 4H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.5, 158.9, 156.9, 152.5, 145.6, 139.2, 128.7, 123.2, 119.5, 119.2, 110.2, 109.3, 98.3, 89.2, 61.6, 49.0, 29.4, 25.5 ppm. MS (ESI) *m/z* [M + Na$^+$] calcd for $C_{20}H_{18}$ N_2O_5 389.11 (100%), 390.11 (21.6%), 391.12 (2.2%), found 398.09 (100%), 390.09 (21.3%), 391.10 (1.2%); purity (HPLC) = 95%

(*S*)-1-(6,7-Dihydroxy-2-naphthoyl)-*N*-phenylpyrrolidine-2-carboxamide (36m): 90 mg, 0.24 mmol, yield: 96%. mp: 200–202 °C. R_f =

0.50 (DCM/MeOH 19:1 + 0.1% TFA). $[\alpha]_D^{20} = -40$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 10.29–9.89 (m, 1H), 8.20–6.95 (m, 11*H*), 3.75–3.53 (m, 2H), 2.38–2.15 (m, 1H), 2.04–1.79 (m, 3H) ppm.^{13}C NMR: (75 MHz, DMSO-d_6): δ = 170.8, 169.0, 148.2, 147.5, 139.3, 129.5 128.8, 127.7, 125.4, 125,1, 123.2, 121.8, 119.3, 110.3, 109.4, 60.9, 50.2, 29.8, 25.2 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{22}H_{20}N_2O_4$ 377.14 (100%), 378.15 (23.8%), 379.15 (2.7%), found 377.13 (100%), 378.13 (20.8%), 379.12 (1.4%) and [M−$C_6H_6N^{+\cdot}$] 284.11; purity (HPLC) = 99%

(*S*)-1-(5,6-Dihydroxybenzofuran-2-carbonyl)-*N*-phenylpiperidine-2-carboxamide (36n): 60 mg, 0.16 mmol, yield: 64%. mp: 81–83 °C. $R_f = 0.13$ (DCM/MeOH 19:1 + 0.1% TFA). $[\alpha]_D^{20} = +94$ (c = 5 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6): δ = 9.99 (s, 1H), 9.46 (s, 1H), 9.06 (s, 1H), 7.65–7.54 (m, 2H), 7.32 (t, J = 7.6 Hz, 2H), 7.19 (s, 1H), 7.06 (t, J = 7.6 Hz, 1H), 6.99 (s, 1H), 6.92 (s, 1H), 5.18–5.09 (m, 1H), 4.41–4.26 (m, 1H), 2.27–2.15 (m, 1H), 1.84–1.64 (m, 3H), 1.58–1.40 (m, 2H), 1.28–1.10 (m, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6): δ = 169.9, 160.7, 148.8, 146.8, 143.7, 138.9, 128.7, 123.5, 119.7, 118.0, 111.9, 105.8, 97.7, 20.2 ppm. MS (ESI) *m/z* [M + Na$^+$] calcd for $C_{21}H_{20}N_2O_5$ 403.13 (100%), 404.13 (22.7%), 405.13 (2.5%), found 403.22 (100%), 404.22 (22.1%), 405.20 (1.9%); purity (HPLC) = 97%

4.9. *General procedure for the formation of sulfonamides*

The respective amine educt (1.0 eq.) was dissolved in pyridine and the corresponding sulfonyl chloride (1.0 eq.) was added portion-wise at room temperature. After stirring for 16 h, the solvent was removed, and the residue was purified by silica column chromatography.

(*R*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide (1b): 86 mg, 0.20 mmol, yield: 71%. mp: 205–209 °C. $R_f = 0.25$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +113$ ($c = 10$ in MeOH). ^{1}H NMR: (300 MHz, MeOD-d_4) δ = 8.69 (d, J = 8.1 Hz, 2H), 8.31 (d, J = 8.1 Hz, 2H), 8.18–8.03 (m, 2H), 5.37–5.19 (m, 1H), 4.99 (q, J = 7.1 Hz, 1H), 4.62–4.41 (m, 1H), 4.32–4.10 (m, 2H), 3.04–2.76 (m, 4H), 2.66–2.44 (m, 1H) ppm. ^{13}C NMR: (75 MHz, MeOD-d_4) δ = 172.5, 147.4, 145.8, 141.5, 134.9, 131.1, 128.9, 123.5, 107.0, 106.7, 63.1, 50.6, 32.0, 25.7, 21.5 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{19}H_{19}N_3O_5S_2$ 434.08 (100%), 435.08 (20.5%), 436.07 (9.0%), found 434.07 (100%), 435.03 (24.1%), 436.05 (12.7%); purity (HPLC) = 98%.

(*S*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2-(4-methylphenylsulfonamido)propionamide (8a): 35 mg, 0.09 mmol, yield: 29%. mp: 185–190 °C. $R_f = 0.45$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -32$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.97 (s, 1H), 9.21 (s, 2H), 8.17 (d, J = 8.2 Hz, 1H), 7.65 (d, J = 8.1 Hz, 2H), 7.27 (d, J = 8.1 Hz, 2H), 7.20 (s, 1H), 7.09 (s, 1H), 4.16–3.99 (m, 1H), 2.22 (s, 3H), 1.18 (d, J = 7.0 Hz, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 170.4, 158.1, 157.7, 145.6, 143.9, 142.7, 137.8, 129.4, 126.5, 121.9, 106.5, 106.3, 51.6, 20.8, 18.6 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{17}H_{17}N_3O_5S_2$ 408.06 (100%), 409.06 (18.4%), 410.06 (9.0%), found 408.07 (100%), 409.06 (22.7%), 410.06 (10.6%); purity (HPLC) = 99%.

(*R*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2-(4-methylphenylsulfonamido)propionamide (8b): 85 mg, 0.21 mmol, yield: 70%. mp: 196–199 °C. $R_f = 0.45$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +30$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 7.69–7.59 (m, 2H), 7.27 (d, J = 8.1 Hz, 2H), 7.20 (s, 1H), 7.08 (s, 1H), 4.07 (q, J = 7.0 Hz, 1H), 2.21 (s, 3H), 1.17 (d, J = 7.0 Hz, 3H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 172.1, 170.4, 155.0, 145.5, 143.9, 142.7, 141.7, 137.8, 129.4, 126.5, 121.9, 106.4, 106.3, 51.5, 21.1, 18.7 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{17}H_{17}N_3O_5S_2$ 408.06 (100%), 409.06 (18.4%), 410.06 (9.0%), found 408.15 (100%), 409.08 (20.6%), 410.14 (10.6%); purity (HPLC) = 99%.

(*S,S*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-3-methyl-2-(4-methylphenylsulfonylamido)pentanoicamid (8c): 47 mg, 0.10 mmol, yield: 39%. mp: 253–256 °C. $R_f = 0.35$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -30$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.93 (s, 1H), 9.14 (s, 2H), 8.05 (d, J = 9.2 Hz, 1H), 7.62 (d, J = 8.2 Hz, 2H), 7.21–7.12 (m, 3H), 7.08 (s, 1H), 3.92–3.72 (m, 1H), 2.07 (s, 3H), 1.73–1.62 (m, 1H), 1.56–1.39 (m, 1H), 1.15–0.97 (m, 1H), 0.83–0.63 (m, 6H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 169.6, 154.7, 145.5, 143.9, 142.5, 141.8, 137.7, 129.2, 126.6, 122.0, 106.5, 106.2, 60.1, 36.7, 24.2, 20.7, 14.9, 10.2 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{23}N_3O_5S_2$ 450.11 (100%), 451.11 (21.6%), 452.10 (9.0%), found 450.22 (100%), 451.10 (22.7%), 452.17 (12.0%); purity (HPLC) = 96%.

(*S*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-3,3-dimethyl-2-(4-methylphenylsulfonamido)butyricamide (8d): 80 mg, 0.18 mmol, yield: 34%. mp: 254–256 °C. $R_f = 0.43$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = -29$ (c = 10 in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 11.83 (s, 1H), 9.13 (s. 2H), 7.92 (s. 1H), 7.63 (d, J = 8.1 Hz, 2H), 7.17 (s, 1H), 7.11 (d, J = 8.1 Hz, 2H), 7.07 (s, 1H), 3.80 (s, 1H), 1.95 (s, 3H), 0.89 (s, 9H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 168.4, 165.6, 145.5, 143.8, 142.5, 141.8, 137.4, 129.0, 126.8, 121.9, 106.4, 106.2, 63.6, 34.3, 26.2, 20.6 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{20}H_{23}N_3O_5S_2$ 450.11 (100%), 451.11 (21.6%), 452.10 (9.0%), found 450.10 (100%), 451.07 (23.7%), 452.10 (11.5%); purity (HPLC) = 98%.

(*S*)-*N*-(5,6-Dihydroxybenzo [*d*] thiazol-2-yl)-2-(4-methylphenylsulfonamido)-3-phenylpropionamide (8e): 40 mg, 0.09 mmol, yield: 35%. mp: 209–213 °C. $R_f = 0.28$ (CH/EtOAc 1:1). $[\alpha]_D^{20} = +8$ ($c = 5$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 12.10 (s, 1H), 8.32 (d, J = 9.2 Hz, 1H), 7.41 (d, J = 8.2 Hz, 2H), 7.23–7.12 (m, 6H), 7.12–7.05 (m, 3H), 4.42–4.20 (m, 1H), 3.02–2.87 (m, 1H), 2.81–2.68 (m, 1H), 2.13 (s, 3H) ppm.^{13}C NMR: (75 MHz, DMSO-d_6) δ = 169.6, 154.9, 145.6, 144.0, 142.4, 141.8, 137.7, 136.5, 129.3, 129.2, 128.1, 126.6, 126.4, 122.0, 106.5, 106.3, 57.4, 38.1, 20.8 ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{23}H_{21}N_3O_5S_2$ 484.09 (100%), 485.10 (24.9%), 486.09 (9.0%), found 484.22 (100%), 485.10 (26.3%), 486.17 (13.2%); purity (HPLC) = 97%.

(*S*)-3-(Benzylthio)-*N*-(5,6-dihydroxybenzo [*d*] thiazol-2-yl)-2-(4-methylphenylsulfonamido)propionamide (8f): 30 mg, 0.05 mmol, yield: 32%. mp: 137 °C. $R_f = 0.26$ (CH/EtOAc 1:3). $[\alpha]_D^{20} = -96$ ($c = 1$ in MeOH). ^{1}H NMR: (300 MHz, DMSO-d_6) δ = 8.39 (d, J = 8.4 Hz, 1H), 7.68 (d, J = 7.7 Hz, 2H), 7.35–7.18 (m, 5H), 7.13 (dd, J = 7.7, 2.1 Hz, 4H), 4.32–4.19 (m, 1H), 4.08–3.92 (m, 1H), 3.68 (s, 2H), 3.19 (s, 1H), 2.44 (s, 3H), 2.17 (s, 1H) ppm. ^{13}C NMR: (75 MHz, DMSO-d_6) δ = 174.2, 168.6, 164.5, 154.7, 153.3, 145.5, 144.6, 144.0, 142.9, 142.7, 137.7, 129.3, 128.8, 128.3, 126.9, 126.6, 126.1, 121.9, 106.8, 106.3, 58.0, 55.2, 34.4, 20.8. ppm. MS (ESI) *m/z* [M + H$^+$] calcd for $C_{24}H_{23}N_3O_5S_3$ 530.08 (100%), 531.08 (26.0%), 532.08 (9.0%), found 530.08 (100%), 531.08 (29.6%), 532.08 (17.6%); purity (HPLC) = 99%.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Janina Metzler, Nabil Farooq, and Jan Kwasnicki for their assistance during the syntheses of the inhibitors.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmc.2021.116392.

References

1 Stanaway JD, Shepard DS, Undurraga EA, et al. The Global Burden of Dengue: An Analysis from the Global Burden of Disease Study 2013. *Lancet Infect. Dis.* 2016;16 (6):712–723. https://doi.org/10.1016/S1473-3099(16)00026-8.

2 Bhatt S, Gething PW, Brady OJ, et al. The Global Distribution and Burden of Dengue. *Nature.* 2013;496(7446):504–507. https://doi.org/10.1038/nature12060.

3 Coluccia A, Puxeddu M, Nalli M, et al. Discovery of Zika Virus NS2B/NS3 Inhibitors That Prevent Mice from Life-Threatening Infection and Brain Damage. *ACS Med. Chem. Lett.* 2020;11(10):1869–1874. https://doi.org/10.1021/acsmedchemlett.9b00405.

4 Lowe R, Barcellos C, Brasil P, et al. The Zika Virus Epidemic in Brazil: From Discovery to Future Implications. *Int. J. Environ. Res. Public Health.* 2018;15(1):96. https://doi.org/10.3390/ijerph15010096.

5 Kraemer MUG, Reiner RC, Brady OJ, et al. Past and Future Spread of the Arbovirus Vectors Aedes Aegypti and Aedes Albopictus. *Nat. Microbiol.* 2019;4(5):854–863. https://doi.org/10.1038/s41564-019-0376-y.

6 Bonizzoni M, Gasperi G, Chen X, James AA. The Invasive Mosquito Species Aedes Albopictus: Current Knowledge and Future Perspectives. *Trends Parasitol.* 2013;29 (9):460–468. https://doi.org/10.1016/j.pt.2013.07.003.

7 Rogers DJ, Suk JE, Semenza JC. Using Global Maps to Predict the Risk of Dengue in Europe. *Acta Trop.* 2014;129(1):1–14. https://doi.org/10.1016/j.actatropica.2013.08.008.

8 Gjenero-Margan, I.; Aleraj, B.; Krajcar, D.; Lesnikar, V.; Klobučar, A.; Pem-Novosel, I.; Kurečić-Filipović, S.; Komparak, S.; Martić, R.; Duričić, S.; Betica-Radić, L.; Okmadžić, J.; Vilibić-Čavlek, T.; Babić-Erceg, A.; Turković, B.; Avšić-Županc, T.; Radić, I.; Ljubić, M.; Šarac, K.; Benić, N.; Mlinarić-Galinović, G. Autochthonous Dengue Fever in Croatia, August- September 2010. Eurosurveillance 2011, 16 (9), 1–4. https://doi.org/10.2807/ese.16.09.19805-en.

9 La Ruche, G.; Souarès, Y.; Armengaud, A.; Peloux-Petiot, F.; Delaunay, P.; Desprès, P.; Lenglet, A.; Jourdain, F.; Leparc-Goffart, I.; Charlet, F.; Ollier, L.; Mantey, K.; Mollet, T.; Fournier, J. P.; Torrents, R.; Leitmeyer, K.; Hilairet, P.; Zeller, H.; Van Bortel, W.; Dejour-Salamanca, D.; Grandadam, M.; Gastellu-Etchegorry, M. First Two Autochthonous Dengue Virus Infections in Metropolitan France, September 2010. Eurosurveillance 2010, 15 (39), 1–5. https://doi.org/10.2807/ese.15.39.19676-en.

10 ECDC (European Centre for Disease Prevention and Control). Dengue Outbreak in Madeira, Portugal, October-November 2012.; 2014. https://doi.org/10.2900/75830.

11 Counotte MJ, Kim CR, Wang J, et al. Sexual Transmission of Zika Virus and Other Flaviviruses: A Living Systematic Review. *PLoS Med.* 2018;15(7):e1002611. https://doi.org/10.1371/journal.pmed.1002611.

12 Mlakar J, Korva M, Tul N, et al. Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* 2016;374(10):951–958. https://doi.org/10.1056/NEJMoa1600651.

13 Cao-Lormeau V, Blake A, Mons S, et al. Guillain-Barré Syndrome Outbreak Caused by ZIKA Virus Infection in French Polynesia. *Lancet.* 2016;387(10027):1531–1539. https://doi.org/10.1016/S0140-6736(16)00562-6.

14 Duffy MR, Chen T-H, Hancock WT, et al. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* 2009;360(24):2536–2543. https://doi.org/10.1056/NEJMoa0805715.

15 Aubry M, Teissier A, Huart M, et al. Zika Virus Seroprevalence, French Polynesia, 2014–2015 Maite. *Emerg. Infect. Dis.* 2017;23(4):669–672. https://doi.org/10.3201/eid2304.161549.

16 Kirkpatrick, B. D.; Whitehead, S. S.; Pierce, K. K.; Tibery, C. M.; Grier, P. L.; Hynes, N. A.; Larsson, C. J.; Sabundayo, B. P.; Talaat, K. R.; Janiak, A.; Carmilli, M. P.; Luke, C. J.; Diehl, S. A.; Durbin, A. P. The Live Attenuated Dengue Vaccine TV003 Elicits Complete Protection against Dengue in a Human Challenge Model. Sci. Transl. Med. 2016, 8 (330), 1–8. https://doi.org/10.1126/scitranslmed.aaf1517.

17 Yauch LE, Shresta S. Dengue Virus Vaccine Development. *Adv. Virus Res.* 2014;88: 315–372. https://doi.org/10.1016/B978-0-12-800098-4.00007-6.

18 Katzelnick LC, Gresh L, Halloran ME, et al. Antibody-Dependent Enhancement of Severe Dengue Disease in Humans. *Science.* 2017;358(6365):929–932. https://doi.org/10.1126/science.aan6836.

19 WHO. Dengue and severe dengue https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.

20 Mustafa MS, Rasotgi V, Jain S, Gupta V. Discovery of Fifth Serotype of Dengue Virus (DENV-5): A New Public Health Dilemma in Dengue Control. *Med. J. Armed Forces India.* 2015;71(1):67–70. https://doi.org/10.1016/j.mjafi.2014.09.011.

21 Normile, D. Surprising New Dengue Virus Throws A Spanner in Disease Control Efforts. Science 2013, 342 (6157), 415. https://doi.org/10.1126/science.342.6157.415.

22 Preugschat F, Yao CW, Strauss JH. In Vitro Processing of Dengue Virus Type 2 Nonstructural Proteins NS2A, NS2B, and NS3. *J. Virol.* 1990;64(9):4364–4374. https://doi.org/10.1128/jvi.64.9.4364-4374.1990.

23 Falgout B, Markoff L. Evidence That Flavivirus NS1-NS2A Cleavage Is Mediated by a Membrane-Bound Host Protease in the Endoplasmic Reticulum. *J. Virol.* 1995;69 (11):7232–7243. https://doi.org/10.1128/jvi.69.11.7232-7243.1995.

24 Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus Genome Organization, Expression and Replication. *Annu. Rev. Microbiol.* 1990;44(1):649–688. https://doi.org/10.1146/annurev.mi.44.100190.003245.

25 Stadler K, Allison SL, Schalich J, Heinz FX. Proteolytic Activation of Tick-Borne Encephalitis Virus by Furin. *J. Virol.* 1997;71(11):8475–8481. https://doi.org/10.1128/jvi.71.11.8475-8481.1997.

26 Yamshchikov VF, Compans RW. Processing of the Intracellular Form of the West Nile Virus Capsid Protein by the Viral NS2B-NS3 Protease: An in Vitro Study. *J. Virol.* 1994;68(9):5765–5771. https://doi.org/10.1128/jvi.68.9.5765-5771.1994.

27 Aguirre S, Maestre AM, Pagni S, et al. DENV Inhibits Type I IFN Production in Infected Cells by Cleaving Human STING. *PLoS Pathog.* 2012;8(10):e1002934. https://doi.org/10.1371/journal.ppat.1002934.

28 Luo D, Vasudevan SG, Lescar J. The Flavivirus NS2B-NS3 Protease-Helicase as a Target for Antiviral Drug Development. *Antiviral Res.* June 2015;2015(118): 148–158. https://doi.org/10.1016/j.antiviral.2015.03.014.

29 Nitsche C, Holloway S, Schirmeister T, Klein CD. Biochemistry and Medicinal Chemistry of the Dengue Virus Protease. *Chem. Rev.* 2014;114(22):11348–11381. https://doi.org/10.1021/cr500233q.

30 da Silva Rodrigues, É. E.; Maus, H.; Hammerschmidt, S. J.; Ruggieri, A.; dos Santos, E. C.; Bassi, Ê. J.; Anderson, L.; Aquino, P. G. V.; de Araújo-Júnior, J. X.; Wei, F.; Liu, X.; Zhan, P.; Schirmeister, T.; da Silva-Júnior, E. F. The Medicinal Chemistry of Zika Virus. In Human Viruses: Diseases, Treatments and Vaccines; Ahmad, S. I., Ed.; Springer International Publishing, 2021; pp 233–295. https://doi.org/10.1007/978-3-030-71165-8_13.

31 Silva-Júnior EF, Schirmeister T, Araújo-Júnior JX. Recent Advances in Inhibitors of Flavivirus NS2B-NS3 Protease from Dengue, Zika, and West Nile Viruses. *Vector-Borne Dis. Treat.* 2018;No. November:1–25.

32 Noble CG, Seh CC, Chao AT, Shi PY. Ligand-Bound Structures of the Dengue Virus Protease Reveal the Active Conformation. *J. Virol.* 2012;86(1):438–446. https://doi.org/10.1128/JVI.06225-11.

33 Hill ME, Yildiz M, Hardy JA. Cysteine Disulfide Traps Reveal Distinct Conformational Ensembles in Dengue Virus NS2B-NS3 Protease. *Biochemistry*. 2019;58(6):776–787. https://doi.org/10.1021/acs.biochem.8b00978.

34 Yusof R, Clum S, Wetzel M, Murthy HMK, Padmanabhan R. Purified NS2B/NS3 Serine Protease of Dengue Virus Type 2 Exhibits Cofactor NS2B Dependence for Cleavage of Substrates with Dibasic Amino Acids in Vitro. *J. Biol. Chem.* 2000;275 (14):9963–9969. https://doi.org/10.1074/jbc.275.14.9963.

35 Erbel P, Schiering N, D'Arcy A, et al. Structural Basis for the Activation of Flaviviral NS3 Proteases from Dengue and West Nile Virus. *Nat. Struct. Mol. Biol.* 2006;13(4): 372–373. https://doi.org/10.1038/nsmb1073.

36 Lei J, Hansen G, Nitsche C, Klein CD, Zhang L, Hilgenfeld R. Crystal Structure of Zika Virus Ns2b-Ns3 Protease in Complex with a Boronate Inhibitor. *Science.* 2016;353 (6298):503–505. https://doi.org/10.1126/science.aag2419.

37 Millies B, von Hammerstein F, Gellert A, et al. Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases. *J. Med. Chem.* 2019;62(24): 11359–11382. https://doi.org/10.1021/acs.jmedchem.9b01697.

38 Yildiz M, Ghosh S, Bell JA, Sherman W, Hardy JA. Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. *ACS Chem. Biol.* 2013;8(12):2744–2752. https://doi.org/10.1021/cb400612h.

39 Nitsche C, Zhang L, Weigel LF, et al. Peptide-Boronic Acid Inhibitors of Flaviviral Proteases: Medicinal Chemistry and Structural Biology. *J. Med. Chem.* 2017;60(1): 511–516. https://doi.org/10.1021/acs.jmedchem.6b01021.

40 Lee H, Ren J, Nocadello S, et al. Identification of Novel Small Molecule Inhibitors against NS2B/NS3 Serine Protease from Zika Virus. *Antiviral Res.* 2017;139:49–58. https://doi.org/10.1016/j.antiviral.2016.12.016.

41 Behnam MAM, Graf D, Bartenschlager R, Zlotos DP, Klein CD. Discovery of Nanomolar Dengue and West Nile Virus Protease Inhibitors Containing a 4-Benzyloxyphenylglycine Residue. *J. Med. Chem.* 2015;58(23):9354–9370. https://doi.org/10.1021/acs.jmedchem.5b01441.

42 Gruba N, Rodriguez Martinez JI, Grzywa R, et al. Substrate Profiling of Zika Virus NS2B-NS3 Protease. *FEBS Lett.* 2016;590(20):3459–3468. https://doi.org/10.1002/1873-3468.12443.

43 Kühl N, Graf D, Bock J, Behnam MAM, Leuthold M-M, Klein CD. A New Class of Dengue and West Nile Virus Protease Inhibitors with Submicromolar Activity in Reporter Gene DENV-2 Protease and Viral Replication Assays. *J. Med. Chem.* 2020;63 (15):8179–8197. https://doi.org/10.1021/acs.jmedchem.0c00413.

44 Kühl N, Leuthold MM, Behnam MAM, Klein CD. Beyond Basicity: Discovery of Nonbasic DENV-2 Protease Inhibitors with Potent Activity in Cell Culture. *J. Med. Chem.* 2021;64(8):4567–4587. https://doi.org/10.1021/acs.jmedchem.0c02042.

45 Chen X, Yang K, Wu C, et al. Mechanisms of Activation and Inhibition of Zika Virus NS2B-NS3 Protease. *Cell Res.* 2016;26(11):1260–1263. https://doi.org/10.1038/cr.2016.116.

46 Wu H, Bock S, Snitko M, et al. Novel Dengue Virus NS2B/NS3 Protease Inhibitors. *Antimicrob. Agents Chemother.* 2015;59(2):1100–1109. https://doi.org/10.1128/AAC.03543-14.

47 Abd-Allah WH, Aboutabl ME, Aboul-Enein MN, El-Azzouny AAS. Synthesis, Molecular Modeling Studies and Anticonvulsant Activity of Certain (1-(Benzyl (Aryl) Amino) Cyclohexyl) Methyl Esters. *Bioorg. Chem.* April 2017;2017(71):135–145. https://doi.org/10.1016/j.bioorg.2017.01.021.

48 Sandeep, Venugopalan P, Kumar A. Metal Free, Direct and Selective Deoxygenation of α-Hydroxy Carbonyl Compounds: Access to α, α-Diaryl Carbonyl Compounds. *European. J. Org. Chem.* 2020;2020(17):2530–2536. https://doi.org/10.1002/ejoc.202000142.

49 Moore DL, Denton AE, Kohinke RM, Craig BR, Brenzovich WE. Silica Sulfuric Acid as a Highly Efficient Catalyst for the Synthesis of Diarylacetic Acids. *Synth. Commun.* 2016;46(7):604–612. https://doi.org/10.1080/00397911.2016.1158269.

50 Gbilimou A, Darmanin T, Godeau G, Guittard F. A Templateless Electropolymerization Approach to Nanorings Using Substituted 3,4-Naphthalenedioxythiophene (NaPhDOT) Monomers. *ChemNanoMat*. 2018;4(1):140–147. https://doi.org/10.1002/cnma.201700269.

51 Göksu S, Kazaz C, Sütbeyaz Y, Seçen H. A Concise Synthesis of 2-Amino-1,2,3,4-Tetrahydronaphthalene-6,7-Diol ('6,7-ADTN') from Naphthalene-2,3-Diol. *Helv. Chim. Acta*. 2003;86(10):3310–3313. https://doi.org/10.1002/hlca.200390272.

52 Zhao S, Zhao L, Zhang X, et al. Design, Synthesis, and Structure-Activity Relationship Studies of Benzothiazole Derivatives as Antifungal Agents. *Eur. J. Med. Chem.* 2016; 123:514–522. https://doi.org/10.1016/j.ejmech.2016.07.067.

53 Boger, D. L. CBI Analogues of the Duocarmycins and CC-1065. US 2005/0026987 A1, 2003.

54 Wang S, Jia K, Cheng J, Chen Y, Yuan Y. Dual Roles of Substituted Thiourea as Reductant and Ligand in CuAAC Reaction. *Tetrahedron Lett.* 2017;58(38):3717–3721. https://doi.org/10.1016/j.tetlet.2017.08.029.

55 Jakkaraj SR, Young VG, Georg GI. Syntheses of PDE3A Inhibitor ORG9935 and Determination of the Absolute Stereochemistries of Its Enantiomers by X-Ray

Crystallography. *Tetrahedron.* 2018;74(22):2769–2774. https://doi.org/10.1016/j.tet.2018.04.045.

56 Laplace DR, Verbraeken B, Van Hecke K, Winne JM. Total Synthesis of (+/-)-Frondosin B and (+/-)-5-Epi-Liphagal by Using a Concise (4+3) Cycloaddition Approach. *Chem. - A Eur. J.* 2014;20(1):253–262. https://doi.org/10.1002/chem.201303273.

57 Tew, D. G.; Thompson, S. K.; Veber, D. F. WO20/17924 A1, 2002.

58 Santangelo F, Casagrande C, Miragoli G, Vecchietti V. Synthesis and Positive Inotropic Effect of 1-Alkyl- and 1-Acyl-6,7-Dimethoxy-3-Dimethylamino-1,2,3,4-Tetrahydroquinolines. *Eur. J. Med. Chem.* 1994;29(11):877–882. https://doi.org/10.1016/0223-5234(94)90110-4.

59 Gavara L, Boisse T, Hénichart JP, Daïch A, Rigo B, Gautret P. Toward New Camptothecins. Part 6: Synthesis of Crucial Ketones and Their Use in Friedländer Reaction. *Tetrahedron.* 2010;66(38):7544–7561. https://doi.org/10.1016/j.tet.2010.07.048.

60 Gryko D, Lipiński R. Asymmetric Direct Aldol Reaction Catalysed by L-Prolinethioamides. *European J. Org. Chem.* 2006;2006(17):3864–3876. https://doi.org/10.1002/ejoc.200600219.

61 von Hammerstein F, Lauth LM, Hammerschmidt S, Wagner A, Schirmeister T, Hellmich UA. Cis Autocatalytic Cleavage of Glycine-Linked Zika Virus NS2B-NS3 Protease Constructs. *FEBS Lett.* 2019;593(16):2204–2213. https://doi.org/10.1002/1873-3468.13507.

62 Steuer C, Heinonen KH, Kattner L, Klein CD. Optimization of Assay Conditions for Dengue Virus Protease: Effect of Various Polyols and Nonionic Detergents. *J. Biomol. Screen.* 2009;14(9):1102–1108. https://doi.org/10.1177/1087057109344115.

63 Maier GP, Bernt CM, Butler A. Catechol Oxidation: Considerations in the Design of Wet Adhesive Materials. *Biomater. Sci.* 2018;6(2):332–339. https://doi.org/10.1039/c7bm00884h.

64 Brecher M, Li Z, Liu B, et al. A Conformational Switch High-Throughput Screening Assay and Allosteric Inhibition of the Flavivirus NS2B-NS3 Protease. *PLoS Pathog.* 2017;13(5):e1006411. https://doi.org/10.1371/journal.ppat.1006411.

65 Othman R, Kiat TS, Khalid N, et al. Docking of Noncompetitive Inhibitors into Dengue Virus Type 2 Protease: Understanding the Interactions with Allosteric Binding Sites. *J. Chem. Inf. Model.* 2008;48(8):1582–1591. https://doi.org/10.1021/ci700388k.

66 BioSolveIT GmbH. LeadIT Version 2.3.2; BioSolveIT GmbH, Sankt Augustin, Germany, 2017, Www.Biosolveit.de/LeadIT. BioSolveIT GmbH: Sankt Augustin, Germany.

67 Zimmermann MO, Lange A, Zahn S, Exner TE, Boeckler FM. Using Surface Scans for the Evaluation of Halogen Bonds toward the Side Chains of Aspartate, Asparagine, Glutamate, and Glutamine. *J. Chem. Inf. Model.* 2016;56(7):1373–1383. https://doi.org/10.1021/acs.jcim.6b00075.

68 Kumar A, Buhler S, Selisko B, et al. Nuclear Localization of Dengue Virus Nonstructural Protein 5 Does Not Strictly Correlate with Efficient Viral RNA Replication and Inhibition of Type I Interferon Signaling. *J. Virol.* 2013;87(8):4545–4557. https://doi.org/10.1128/JVI.03083-12.

69 Münster M, Płaszczyca A, Cortese M, et al. A Reverse Genetics System for Zika Virus Based on a Simple Molecular Cloning Strategy. *Viruses.* 2018;10(7):368. https://doi.org/10.3390/v10070368.

70 Cortese M, Goellner S, Acosta EG, et al. Ultrastructural Characterization of Zika Virus Replication Factories. *Cell Rep.* 2017;18(9):2113–2123. https://doi.org/10.1016/j.celrep.2017.02.014.

71 Johansson M, Brooks AJ, Jans DA, Vasudevan SG. A Small Region of the Dengue Virus-Encoded RNA-Dependent RNA Polymerase, NS5, Confers Interaction with Both the Nuclear Transport Receptor Importin-β and the Viral Helicase, NS3. *J. Gen. Virol.* 2001;82(4):735–745. https://doi.org/10.1099/0022-1317-82-4-735.

72 Baronti C, Piorkowski G, Charrel RN, Boubis L, Leparc-Goffart I, de Lamballerie X. Complete Coding Sequence of Zika Virus from a French Polynesia Outbreak in 2013. *Genome Announc.* 2014;2(3):e00500–e514. https://doi.org/10.1128/genomeA.00500-14.

73 Rarey M, Kramer B, Lengauer T. Multiple Automatic Base Selection: Protein-Ligand Docking Based on Incremental Construction without Manual Intervention. *J. Comput. Aided. Mol. Des.* 1997;11(4):369–384. https://doi.org/10.1023/A:1007913026166.

74 Halgren TA. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94*. *J. Comput. Chem.* 1996;17(5–6):490–519. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

75 Leatherbarrow RJ. *GraFit 6*. Erithacus Software Limited: East Grinstead, West Sussex, UK; 2007.

76 Ludewig S, Kossner M, Schiller M, Baumann K, Schirmeister T. Enzyme Kinetics and Hit Validation in Fluorimetric Protease Assays. *Curr. Top. Med. Chem.* 2010;10(3):368–382. https://doi.org/10.2174/156802610790725498.

77 Klein P, Barthels F, Johe P, et al. Naphthoquinones as Covalent Reversible Inhibitors of Cysteine Proteases—Studies on Inhibition Mechanism and Kinetics. *Molecules.* 2020;25(9):2064. https://doi.org/10.3390/molecules25092064.

78 Amendola G, Ettari R, Previti S, et al. Lead Discovery of SARS-CoV-2 Main Protease Inhibitors through Covalent Docking-Based Virtual Screening. *J. Chem. Inf. Model.* 2021;61(4):2062–2073. https://doi.org/10.1021/acs.jcim.1c00184.

79 Angelini A, Cendron L, Chen S, et al. Bicyclic Peptide Inhibitor Reveals Large Contact Interface with a Protease Target. *ACS Chem. Biol.* 2012;7(5):817–821. https://doi.org/10.1021/cb200478t.

80 Barthels F, Marincola G, Marciniak T, et al. Asymmetric Disulfanylbenzamides as Irreversible and Selective Inhibitors of Staphylococcus Aureus Sortase A. *ChemMedChem.* 2020;15(10):839–850. https://doi.org/10.1002/cmdc.201900687.

81 Poreba M, Szalek A, Rut W, et al. Highly Sensitive and Adaptable Fluorescence-Quenched Pair Discloses the Substrate Specificity Profiles in Diverse Protease Families. *Sci. Rep.* 2017;7(1). https://doi.org/10.1038/srep43135.

82 van den Hoff, M. J. B.; Christoffels, V. M.; Labruyère, W. T.; Moorman, A. F. M.; Lamers, W. H. Electrotransfection with "Intracellular" Buffer. In Animal Cell Electroporation and Electrofusion Protocols; Nickoloff, J. A., Ed.; Humana Press, 1995; Vol. 48, pp 185–197. https://doi.org/10.1385/0-89603-304-X:185.

83 GraphPad Prism 7.0.4 for Windows. GraphPad Software: San Diego, California USA.

84 Trapani G, Franco M, Latrofa A, Reho A, Liso G. Synthesis, in Vitro and in Vivo Cytotoxicity, and Prediction of the Intestinal Absorption of Substituted 2-Ethoxycarbonyl-Imidazo[2,1-b]Benzothiazoles. *Eur. J. Pharm. Sci.* 2001;14(3):209–216. https://doi.org/10.1016/S0928-0987(01)00173-7.

85 Maolanon AR, Risgaard R, Wang S-Y, et al. Subtype-Specific Agonists for NMDA Receptor Glycine Binding Sites. *ACS Chem. Neurosci.* 2017;8(8):1681–1687. https://doi.org/10.1021/acschemneuro.7b00117.

86 Götz C, Hinze A, Gellert A, et al. Conformational dynamics of the Dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies. *J. Phys. Chem. Part B.* 2021;125:6837–6846. https://doi.org/10.1021/acs.jpcb.1c01797.

SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases

Hannah Maus‡[a], Fabian Barthels‡[a], Stefan Josef Hammerschmidt[a], Katja Kopp[b], Benedikt Millies[a], Andrea Gellert[a], Alessia Ruggieri[b], Tanja Schirmeister[a]*

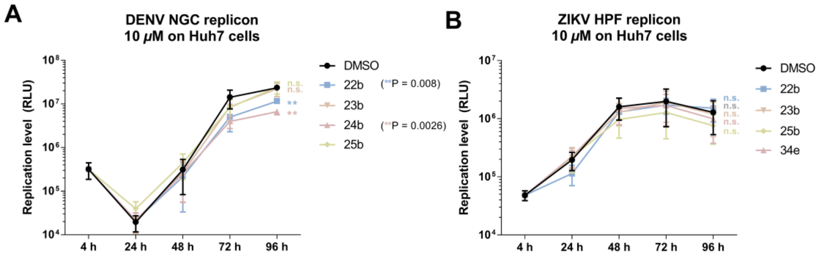
‡authors contributed equally

[a] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz.

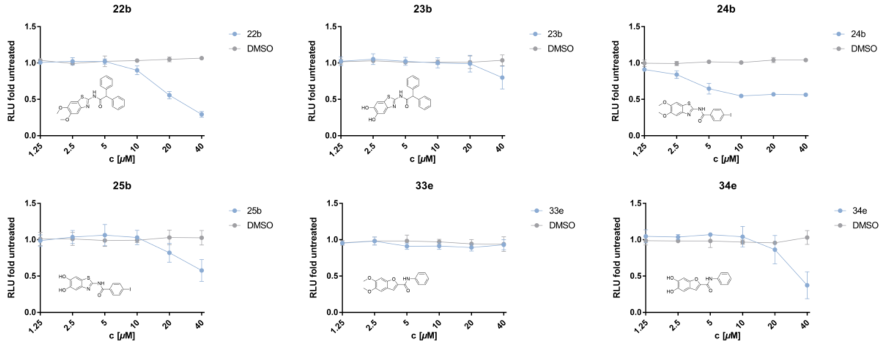
[b] Department of Infectious Diseases, Molecular Virology, Center for Integrative Infectious Disease Research, University of Heidelberg, Im Neuenheimer Feld 344, 69120 Heidelberg.

Table of contents

Supporting data for cellular experiments



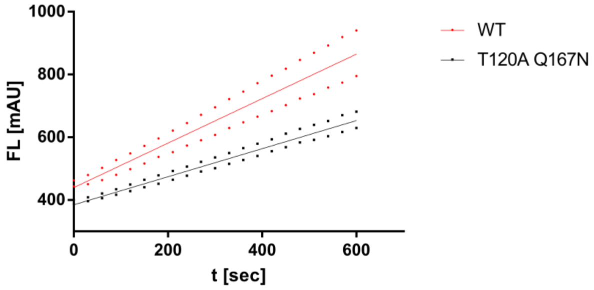
SI Figure 1. DENV/ZIKV RNA replication time curves for the most active inhibitors. Huh7 cells were treated with 10 μM test compound after electroporation and replicons were determined by the luciferase assay for (**A**) DENV NGC replicon and (**B**) ZIKV HPF replicon. Shown are means ± standard deviation (n = 3). **, P < 0.01; n.s., not significant. Statistical significance was calculated by using 2way ANOVA for replication curves.



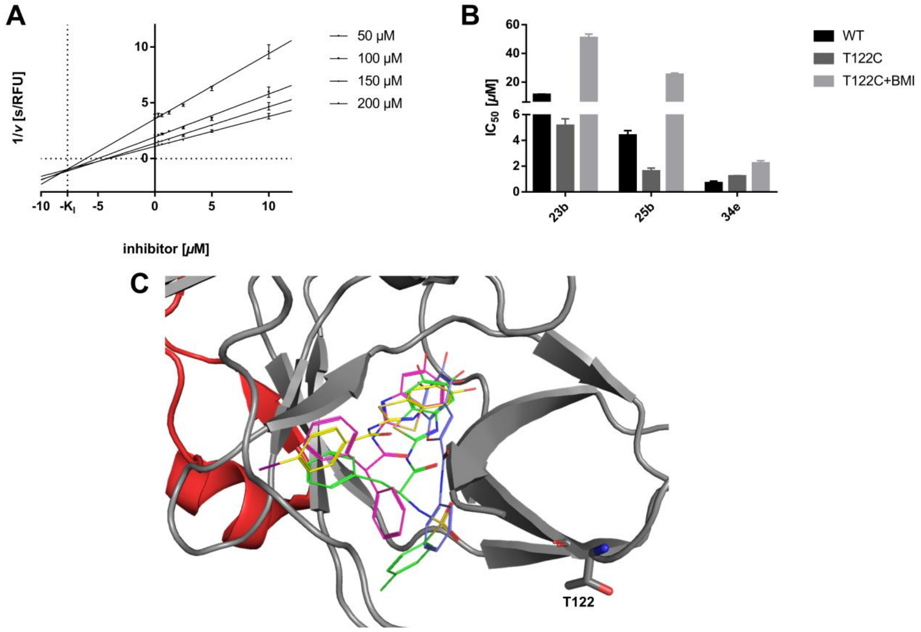
SI Figure 2. Viability of Huh7 cells treated with the respective test compounds as determined by CellTiter-Glo Luminescent assay (Promega). Shown are means ± standard deviation (n = 3). **, P < 0.01; n.s., not significant.

```
DENV-NGC  13   SADLELERAADVKWEDQAEISGSSPILSITISEDGSMSIKNEEEEQTLTILIRTGLLVIS  72
               +ADLELERAADV+WE+QAEISGSSPILSITISEDGSMSIKNEEEEQTL
Protein    3   AADLELERAADVRWEEQAEISGSSPILSITISEDGSMSIKNEEEEQTLGGGGSG------  56

DENV-NGC  73   GLFPVSIPITAAAWYLWEVKKQRAGVLWDVPSPPPMGKAELEDGAYRIKQKGILGYSQIG  132
                                      AGVLWDVPSPPP+GKAELEDGAYRIKQKG  GYSQIG
Protein   57   --------------------GGGAGVLWDVPSPPPVGKAELEDGAYRIKQKGAAGYSQIG  96

DENV-NGC 133   AGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQV  192
               AGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQV
Protein   97   AGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWADVKKDLISYGGGWKLEGEWKEGEEVQV  156

DENV-NGC 193   LALEPGKNPRAVQTKPGLFKTNAGTIGAVSLDFSPGTSGSPIIDKKGKVVGLYGNGVVTR  252
               LALEPGKNPRAVQTKPGLFKTN GTIGAVSLDFSPGTSGSPI+DKKGKVVGLYGNGVVTR
Protein  157   LALEPGKNPRAVQTKPGLFKTNTGTIGAVSLDFSPGTSGSPIVDKKGKVVGLYGNGVVTR  216

DENV-NGC 253   SGAYVSAIAQTEKSIEDNPEIEDDIFRK  280
               SGAYVSAIA TEKSIEDNPEIEDDIFRK
Protein  217   SGAYVSAIANTEKSIEDNPEIEDDIFRK  244
```

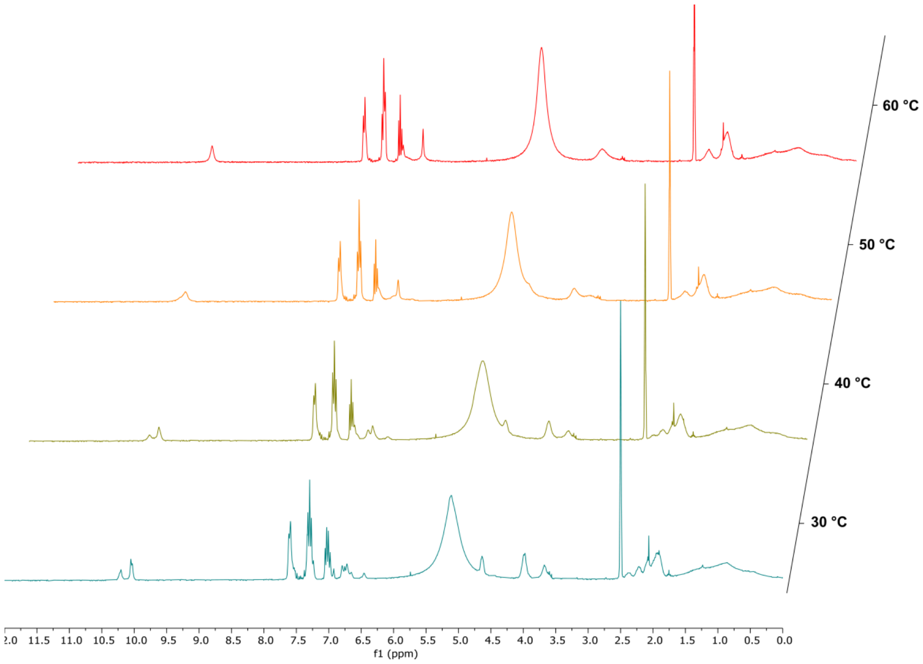
SI Figure 3. Sequence alignment (Blastp) of the cellular NS2B/NS3 protein (DENV2 NGC) vs. the recombinantly expressed protein construct. Two mutations in the allosteric binding region were identified (T120A & N167Q). Sequence identity: 224/268(84%).



SI Figure 4. Fluorometric activity determination of the DENV2 NS2B NS3 double mutant T120A Q167N in comparison to the DENV2 wild type. The enzymatic activity at an enzyme concentration of 125 nM and a substrate concentration of 100 μM (Boc-Gly-Arg-Arg-AMC) is k_{cat} = 0.709±0.039 mAU s^{-1} for the DENV wild type and k_{cat} = 0.447 ± 0.028 mAU s^{-1} for the double mutant.



SI Figure 5. (**A**) Enzyme inhibition of the DENV2 protease treated with variable concentrations of the inhibitor **25b** analyzed by a Dixon-plot at four different substrate concentrations (Boc-Gly-Arg-Arg-AMC: 50 µM, 100 µM, 150 µM, 200 µM). The apparent non-competitive inhibition constant was determined to be K_I = 7.67 µM. (**B**) IC_{50} values of **23b**, **25b**, or **34e** on wild-type (WT) and mutant T122C DENV2 proteases in the absence or treated with *N*-benzylmaleimide (BMI, 250 nM). IC_{50} values are indicated as means ± standard deviation from three independent measurements. (**C**) Docking pose superposition of the top inhibitors of each SAR series (**8e**: green, **23b**: purple, **25b**: yellow, **34e**: blue) in the putative allosteric binding pocket of the DENV2 NS2B-NS3 protease (pdb: 2FOM). The muted residue T122C in close proximity to the binding pocket is shown by the stick representation.



SI Figure 6. Variable temperature (VT)-NMR of compound **36e** (30–60 °C). Duplication of the NMR signal set (30 °C) was due to different energetic states (rotamers). The resolution of the different energetic states at higher temperatures is exemplified.

“

4.2 Fragment-based Design of α-Cyanoacrylates and α-Cyanoacrylamides Targeting Dengue and Zika NS2B-NS3 Proteases.

4.2.1 Context, Project Summary, and Own Contributions

Competitive inhibitors of flaviviral proteases frequently incorporate highly charged guanidino or similar functional groups because of the recognition preference of the enzyme for basic residues. However, most competitive inhibitors show poor membrane permeability due to their charged nature and high molecular weight.[86,253,254] Therefore, a challenge arises to reconcile the two opposing development aims of pharmacodynamic and pharmacokinetic optimization, where basicity can be considered a "friend and foe", respectively.[256] Consequently, designing compounds with reduced molecular size and polarity might help to approach the expected criteria for oral bioavailability, as described in the Lipinski rule of five and similar concepts.[256–258]

Fragment-based drug design (FBDD) has been previously used to discover inhibitors against DENV2 NS5[259] and against ZIKV and DENV2 NS2B-NS3 protease[260,261]. Based on this strategy, cyanoacrylates and cyanoacrylamides were designed as potential antiviral agents targeting DENV2 and ZIKV NS2B-NS3 proteases (Figure 20). We investigated their binding affinities, modes, and intermolecular interactions using fluorometric assay, molecular docking, dynamics simulations, and MM/PBSA calculations. Initially, a virtual library of 340 chemical fragments was screened, and promising fragments were selected according to their FitScore. Previous research has highlighted the efficacy of cyano groups as linker groups against DENV2 and West Nile virus (WNV).[262] Additionally, α-ketoamides and cinnamyl moieties have shown activity against DENV2 and WNV NS2B/NS3 proteases.[263] Considering these findings, aryl cyanoacrylamides were integrated to enhance the antiviral activity of the selected fragments. The most promising molecules were selected for synthesis and biological evaluation. A total of 27 compounds were synthesized and screened against ZIKV and DENV2 NS2B-NS3 protease. We investigated the binding mode of the best inhibitors for DENV2 NS2B-NS protease. The experimental data confirmed that the designed compounds are competitive inhibitors binding to the protease's active site. Dynamics simulations indicated good stability of NS2B-NS3-ligand complexes, and MM/PBSA calculations showed similar binding energies for the analyzed compounds. Similar binding modes were suggested in docking studies, indicating interactions with key protease residues.

One compound exhibited good antiviral activity against ZIKV-infected cells, reducing viral RNA copy numbers, consequently representing a potential inhibitor of viral replication. Overall, this study provided an alternative approach to design hit compounds against DENV2 and ZIKV NS2B-NS3 protease, with the potential for further optimization to enhance their inhibitory activities and expand the therapeutic arsenal against these flaviviruses.

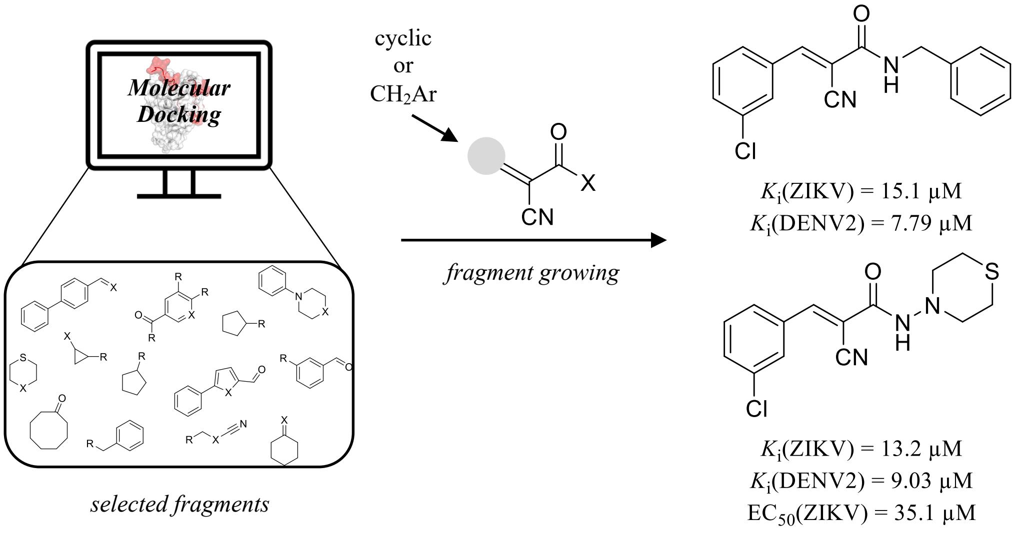


Figure 20. Potent competitive inhibitors of the ZIKV and DENV2 protease discovered by fragment-based drug design.

Own contributions: DENV and ZIKV inhibition assays, protease selectivity assays, and manuscript editing.

Contribution from other authors: molecular docking, inhibitor synthesis, protein expression & purification, molecular dynamics simulations, MM/PBSA calculations, cytotoxicity assays, evaluation of *in vitro* antiviral activity, quantification of viral load, and writing of the original draft & editing of the manuscript.

This work has been published in New Journal of Chemistry (impact factor: 3.93).

Article reprinted with permission from *New Journal of Chemistry* **2022**, 46 (42): 20322–20346 “Fragment-based design of α-cyanoacrylates and α-cyanoacrylamides targeting Dengue and Zika NS2B/NS3 proteases.” © 2022 Royal Society of Chemistry (United Kingdom).

The Supporting Information can be accessed online at doi: 10.1039/D2NJ01983C.

4.2.2 Publication

The following publication quoted within “” from page 75 to page 99 is the same as the manuscript cited on page 74.

“

NJC

ROYAL SOCIETY OF CHEMISTRY

PAPER

Check for updates

Cite this: DOI: 10.1039/d2nj01983c

Fragment-based design of α-cyanoacrylates and α-cyanoacrylamides targeting Dengue and Zika NS2B/NS3 proteases†

Gabriel Gomes Vilela,[‡ab] Wadja Feitosa dos Santos Silva,[‡ab] Vitoria de Melo Batista,[b] Leandro Rocha Silva,[ab] Hannah Maus,[c] Stefan Josef Hammerschmidt,[c] Clara Andrezza Crisóstomo Bezerra Costa,[d] Orlando Francisco da Silva Moura,[d] Johnnatan Duarte de Freitas,[d] Grazielle Lobo Coelho,[e] Júlia de Andrade Brandão,[e] Leticia Anderson,[ef] Ênio José Bassi,[e] João Xavier de Araújo-Júnior,[b] Tanja Schirmeister[c] and Edeildo Ferreira da Silva-Júnior[*ab]

Viruses belonging to the *Flaviviridae* family account for millions of cases of infections worldwide, mainly Dengue (DENV) and Zika (ZIKV) viruses. Both mosquito-borne flaviviruses are transmitted by bites of infected *Aedes* vectors, which are mainly endemic in tropical and sub-tropical countries. DENV and ZIKV are composed of structural and non-structural (NS) proteins that are initially expressed as a single precursor polyprotein. NS2B/NS3 is a serine protease complex that has a critical role during the viral replication cycle since the polyprotein processing is also accomplished by this enzyme. As an approach for developing novel NS2B/NS3 inhibitors, a virtual fragment-based drug design protocol was developed to select small fragments for new inhibitors. Based on this, we developed a series of α-cyanoacrylates and α-cyanoacrylamides targeting ZIKV and DENV NS2B/NS3 proteases. Among them, **LQM467**, **471**, **472**, and **474** were found to be active against both proteases, with K_i values ranging from 7.8 to 20 μM. Dynamics simulations revealed good stability during 200 ns for all enzyme inhibitor-complexes. Further, the MM/BPSA approach revealed that these active compounds present similar energetic profiles, corroborating our experimental data. Additionally, docking analyses showed that the compounds share similar binding poses involving the catalytic residues. Subsequently, these compounds were screened for their cytotoxic effects, showing that only **LQM474** is toxic to *Vero E6* cells. Thus, **LQM467**, **LQM471**, and **LQM472** were evaluated on ZIKV-infected cells to determine their antiviral potential. As a result, it was verified that **LQM471** is the most promising compound, exhibiting an EC_{50} value of 35.07 μM. Finally, in order to obtain insights into its mechanism of action, a qPCR was performed, exhibiting that **LQM471** is able to reduce the viral RNA copies compared to the untreated ZIKV-infected cells, suggesting that this compound can be a promising inhibitor of virus replication. This study presents an interesting alternative to design new hit compounds against DENV-2 and ZIKV NS2B/NS3 proteases contributing to the development of new therapeutic agents against these flaviviruses.

Received 22nd April 2022,
Accepted 7th September 2022

DOI: 10.1039/d2nj01983c

rsc.li/njc

---

[a] *Institute of Chemistry and Biotechnology, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões campus, 57072-970, Maceió, Alagoas, Brazil. E-mail: edeildo.junior@iqb.ufal.br; Tel: +55-82-9-9610-8311*

[b] *Laboratory of Medicinal Chemistry, Post-Graduation Program in Pharmaceutical Sciences, Institute of Pharmaceutical Sciences, Federal University of Alagoas, Lourival Melo Mota Avenue, AC. Simões Campus, 57072-970, Maceió, Alagoas, Brazil*

[c] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany*

[d] *Department of Chemistry, Federal Institute of Alagoas, Maceió campus, Mizael Domingues Street, 57020-600, Maceió, Alagoas, Brazil*

[e] *Immunoregulation Research Group, Laboratory of Research in Virology and Immunology, Institute of Biological and Health Sciences, Lourival Melo Mota Avenue, AC. Simões Campus, 57072-970, Maceió, Alagoas, Brazil*

[f] *CESMAC University Center, 57051-160, Maceió, Alagoas, Brazil*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2nj01983c

‡ Authors contributed equally.

1. Introduction

In recent years, arthropod-borne viruses (or arboviruses) have spread significantly around the world.[1–5] Flaviviruses belong to the *Flaviviridae* family, composed of 53 different species of viruses that are related to several medicinally important pathogens, including Yellow fever viruses (YFV), the Japanese encephalitis virus (JEV), the West Nile fever virus (WNV), the Tick-borne encephalitis virus (TBEV),[6,7] the Hepatitis C virus (HCV),[8,9] and the Pestis virus,[9] as well as Dengue viruses (DENV), and the Zika virus (ZIKV).[10] Interestingly, flaviviruses are mainly transmitted by mosquitoes from the genus *Aedes* spp.[11,12] In contrast to other arboviruses, ZIKV can also be transmitted from person to person through sexual and non-sexual bodily fluids. In addition, mother-to-child transmission is possible both during prenatal, perinatal (during birth) and postnatal (during breastfeeding) periods.[13–15] The disease is associated with a variety of flu-like and neurological symptoms, such as Guillain–Barré syndrome, microcephaly in newborns, cognitive disfunctions, as well as epileptic episodes.[16,17] By the end of July 2019, 87 countries had reported ZIKV cases, including Africa, America, Southeast Asia, and the Western Pacific regions.[18] Additionally, thousands of DENV cases have been reported by the Pan American Health Organization (PAHO), where only in 2021, more than 55 800 cases were registered, with 4297 reported in Nicaragua and 4118 in Colombia. Paraguay, Mexico, and Ecuador had 3790, 1951, and 1670 reported cases, respectively. Among those, about 210 patients developed Dengue Shock Syndrome (DSS), being responsible for 13 fatal cases.[19]

The members of the *Flaviviridae* family have high similarities in their life cycles. Typically, diverse cellular receptors are responsible for binding to virus superficial glycoproteins, such as DC-SIGN, L-SIGN, Tyro3, MertK, and AXL.[20,21] The virus internalization occurs by a clathrin-dependent endocytosis mechanism, leading to the membranes' fusion (pH ∼ 6) and then, releasing the viral genome.[22] Subsequently, genome replication occurs in the Golgi apparatus membrane.[23] The genomic RNA is used to produce a single polyprotein by the action of host ribosomes, which is translocated to the endoplasmic reticulum membrane. Posteriorly, this polyprotein is cleaved in different viral components, namely three structural proteins (pre-membrane (prM), envelope (E), and capsid (C)) and seven non-structural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5).[7,24] Then, the mature virions are transported and released from the host cell.[10,25] The three structural proteins are involved in the virions formation. The C protein is highly basic (11 kDa) and encapsulates the viral genome. Moreover, prM comprises a segment of 20 amino acids at the N-terminus, which involves three pH-sensitive histidines. The M protein (8 kDa) is a small proteolytic fragment of the prM, produced after being cleaved by the protease furin of host cells.[26] The E protein (50 kDa) mediates the virus binding and membrane fusion during virus entry.[27] Regarding the NS proteins, the NS1 is a glycoprotein that has been associated with the E protein for the detection of different DENV serotypes.[28] The NS2A is a hydrophobic protein (∼22 kDa) that exhibits a critical role in the virus organization.[29] In contrast, NS2B (27 kDa) is associated with the virus membrane, which constitutes a complex with NS3. This protease complex is responsible for the polyprotein processing.[30,31] NS4A (16 kDa) and NS4B (27 kDa) proteins are hydrophobic and associated with the organization of the RNA replication.[32] Furthermore, NS5 is a multifunctional protein (103 kDa), which displays methyltransferase and RNA-polymerase activities.[33,34]

In flaviviruses, the viral replication depends on the non-structural protein 3 protease (NS3^{pro}, 70 kDa), which is only fully active when associated with NS2B, acting as a cofactor to increase the efficiency of enzymatic activity against peptide substrates 300 to 7600 fold.[24,35–37] NS3^{pro} harbors a trypsin-like serine protease domain with two β-barrels consisting of six β-sheets, comprising 180 amino acids. Between both lobes of these barrels, a catalytic cleft is formed, featuring four specifically well-defined pockets.[38] NS3^{pro} contains a functional catalytic triad consisting of His^{51}, Asp^{75}, and Ser^{135} residues.[39] In general, cleavage of its peptide substrate starts with a nucleophilic attack of Ser^{135} on P1 carbonyl from the substrate. Since the hydroxyl group from Ser^{135} is inherently a weak nucleophile, it needs to be activated by the action of the adjacent His^{51} in an acid–base mechanism. The decomposition of the tetrahedral intermediate, which is stabilized by Asp^{75}, results in the cleavage of the C-terminal portion and the release of an amine fragment. The N-terminal region remains covalently bound to the protease *via* an ester bond, which is subsequently hydrolyzed. In this step, His^{51} again acts as a base to elevate the nucleophilic character of the involved water molecule. Finally, the N-terminal portion of the cleaved product is released *via* re-protonation of the carboxylic acid, starting a new catalytic cycle.[10,31,39] The cocrystallization of specific inhibitors with the protein is a great challenge, mainly due to the protein's conformational flexibility.[40] Different orientations of the NS2B cofactor to NS3^{pro} are referred to as "*open*" or "*closed*".[41] In the closed conformation of the DENV-2 protease, often observed in presence of active site inhibitors, NS2B surrounds almost the entire protease domain, allowing the deduction of specific S1–S3 pockets.[42,43] In the open conformation of the DENV-2 protease, the N-terminal domain of NS2B is only partially wrapped around the protease domain whereas its C-terminal domain remains largely unstructured.[44] Initially, two allosteric inhibition mechanisms for the protease were postulated. In the first, a ligand is able to stabilize the open state of NS2B/NS3^{pro}.[45] In the second, the mobility of NS2B is restricted or a disruption of the interaction between NS2B and NS3^{pro} is observed, preventing the conformational change of this protease to its active form.[46] Allosteric inhibition is thought to be related to the bifunctional Asn^{152} residue, which is involved in both substrate recognition, *via* hydrogen interaction with Gly^{82} residue from the NS2B β-hairpin in the closed conformation and the formation of an allosteric site in the open conformation. Additionally, it was identified that the most commonly found allosteric site is characterized by a predominantly hydrophobic cleft anterior to the classical catalytic site, which also includes the Ala^{125} residue.[47] The importance of Asn^{152} residue is underlined by its conservation in all flaviviral proteases. Furthermore, it is known that mutations of Asn^{152} lead to the complete inactivation of these proteases.[48] Consequently, the rational design of inhibitors targeting the Asn[152] residue could stabilize the open conformation and inhibit the formation of a proteolytic β-hairpin structure.[49] On the other hand, interactions with Lys[74], which is immediately adjacent to Asp[75] from the catalytic triad, could lead to a decreased activity by a ligand-induced conformational change.[45,50] Furthermore, Val[78] and Met[84] residues from NS2B, and Lys[73], Glu[88], Trp[89], Thr[118], Thr[120], Thr[122], Gly[124], and Val[147] from NS3 are potentially associated with the binding of allosteric protease inhibitors.[51] Considering these structural and functional aspects of the NS2B/NS3^{pro}, our research team has published several reviews[10,31,52] and original studies focused on the development of DENV and ZIKV NS2B/NS3^{pro} inhibitors.[53–56]

The protease complex represents an attractive target for drug design and development of new antiviral agents.[10,51] Herein, we reported the rational fragment-based design of cyanoacrylates and cyanoacrylamides against DENV-2 and ZIKV NS2B/NS3 proteases. Additionally, their binding affinities and binding modes were experimentally investigated. Finally, *in silico* studies were performed by applying molecular docking, dynamics simulations (at 200 ns), and MM/PBSA calculations to obtain information on the main intermolecular interactions, complex stability, and energetic terms, such as van der Waals, electrostatic, and total energies.

2. Results and discussion

2.1. Rational fragment-based design of cyanocinnamyl derivatives targeting DENV-2 and ZIKV NS2/NS3 proteases

Regarding that fragment-based drug design (FBDD) has been used to discover inhibitors against DENV-2 NS5,[57] and also ZIKV and DENV-2 NS2B/NS3^{pro},[58,59] we decided to utilize this strategy to identify small chemical fragments that could become promising NS2B/NS3^{pro} inhibiting compounds. Initially, an in-house virtual library of 340 chemical fragments was screened by molecular docking simulations on DENV NS2B/NS3^{pro}. These small fragments presented FitScore values ranging from 14.0 (lowest affinity) to 53.9 (highest affinity). Different aldehydes, acids, cyanoesters, and primary and secondary amines were retrieved as promising fragments. Since the FBDD is based on the identification of small chemical fragments that typically bind with weak affinity to the macromolecular target and which are further used to generate molecules with higher affinities than their precursors,[60] we decided to perform a fragment growth approach for combining fragments with FitScore values higher than 35.0 to obtain molecules with FitScores higher than 65.0.

Previously, cyano groups acting as linker groups have been reported as an interesting chemical fragment against DENV-2 and WNV. Considering that serine proteases can be targeted by different electrophile warheads, leading to covalent-reversible or covalent-irreversible binding modes, electrophiles promise to be interesting scaffolds. Moreover, their reactivity can be adjusted to match the requirements for a specific target, by retaining their selectivity and stability.[61] Additionally, α-ketoamides have been described as active compounds against DENV-2 and WNV NS2B/NS3 proteases and were effective in DENV-infected cell culture.[62] A cinnamyl moiety has been identified as a promising fragment, which can provide potent inhibitors of NS2B/NS3^{pro} when bound to electrophilic groups.[61,62] In this context, the nitrile group can be an interesting candidate for improving the antiviral activity of the cinnamyl moiety.[61] Since it may act as an electrophilic trap for the nucleophilic catalytic Ser[135], it can perform dipole–dipole interactions or even enable a reversible-covalent binding mode between the inhibitor and the proteases.[61,63–65] Additionally, aryl and amide moieties surrounding the nitrile group can interact with S1 and S1′ subsites from NS2B/NS3^{pro}, respectively.[61] Furthermore, cyanohydrazone groups in biphenyl and benzofuran derivatives have demonstrated broad-spectrum *in vivo* activity against ZIKV, DENV, and JEV.[66] However, hydrazone moieties are associated with pan-assay interference scaffolds (PAINS).[67–73] Therefore, we decided to develop aryl cyanoacrylates and cyanoacrylamides as Michael acceptors, avoiding a hydrazone moiety by performing retroisosterism. This approach is based on the inversion of a determined functional group present in the active compound structure, producing a new isostere with the same function and similar activities.[74,75] These compounds exhibited FitScore values higher than 65. Thereof, the most promising molecules were selected for synthesis and biological evaluation. Fig. 1 depicts the workflow performed in this study to design compounds targeting DENV NS2B/NS3^{pro}.

2.2 Chemistry

First, cinnamoyl analogs (**LQM448**, **449**, and **452–455**) were prepared *via* Knoevenagel condensation in 25–99% yields.[76–79] Herein, 3-chlorocarboxaldehyde or cyclic ketones and cyanoesters were dissolved in ethanol at room temperature, then triethylamine was added as catalyst base. Alternatively, ethyl-2-cyanoacetate was converted into secondary or tertiary cyanoamides by reaction with either primary or cyclic secondary amines at room temperature and solvent-free conditions, *via* an aminolysis mechanism. Then, these amides were used without additional purification to obtain **LQM450**, **451**, and **456–474** also *via* Knoevenagel condensation, reacting with 3-chlorocarboxaldehyde (72–85%), biphenyl-4-carboxaldehyde (24–85%) and phenyl-5-furfural (29–89%), respectively. These biphenyl aldehydes were prepared *via* Suzuki–Miyaura cross-coupling reaction,[80–82] involving different boronic acids and 4-bromocarboxaldehyde. All synthetic routes are presented in Scheme 1.

2.3 Enzymatic inhibition assays toward DENV-2 and ZIKV proteases

After the successful synthesis of all 27 rationally designed compounds, they were screened against ZIKV and DENV-2 NS2B/NS3^{pro} at 20 μM compound concentration. The mean inhibitory concentration (IC_{50}) was determined for compounds that exhibited at least 40% inhibition at 20 μM. The results at the different targets are shown in Table 1. Moreover, the IC_{50} determination curves are shown in Fig. 2. Furthermore, the K_i values were calculated using the Cheng Prusoff equation.

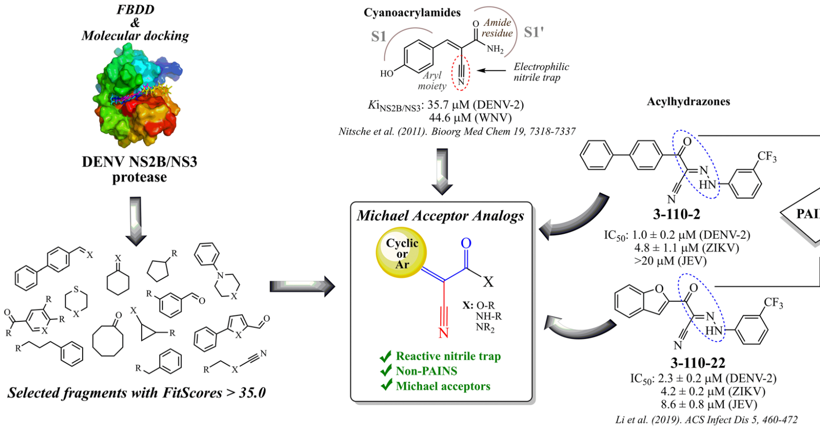
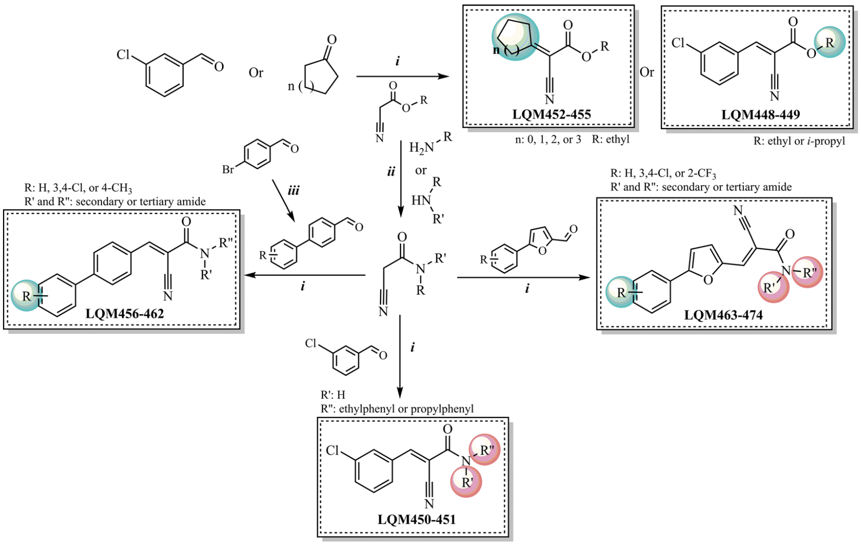


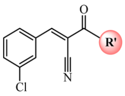
Fig. 1 Workflow for designing compounds targeting DENV NS2B/NS3^{pro}. Initially, 340 small chemical fragments were investigated toward the open and closed conformation of the DENV-2 NS2B/NS3^{pro} (PDB's: 2FOM and 2M9P, respectively) by using molecular docking studies. Then, an FBDD approach was used to combine the chemical fragments with FitScore values higher than 35. The final compounds (FitScore ≥ 65) were designed to have a reactive nitrile trap, and a Michael acceptor, and to exclude PAINS.

To evaluate the binding mode of the best inhibitors, the hit compound **LQM467** was used as an example. Its binding mode was investigated for DENV-2 NS2B/NS3^{pro}. Inhibitors that bind to the active binding site of the enzyme compete with the substrate for this binding site. The IC_{50} value of these compounds is thus dependent on the substrate concentration used. Therefore, to investigate the binding mode in more detail, the IC_{50} of the model compound **LQM467** was first determined at different substrate concentrations (50, 100, and 150 μM). The IC_{50} was found to increase with increasing substrate concentration (Table 2), indicating that the active site of the protease is targeted. Plotting the IC_{50} values against the substrate



Scheme 1 Synthetic routes used in this study to obtain the targeted compounds. Reaction conditions: (i) aldehyde or ketone, ethyl or propyl cyanoacetate, ethanol, triethylamine, room temperature, overnight; (ii) Suzuki–Miyaura method: 4-bromobenzaldehyde, palladium(II) acetate, boronic acids, potassium hydroxide, dioxane : water (2 : 1), 100 °C, 48 h.

Table 1 Inhibition of ZIKV and DENV-2 NS2B/NS3 proteases by α-cyanoacrylates and α-cyanoacrylamides



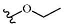
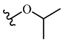
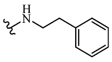
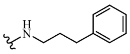
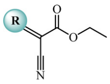
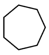
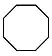
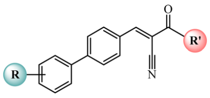
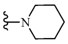
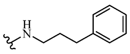
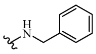
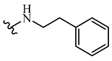
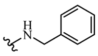
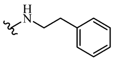
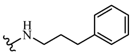
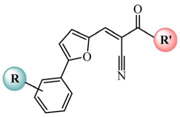
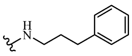
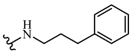
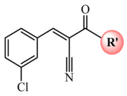
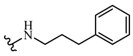
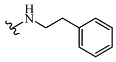
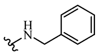
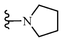
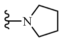
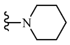
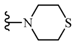
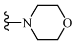
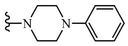
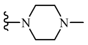
Code	R	R′	ZIKV IC_{50} ± SD^a (μM)/ % inh. ± SD^a/K_i ± SD^a (μM)	DENV-2 IC_{50} ± SD (μM)/ % inh. ± SD^a/K_i ± SD^a (μM)
LQM448	—		N.I.	N.I.
LQM449	—		N.I.	N.I.
LQM450	—		N.I.	N.I.
LQM451	—		N.I.	N.I.
				
LQM452		—	N.I.	N.I.
LQM453		—	N.I.	N.I.
LQM454		—	N.I.	N.I.
LQM455		—	N.I.	N.I.
				
LQM456	H		N.I.	N.I.
LQM457	H		N.I.	N.I.
LQM458	3,4-Di-Cl		15 ± 11%	N.I.
LQM459	3,4-Di-Cl		N.I.	N.I.
LQM460	4-CH_3		N.I.	N.I.
LQM461	4-CH_3		11 ± 8%	N.I.
LQM462	4-CH_3		N.I.	N.I.
				
LQM463	H		17 ± 9%	N.I.
LQM464	2-CF_3		N.I.	N.I.

Table 1 (*continued*)



Code	R	R′	ZIKV IC$_{50}$ ± SD^a (µM)/ % inh. ± SD^a/K_i ± SD^a (µM)	DENV-2 IC$_{50}$ ± SD (µM)/ % inh. ± SD^a/K_i ± SD^a (µM)
LQM465	3,4-Di-Cl		N.I.	N.I.
LQM466	3,4-Di-Cl		N.I.	N.I.
LQM467	3,4-Di-Cl		**20.0 ± 1.3 /15.1 ± 0.9**	**18.2 ± 1.0/7.79 ± 0.43**
LQM468	2-CF$_3$		17 ± 2%	N.I.
LQM469	3,4-Di-Cl		36 ± 2%	30 ± 4%
LQM470	3,4-Di-Cl		32 ± 6%	N.I.
LQM471	3,4-Di-Cl		**17.4 ± 1.2 /13.2 ± 0.8**	**21.1 ± 1.4/9.03 ± 0.60**
LQM472	3,4-Di-Cl		**20.9 ± 2.6/15.8 ± 2.0**	**24.4 ± 2.4/10.4 ± 1.0**
LQM473	3,4-Di-Cl		26 ± 2%	11 ± 3%
LQM474	3,4-Di-Cl		**26.8 ± 2.8 /20.3 ± 2.1**	**39.9 ± 2.4/17.0 ± 1.0**

^{a} IC$_{50}$ and K_i values are indicated by mean ± standard deviation (SD) from three independent measurements. Percentage inhibition was determined at 20 µM inhibitor concentration; N.I. means no inhibition.

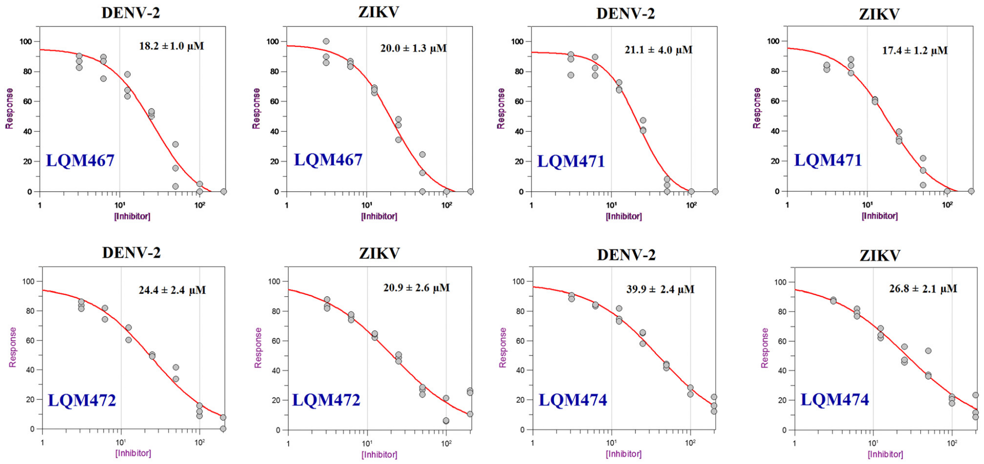


Fig. 2 Mean inhibitory concentration (IC$_{50}$) determination curves for **LQM467**, **471**, **472**, and **474** toward DENV-2 and ZIKV NS2B/NS3^{pro}. IC$_{50}$ values are indicated by mean ± standard deviation (SD) from three independent measurements.

concentration yields a straight line at whose intersection with the *y* axis the K_i can be determined as 9.1 µM.

To support this observation, a Dixon Plot was generated (Fig. 3). The position of the interface of the straight lines

Table 2 IC_{50} values for **LQM467** with different substrate concentrations

Substrate concentration [μM]	IC_{50} (DENV) [μM]
50	13.4 ± 0.9
100	18.2 ± 1.0
150	22.4 ± 2.0

resulting from the different substrate concentrations provides information about the binding mode of ligands. For the model compound **LQM467**, the intersection point is located in the 2nd quadrant of the coordinate system. This supports the assumption that **LQM467** is a competitive inhibitor. Thus, its K_i value could be determined from the Dixon plot as the intersection of trend lines, being 6.27 ± 0.72 μM. By applying the Cheng–Prusoff equation, the K_i value was found to be 7.79 ± 0.43 μM, which is in good accordance. Both the substrate dependence of inhibitory activity and the Dixon plot show that **LQM467** is a competitive inhibitor. To further confirm these findings, IC_{50} values of **LQM467** upon DENV NS2B/NS3^{pro} mutant T122C were determined before and after blockade with *N*-benzylmaleimide (BMI). These experiments are based on the fact that T122C is found located into the allosteric binding pocket of DENV-2 NS2B/NS3 protease. The blockade with BMI consequently leads to an attenuation of the inhibitory effect of allosteric inhibitors.[54] However, this was not observed for **LQM467**. Finally, IC_{50} values for DENV/T122C and DENV/T122C+BMI were found to be 10.3 ± 1.0 and 8.78 ± 0.68 μM, respectively. In summary, all experimental data confirm that **LQM467** is a competitive inhibitor binding to the active and not the allosteric site of the protease.

2.4 Structure–activity relationship (SAR)

The cyanocinnamyl derivatives turned out to have similar activities toward ZIKV NS2B/NS3^{pro} since DENV-2 and ZIKV proteases share 43.2% sequence identity (see Fig. S1 and S2 in ESI†). In general, all 3-chlorophenyl cyanoacrylates **LQM448–451**, and cyclic cyanoacrylates **LQM452–455** are completely inactive against both proteases. Also biphenyl analogs are completely inactive, except for compounds **LQM458** and **LQM461**, which displayed very weak inhibition. Concerning the 5-phenylfuran analogs, three of them were inactive against ZIKV NS2B/NS3^{pro} (**LQM464**, **LQM465**, and **LQM466**), and five against DENV-2 NS2B/NS3^{pro} (**LQM463–466**, **LQM468**, and **LQM470**). Compound **LQM463** presented poor inhibition of ZIKV NS2B/NS3^{pro} (17 ± 9% at 20 μM), while it was inactive against DENV-2 protease. When 2-trifluoromethyl or 3,4-dichloro groups are introduced at the 5-phenylfuran moiety, inactive molecules towards both proteases are obtained (**LQM464** and **LQM465**). When maintaining the 3,4-dichloro substituents and reducing the spacer group (methylene) between the terminal phenyl and amide function, a completely inactive molecule is generated as well (LQM466). Surprisingly, when this spacer group is even more reduced by removing another methylene group, a very active analog is obtained (**LQM467**), exhibiting K_i values of 15.1 ± 0.9 μM and 7.79 ± 0.43 μM against ZIKV and DENV-2 NS2B/NS3^{pro}, respectively. It was seen that **LQM467** presents hydrophobic interactions with Ser^{135} residue *via* closed contact with its benzyl group (see Section 2.8). In contrast, when this *N*-ethylphenyl group is replaced with a pyrrolidine (**LQM469**), its activity is drastically reduced (36 ± 2 and 30 ± 4% inhibition of ZIKV and DENV-2 NS2B/NS3^{pro} at 20 μM, respectively). Furthermore, the activity is even more reduced when the 3,4-dichloro groups are substituted by 2-trifluoromethyl (**LQM468**), inhibiting ZIKV NS2B/NS3^{pro} only by 17 ± 2%, while being completely inactive against DENV-2 protease at 20 μM. Modifying **LQM469** by replacing the pyrrolidine group with a piperidine substituent, a molecule (**LQM470**) with similar activity against ZIKV NS2B/NS3^{pro} is obtained (32 ± 6% at 20 μM), although it is completely inactive against DENV-2 NS2B/NS3^{pro}. By replacing the piperidine with an *N*-phenylpiperazine group, a less active analog (**LQM473**) is produced (26 ± 2% at 20 μM). The compound displayed moderate inhibition of DENV-2 NS2B/NS3^{pro} (11 ± 3% at 20 μM) when compared to its precursor (**LQM470**). When the *N*-phenylpiperazine moiety is replaced with an *N*-methylpiperazine group, an active compound is obtained (**LQM474**), exhibiting K_i values of 20.3 ± 2.1 μM and 17.0 ± 1.0 μM against ZIKV and DENV-2 NS2B/NS3^{pro}, respectively. Moreover, it was revealed that its methylpiperazine ring is found protonated at physiological conditions, allowing an electrostatic interaction with the negatively-charged Asp^{75} residue, which is associated with active compounds (see Section 2.8). Finally, when this *N*-methylpiperazine group is replaced with a thiomorpholine, an even more active compound is obtained (**LQM471**), exhibiting K_i values of 13.2 ± 0.8 μM and 9.03 ± 0.60 μM towards ZIKV and DENV-2 NS2B/NS3^{pro}, respectively. Its thiomorpholine ring is placed close to the Asp^{75} and His^{51} residues, performing hydrophobic interactions, while its furan ring is found near to Ser^{135} residue (see Section 2.8). When a classical bioisoster replacement is performed by using a morpholine ring, a compound with equivalent activity is generated (**LQM472**), displaying K_i values of 15.8 ± 2.0 μM and 10.4 ± 1.0 μM toward ZIKV and DENV-2 NS2B/NS3^{pro}, respectively. Moreover, this new ring is also found to be oriented to the Asp^{75} and His^{51} residues, as observed for its precursor (**LQM471**). In contrast, the morpholine

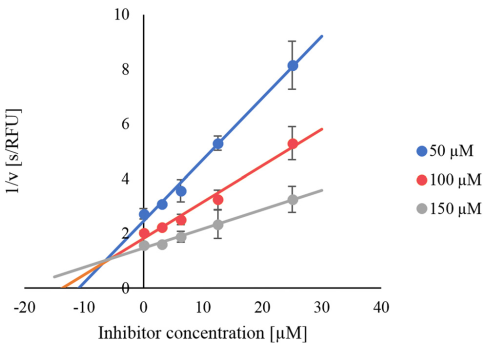


Fig. 3 Dixon plot for DENV-2 NS2B/NS3^{pro} with **LQM467**. The reciprocal turnover rate is blotted against the inhibitor concentration. Each linear regression fit represents the turnover rate for a single substrate concentration. The intersection of these trend lines shows the K_i value of 6.27 ± 0.72 μM.

ring prevents the interaction with Ser^{135} residue (see Section 2.8). Alltogether, it could be shown that cyclic non-aromatic amine substituents are promising fragments for the design of NS2B/NS3$^{\text{pro}}$ inhibitors.

2.5 Off-target selectivity

To investigate how the phenyl-5-furan inhibitors affect the activity of other serine and cysteine proteases, a selectivity panel was constructed involving **LQM467**, **471**, **472**, and **474** inhibitors and six different proteases, namely the serine proteases trypsin, urokinase plasminogen activator (uPA), human matriptase (MT), and the cysteine proteases cathepsin-L, SARS-CoV-2 M$^{\text{pro}}$, and PL$^{\text{pro}}$. Since NS2B/NS3$^{\text{pro}}$ is a trypsin-like protease, it is structurally mostly related to trypsin, but the latter does not have an NS2B cofactor and thus does not feature an allosteric binding pocket.[53] This fact explains the activity of all these acrylamide analogs upon trypsin since they can address the NS3$^{\text{pro}}$ active site.[83] Additionally, **LQM471** was active towards the cysteine protease cathepsin-L, although with a K_i value higher than that for ZIKV and DENV NS2B/NS3$^{\text{pro}}$. Besides, **LQM472** and **LQM474** also exhibited activity against uPA. Furthermore, none of these compounds displayed appreciable inhibition values at 20 μM concentration toward MT, SARS-CoV-2 M$^{\text{pro}}$, and PL$^{\text{pro}}$. All results associated with the protease inhibition selectivity are summarized in Table 3.

2.6 Molecular dynamics simulations

Molecular dynamics (MD) simulations can be utilized to study the time-dependent interaction and behavior of small molecules with the target.[84] In this context, this *in silico* approach has been frequently used in studies involving structural aspects and active compounds toward NS2B/NS3$^{\text{pro}}$.[85–90] Considering that **LQM467**, **471**, **472**, and **474** are competitive inhibitors, we decided to investigate their interactions with the catalytic site of DENV-2 NS2B/NS3$^{\text{pro}}$ in the closed conformation (PDB: 2M9P). In all MD simulations, the complexes were initially heated at 0 K and then continue up to 300 K for 2 ns. Then, these complexes were equilibrated at constant pressure (1 atm). Subsequently, 200 ns MD simulations for all inhibitors in complex with NS2B/NS3$^{\text{pro}}$ were performed to investigate their structural stability *via* the root mean square deviation (RMSD). Ramachandran plots were also generated to check the stereochemical quality of these complexes. Additionally, other parameters such as root mean square fluctuation (RMSF) and hydrogen-bonding interactions (H-bonds) were calculated to obtain insights into changes in the structural stability of the complexes at 300 K. Therefore, RMSF was used to understand the flexibility of NS2B/NS3$^{\text{pro}}$ by analyzing the alpha carbon (C$_\alpha$) atoms after interaction with the ligands, by analyzing individual residue flexibilities (or fluctuations).

In general, RMSD and RMSF values are interesting data that are used to check the stability and conformation of the protein and protein–ligand complex within the simulation time. Regarding the native protein, low RMSD values for the protein–ligand complex of less than 1.5 Ångström (Å) or even better less than 1 Å are acceptable, representing a trustworthy binding mode of a ligand into a cavity (or binding site) from a native protein.[84] Since all acrylamide inhibitors exhibited a competitive binding mode as observed by our experimental approaches, their results towards 2M9P will be discussed in detail. However, all results for Ramachandran plots for all protease–ligand complexes can be found in Fig. S3–S6 in the ESI.† For the 2M9P–**LQM467** complex, the Ramachandran plot revealed that the NS2B/NS3$^{\text{pro}}$ structure has 85.8% of its amino acids in the most favored regions, while 12.6% are found in additional allowed regions. Similar results were obtained for the 2M9P–**LQM471** complex, having 87.9% of its amino acids in the most favored regions, while 11.1% and 1.1% are found in the additional allowed and generously allowed regions, respectively. The 2M9P–**LQM472** complex displayed 88.9% of its amino acids in the most favored regions, while the other 10.5% amino acids are found in additional allowed regions. Moreover, the 2M9P–**LQM474** complex presented 80% of its amino acids placed in the most favored regions and 17.9% in additional allowed regions. Finally, none of these inhibitors complexed with 2M9P present any residues found in disallowed regions (outliers), suggesting that the force field used was efficient to simulate the native environment. RMSD plots for the 2M9P–**LQM467** complex (Fig. 4A) predict structural stability during the whole simulation time, with no meaningful changes observed. Initially, the protein RMSD value gradually increases at the first 35–40 ns and slightly decreases during 50–55 ns. Then, it continuously raises at 56–100 ns and maintains its stability from 110 to 190 ns. However, after 175 ns it increases again and maintains a variation of ∼1.3 Å till the end of the simulation time, similar

Table 3 Protease inhibition selectivity of **LQM467**, **471**, **472**, and **474** towards other serine and cysteine proteases

	% inhibition at 20 μM/IC$_{50}$ ± SD^a (μM)/K_i ± SD^a (μM)			
Protease	**LQM467**	**LQM471**	**LQM472**	**LQM474**
Trypsin	80%/**12.1 ± 0.9/7.8 ± 0.09**	54%/**30.2 ± 3.3/19.5 ± 2.1**	67%/**22.7 ± 2.2/14.6 ± 1.4**	74%/**27.3 ± 3.1/17.6 ± 2.0**
uPA	31%	27%	48%/**23.8 ± 2.0/13.1 ± 1.1**	68%/**22.1 ± 2.9/12.2 ± 1.6**
MT	31%	33%	24%	23%
Cathepsin-L	33%	57%/**24.5 ± 2.3/12.4 ± 1.2**	29%	32%
SARS-CoV-2 M$^{\text{pro}}$	N.I.	N.I.	7%	8%
SARS-CoV-2 PL$^{\text{pro}}$	24%	29%	26%	38%

[a] Enzymatic inhibition values at 20 μM compound concentrations are indicated by mean ± standard deviation from three independent measurements. IC$_{50}$ values were only determined when the compound presented % inhibition ≥ 40. The enzyme activity is expressed relative to the DMSO control in [%]. N.I. means no inhibition.

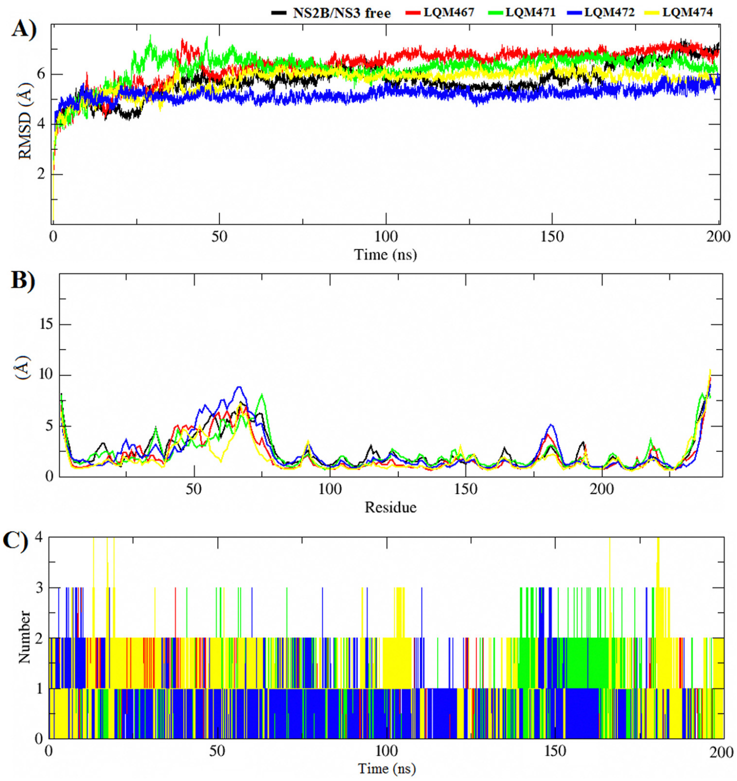


Fig. 4 Trajectories for (A) RMSD, (B) RMSF, and (C) H-bonds obtained by molecular dynamics simulations for 200 ns at 300 K of free NS2/NS3 protease and protease complexed with **LQM467**, **471**, **472**, and **474** inhibitors.

to a study performed by Nutho *et al.* (2019).[85] However, this value has been considered acceptable according to Kumar *et al.* (2021).[84] The 2M9P–**LQM467** RMSD presented a similar stability profile as the free NS2B/NS3^{pro} (Fig. 4A), except that no increase was observed at the first 35–40 ns. These findings suggest that this complex is likely to be stable ranging from 100 to 175 ns. Moreover, 2M9P–**LQM471** (Fig. 4A) displayed a similar stability profile as the corresponding **LQM467** complex during the whole simulation time (variation of ~1.5 Å). However, its higher stability was observed from 75 to 180 ns. Regarding the 2M9P–**LQM472** and 2M9P–**LQM474** complexes, which exhibit similar IC_{50} values, also similar RMSD profiles (Fig. 4A) were found mainly from 50 to 200 ns. In contrast, 2M9P–**LQM472** exhibited a lower RMSD variation (~0.25 Å) after 50 ns. Regarding structural fluctuations verified by RMSF plots, it is possible to verify that **LQM467**, **471**, **472**, and **474** inhibitors induce meaningful structural changes in residues numbered from 50 to 75 when compared to NS2B/NS3 free protease (Fig. 4B), suggesting that these compounds perform similar binding modes at the catalytic site of NS2B/NS3^{pro}. In contrast, **LQM467** induces minimal structural changes on the NS2B/NS3^{pro}, having a similar profile as free protease. On the contrary, it is observed that **LQM472** induces the highest structural changes in His^{51} and Ser^{135} residues, with 5.3 and 8.8 Å, respectively. This suggests that the steric volume of **LQM472** induces relevant changes at the catalytic site of NS2B/NS3^{pro} when compared with the free protease. Moreover, all compounds perform changes in Trp^{89}, Thr^{120}, and Val^{147} residues, which have been associated with potential antiviral agents.[46,50] For binding recognition and NS2B/NS3-mediated cleavage of individual substrates, intermolecular H-bonds located between the substrate and surrounding amino acid residues at the binding site are extremely important.[85] Concerning the H-bonds constituted in the most stable complexes within the simulation time (100–175 ns), it was verified that **LQM467**, **471**, and **474** preferably perform two H-bonds, while **LQM472** complexes involve only one H-bond (Fig. 4C). However, complexes involving three or four H-bonds can be visualized within 200 ns simulation, although these are scarce. All hydrophobic and H-bond interactions of the inhibitors will be further discussed in detail (see Section 2.8). Finally, the solvent-accessible surface area (SASA) and radius of gyration (R_g) plots revealed a constant aspect for the formation of all complexes, exhibiting values of ~50–75 Å^2 and ~15–16 Å, respectively. This fact suggests that all these compounds are similarly placed into the binding site and their interactions are not capable of altering the structural compactness of the target (see Fig. S7 and S8 for SASA and R_g plots, respectively, in the ESI†).

2.7 Molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) calculations

The MM/PBSA approach combines three energetic terms to obtain changes in the free energy of binding. The first energetic term is associated with a change in the potential energy in the vacuum, including both bonded (such as bonds, angles, and torsion energies) and non-bonded (van der Waals and electrostatic interactions) terms; the second term is related to the desolvation of chemical species, which quantifies the sum of two energetic terms (non-polar and polar solvation energies using implicit solvation model).[91–94] Finally, a third term that is capable of obtaining configurational entropy energy related to the complex formation in the gas phase, is used.[95] Thus, changes in free binding energies or binding affinities are used to determine potential molecules, predicting potential inhibitors based on their binding strength with the target. MM/PBSA is a computationally efficient approach, which is broadly used for the estimation of the accurate relative free binding energies.[84,95] It incorporates conformational fluctuations and entropic contributions to binding energies of ligands, after dynamics simulations.[95,96] To estimate the free binding energies of **LQM467**, **471**, **472**, and **474** in complex with the catalytic site of NS2B/NS3^{pro}, the MM/PBSA approach was used in this study, in which the non-polar solvation term is obtained by the solvent-accessible surface area (SASA).[85] All corresponding energy contributions (Table 4) were estimated by using 50 ns from the most stable fragment of their MD trajectories (100–150 ns) after the cluster analyses of all the complexes.

In general, MM/PBSA calculations revealed that the α-acrylamide derivatives exhibit a similar affinity for the catalytic site of NS2B/NS3^{pro}, considering all energetic terms calculated in this study. Concerning the van der Waals energy term (ΔE_{vdW}), **LQM472** presents more favorable hydrophobic interactions with the catalytic site of the NS2B/NS3 protease among the other analogs, even having a different hydrophobic volume than **LQM467**. It suggests that the morpholine group could efficiently perform van der Waals interactions with NS2B/NS3^{pro}. Additionally, negative values of ΔE_{vdW} favor the stabilization of the ligand–protease complex.[85] Differently, **LQM474** demonstrated the most favorable ΔE_{elec} value, suggesting the *N*-methylpiperazine substituent could be more effective to perform interactions with charged residues since it is found protonated at physiological pH. Nevertheless, the highly unfavorable contributions of polar solvation free energy (ΔG_{solv}^{polar}) for all inhibitors were calculated using a PB model. Meanwhile, van der Waals interactions and non-polar solvation free energy ($\Delta G_{solv}^{non\text{-}polar}$) are favorable contributions to the total free binding energy for all complexes. Moreover, it was observed that all of these inhibitors display similar $\Delta G_{solv}^{non\text{-}polar}$ values for the catalytic site, as expected since these compounds have closely-related chemical structures. However, **LQM474** exhibits a slightly higher value of $\Delta G_{solv}^{non\text{-}polar}$ when compared to the other derivatives, suggesting that it is less solvent accessible during the complex formation. The contrary is observed for **LQM471**, which displays the most favorable $\Delta G_{solv}^{non\text{-}polar}$ term. Moreover, the polar contribution (ΔG_{solv}^{polar}) plays a significant role for inhibitors' binding since negatively charged residues can interact with the inhibitors *via* attractive electrostatic interactions and hydrogen bonds.[85] In addition, the higher values for the solvation term (ΔG_{solv}) indicate that 2M9P is highly solvated, with similar values for all complexes. Concerning the binding energy of protease–inhibitor complexes (ΔE_{MM}), it was found that **LQM472** presents the highest preference for the NS2B/NS3^{pro} catalytic site among its analogs. However, for inhibitors investigated by using the MM/PBSA approach, it could be assumed that all these compounds display almost equivalent binding energy values, corroborating our experimental results, which suggest these are competitive inhibitors toward NS2B/NS3^{pro} with similar inhibition values. By considering the ΔE_{MM} term, the main contribution to the total free binding energy of all complexes is the hydrophobic interaction (ΔE_{vdW}). Yotmanee *et al.* (2015)[86] obtained similar values when studying NS2B/NS3 polypeptide substrates by using molecular mechanics energies combined with a generalized Born and surface area continuum solvation (MM/GBSA) approach. Lastly, it is worth noting that all free binding energies obtained in our study by using MM/PBSA calculations do not provide an absolute energy value compared to the experimental results since the free binding energy is related to the difference between the large desolvation energy and the protein–ligand complex energy, but these are still useful for establishing trends of relative binding

Table 4 Different MM/PBSA terms for the relative change in the enthalpy of NS2B/NS3 catalytic site in complex with **LQM467**, **471**, **472**, and **474** inhibitors at 300 K

Energy term	**LQM467** (kcal mol^{-1}) ± SD	**LQM471** (kcal mol^{-1}) ± SD	**LQM472** (kcal mol^{-1}) ± SD	**LQM474** (kcal mol^{-1}) ± SD
ΔE_{vdW}	−28 ± 1	−36 ± 1	−39 ± 2	−34 ± 1
ΔE_{elec}	−3 ± 1	−13 ± 1	−13 ± 3	−14 ± 3
$\Delta G_{solv}^{non\text{-}polar}$	−21 ± 0	−19 ± 0	−20 ± 0	−22 ± 0
ΔG_{solv}^{polar}	2188 ± 37	2576 ± 43	2444 ± 42	1564 ± 37
ΔG_{solv}	2210 ± 37	2595 ± 44	2465 ± 42	1587 ± 38
ΔE_{MM}	−42 ± 1	−52 ± 2	−53 ± 4	−49 ± 2

Catalytic site: 2M9P; vdW: van der Waals; Elec: electrostatic; SD: standard-deviation. ΔG_{solv} values were determined by the equation $\Delta G_{solv} = (\Delta G_{solv}^{polar} + \Delta G_{solv}^{non\text{-}polar})$, G_{polar}, and $G_{non\text{-}polar}$ are the electrostatic and non-electrostatic contributions to the solvation free energy respectively, in accordance with Kumari *et al.*[95]

efficiency of inhibitors towards NS2B/NS3^{pro}.[85] Notwithstanding this challenging task, these large values could be responsible for minimal errors in the method, leading to the possible artifacts on the binding energy values.[85,97,98]

2.8 Molecular docking studies

Molecular docking has been considered an essential and powerful tool for drug design and discovery, which has resulted in promising inhibitors targeting proteins of medicinal interest.[99–103] The goal of the docking studies is to predict the predominant binding pose(s) of a ligand into a protein, performing searches in high-dimensional spaces effectively by applying different scoring functions to rank hit compounds.[104] Unfortunately, the protein targeted is generally considered as a rigid structure. Thus, it is necessary to consider molecular dynamics simulations for this macromolecule in order to improve the results.[105] Regarding this fact, the most stable ligand–NS2B/NS3 complexes obtained after the dynamics simulations were assumed as starting points for our docking studies. The cluster analysis revealed that all acrylamide inhibitors present a similar binding pose into the catalytic site of NS2B/NS3^{pro}, except **LQM467** which is placed in a different orientation (Fig. 5A). Still, **LQM471** and **472** have a similar binding mode, which probably could be associated with their similar K_i values (see Table 1). **LQM467** and **LQM474** seem to be more solvent-accessible, being in accordance with the results obtained by MM/PBSA calculations when considering the $\Delta G_{solv}^{non\text{-}polar}$ values. In general, all inhibitors hydrophobically interact with the Leu^{128} residue, which – in accordance with Qamar *et al.* 2016[106] – is involved in interactions with promising NS2B/NS3^{pro} inhibitors. Also, interactions with Gly^{151}, Gly^{153}, and Pro^{132} have been reported in the interactions of inhibitors with this protease.[107] Regarding the interactions of the inhibitors, **LQM467** presents seven hydrophobic interactions, including the catalytic triad (Fig. 5B). Moreover, **LQM471** performs five hydrophobic interactions (Fig. 5C). Additionally, its thiomorpholine group is involved in hydrophobic interactions, as evidenced by the MM/PBSA calculations. Similar interactions are observed for **LQM472** (Fig. 5D). However, interactions with Ser^{135} were not verified. Surprisingly, **LQM474** also does not exhibit interactions with Ser^{135}. Electrostatic interactions between the protonated nitrogen at the piperazine ring and the Asp^{75} residue were observed at a distance of 2.61 Å (Fig. 5E), corroborating with the MM/PBSA calculations.

2.9 Evaluation of *in vitro* cytotoxicity and antiviral activity against ZIKV

The evaluation of *in vitro* cytotoxicity of **LQM467**, **471**, **472**, and **474** was performed on *Vero E6* cells at 25, 50, 100, and 200 μM concentrations to find the maximum non-toxic concentrations (MNTCs) of these compounds. The MNTCs for **LQM467**, **471**, and **472** were found to be 50, 100, and 25 μM, respectively. In contrast, **LQM474** was significantly cytotoxic at the lowest tested concentration (25 μM), in addition to solubility issues at conditions for testing. Therefore, further investigations were not performed for this compound. Thus, the 50% cytotoxic concentration (CC_{50}) was found to be $>$ 200 μM for **LQM467** and $>$ 100 μM for both **LQM471** and **LQM472**. On ZIKV-infected cell-based antiviral assays, virus adsorption was performed on *Vero E6* cells followed by the treatment with **LQM467**, **471**, and **472** in several concentrations starting from their corresponding MNTCs, in which their percentage of viral inhibition was assessed after 52 hours. **LQM467** and **LQM472** did not achieve a 50% inhibition value for the MNTC tested for each compound (22.2 ± 8.6% at 50 μM/MNTC for **LQM467**, while 42.0 ± 7.2% at 25 μM/MNTC for **LQM472**). In contrast, significant antiviral activity was detected for **LQM471** (80.0 ± 10.9% at 100 μM/MNTC) (Fig. 6A). Moreover, **LQM471** (100 μM) remarkably reduced the ZIKV-induced cytopathogenic effect when compared

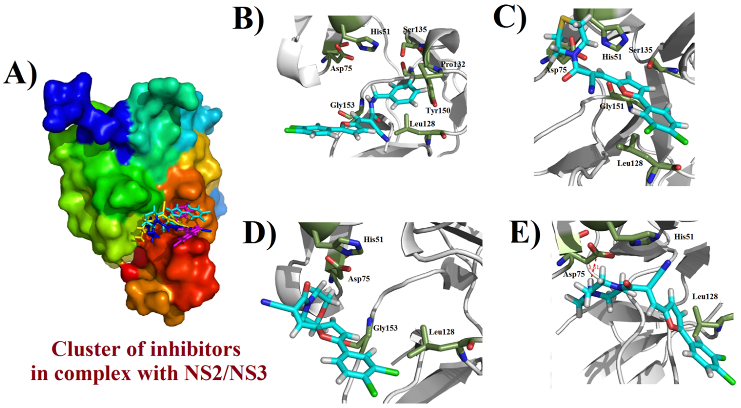


Fig. 5 The cluster of inhibitors placed into the catalytic site of NS2B/NS3^{pro} (A) and their interactions with the amino acid residues (B–E). The structure of the protease is represented using the surface model and colored in rainbow style. In (A): **LQM467** (magenta), **LQM471** (cyan), **LQM472** (blue), and **LQM474** (yellow). In (B): **LQM467** interactions; In (C): **LQM471** interactions; In (D): **LQM472** interactions; In (E): **LQM474** interactions. Electrostatic interaction is represented as a red dashed line, with distance computed in Ångström (Å). Illustration elaborated by using PyMol® v.0.99.

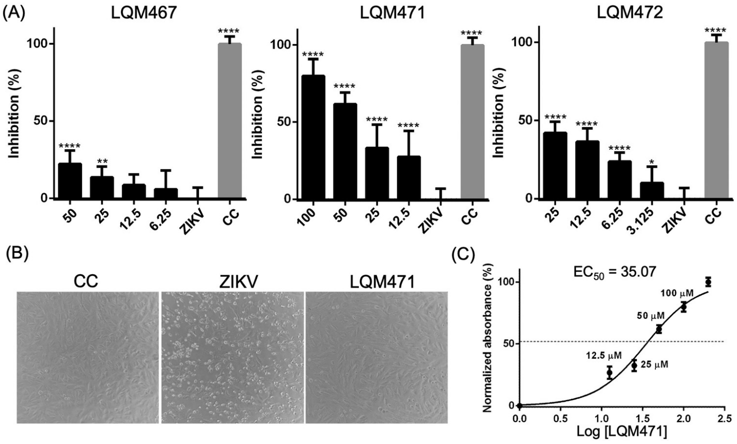


Fig. 6 *In vitro* evaluation of antiviral activity of **LQM467**, **LQM471**, and **LQM472** on ZIKV-infected Vero E6 cells. The virus adsorption was performed for 90 min followed by the addition of each compound in several concentrations starting at MNTCs. (A) The viral inhibition (%) was assessed for each compound after 52 h. The plotted values were the mean ± SD of three independent experiments performed in triplicate. * $p \leq 0.05$; ** $p \leq 0.01$; **** $p \leq 0.0001$ compared to ZIKV. (B) Micrographs showing the ZIKV-induced cytopathogenic effect, such as cell refringence changes and aggregates (200× magnification). ZIKV = infected untreated cells. CC = uninfected cellular control. (C) The EC_{50} (concentration of the compound that promotes 50% of viral inhibition in the MTT assay) for **LQM471** was calculated by using nonlinear regression curve fitting.

to untreated cells. The EC_{50} (concentration of the compound that promotes 50% of viral inhibition) value was found to be 35.1 μM. To confirm this promising anti-ZIKV activity of **LQM471**, the viral load was quantified in the supernatants of both treated and untreated ZIKV-infected *Vero E6* cells by qPCR. As result, untreated ZIKV-infected cells showed a viral RNA (vRNA) concentration of 3.9×10^4 copies per μL in the qPCR reaction, while the treatment with **LQM471** (100 μM/MNTC) notably reduced the viral load under the lowest point of the standard curve (lower than 7.5×10^2 copies per μL, Fig. S9 in the ESI†), reinforcing that its mechanism of action could be the inhibition of viral replication.

3. Conclusions

In total, 27 compounds were virtually designed by a fragment-based drug design approach, targeting the ZIKV and DENV NS2B/NS3^{pro}. The target structures were then synthesized by aminolysis and/or Knoevenagel reactions. In general, four 5-phenylfuran-containing α-cyanoacrylamides (**LQM467**, **471**, **472**, and **474**) were found to be active against both serine proteases. Moreover, the Dixon plot revealed that these chemical entities exhibit a competitive inhibition mode. Posteriorly, it was observed that these active compounds could be active against other proteases, primarily against trypsin, which has a similar binding pocket. Cathepsin-L and uPA proteases were also inhibited by **LQM471** and **474**. Dynamics simulations at 200 ns showed that all NS2B/NS3^{pro}–ligand complexes present good stabilities. **LQM472** induced more conformational alterations in the binding site, verified in the RMSF plot. MM/PBSA calculations suggested that these compounds present similar binding energies toward the NS2B/NS3^{pro} catalytic site. In this context, it was verified that the ΔE_{vdW} term represents the main energetic contribution to the complexes' stabilization. Docking studies displayed that these compounds exhibit similar binding modes and also similar interactions. Further investigations demonstrated interactions with His^{51}, Asp^{75} as well as Leu^{128} for all inhibitors. For **LQM472**, it was observed that the *N*-methylpiperazine group is found protonated at physiological pH and it can perform an electrostatic interaction with Asp^{75}. Cytotoxicity assay demonstrated that **LQM474** is toxic even at the minimal tested concentration (25 μM). In contrast, **LQM467**, **471**, and **472** were investigated for their antiviral activities toward ZIKV-infected cells, in which was displayed that only **LQM471** exhibits promising results, showing an EC_{50} value of 35.1 μM. Further investigation by using qPCR revealed that this compound is capable of decreasing the viral RNA copy number compared to the untreated ZIKV-infected cells. These findings suggest that this compound can be a promising inhibitor of viral replication into cells. In summary, this study presents an interesting alternative to design new hit compounds against DENV-2 and ZIKV NS2B/NS3^{pro}, which could be further optimized to improve their inhibitory activities, increasing the potential therapeutic arsenal against these flaviviruses in the future.

4. Experimental section

4.1 Protein preparation

DENV-2 and ZIKV proteases were expressed in *Escherichia coli* BL21-Gold (DE3) (*Agilent Technologies*) cells, as reported by Hammerstein *et al.* (2019).[108] Cells were grown in LB media containing 100 µg mL^{-1} ampicillin at 37 °C. The overexpression was induced by adding 1 mM IPTG at an OD_{600} of ~0.8 nm and incubation at 20 °C for 16 h. Cells were harvested by centrifugation, shock frozen in liquid nitrogen, and stored at −20 °C. For purification, cells were resuspended in lysis buffer (20 mM Tris–HCl pH 8, 300 mM NaCl, 2 mM imidazole, 0.1% (v/v) TritonX-100, RNase, and lysozyme) and lysed by sonication (*Sonopuls, Bandelin*). After centrifugation to remove cell debris, affinity chromatography using a HisTrap HP column (*GE Healthcare*, Chicago, USA) was used to isolate hexahistidine-tagged protein from the supernatant. Subsequently increased amounts of elution buffer (20 mM Tris–HCl pH 8, 300 mM NaCl, 250 mM imidazole) were used to remove histidine-rich proteins and other impurities. For further purification, fractions containing DENV-2 or ZIKV protein were subjected to size exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75 pg column (GE Healthcare, Chicago, USA) in SEC buffer (20 mM Tris–HCl pH 8, 150 mM NaCl). Lastly, prior to storage at −80 °C, proteins were concentrated, and flash-frozen with liquid nitrogen. All these procedures are in agreement with our recent publications.[53,54]

4.2 Enzyme inhibition assays

To investigate the inhibitory activity of the compounds against DENV-2 and ZIKV proteases, an assay based on the fluorogenic substrate Boc-Gly-Arg-Arg-AMC (Bachem, Bubendorf, Switzerland) was performed. The tested compounds and substrates were prepared as DMSO-stock solutions. Subsequently, these compounds were plated in white flat-bottom 96-well microtiter plates, being 90 µL buffer, 2.5 µL enzyme solution, 5 µL inhibitor in DMSO (or pure DMSO as control), and 2.5 µL solution of the corresponding substrate, totalizing 100 µL per well. All fluorescence values were measured using a Tecan Infinite F2000 PRO plate reader. In general, all results were obtained and analyzed from technical triplicates. A concentration of 20 µM for the inhibitors was chosen for initial screenings. Then, IC_{50} values were determined with dilution series ranging from 0.01 and 100 µM. The fluorescence was measured every 30 seconds for 10 min at 25 °C with 380 nm for excitation and 460 nm as the emission wavelength. IC_{50} values were calculated using GraFit® v. 6.0.12 (Erithacus Software Limited, East Grinstead, West Sussex, UK)[109] by fitting the remaining enzymatic activity to four-parameters IC_{50} eqn (1):

$$Y = \frac{Y_{\max} - Y_{\min}}{1 + \left(\frac{[\mathrm{I}]}{\mathrm{IC}_{50}}\right)^{s}} + Y_{\min} \qquad (1)$$

where Y [ΔF/min] is the substrate hydrolysis rate, $Y_{\max}$ represents the maximum value of the dose–response curve, measured at inhibitor concentrations of [I] = 0 µM, $Y_{\min}$ is the minimum value, obtained at high inhibitor concentrations, and s representing the Hill coefficient.[110] Furthermore, in order to elucidate if the compounds are competitive or non-competitive, we performed IC_{50} measurements of **LQM467** in presence of several substrate concentrations (50, 100, 150, and 200 µM), generating a Dixon plot. To verify the findings of the Dixon plot, we compared the K_i obtained from the dissection point of the Dixon plot with the K_i value determined by the Cheng–Prusoff equation (eqn (2)) (K_M(Boc-Gly-Arg-Arg-AMC) = 74.8 µM); where K_i is the constant of inhibition, K_m is the Michaelis–Menten constant for the substrate, and [S] represents the concentration of the substrate.[111–113] Additionally, its IC_{50} value against DENV mutant T122C was evaluated before and after blockade with *N*-benzylmaleimide (BMI). Finally, all these procedures are in agreement with our recent publications.[53,54]

$$K_i = \frac{\mathrm{IC}_{50}}{1 + \frac{[\mathrm{S}]}{K_m}} \qquad (2)$$

For off-target selectivity, similar approaches for determining inhibitory potency at 20 µM were applied. Noteworthy, in the case of SARS-CoV-2 M^{pro}, a FRET-based substrate was applied.

Buffers and substrates. SARS-CoV-2 M^{pro} (125 nM enzyme, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 µM Dabcyl-KTSAVLQSGFRKME-Edans),[114] DENV-2 NS2B/NS3 (250 nM enzyme) and ZIKV NS2B/NS3 (50 nM enzyme) (50 mM Tris pH 9.0, 20% (v/v) glycerol, 1 mM Chaps, 100 µM Boc-Gly-Arg-Arg-AMC),[115] uPA (2.5 U enzyme, 50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 µM *Z*-Gly-Gly-Arg-AMC),[116] trypsin (250 ng mL^{-1} enzyme, 50 mM Tris pH 8.0, 100 mM NaCl, 5 mM EDTA, 40 µM *Z*-Phe-Arg-AMC),[117] MT2 (1.25 nM enzyme, 25 mM Tris pH 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 0.01% Triton-X100, 100 µM Boc-Leu-Arg-Arg-AMC),[118] and cathepsin-L (220 ng mL^{-1} enzyme, 50 mM Tris pH 6.5, 5 mM EDTA, 200 mM NaCl, 2 mM DTT, 6.25 µM *Z*-Phe-Arg-AMC),[119] and SARS-CoV-2-PL^{pro} (75 nM enzyme, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 µM *Z*-Arg-Leu-Arg-Gly-Gly-AMC).[120]

4.3 Fragment-based design of inhibitors targeting NS2B/NS3 and molecular docking procedures

All molecular docking simulations were performed in a Dell® notebook, (Texas, USA), model 5500U, with an Intel® CoreTM 5th generation *i*-7 processer, CPU 2.4 GHz, 16 GB RAM, and running at Windows® 8.1 platform (Redmond, USA). Fragment-based drug design (FBDD) has emerged as an interesting, promising, and effective alternative to high-throughput screening (HTS) for the identification of bioactive compounds in drug discovery in the past decades.[60,121,122]

In total, 340 small in-house chemical fragments were drawn, converted into three-dimensional structures, and energetically minimized by the application of the semi-empiric method Austin Model 1 (AM1), by using ArgusLab® v. 4.0.1 (Richland, USA) (**https://www.arguslab.com**).[123] 3D-structures of open-conformation DENV-2 NS2B/NS3 (PDB: 2FOM),[44] and closed-conformation of DENV-2 NS2B/NS3 (PDB: 2M9P)[124] proteins were obtained at the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, San Diego, USA), website (**https://www.rcsb.org**). Pretreatment of these macromolecules and molecular docking simulations were performed using GOLD® v. 5.8.1 software (Cambridge, UK) (**https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/**).[125] Subsequently, all co-crystallized ligands were removed, including water molecules and ions. Then, H-bond acceptors and donors were assumed as solvent accessible. Additionally, Chemical Piecewise Linear Potential (CHEMPLP) was used as a scoring function for our protocol. Furthermore, all intermolecular interactions (H-bond, hydrophobic, and van der Waals) were individually analyzed by using AutoDock Tools v. 1.5.6 (San Diego, USA) (**https://ccsb.scripps.edu/mgltools/**).[126] All illustrations were generated by using PyMol® software, v. 0.99 (**https://pymol.org/2/**).[127] Moreover, ligand–NS2B/NS3 docking complexes were used in molecular dynamics simulations to determine the complexes' stabilities. Lastly, all these procedures performed in this study are in accordance with recently published works by our research team involving *in silico* approaches.[101,128–134] Primary and secondary amines, as well as, aldehydes, carboxylic acids, and cyano analogs emerged as promising fragments. Thus, those fragments presenting FitScore values ≥ 35 were selected for performing fragment growing approach, by combining them into molecules. Also, in full agreement with the rule of three for fragments ($M_W \leq 300$ Da, H-bond donor ≤ 3, H-bond ≤ 3, and $\log P \leq 3$),[135,136] and some additional parameters, such rotatable bonds ≤ 3 and polar surface are (PSA) ≤ 60 Å^2.[121] In sense, the molecules with FitScore values ≥ 65.0 were filtered by applying a step using a Pan-Assay Interference Scaffold (PAINS) filter, implemented by *SWISSADME* (**https://www.swissadme.ch/index.php**).[137] Finally, the most promising non-PAIN molecules were selected for synthesizing and further investigations.

4.4 Molecular dynamics simulations

All molecular dynamics simulations were performed in an Asus® desktop (Taipei, Taiwan), with a Xeon® Core 2nd generation processer, CPU 3.2 GHz, 8 GB RAM, Video card Asus TUF NVIDIA® GeForce RTX 2060 – 6GB GDDR6 192 bits Dual, 1920 NVIDIA CUDA® cores, and running at Linux® platform. Then, NS2B/NS3 in complex with **LQM467**, **471**, **472**, and **474** complexes were initially obtained by molecular docking studies, as aforementioned. Subsequently, all water molecules were removed, while charges and hydrogens were added, using the *DockPrep* module from the *Chimera*® software.[138] Then, the CHARMM27 force field was employed, while TIP3P was selected for the solvation model.[139] Ligands' topologies were generated by using the *SwissParam*® (**https://www.swissparam.ch/**) webtool.[140] Thus, a 1.0 nm triclinic box was created, adding water, Na^+ and Cl^- ions at the physiological concentration (0.15 M). The system was then initially equilibrated in 20 000 steps by the conjugate gradient method, followed by the system's total minimization in 100 000 steps. Thus, the initial system was energetically minimized to relax potential steric clashes in the structure, and then it was equilibrated. *NVT* (constant Number of particles, volume, and temperature) and *NPT* (constant number of particles, pressure, and temperature) equilibrations were performed at 300 K, during 2 ns. Posteriorly to the system assembled, 200 ns molecular dynamics (MD) simulations were performed using the *GROMACS*® v. 2018.3 software,[141] with the free-protein and then ligands-complexed. Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), solvent-accessible surface area (SASA), radius of gyration (R_g), and hydrogen bonds (H-bond) were calculated, in which plots were generated using the *Xmgrace*® software.[142] In this context, RMSD plots were calculated by using eqn (3), as shown below:

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i^m - X_i^1)^2 + (Y_i^m - Y_i^l)^2 + (Z_i^m - Z_i^l)^2} \tag{3}$$

where N is the number of atoms, X^m, Y^m, and Z^m are Cartesian coordinates of the initial structure, while X^l, Y^l, and Z^l are the coordinates for the trajectory at a frame t. Regarding the RMSF plot, it is obtained by the following eqn (4):

$$\mathrm{RMSF} = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(X_i - \bar{X})^2} \tag{4}$$

where T is the number of trajectory frames and $\bar{X}$ is the time of averaged position. SASA is a widely used non-polar method. In that case, it is based on the fact that the SASA parameter depends on linearly on the $G_{\text{non-polar}}$ term, calculated by eqn (5):

$$G_{\text{non-polar}} = \gamma A + b \tag{5}$$

where, γ is a coefficient associated with the surface tension of the solvent, while A is SASA and b is the fitting parameter.[93,94,143–146] The R_g parameter is a rough measure for the compactness of a structure, which can be calculated as shown by eqn (6):

$$R_g = \left(\frac{\sum_i \|r_i\|^2 m_i}{\sum_i m_i}\right)^{\frac{1}{2}} \tag{6}$$

where m_i represents the mass of atom i and r_i the position of this atom with respect to the center of mass of the molecule.[147] Furthermore, all specified donor–hydrogen acceptor triplets or H-bonds observed during the complex formation between NS2B/NS3 proteases and the most active ligands within the simulation time were determined. The lifetime of H-bonds was estimated from the average over all autocorrelation functions of the existing functions of all H-bonds,[147] as described in eqn (7):

$$C(\tau) = \langle S_i(t)S_i(t+\tau)\rangle \tag{7}$$

where $S_i(t) = \{0,1\}$ for H-bond i at time t, while $C(\tau)$ represents the complete autocorrelation function, in which it provides a rough estimate of the average H-bond lifetime τ_{HB},[147] as shown in eqn (8):

$$\tau\mathrm{HB} = \int_0^{\infty} C(\tau)\mathrm{d}\tau \tag{8}$$

Moreover, *UCSF Chimera*® software was used to perform the cluster analyzes for each complex to obtain the most stable complex structures for further investigations (see ESI†). This protocol is in agreement with others previously published works from our research team.[148–150] Finally, the stereochemical quality of NS2B/NS3 was determined by Ramachandran plots using the web software *SAVES* (**https://saves.mbi.ucla.edu/**) by applying the *PROCHECK* tool.[151]

4.5 MM/PBSA calculations

The MM/PBSA method has been reported in studies involving the prediction of binding free energies, allowing the evaluation of the relative stabilities of different ligands and targets.[144,145,152–155] In this respect, MM/PBSA has been applied in high-throughput molecular docking and MD simulations studies in order to remove false positives and to distinguish between inactive and active lead compounds by re-scoring their energetic profiles.[156–161] Herein, MD simulations were used to obtain an ensemble of binding modes for **LQM467**, **471**, **472**, and **474** in the presence of explicit water, in which the MM/PBSA method was employed to estimate their accurate binding energies. To perform MM/PBSA calculations, *g_mmpbsa* package was used *via GROMACS*, which includes several python scripts to perform all final statistical analyzes of energetic terms, being able to determine the energetic contributions of each amino acid residue in the protein–ligand complex. More detail associated with *g_mmpbsa* package can be found available at **https://rashmikumari.github.io/g_mmpbsa**. The binding energy can be decomposed on a per residue basis by using *g_mmpbsa* tool, in which the energy components E_{MM} (eqn (9)),[162–164] G_{polar} (eqn (10)),[155,165,166] and $G_{\text{non-polar}}$ (eqn (11))[167–169] of individual atoms are initially calculated in bonded and unbonded forms, as well as, their contribution to the binding energy ΔR_x^{BE} of residue *x*. These correlations are displayed in eqn (12), where, A_i^{bound} and A_i^{free} are the energetic terms of *i* atom from residue *x* in the bonded and unbonded forms, respectively.[95]

$$E_{\mathrm{MM}} = E_{\mathrm{bonded}} + E_{\mathrm{unbonded}} \tag{9}$$

$$\nabla[\varepsilon(r)\nabla\varnothing(r)] - \varepsilon(r)\kappa(r)^2\sinh[\varnothing(r)] + 4\pi\rho^f(r)/kT = 0 \tag{10}$$

$$G_{\text{non-polar}} = G_{\mathrm{cavity}} + G_{\mathrm{vdW}} \tag{11}$$

$$\Delta R_x^{\mathrm{BE}} = \sum_{i=0}^{n}\left(A_i^{\mathrm{bound}} - A_i^{\mathrm{free}}\right) \tag{12}$$

Then, the determination of energy components (van der Waals and electrostatic energies) of each complex, including also free protease and ligands energies, were calculated for snapshots extracted every 0.5 ns from the most stable complex formed during the production trajectories.

4.6 Cytotoxicity assay

Vero E6 cells were seeded at 2 × 10^4 cells per well in 96-well plates and maintained at 37 °C and 5% CO_2 atmosphere. After reaching ∼80–90% confluency and, to find the maximum non-toxic concentrations (MNTCs), monolayers were treated with each of the compounds (**LQM467**, **471**, **472**, and **474**) at concentrations from 25 to 200 μM in DMEM low-glucose (Sigma-Aldrich) culture media supplemented with 2% fetal bovine serum (Gibco), and 1% antibiotic antimycotic solution (Gibco) for 52 hours. Cell viability was then assessed by MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium) (Sigma-Aldrich) cytotoxicity assay by adding MTT solution to a final concentration of 0.5 mg mL^{-1}, followed by incubation for 3 hours. The supernatant was removed and then 150 μL of dimethyl sulfoxide (DMSO) was added and incubated for 15 min. The absorbance of each well was measured using ThermoPlate TP-READER at 492 nm wavelength and the cellular viability percentage was calculated as previously detailed. The absorbance value of blank control (only the culture media) was subtracted from the absorbance values obtained for all the samples.[101]

4.7 Evaluation of *in vitro* antiviral activity against ZIKV

Firstly, a serial dilution of ZIKV stock was performed and the viral dilution that reduced cell viability by at least 50% after 52 hours was used in antiviral assays. Therefore, *Vero E6* cells were submitted to virus adsorption by incubation with ZIKV at a 1 : 7.5 ratio for 90 min, with homogenization every 15 min. Afterward, the viral inoculum was removed and monolayers were washed with phosphate-buffered saline followed by treatment with several concentrations starting at the MNTC of each compound (**LQM467** from 50 to 6.25 μM; **LQM471** from 100 to 12.5 μM; **LQM472** from 25 to 3.125 μM). Cell viability was assessed after 52 h by MTT assay and the percentage of viral inhibition was calculated as previously described.[101]

4.8 Quantification of viral load by qPCR

Viral RNA (vRNA) was isolated from cell supernatant by using the *ReliaPrep™ Viral TNA Miniprep System* (Promega), in accordance with the manufacturer's instructions. Moreover, the complementary DNA (cDNA) was synthesized with 2 μL of vRNA and 1 μM of specific reverse primer for the ZIKV genome (*5′-CCAGTGCTTCTTTGTTGTTCC*) with a *High-Capacity cDNA Reverse Transcription kit* (Applied Biosystems), in agreement with the manufacturer's instructions. The obtained cDNA was diluted at 1 : 5 in nuclease-free water. A standard curve was prepared using a 94 bp cDNA amplicon of the ZIKV genome (forward primer: *5′-TTGGTTCACAAGGAGTGGTTC* and reverse primer: *5′-CCAGTGCTTCTTTGTTGTTCC*) obtained by a PCR reaction using *GoTaq Master Mix* (Promega). The amplicon was purified from 2% agarose gel with *Wizard® SV Gel and PCR Clean-Up System* (Promega), as recommended by the manufacturer's instructions. The purified amplicon concentration was quantified with Qubit™ 4 (Invitrogen) using the *Qubit™ dsDNA BR Assay kit* (Invitrogen). Nucleic acid molecular weight (M.W.) conversion was obtained through the formula M.W. of dsDNA = (# nucleotides × 607.4) + 157.9 (according to Thermo Fisher Scientific instructions), then resulting in the dsDNA copy number used to be serially diluted obtaining the seven standard curve points (from 7.5 × 10^8 copies per μL to 7.5 × 10^2 copies per μL). Furthermore, the qPCR assay for absolute quantification was performed in a 10 μL reaction mix containing 5 μL of *Power SYBR™ Green PCR Master Mix* (Applied Biosystems), 2.5 μL of forward and reverse primers (0.8 μM each), and 2.5 μL of diluted sample cDNA or standard curve points. The qPCR was run in QuantStudio™ 3 Real-Time PCR System (Applied Biosystems). All data were analyzed at Thermo Fisher Connect software (Applied Biosystems) in the standard curve mode converting cycle threshold (Ct) values into vRNA copy numbers. The efficiency of the qPCR standard curve presented a slope of −3.4 and a coefficient of correlation of 0.987% (Fig. S9 in the ESI†).

4.9 Statistical analysis

The statistical analyses were performed in the GraphPad Prism® v.6.0 software (San Diego, CA, USA), using One-Way ANOVA followed by Dunnett multiple comparison tests, in which the $p \leq 0.05$ value was considered statistically significant.

4.10 Synthesis of α-cyanoacrylates and α-cyanoacrylamides

Initially, the α-cyanoacrylates were prepared by a direct reaction involving cyanoesters and a 3-chlorobenzaldehyde (**LQM448** and **449**) or cyclic ketones (**LQM452–455**), *via* the Knoevenagel condensation mechanism of addition.[78,170] In a tube was added 100 mg ethyl or propyl-2-cyanoacetate (94 or 97 μL, respectively) and 1 eq. triethylamine (123 μL), under stirring, room temperature, and solvent-free. After 30 min, 1 eq. of corresponding aldehyde or ketone diluted in 1 mL ethanol was added into the reactional mixture dropwise. Then, the reaction was kept under stirring and at room temperature overnight. Posteriorly to the reaction completion (verified by thin-layer chromatography – TLC, using hexane : ethyl acetate 8 : 2 as eluent), the workup was initiated by removing the ethanol. The crude reaction was then extracted by using 20 mL dichloromethane (DCM) and 3 × 10 mL distilled water (added by 10 μL glacial acetic acid). Subsequently, all fractions were put together, filtrated using anhydrous Na_2SO_4, and then dried under reduced pressure at 30 °C to obtain the pure products. For the biphenyl cyanoacrylamides (**LQM456–462**), some biphenyl-4-carboxaldehydes were prepared by cross-coupling Suzuki–Miyaura reactions.[171–173] Thus, 100 mg 4-bromobenzaldehyde, 0.2 eq. palladium(II) acetate (24 mg), and a mixture of dioxane : water 2 : 1 (1.5 mL) were added to a tube for reaction. This reactional mixture was kept under stirring and at room temperature for 15 min. Then, 1 eq. of corresponding boronic acids and 2 eq. potassium hydroxide (116 mg) were also added to this mixture and, then, the tube was sealed. The reaction was kept under stirring and at 100 °C for 48 h. After the reaction completion (verified by TLC, using hexane as eluent), the workup was initiated for obtaining the desired aldehyde. Then, an extraction using 30 mL DCM and 3 × 10 mL distilled water was performed. All fractions were put together, filtrated using anhydrous Na_2SO_4, and then dried under reduced pressure at 60 °C. This residual material was then purified by column chromatography using a mixture of hexane : chloroform 9 : 1 (60 mL), monitoring by TLC. Finally, all pure fractions were put together and dried under reduced pressure, yielding pure intermediates. These aldehydes were confirmed by High-Performance Liquid Chromatogram (HPLC) using standard samples. For the preparation of α-cyanoacrylamides (**LQM456–474**), the cyanoamide intermediates were initially synthesized, reacting ethyl-2-cyanoacetate and primary or secondary amines by aminolysis of esters under free-solvent conditions.[174,175] Thus, 100 mg ethyl-2-cyanoacetate (94 μL) and 1.1 eq. primary or secondary amines were added to a bottom flask. This reactional mixture was stirred at room temperature for 48 h. The reaction completion was verified by HPLC detecting at 200 nm for the wave length. Subsequently, the resulting cyanoamide intermediates were not isolated, being utilized directly in the next step of the α-cyanoacrylamide synthetic route. Lastly, the α-cyanoacrylamides were then obtained by reacting these cyanoamide intermediates with 1 eq. of corresponding aldehydes in presence of 1 eq. triethylamine, *via* Knoevenagel condensation, as aforementioned.

4.11 Chemical characterization of α-cyanoacrylates and α-cyanoacrylamides

4.11.1 Reagents and solvents. All solvents and starting reagents were acquired from Merck/Sigma-Aldrich® Company (St. Louis, MO, USA), which were commercial supplies of high purity (> 98%), as well as, the deuterated solvents utilized in the NMR analyses. For the synthetic routes and column chromatography, all solvents were subjected to distillation prior to use for removing possible impurities. Besides, all HPLC runs were performed by using methanol HPLC degree from Tedia® Company (Fairfield, OH, USA) was used as eluent.

4.11.2 High-performance liquid chromatography with diode-array detection – (HPLC-DAD). To determine the retention time (R_T) and purity degree (%), a Shimadzu® (Kyoto, Japan) LC equipment model SIL-20AHT, with a Luna® 5 μm C18(2) 100 Å column (250 × 4.6 mm) was used in wavelengths (λ) of 254 nm. As the mobile phase, methanol HPLC degree (≥ 99%) was utilized in all HPLC-DAD runs. In this context, a specific parametrization was assumed for analysis, being (a) 1 mg mL^{-1} for the sample concentration; (b) 1 mL min^{-1} for the flowrate; (c) 10 min total time; and (d) 5 μL for the volume injected. Finally, all HPLC R_T values were computed in minutes (min), while the absorbance was computed as milli-absorbance unities (mAU).[101,176]

4.11.3 Melting point determination – (mp). An MSTecnopon® (Piracicaba, Brazil), model PFMII Digital, equipment was used to determine the melting point (mp) for each final compound. All measurements were performed by using glass capillaries containing the samples. For the initial temperature, 30 °C was admitted and then a temperature increase by 1 °C min^{-1} was utilized. Furthermore, all mp are uncorrected and a range of 1–2 °C was assumed.[101,177]

4.11.4 Attenuated total reflectance-Fourier transformed infrared spectroscopy – (ATR-FTIR). All infrared spectra were obtained using a spectrophotometer from Shimadzu® (Kyoto, Japan), model IRPrestige-21, employing the attenuated total reflectance (ATR) technique, in the range from 4000 to 400 cm^{-1}.[101] All spectra were treated by using Shimadzu IRsolution® software, version 1.50, 2008. Finally, all bond stretches (ν) and angle deformations (δ) for the main functional group from the final compounds were computed in transmittance (T%) and wavenumber (cm^{-1}).[178]

4.11.5 ^{1}H and ^{13}C nuclear magnetic resonance spectroscopy – (NMR). To correctly characterize all chemical compounds synthesized in this study, a Bruker® equipment (Billerica, USA), model UltraShield 600 MHz, was used to obtain all ^{1}H and ^{13}C NMR spectra. Moreover, deuterated dimethyl sulfoxide (DMSO-d_6) and chloroform ($CDCl_3$) were employed as analytical solvents for the solubilization of the samples. For hydrogen and carbon nuclei was used a total of 16 and

2048 scans, respectively. Thus, the chemical shifts (δ) were recorded in parts per million (ppm). Still, coupling constants (J) for ^{1}H signals were computed in hertz (Hz). In the spectra, signal multiplicities were attributed as singlet (s), broad singlet (br s), doublet (d), double-doublets (dd), triplet (t), quartet (q), quintent (qi), septet (sep), and multiplet (m).[101,102,179,180] Finally, all NMR spectra were treated and analyzed by using the academic licensed Bruker TopSpin® software, version 4.0.8, 2019 (**https://www.bruker.com/de/products-and-solutions/mr/nmr-software/topspin.html**).

4.11.6 Gas chromatography-mass spectrometry – (GC-MS). All samples were dissolved in methanol HPLC degree (1 mg mL^{-1}), into a glass vial, and then analyzed on a gas chromatograph coupled to a mass spectrometer detector GC/MS-GCMS-QP2010CN Ultra (Shimadzu, Kyoto, Japan) equipped with a split/splitless injector (250 °C), using a split ratio of 1 : 10. Posteriorly, an injection volume of 1 µL was used for all analyzes. A capillary SH-RXI-5Ms column (Shimadzu, Kyoto, Japan) was used, with a stationary phase composed of (5%-phenyl)-methylpolysiloxane, 30 m length, 0.25 mm internal diameter, and 0.25 µm film thickness. The temperature of the program used was 40 to 150 °C, increasing by 15 °C min^{-1}. Posteriorly, the temperature was increased to 300 °C, during 20 min. Both ionization chamber and interface temperatures were established as 280 °C. He carrier gas (145.2 kPa), the flow of 2.67 mL min^{-1}. The mass spectrometer operated with ionization energy of 70 eV in *Full Scan* mode, using the chemical ionization method. Finally, all data obtained were analyzed by using the Lab Solution software (Shimadzu, Kyoto, Japan), in which all GC R_T values were registered in minutes (min). All procedures and methods were adapted from Santos-Júnior *et al.* (2021).[181]

4.11.7 Chemical description of the compounds. In general, all final compounds were obtained with purities ranging from 95.0 to 99.9%, verified by HPLC experiments. Attenuated total reflectance (ATR) analyses were performed for all final compounds. In this context, it was verified that the vibrational stretching (ν) for nitrile group (C≡N) is found ranging from 2573 to 2198 cm^{-1}, while carbonyl signal ($\nu_{C=O}$) was observed from 1867 to 1556 cm^{-1}. Additionally, angular deformations (δ) for aromatic hydrogens (Ar–H) were identified between 1618 and 1406 cm^{-1}, while methine deformation ($\delta_{C=C_{ene}}$) was found varying from 831 to 771 cm^{-1}. Furthermore, for the secondary amide-containing derivatives, the characteristic stretching (ν_{N-H}) was exhibited ranging from 3658 to 3332 cm^{-1}, while the tertiary amide-containing analogs displayed a signal varying from 2976 to 2846 cm^{-1} for the carbonyl-nitrogen stretching ($\nu_{N_{ter}-C}$). Concerning ^{1}H NMR spectra, a singlet in 7.97–9.03 ppm for methine (CH) was assumed to confirm the products of Knoevenagel. In ^{13}C NMR spectra, this methine carbon signal was observed in 148–158 ppm. In contrast, for compounds **LQM452–455** has been verified the absence of acidic hydrogens in ^{1}H NMR spectra and the presence of a signal in 179–187 ppm for the quaternary carbon from the fused ring, corroborating these Knoevenagel products. Furthermore, the carbonyl carbon signal from the amide in ^{13}C NMR spectra was found ranging from 161 to 162 ppm for acrylates and from 159 to 166 ppm for acrylamides. Then, the nitrile carbon signal was observed varying from 114 to 115 ppm for acrylates and from 115 to 118 ppm for acrylamides. Finally, GC/MS experiments were able to identify the corresponding molecular mass (m/z) for all compounds.

Furthermore, all physicochemical data obtained to characterize the molecules synthesized in this study are presented below:

*Ethyl (E)-3-(3-chlorophenyl)-2-cyanoacrylate (**LQM448**).* Yield: 99%; aspect: white crystalline powder; HPLC R_T: 3.26 min; purity: 99.3%; mp: 107–110 °C; ATR-IR (cm^{-1}): 2222 ν(C≡N); 1722 ν(C=O); 1606, 1564, 1481, and 1433 δ(Ar–H); 792 δ(C=C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 1.43 (t, 3H, CH$_3$, J = 7.0 Hz), 4.41 (q, 2H, CH$_2$, J = 7.1 Hz), 7.47 (t, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.54 (d, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.92 (d, 2H, CH$_{Ar}$, J = 6.7 Hz), 8.2 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 14.12, 62.97, 104.75, 114.91, 128.60, 128.68, 130.53, 130.84, 133.04, 135.39, 153.17, 162.00. GC R_T: 13.92 min; MS (m/z): $C_{12}H_{10}ClNO_2$ calc.: 235.67; found: 235.60.

*Isopropyl (E)-3-(3-chlorophenyl)-2-cyanoacrylate (**LQM449**).* Yield: 95%; aspect: white amorphous powder; HPLC R_T: 3.37 min; purity: 99.9%; mp: 125–127 °C; ATR-IR (cm^{-1}): 2218 ν(C≡N); 1726 ν(C=O); 1608, 1558, and 1471 δ(Ar-H); 796 δ(C=C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 1.40 (d, 6H, (CH$_3$)$_2$, J = 6.2 Hz), 1.41 (s, 3H, CH$_3$), 5.23 (sep, 1H, CH, J = 6.2 Hz), 7.47 (t, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.54 (d, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.91 (s, 1H, CH$_{Ar}$), 7.94 (d, 1H, CH$_{Ar}$, J = 7.8 Hz), 8.18 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 21.69, 71.11, 105.27, 114.92, 128.52, 130.50, 130.82, 132.92, 133.16, 135.35, 152.89, 161.47. GC R_T: 14.13 min; MS (m/z): $C_{13}H_{12}ClNO_2$ calc.: 249.69; found: 249.50.

*(E)-3-(3-Chlorophenyl)-2-cyano-N-phenethylacrylamide (**LQM450**).* Yield: 85%; aspect: white crystalline powder; HPLC R_T: 3.35 min; purity: 99.9%; mp: 185–186 °C; ATR-IR (cm^{-1}): 3361 ν(N–H); 2222 ν(C≡N); 1676 ν(C=O); 1602, 1529, 1498, and 1450 δ(Ar-H); 790 δ(C=C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 2.94 (t, 2H, CH$_2$, J = 7.0 Hz), 3.71 (q, 2H, CH$_2$, J = 6.6 Hz), 6.44 (br s, 1H, NH), 7.25–7.29 (m, 3H, CH$_{Ar}$), 7.36 (t, 2H, CH$_{Ar}$, J = 7.6 Hz), 7.45 (t, 1H, CH$_{Ar}$, J = 7.8 Hz), 7.51 (d, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.82 (d, 1H, CH$_{Ar}$, J = 7.7 Hz), 7.88 (s, 1H, CH$_{Ar}$), 8.27 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 35.48, 41.84, 105.67, 116.32, 126.87, 128.34, 128.71, 128.86, 130.29, 130.46, 132.52, 133.47, 135.33, 138.07, 151.11, 159.60. GC R_T: 21.29 min; MS (m/z): $C_{18}H_{15}ClN_2O$ calc.: 310.78; found: 310.65.

*(E)-3-(3-Chlorophenyl)-2-cyano-N-(3-phenylpropyl)acrylamide (**LQM451**).* Yield: 72%; aspect: pale yellow crystalline powder; HPLC R_T: 3.41 min; purity: 95.4%; mp: 195–197 °C; ATR-IR (cm^{-1}): 3348 ν(N–H); 2222 ν(C≡N); 1664 ν(C=O); 1521, 1473, and 1452 δ(Ar-H); 790 δ(C=C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 1.97 (qi, 2H, CH$_2$, J = 7.3 Hz), 2.72 (t, 2H, CH$_2$, J = 7.5 Hz), 3.47 (q, 2H, CH$_2$, J = 6.7 Hz), 6.38 (br s, 1H, NH), 7.20–7.22 (m, 3H, CH$_{Ar}$), 7.31 (t, 2H, CH$_{Ar}$, J = 7.5 Hz), 7.44 (t, 1H, CH$_{Ar}$, J = 7.9 Hz), 7.50 (d, 1H, CH$_{Ar}$, J = 8.0 Hz), 7.82 (d, 1H, CH$_{Ar}$, J = 7.7 Hz), 7.87 (s, 1H, CH$_{Ar}$), 8.26 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 30.77, 33.15, 40.25, 105.68, 116.47,

126.19, 128.33, 128.58, 130.28, 130.46, 132.51, 133.49, 135.33, 140.90, 151.11, 159.61; GC R_T: 23.00 min; MS (m/z): $C_{19}H_{17}ClN_2O$ calc.: 324.81; found: 324.75.

*Ethyl 2-cyano-2-cyclopentylideneacetate (**LQM452**).* Yield: 65%; aspect: dark brown powder; HPLC R_T: 3.16 min; purity: 99.9%; mp: 49–50 °C; ATR-IR (cm^{-1}): 2227 $\nu(C{\equiv}N)$; 1728 $\nu(C{=}O)$; 1618, 1589, and 1440 δ(Ar-H); 813 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 1.36 (t, 3H, CH_3, J = 7.0 Hz), 1.83 (q, 2H, CH_2, J = 6.9 Hz), 1.88 (q, 2H, CH_2, J = 6.8 Hz), 2.82 (t, 2H, CH_2, J = 7.1 Hz), 3.01 (t, 2H, CH_2, J = 7.1 Hz), 4.28 (q, 2H, CH_2, J = 7.0 Hz). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 14.15, 25.09, 26.56, 35.42, 37.77, 61.52, 100.89, 115.56, 161.94, 187.28. GC R_T: 9.07 min; MS (m/z): $C_{10}H_{13}NO_2$ calc.: 179.22; found: 179.16.

*Ethyl 2-cyano-2-cyclohexylideneacetate (**LQM453**).* Yield: 25%; aspect: dark brown oil; HPLC R_T: 3.23 min; purity: 98.8%; ATR-IR (cm^{-1}): 2227 $\nu(C{\equiv}N)$; 1730 $\nu(C{=}O)$; 1598 and 1446 δ(Ar-H); 775 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 1.36 (t, 3H, CH_3, J = 7.0 Hz), 1.67 (qi, 2H, CH_2, J = 5.7 Hz), 1.75 (qi, 2H, CH_2, J = 5.7 Hz), 1.82 (qi, 2H, CH_2, J = 5.7 Hz), 2.68 (t, 2H, CH_2, J = 6.2 Hz), 2.99 (t, 2H, CH_2, J = 6.1 Hz), 4.28 (q, 2H, CH_2, J = 7.0 Hz). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 14.06, 25.63, 28.22, 31.59, 36.87, 41.96, 61.69, 102.10, 115.57, 162.02, 179.86. GC R_T: 10.64 min; MS (m/z): $C_{11}H_{15}NO_2$ calc.: 193.25; found: 193.15.

*Ethyl 2-cyano-2-cycloheptylideneacetate (**LQM454**).* Yield: 30%; aspect: dark brown oil; HPLC R_T: 3.41 min; purity: 98.1%; ATR-IR (cm^{-1}): 2222 $\nu(C{\equiv}N)$; 1722 $\nu(C{=}O)$; 1587 and 1446 δ(Ar-H); 773 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 1.36 (t, 3H, CH_3, J = 7.3 Hz), 1.58 (s, 4H, CH_2), 1.74–1.76 (m, 2H, CH_2), 1.78–1.80 (m, 2H, CH_2), 2.83 (t, 2H, CH_2, J = 6.1 Hz), 3.05 (t, 2H, CH_2, J = 6.1 Hz), 4.28 (q, 2H, CH_2, J = 7.3 Hz). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 13.96, 24.35, 26.21, 28.65, 33.81, 38.01, 43.88, 61.56, 104.23, 115.78, 161.96, 183.33. GC R_T: 11.97 min; MS (m/z): $C_{12}H_{17}NO_2$ calc.: 207.27; found: 207.15.

*Ethyl 2-cyano-2-cyclooctylideneacetate (**LQM455**).* Yield: 45%; aspect: dark brown oil; HPLC R_T: 3.52 min; purity: 99.9%; ATR-IR (cm^{-1}): 2222 $\nu(C{\equiv}N)$; 1722 $\nu(C{=}O)$; 1579 and 1446 δ(Ar-H); 779 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 1.33–1.40 (m, 5H, CH_2 and CH_3), 1.52–1.58 (m, 4H, CH_2), 1.87–1.91 (m, 3H, CH_2), 1.92–1.96 (m, 1H, CH_2), 2.42 (t, 2H, CH_2, J = 6.1 Hz), 2.76 (t, 1H, CH_2, J = 6.0 Hz), 2.98 (t, 1H, CH_2, J = 5.9 Hz), 4.29 (q, 2H, CH_2, J = 7.4 Hz). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 14.08, 24.72, 25.66, 27.18, 32.75, 37.52, 41.93, 63.00, 103.77, 112.95, 115.95, 162.84, 186.27. GC R_T: 12.93 min; MS (m/z): $C_{13}H_{19}NO_2$ calc.: 221.30; found: 221.20.

*(E)-3-([1,1′-Biphenyl]-4-yl)-2-(piperidine-1-carbonyl)acrylonitrile (**LQM456**).* Yield: 85%; aspect: yellow crystalline powder; HPLC R_T: 3.43 min; purity: 98.4%; mp: 70–72 °C; ATR-IR (cm^{-1}): 2860 $\nu(N_{ter}{-}C)$; 2204 $\nu(C{\equiv}N)$; 1647 $\nu(C{=}O)$; 1597, 1552, 1487, and 1433 δ(Ar-H); 779 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 1.71 (s, 6H, CH_2), 3.63–3.64 (m, 4H, CH_2), 7.41 (t, 1H, CH_{Ar}, J = 7.3 Hz), 7.48 (t, 2H, CH_{Ar}, J = 7.6 Hz), 7.64 (dd, 2H, CH_{Ar}, J = 7.3 and 0.9 Hz), 7.71 (d, 2H, CH_{Ar}, J = 8.3 Hz), 7.73 (s, 1H, CH), 7.97 (d, 2H, CH_{Ar}, J = 8.3 Hz). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 24.35, 106.39, 116.27, 127.16, 127.64, 128.30, 129.00, 130.54, 131.27, 139.67, 144.80, 150.94, 162.89. GC R_T: 25.30 min; MS (m/z): $C_{12}H_{20}N_2O$ calc.: 316.40; found: 316.30.

*(E)-3-([1,1′-Biphenyl]-4-yl)-2-cyano-N-(3-phenylpropyl)acrylamide (**LQM457**).* Yield: 74%; aspect: white amorphous powder; HPLC R_T: 3.62 min; purity: 99.2%; mp: 120–122 °C; ATR-IR (cm^{-1}): 3371 $\nu(N{-}H)$; 2212 $\nu(C{\equiv}N)$; 1666 $\nu(C{=}O)$; 1593, 1519, 1487, and 1450 δ(Ar-H); 794 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 2.00 (qi, 2H, CH_2, J = 7.3 Hz), 2.75 (t, 2H, CH_2, J = 6.1 Hz), 3.49 (q, 2H, CH_2, J = 6.4 Hz), 6.41 (br s, 1H, NH), 7.22 (d, 3H, CH_{Ar}, J = 7.2 Hz), 7.33 (t, 2H, CH_{Ar}, J = 7.4 Hz), 7.43 (t, 1H, CH_{Ar}, J = 7.4 Hz), 7.50 (t, 2H, CH_{Ar}, J = 7.6 Hz), 7.66 (d, 2H, CH_{Ar}, J = 7.2 Hz), 7.74 (d, 2H, CH_{Ar}, J = 8.4 Hz), 8.03 (d, 2H, CH_{Ar}, J = 8.3 Hz), 8.37 (s, 1H, CH). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 30.86, 33.16, 40.19, 103.48, 117.24, 126.16, 127.19, 127.75, 128.35, 128.45, 128.57, 129.07, 131.23, 139.55, 140.99, 145.48, 152.30, 154.44, 160.30. GC R_T: 32.18 min; MS (m/z): $C_{25}H_{22}N_2O$ calc.: 366.46; found: 366.35.

*(E)-N-Benzyl-2-cyano-3-(3′,4′-dichloro-[1,1′-biphenyl]-4-yl)acrylamide (**LQM458**).* Yield: 65%; aspect: white amorphous powder; HPLC R_T: 3.86 min; purity: 99.3%; mp: 183–185 °C; ATR-IR (cm^{-1}): 3658 $\nu(N{-}H)$; 2218 $\nu(C{\equiv}N)$; 1556 $\nu(C{=}O)$; 1471 and 1462 δ(Ar-H); 831 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, DMSO-d_6, ppm) δ 4.43 (d, 2H, CH_2, J = 5.8 Hz), 7.26–7.27 (m, 1H, CH_{Ar}), 7.34–7.36 (m, 4H, CH_{Ar}), 7.75 (d, 1H, CH_{Ar}, J = 8.3 Hz), 7.78 (dd, 1H, CH_{Ar}, J = 8.4 and 1.3 Hz), 7.94 (d, 2H, CH_{Ar}, J = 8.3 Hz), 8.05 (d, 3H, CH_{Ar}, J = 8.3 Hz), 8.26 (s, 1H, CH), 9.03 ("t", 1H, NH, J = 5.8 Hz). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 43.72, 106.90, 116.84, 127.43, 127.59, 127.93, 127.98, 128.79, 129.23, 131.23, 131.63, 131.70, 132.26, 132.42, 139.30, 139.77, 141.34, 150.38, 161.50. GC R_T: 37.63 min; MS (m/z): $C_{23}H_{16}Cl_2N_2O$ calc.: 406.06; found: 406.00.

*(E)-2-Cyano-3-(3′,4′-dichloro-[1,1′-biphenyl]-4-yl)-N-phenethylacrylamide (**LQM459**).* Yield: 53%; aspect: white amorphous powder; HPLC R_T: 4.03 min; purity: 95.1%; mp: 226–228 °C; ATR-IR (cm^{-1}): 3369 $\nu(N{-}H)$; 2218 $\nu(C{\equiv}N)$; 1668 $\nu(C{=}O)$; 1598, 1543, 1471, and 1450 δ(Ar-H); 790 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 2.95 (t, 2H, CH_2, J = 6.9 Hz), 3.71 (q, 2H, CH_2, J = 6.4 Hz), 6.45 (br s, 1H, NH), 7.26–7.30 (m, 3H, CH_{Ar}), 7.36–7.38 (m, 2H, CH_{Ar}), 7.47 (dd, 1H, CH_{Ar}, J = 8.3 and 1.8 Hz), 7.56 (d, 1H, CH_{Ar}, J = 8.3 Hz), 7.68 (d, 2H, CH_{Ar}, J = 8.2 Hz), 7.73 (d, 1H, CH_{Ar}, J = 1.7 Hz), 8.02 (d, 2H, CH_{Ar}, J = 8.2 Hz), 8.36 (s, 1H, CH). ^{13}C NMR (150 MHz, $CDCl_3$, ppm) δ 35.53, 41.82, 104.28, 116.85, 126.32, 126.85, 127.60, 128.72, 128.85, 129.01, 131.00, 131.30, 131.52, 132.77, 133.30, 138.14, 139.52, 142.76, 151.84, 160.04. GC R_T: 41.44 min; MS (m/z): $C_{24}H_{18}Cl_2N_2O$ calc.: 421.32; found: 421.25.

*(E)-N-Benzyl-2-cyano-3-(4′-methyl-[1,1′-biphenyl]-4-yl)acrylamide (**LQM460**).* Yield: 34%; aspect: white amorphous powder; HPLC R_T: 3.66 min; purity: 95.2%; mp: 197–201 °C; ATR-IR (cm^{-1}): 3394 $\nu(N{-}H)$; 2218 $\nu(C{\equiv}N)$; 1660 $\nu(C{=}O)$; 1587, 1519, 1452, and 1415 δ(Ar-H); 794 $\delta(C{=}C_{ene})$; ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 2.41 (s, 3H, CH_3), 4.53 (t, 2H, CH_2, J = 6.6 Hz), 6.62 (br s, 1H, NH), 7.29–7.40 (m, 6H, CH_{Ar}), 7.47–7.50 (m, 3H, CH_{Ar}), 7.54 (d, 1H, CH_{Ar}, J = 7.4 Hz), 7.65 (d, 1H, CH_{Ar}, J = 7.8 Hz), 7.71 (d, 1H, CH_{Ar},

J = 7.9 Hz), 8.01 (d, 1H, CH$_{Ar}$, J = 7.8 Hz), 8.40 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 21.07, 44.62, 113.19, 127.00, 127.10, 127.29, 128.02, 128.22, 128.90, 129.04, 129.61, 129.76, 131.29, 137.17, 137.80, 143.37, 152.79, 161.31. GC R_T: 29.45 min; MS (m/z): $C_{24}H_{20}N_2O$ calc.: 352.44; found: 352.35.

*(E)-2-Cyano-3-(4′-methyl-[1,1′-biphenyl]-4-yl)-N-phenethylacrylamide (**LQM461**).* Yield: 68%; aspect: white amorphous powder; HPLC R_T: 3.61 min; purity: 97.1%; mp: 205–207 °C; ATR-IR (cm^{-1}): 3361 ν(N–H); 2214 ν(C≡N); 1660 ν(C═O); 1600, 1589, 1494, and 1446 δ(Ar-H); 771 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 2.44 (s, 3H, CH$_3$), 2.95 (t, 2H, CH$_2$, J = 7.1 Hz), 3.71 (q, 2H, CH$_2$, J = 6.6 Hz), 6.43 (br s, 1H, NH), 7.26–7.28 (m, 3H, CH$_{Ar}$), 7.30 (d, 2H, CH$_{Ar}$, J = 8.1 Hz), 7.37 (t, 2H, CH$_{Ar}$, J = 7.4 Hz), 7.56 (d, 2H, CH$_{Ar}$, J = 8.0 Hz), 7.22 (d, 2H, CH$_{Ar}$, J = 8.3 Hz), 8.01 (d, 2H, CH$_{Ar}$, J = 8.2 Hz), 8.36 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 21.14, 35.57, 41.78, 103.22, 117.11, 126.80, 127.00, 127.46, 128.71, 128.82, 129.76, 130.44, 131.24, 136.64, 138.22, 138.52, 145.48, 152.33, 160.34. GC R_T: 31.31 min; MS (m/z): $C_{25}H_{22}N_2O$ calc.: 366.46; found: 366.35.

*(E)-2-Cyano-3-(4′-methyl-[1,1′-biphenyl]-4-yl)-N-(3-phenylpropyl)-acrylamide (**LQM462**).* Yield: 24%; aspect: white amorphous powder; HPLC R_T: 3.93 min; purity: 97.3%; mp: 162–164 °C; ATR-IR (cm^{-1}): 3356 ν(N–H); 2218 ν(C≡N); 1662 ν(C═O); 1597, 1516, 1498, and 1450 δ(Ar-H); 788 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 2.00 (qi, 2H, CH$_2$, J = 7.2 Hz), 2.44 (s, 3H, CH$_3$), 2.75 (t, 2H, CH$_2$, J = 7.3 Hz), 3.49 (q, 2H, CH$_2$, J = 6.6 Hz), 6.39 (br s, 1H, NH), 7.22–7.23 (m, 3H, CH$_{Ar}$), 7.30–7.34 (m, 4H,CH$_{Ar}$), 7.56 (d, 2H, CH$_{Ar}$, J = 8.0 Hz), 7.72 (d, 2H, CH$_{Ar}$, J = 8.3 Hz), 8.01 (d, 2H, CH$_{Ar}$, J = 8.3 Hz), 8.36 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 21.14, 30.85, 33.16, 40.17, 103.24, 117.28, 126.14, 127.01, 127.46, 128.33, 128.56, 129.76, 130.46, 131.22, 136.65, 138.51, 140.98, 145.45, 152.32, 160.35. GC R_T: 35.87 min; MS (m/z): $C_{26}H_{24}N_2O$ calc.: 380.49; found: 380.30.

*(E)-2-Cyano-3-(5-phenylfuran-2-yl)-N-(3-phenylpropyl)acrylamide (**LQM463**).* Yield: 35%; aspect: yellow amorphous powder; HPLC R_T: 3.46 min; purity: 98.8%; mp: 160–162 °C; ATR-IR (cm^{-1}): 3332 ν(N–H); 2214 ν(C≡N); 1867 ν(C═O); 1604, 1543, 1492, and 1454 δ(Ar-H); 792 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, DMSOd_{-6}, ppm): δ 1.81 (qi, 2H, CH$_2$, J = 7.38 Hz), 2.60 (t, 2H, CH$_2$, J = 7.61 Hz), 3.23 (q, 2H, CH$_2$, J = 6.6 Hz), 7.17 (t, 1H, CH$_{Ar}$, J = 7.27 Hz), 7.21 (d, 2H, CH$_{Ar}$, J = 7.0 Hz), 7.26 (s, 1H, CH$_{Ar}$), 7.28 (d, 1H, CH$_{Ar}$, J = 7.7 Hz), 7.33 (d, 1H, CH$_{Ar}$, J = 3.7 Hz), 7.42–7.44 (m, 2H, CH$_{Ar}$), 7.52 (t, 2H, CH$_{Ar}$, J = 7.52 Hz), 7.91 (d, 1H, CH$_{Ar}$, J = 7.27 Hz), 7.95 (s, 1H, CH), 8.38 (t, 1H, NH, J = 5.59 Hz). ^{13}C NMR (150 MHz, CDCl$_3$, ppm): δ 30.90, 31.03, 32.99, 100.63, 110.12, 117.23, 124.93, 125.16, 126.21, 128.75, 129.18, 129.67, 130.01, 135.36, 142.11, 148.26, 158.07, 158.12, 161.48. GC R_T: 28.89 min; MS (m/z): $C_{23}H_{20}N_2O_2$ calc.: 356.43; found: 356.35.

*(E)-2-Cyano-N-(3-phenylpropyl)-3-(5-(2-(trifluoromethyl)phenyl)-furan-2-yl)acrylamide (**LQM464**).* Yield: 85%; aspect: yellow amorphous powder; HPLC R_T: 3.47 min; purity: 99.2%; mp: 153–155 °C; ATR-IR (cm^{-1}): 3381 ν(N–H); 2208 ν(C≡N); 1666 ν(C═O); 1608, 1521, and 1442 δ(Ar-H); 806 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 1.98 (qi, 2H, CH$_2$, J = 7.3 Hz), 2.73 (t, 2H, CH$_2$, J = 7.6 Hz), 3.47 (q, 2H, CH$_2$, J = 6.6 Hz), 6.37 (br s, 1H, NH), 7.02 (d, 1H, CH$_{Ar}$, J = 3.5 Hz), 7.22–7.23 (m, 3H, CH$_{Ar}$), 7.26 (d, 1H, CH$_{Ar}$, J = 3.7 Hz), 7.31–7.33 (m, 2H, CH$_{Ar}$), 7.54 (t, 1H, CH$_{Ar}$, J = 7.6 Hz), 7.71 (t, 1H, CH$_{Ar}$, J = 7.6 Hz), 7.81 (d, 1H, CH$_{Ar}$, J = 7.8 Hz), 7.88 (d, 1H, CH$_{Ar}$, J = 7.8 Hz), 8.09 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 30.91, 33.13, 40.13, 100.10, 113.86, 113.89, 113.93, 116.98, 122.90, 126.14, 126.89, 12.78, 128.34, 128.55, 129.25, 130.25, 132.32, 136.50, 140.99, 148.74, 155.00, 160.33. GC R_T: 26.31 min; MS (m/z): $C_{24}H_{19}F_3N_2O_2$ calc.: 424.42; found: 424.35.

*(E)-2-Cyano-3-(5-(3,4-dichlorophenyl)furan-2-yl)-N-(3-phenylpropyl)-acrylamide (**LQM465**).* Yield: 72%; aspect: yellow amorphous powder; HPLC R_T: 3.84 min; purity: 99.9%; mp: 189–190 °C; ATR-IR (cm^{-1}): 3365 ν(N–H); 2212 ν(C≡N); 1712 ν(C═O); 1608, 1527, 1492, and 1450 δ(Ar-H); 794 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, DMSO-d_6, ppm) δ 1.82 (qi, 2H, CH$_2$, J = 7.2 Hz), 2.61 (t, 2H, CH$_2$, J = 7.5 Hz), 3.23 (q, 2H, CH$_2$, J = 6.5 Hz), 7.22 (d, 1H, CH$_{Ar}$, J = 7.4 Hz), 7.28 (d, 1H,CH$_{Ar}$, J = 7.4 Hz), 7.55 (d, 1H, CH$_{Ar}$, J = 3.7 Hz), 7.6 (d, 1H, CH$_{Ar}$, J = 3.7 Hz), 7.80–7.83 (m, 3H, CH$_{Ar}$), 7.84–7.88 (m, 2H, CH$_{Ar}$), 7.97 (s, 1H, CH$_{Ar}$), 8.16 (s, 1H, CH), 8.40 (t, 1H, NH, J = 6.5 Hz). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 30.99, 32.98, 62.62, 97.48, 101.78, 112.25, 116.41, 125.23, 126.21, 127.04, 128.75, 129.41, 132.07, 132.66, 138.75, 142.10, 148.63, 155.32, 156.69, 162.73. GC R_T: 39.75 min; MS (m/z): $C_{23}H_{18}Cl_2N_2O_2$ calc.: 425.31; found: 425.25.

*(E)-2-Cyano-3-(5-(3,4-dichlorophenyl)furan-2-yl)-N-phenethylacrylamide (**LQM466**).* Yield: 89%; aspect: yellow amorphous powder; HPLC R_T: 3.90 min; purity: 99.1%; mp: 172–174 °C; ATR-IR (cm^{-1}): 3361 ν(N–H); 2218 ν(C≡N); 1656 ν(C═O); 1612, 1523, 1498, and 1452 δ(Ar-H); 802 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 2.94 (t, 2H, CH$_2$, J = 6.6 Hz), 3.71 (q, 2H, CH$_2$, J = 6.3 Hz), 6.41 (br s, 1H, NH), 6.90 (d, 1H, CH$_{Ar}$, J = 3.5 Hz), 7.17 (d, 1H, CH$_{Ar}$, J = 3.2 Hz), 7.25–7.29 (m, 3H, CH$_{Ar}$), 7.35–7.38 (m, 2H, CH$_{Ar}$), 7.54 (d, 1H, CH$_{Ar}$, J = 8.3 Hz), 7.69 (dd, 1H, CH$_{Ar}$, J = 8.3 and 1.8 Hz), 7.89 (d, 1H, CH$_{Ar}$, J = 1.8 Hz), 8.04 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 35.59, 41.72, 109.60, 117.02, 123.34, 124.18, 126.75, 126.82, 128.72, 128.80, 128.83, 131.24, 133.49, 133.53, 135.99, 138.20, 148.72, 156.41. GC R_T: 35.85 min; MS (m/z): $C_{22}H_{16}Cl_2N_2O_2$ calc.: 411.28; found: 411.20.

*(E)-N-Benzyl-2-cyano-3-(5-(3,4-dichlorophenyl)furan-2-yl)acrylamide (**LQM467**).* Yield: 84%; aspect: yellow amorphous powder; HPLC R_T: 3.73 min; purity: 99.1%; mp: 257–260 °C; ATR-IR (cm^{-1}): 3365 ν(N–H); 2202 ν(C≡N); 1666 ν(C═O); 1608, 1562, 1521, and 1456 δ(Ar-H); 802 δ(C═C$_{ene}$); ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 4.63 (d, 2H, CH$_2$, J = 5.7 Hz), 6.67 ("t", 1H, NH, J = 5.6 Hz), 6.92 (d, 1H, CH$_{Ar}$, J = 3.7 Hz), 7.20 (d, 1H, CH$_{Ar}$ J = 3.4 Hz), 7.33–7.41 (m, 5H, CH$_{Ar}$), 7.56 (d, 1H, CH$_{Ar}$, J = 8.4 Hz), 7.70 (dd, 1H, CH$_{Ar}$, J = 8.4 and 1.9 Hz), 7.90 (d, 1H, CH$_{Ar}$, J = 1.9 Hz), 8.10 (s, 1H, CH). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 44.57, 109.68, 117.13, 123.62, 124.20, 126.01, 126.80, 127.86, 127.92, 128.79, 128.91, 131.27, 133.66, 137.18, 147.74, 148.76, 156.53, 158.58, 160.25. GC R_T: 33.51 min; MS (m/z): $C_{21}H_{14}Cl_2N_2O_2$ calc.: 397.26; found: 397.16.

*(E)-2-(Pyrrolidine-1-carbonyl)-3-(5-(2-(trifluoromethyl)phenyl)-furan-2-yl)acrylonitrile (**LQM468**).* Yield: 35%; aspect: yellow amorphous powder; HPLC R_{T}: 3.31 min; purity: 95.0%; mp: 144–146 °C; ATR-IR (cm^{-1}): 2208 $\nu(\text{C}\equiv\text{N})$; 1643 $\nu(\text{C}{=}\text{O})$; 1602, 1543, 1469, and 1425 δ(Ar-H); 779 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, CDCl$_3$, ppm) δ 1.71 (s, 4H, CH$_2$), 3.65 (s, 4H, CH$_2$), 6.99 (d, 1H, CH$_{\text{Ar}}$, J = 3.5 Hz), 7.24 (d, 1H, CH$_{\text{Ar}}$, J = 3.7 Hz), 7.52 (t, 1H, CH$_{\text{Ar}}$, J = 7.6 Hz), 7.63 (s, 1H, CH), 7.69 (t, 1H, CH$_{\text{Ar}}$, J = 7.6 Hz), 7.8 (d, 1H, CH$_{\text{Ar}}$, J = 7.9 Hz), 8.05 (d, 1H, CH$_{\text{Ar}}$, J = 7.8 Hz). ^{13}C NMR (150 MHz, CDCl$_3$, ppm) δ 24.37, 25.75, 102.65, 113.59, 116.34, 120.96, 122.94, 124.75, 126.78, 127.95, 129.02, 130.25, 132.25, 136.98, 148.98, 148.98, 154.14, 162.58. GC R_{T}: 27.83 min; MS (m/z): C$_{19}$H$_{15}$F$_3$N$_2$O$_2$ calc.: 360.34; found: 360.26.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(pyrrolidine-1-carbonyl)-acrylonitrile (**LQM469**).* Yield: 29%; aspect: yellow amorphous powder; HPLC R_{T}: 3.79 min; purity: 97.7%; mp: 222–224 °C; ATR-IR (cm^{-1}): 2573 $\nu(\text{C}\equiv\text{N})$; 1714 $\nu(\text{C}{=}\text{O})$; 1629, 1587, and 1481 δ(Ar-H); 815 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, CDCl$_3$, ppm): δ 1.25 (s, 2H, CH$_2$), 1.4 (s, 2H, CH$_2$), 1.58 (s, 2H, CH$_2$), 4.38 (s, 2H, CH$_2$), 6.92 (s, 1H, CH), 7.33 (s, 1H, CH$_{\text{Ar}}$), 7.55 (s, 1H, CH$_{\text{Ar}}$), 7.71 (s, 1H, CH$_{\text{Ar}}$), 7.93 (d, 2H, CH$_{\text{Ar}}$, J = 49.7 Hz). ^{13}C NMR (150 MHz, CDCl$_3$, ppm): δ 14.19, 62.60, 98.59, 109.92, 115.67, 124.18, 124.36, 126.86, 128.62, 131.30, 133.51, 133.78, 138.07, 148.43, 157.08, 162.72. GC R_{T}: 37.30 min; MS (m/z): C$_{18}$H$_{14}$Cl$_2$N$_2$O$_2$ calc.: 361.22; found: 361.15.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(piperidine-1-carbonyl)-acrylonitrile (**LQM470**).* Yield: 62%; aspect: yellow amorphous powder; HPLC R_{T}: 3.80 min; purity: 99.8%; mp: 188–190 °C; ATR-IR (cm^{-1}): 2198 $\nu(\text{C}\equiv\text{N})$; 1629 $\nu(\text{C}{=}\text{O})$; 1591, 1539, 1467, and 1406 δ(Ar-H); 815 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, CDCl$_3$, ppm): δ 1.63 (s, 2H, CH$_2$), 1.94–1.98 (m, 2H, CH$_2$), 2.01–2.05 (m, 2H, CH$_2$), 3.63 (t, 2H, CH$_2$, J = 6.8 Hz), 3.86 (t, 2H, CH$_2$, J = 6.5 Hz), 6.89 (d, 1H, CH$_{\text{Ar}}$, J = 3.6 Hz), 7.20 (d, 1H, CH$_{\text{Ar}}$, J = 3.5 Hz), 7.53 (d, 1H, CH$_{\text{Ar}}$, J = 8.4 Hz), 7.69 (dd, 1H, CH$_{\text{Ar}}$, J = 8.3 and 1.8 Hz), 7.89 (s, 1H, CH), 7.91 (d, 1H, CH$_{\text{Ar}}$, J = 1.8 Hz). ^{13}C NMR (150 MHz, CDCl$_3$, ppm): δ 24.01, 26.77, 47.80, 48.62, 101.67, 109.45, 116.80, 122.36, 124.11, 126.68, 128.96, 131.16, 133.27, 133.43, 137.75, 149.00, 155.88, 160.98. GC R_{T}: 22.04 min; MS (m/z): C$_{19}$H$_{16}$Cl$_2$N$_2$O$_2$ calc.: 375.25; found: 375.20.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(thiomorpholine-4-carbonyl)acrylonitrile (**LQM471**).* Yield: 36%; aspect: yellow amorphous powder; HPLC R_{T}: 3.82 min; purity: 95.0%; mp: 225–227 °C; ATR-IR (cm^{-1}): 2208 $\nu(\text{C}\equiv\text{N})$; 1633 $\nu(\text{C}{=}\text{O})$; 1608, 1566, 1546, and 1463 δ(Ar-H); 800 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, DMSO-d_6, ppm) δ 2.81 (s, 4H, CH$_2$), 3.24 (s, 4H, CH$_2$), 7.26 (d, 1H, CH$_{\text{Ar}}$, J = 3.21 Hz), 7.39 (d, 1H, CH$_{\text{Ar}}$, J = 3.20 Hz), 7.76–7.79 (m, 2H, CH$_{\text{Ar}}$), 7.82–7.84 (m, 1H, CH$_{\text{Ar}}$), 8.11 (s, 1H,CH). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 24.71, 45.47, 111.24, 119.11, 121.32, 124.68, 126.39, 130.14, 131.54, 131.87, 132.56, 134.07, 135.87, 149.78, 153.93, 163.74. GC R_{T}: 32.90 min; MS (m/z): C$_{18}$H$_{14}$Cl$_2$N$_2$O$_2$S calc.: 393.28; found: 393.15.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(morpholine-4-carbonyl)-acrylonitrile (**LQM472**).* Yield: 36%; aspect: yellow amorphous powder; HPLC R_{T}: 3.49 min; purity: 98.0%; mp: 227–229 °C; ATR-IR (cm^{-1}): 2218 $\nu(\text{C}\equiv\text{N})$; 1714 $\nu(\text{C}{=}\text{O})$; 1610, 1558, 1539, and 1477 δ(Ar-H); 796 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, DMSO-d_6, ppm) δ 2.78 (s, 4H, CH$_2$), 3.21 (s, 4H, CH$_2$), 7.23 (d, 1H, CH$_{\text{Ar}}$, J = 3.1 Hz), 7.36 (d, 1H, CH$_{\text{Ar}}$, J = 3.2 Hz), 7.73–7.77 (m, 2H, CH$_{\text{Ar}}$), 7.79–7.81 (m, 1H, CH$_{\text{Ar}}$), 8.09 (s, 1H, CH). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 24.61, 24.69, 45.37, 45.45, 111.14, 119.01, 121.21, 124.58, 124.82, 126.29, 126.55, 129.64, 131.76, 132.46, 133.96, 149.67, 153.83, 163.64. GC R_{T}: 27.60 min; MS (m/z): C$_{18}$H$_{14}$Cl$_2$N$_2$O$_3$ calc.: 377.22; found: 377.15.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(4-phenylpiperazine-1-carbonyl)acrylonitrile (**LQM473**).* Yield: 52%; aspect: yellow amorphous powder; HPLC R_{T}: 3.91 min; purity: 97.5%; mp: 105–107 °C; ATR-IR (cm^{-1}): 2846 $\nu(\text{N}_{\text{ter}}\text{–C})$; 2212 $\nu(\text{C}\equiv\text{N})$; 1643 $\nu(\text{C}{=}\text{O})$; 1602, 1494, and 1458 δ(Ar-H); 796 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, DMSO-d_6, ppm) δ 3.12 (t, 2H, CH$_2$, J = 5.0 Hz), 3.18 (t, 2H, CH$_2$, J = 4.9 Hz), 3.5 (t, 2H, CH$_2$, J = 4.9 Hz), 3.60 (t, 2H, CH$_2$, J = 5.0 Hz), 6.81 (t, 1H, CH$_{\text{Ar}}$, J = 7.3 Hz), 6.95 (d, 2H, CH$_{\text{Ar}}$, J = 8.1 Hz), 7.23 (t, 2H, CH$_{\text{Ar}}$, J = 7.8 Hz), 7.55–7.56 (m, 1H, CH$_{\text{Ar}}$), 7.61 (d, 1H, CH$_{\text{Ar}}$, J = 3.7 Hz), 7.84 (d, 1H, CH$_{\text{Ar}}$, J = 8.4 Hz), 7.88 (dd, 1H, CH$_{\text{Ar}}$, J = 8.4 and 1.6 Hz), 8.17 (s, 1H, CH), 8.19 (t, 1H, CH$_{\text{Ar}}$, J = 1.6 Hz). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 14.51, 25.19, 48.50, 53.64, 104.17, 112.28, 116.33, 119.84, 125.26, 127.07, 127.63, 129.46, 132.09, 132.75, 138.76, 148.66, 151.09, 156.75, 160.08, 166.97. GC R_{T}: 23.87 min; MS (m/z): C$_{24}$H$_{19}$Cl$_2$N$_3$O$_2$ calc.: 452.34; found: 452.10.

*(E)-3-(5-(3,4-Dichlorophenyl)furan-2-yl)-2-(4-methylpiperazine-1-carbonyl)acrylonitrile (**LQM474**).* Yield: 40%; aspect: yellow amorphous powder; HPLC R_{T}: 2.88 min; purity: 98.1%; mp: 213–215 °C; ATR-IR (cm^{-1}): 2976 $\nu(\text{N}_{\text{ter}}\text{–C})$; 2231 $\nu(\text{C}\equiv\text{N})$; 1726 $\nu(\text{C}{=}\text{O})$; 1618 and 1585 δ(Ar-H); 819 $\delta(\text{C}{=}\text{C}_{\text{ene}})$; ^{1}H NMR (600 MHz, DMSO-d_6, ppm) δ 2.97 (s, 4H, CH$_2$), 3.30 (s, 3H, CH$_3$), 3.31 (s, 4H, CH$_2$), 7.16 (d, 1H, CH$_{\text{Ar}}$, J = 3.5 Hz), 7.25 (d, 1H, CH$_{\text{Ar}}$, J = 15.6 Hz), 7.35 (d, 1H, CH$_{\text{Ar}}$, J = 3.5 Hz), 7.57 (d, 1H, CH$_{\text{Ar}}$, J = 15.6 Hz), 7.76 (d, 1H, CH$_{\text{Ar}}$, J = 8.41), 7.89 (d, 1H, CH$_{\text{Ar}}$, J = 8.4 Hz), 8.18 (s, 1H, CH). ^{13}C NMR (150 MHz, DMSO-d_6, ppm) δ 23.61, 23.69, 25.98, 44.37, 44.45, 110.14, 118.01, 120.21, 123.58, 123.82, 125.29, 125.55, 128.64, 130.76, 131.46, 132.96, 148.67, 152.83, 162.64. GC R_{T}: 18.90 min; MS (m/z): C$_{19}$H$_{17}$Cl$_2$N$_3$O$_2$ calc.: 389.26; found: 389.10.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All reagents and solvents used to synthesize all intermediate and final compounds in this research were acquired *via* financial support by CNPq/INCT/RENNOFITO, grant number: N465516/2014-0. Moreover, some cell-based assays were performed *via* financial support by CNPq, grant number: 437407/2018-7. The authors thank Prof. Thiago Mendonça de Aquino and PhD.

Alessandre Carmo Crispim for performing our NMR experiments at the Nucleus of NMR Analyzes and Research (NAPRMN). Additionally, we thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, National Council for scientific and Technological Development – CNPq, Financiadora de Estudos e Projetos – FINEP, Fundação de Amparo à Pesquisa do Estado de Alagoas – FAPEAL for their scientific support to the Brazilian Post-Graduation Programs.

References

1 S. Bala Murugan and R. Sathishkumar, *Asian Pac. J. Trop. Med.*, 2016, **9**, 933–937.

2 M. Govindarajan and G. Benelli, *J. Asia-Pac. Entomol.*, 2016, **19**, 377–385.

3 F. Barthels, *Synthesis and evaluation of novel flaviviral NS2B/NS3 protease inhibitors*, Eberhard Karls Universität Tübingen, 2018.

4 B. Millies, F. von Hammerstein, F. Barthels, U. Goppel, U. A. Hellmich and T. Schirmeister, *20. JungChemikerForum Frühjahrssymposium*, Konstanz, 2018.

5 J. D. Martinez, J. A. C. de la Garza and A. Cuellar-Barboza, *Dermatol. Clin.*, 2019, **37**, 95–105.

6 O. Choksupmanee, K. Hodge, G. Katzenmeier and S. Chimnaronk, *Biochemistry*, 2012, **51**, 2840–2851.

7 J. Deng, N. Li, H. Liu, Z. Zuo, O. W. Liew, W. Xu, G. Chen, X. Tong, W. Tang, J. Zhu, J. Zuo, H. Jiang, C. G. Yang, J. Li and W. Zhu, *J. Med. Chem.*, 2012, **55**, 6278–6293.

8 D. Aguilera-Pesantes and M. a Méndez, *Biochem. Biophys. Res. Commun.*, 2017, **492**, 659–667.

9 D. Lu, J. Liu, Y. Zhang, F. Liu, L. Zeng, R. Peng, L. Yang, H. Ying, W. Tang, W. Chen, J. Zuo, X. Tong, T. Liu and Y. Hu, *Eur. J. Med. Chem.*, 2018, **145**, 328–337.

10 E. F. Silva-Júnior, T. Schirmeister and J. X. Araújo-Júnior, *Vector-Borne Diseases & Treatment*, Open Access eBooks, Las Vegas, 1st edn, 2018, pp. 1–25.

11 K. G. Byler, I. V. Ogungbe and W. N. Setzer, *J. Mol. Graphics Modell.*, 2016, **69**, 78–91.

12 Z. Weng, X. Shao, D. Graf, C. Wang, C. D. Klein, J. Wang and G.-C. Zhou, *Eur. J. Med. Chem.*, 2017, **125**, 751–759.

13 T. B. Casale, M. N. Teng, J. P. Morano, T. Unnasch and C. J. Lockwood, *J. Allergy Clin. Immunol.*, 2018, **141**, 482–490.

14 P. Gérardin, G. Barau, A. Michault, M. Bintner, H. Randrianaivo, G. Choker, Y. Lenglet, Y. Touret, A. Bouveret, P. Grivard, K. Le Roux, S. Blanc, I. Schuffenecker, T. Couderc, F. Arenzana-Seisdedos, M. Lecuit and P. Y. Robillard, *PLoS Med.*, 2008, **5**, 0413–0423.

15 M. Chhabra, V. Mittal, D. Bhattacharya, U. Rana and S. Lal, *Indian J. Med. Microbiol.*, 2008, **26**, 5–12.

16 SESA, *Protocolo de Vigilância e Resposta à Ocorrência de Microcefalia e/ou Alterações do Sistema Nervoso Central Relacionadas a Infecções Congênitas*, Fortaleza, 2016.

17 J. Mlakar, M. Korva, N. Tul, M. Popović, M. Poljšak-Prijatelj, J. Mraz, M. Kolenc, K. Resman Rus, T. Vesnaver Vipotnik, V. Fabjan Vodušek, A. Vizjak, J. Pižem, M. Petrovec and T. Avšič Županc, *N. Engl. J. Med.*, 2016, **374**, 951–958.

18 U. Bhardwaj, N. Pandey, M. Rastogi and S. K. Singh, *Virology*, 2021, **560**, 86–95.

19 M. J. Sabir, N. B. S. Al-Saud and S. M. Hassan, *Saudi J. Biol. Sci.*, 2021, **28**, 5074–5080.

20 M. Perera-lecoin, L. Meertens, X. Carnec and A. Amara, *Viruses*, 2014, 69–88.

21 H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163–2166.

22 I. A. Rodenhuis-Zybert, J. Wilschut and J. M. Smit, *Cell. Mol. Life Sci.*, 2010, **67**, 2773–2786.

23 S. Mukhopadhyay, R. J. Kuhn and M. G. Rossmann, *Nat. Rev. Microbiol.*, 2005, **3**, 13–22.

24 L. De La Cruz, T. H. D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**, 19205–19215.

25 B. D. Lindenbach and C. M. Rice, *Advances in Virus Research*, 2003, vol. 59, pp. 23–61.

26 I. A. Zybert, H. van der Ende-Metselaar, J. Wilschut and J. M. Smit, *J. Gen. Virol.*, 2008, **89**, 3047–3051.

27 Seema and S. K. Jain, *Indian J. Clin. Biochem.*, 2005, **20**, 92–103.

28 H. Xu, B. Di, Y. Pan, L. Qiu, Y. Wang, W. Hao, L. He, K. Yuen and X. Che, *J. Clin. Microbiol.*, 2006, **44**, 2872–2878.

29 J. Y. Leung, G. P. Pijlman, N. Kondratieva, J. Hyde, J. M. Mackenzie and A. A. Khromykh, *J. Virol.*, 2008, **82**, 4731–4741.

30 W. W. Phoo, Y. Li, Z. Zhang, M. Y. Lee, Y. R. Loh, Y. B. Tan, E. Y. Ng, J. Lescar, C. Kang and D. Luo, *Nat. Commun.*, 2016, **7**, 13410.

31 E. F. Silva-Júnior and J. X. Araújo-Júnior, *Bioorg. Med. Chem.*, 2019, **27**, 3963–3978.

32 F. Preugschat, C. W. Yao and J. H. Strauss, *J. Virol.*, 1990, **64**, 4364–4374.

33 Y. Zhou, D. Ray, Y. Zhao, H. Dong, S. Ren, Z. Li, Y. Guo, K. A. Bernard, P.-Y. Shi and H. Li, *J. Virol.*, 2007, **81**, 3891–3903.

34 I. J. dos, S. Nascimento, P. F. da, S. Santos-Júnior, T. M. de Aquino, J. X. de Araújo-Júnior and E. F. da Silva-Júnior, *Eur. J. Med. Chem.*, 2021, **224**, 113698.

35 D. Luo, S. G. Vasudevan and J. Lescar, *Antiviral Res.*, 2015, **118**, 148–158.

36 S. M. Tomlinson and S. J. Watowich, *Biochemistry*, 2008, **47**, 11763–11770.

37 B. Falgout, M. Pethel, Y. M. Zhang and C. J. Lai, *J. Virol.*, 1991, **65**, 2467–2475.

38 H. Li, S. Clum, S. You, K. E. Ebner and R. Padmanabhan, *J. Virol.*, 1999, **73**, 3108–3116.

39 L. Hedstrom, *Chem. Rev.*, 2002, **102**, 4501–4524.

40 D. E. Danley, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 569–575.

41 A. Poulsen, C. Kang and T. H. Keller, *Curr. Pharm. Des.*, 2014, **20**, 3422–3427.

42 A. C. Gibbs, R. Steele, G. Liu, B. A. Tounge and G. T. Montelione, *Biochemistry*, 2018, **57**, 1591–1602.

43 S. Voss and C. Nitsche, *Bioorg. Med. Chem. Lett.*, 2020, **30**, 126965.

44 P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan and U. Hommel, *Nat. Struct. Mol. Biol.*, 2006, **13**, 372–373.

45 R. Othman, T. S. Kiat, N. Khalid, R. Yusof, E. I. Newhouse, J. S. Newhouse, M. Alam and N. A. Rahman, *J. Chem. Inf. Model.*, 2008, **48**, 1582–1591.

46 A. Mukhametov, E. I. Newhouse, N. A. Aziz, J. A. Saito and M. Alam, *J. Mol. Graphics Modell.*, 2014, **52**, 103–113.

47 M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman and J. A. Hardy, *ACS Chem. Biol.*, 2013, **8**, 2744–2752.

48 M. Brecher, Z. Li, B. Liu, J. Zhang, C. A. Koetzner, A. Alifarag, S. A. Jones, Q. Lin, L. D. Kramer and H. Li, *PLoS Pathog.*, 2017, **13**, e1006411.

49 D. Aguilera-Pesantes, L. E. Robayo, P. E. Méndez, D. Mollocana, Y. Marrero-Ponce, F. J. Torres and M. A. Méndez, *Biochem. Biophys. Res. Commun.*, 2017, **492**, 631–642.

50 K. J. Chappell, M. J. Stoermer, D. P. Fairlie and P. R. Young, *J. Gen. Virol.*, 2008, **89**, 1010–1014.

51 K. Chappell, M. Stoermer, D. Fairlie and P. Young, *Curr. Med. Chem.*, 2008, **15**, 2771–2784.

52 É. E. da S. Rodrigues, H. Maus, S. J. Hammerschmidt, A. Ruggieri, E. C. Santos, Ê. J. Bassi, L. Anderson, P. G. V. Aquino, J. X. Araújo-Júnior, F. Wei, X. Liu, P. Zhan, T. Schirmeister and E. F. Silva-Júnior, in *Human Viruses: Diseases, Treatments and Vaccines: The New Insights*, ed. S. I. Ahmad, Springer International Publishing, Cham, 2021, pp. 233–295.

53 H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri and T. Schirmeister, *Bioorg. Med. Chem.*, 2021, **47**, 116392.

54 B. Millies, F. von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich and T. Schirmeister, *J. Med. Chem.*, 2019, **62**, 11359–11382.

55 C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister and T. Basché, *J. Phys. Chem. B*, 2021, **125**, 6837–6846.

56 H. Wu, S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W. E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weißbrich, R. Spannaus, E. G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister and J. Bodem, *Antimicrob. Agents Chemother.*, 2015, **59**, 1100–1109.

57 F. Benmansour, I. Trist, B. Coutard, E. Decroly, G. Querat, A. Brancale and K. Barral, *Eur. J. Med. Chem.*, 2017, **125**, 865–880.

58 N. Frimayanti, S. M. Zain, V. S. Lee, H. A. Wahab, R. Yusof and N. A. Rahman, *In Silico Biol.*, 2011, **11**, 29–37.

59 J. P. Quek, S. Liu, Z. Zhang, Y. Li, E. Yihui, Y. Ru, A. W. Hung, D. Luo and C. Kang, *Antiviral Res.*, 2020, **175**, 104707.

60 A. Kumar, A. Voet and K. Y. J. Zhang, *Curr. Med. Chem.*, 2012, **19**, 5128–5147.

61 C. Nitsche, C. Steuer and C. D. Klein, *Bioorg. Med. Chem.*, 2011, **19**, 7318–7337.

62 C. Steuer, C. Gege, W. Fischl, K. H. Heinonen, R. Bartenschlager and C. D. Klein, *Bioorg. Med. Chem.*, 2011, **19**, 4067–4074.

63 Y. Tanaka, S. Niwa, H. Nishioka, T. Yamanaka, M. Torizuka, K. Yoshinaga, N. Kobayashi, Y. Ikeda and H. Arai, *J. Med. Chem.*, 1994, **37**, 2071–2078.

64 J. Li, E. Wilk and S. Wilk, *Arch. Biochem. Biophys.*, 1995, **323**, 148–154.

65 T. E. Hughes, M. D. Mone, M. E. Russell, S. C. Weldon and E. B. Villhauer, *Biochemistry*, 1999, **38**, 11597–11603.

66 P.-C. Li, J. Jang, C.-Y. Hsia, P. V. Groomes, W. Lian, M. de Wispelaere, J. D. Pitts, J. Wang, N. Kwiatkowski, N. S. Gray and P. L. Yang, *ACS Infect. Dis.*, 2019, **5**, 460–472.

67 D. Lagorce, N. Oliveira, M. A. Miteva and B. O. Villoutreix, *Drug Discovery Today*, 2017, **22**, 1131–1133.

68 M. Pouliot and S. Jeanmart, *J. Med. Chem.*, 2016, **59**, 497–503.

69 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719–2740.

70 J. B. Baell and J. W. M. Nissink, *ACS Chem. Biol.*, 2018, **13**, 36–44.

71 J. L. Catrow, Y. Zhang, M. Zhang and H. Ji, *J. Med. Chem.*, 2015, **58**, 4678–4692.

72 J. Baell and M. A. Walters, *Nature*, 2014, **513**, 481–483.

73 C. Aldrich, C. Bertozzi, G. I. Georg, L. Kiessling, C. Lindsley, D. Liotta, K. M. Merz, A. Schepartz and S. Wang, *ACS Med. Chem. Lett.*, 2017, **8**, 379–382.

74 G. Ali, F. Subhan, N. U. Islam, I. Khan, K. Rauf, Samiullah, M. Abbas and A. Rauf, *J. Chem. Soc. Pak.*, 2014, **36**, 150–169.

75 L. Lima and E. Barreiro, *Curr. Med. Chem.*, 2005, **12**, 23–49.

76 K. van Beurden, S. de Koning, D. Molendijk and J. van Schijndel, *Green Chem. Lett. Rev.*, 2020, **13**, 349–364.

77 R. Khare, J. Pandey, S. Smriti and R. Ruchi, *Orient. J. Chem.*, 2019, **35**, 423–429.

78 R. J. Ouellette and J. D. Rawn, *Organic Chemistry*, Elsevier, 2018, pp. 711–762.

79 M. M. Heravi, F. Janati and V. Zadsirjan, *Monatshefte Chem. – Chem. Mon.*, 2020, **151**, 439–482.

80 R. Martin and S. L. Buchwald, *Acc. Chem. Res.*, 2008, **41**, 1461–1473.

81 C. Len, S. Bruniaux, F. Delbecq and V. Parmar, *Catalysts*, 2017, **7**, 146.

82 M. J. Buskes and M.-J. Blanco, *Molecules*, 2020, **25**, 3493.

83 W. W. Phoo, A. El Sahili, Z. Zhang, M. W. Chen, C. W. Liew, J. Lescar, S. G. Vasudevan and D. Luo, *Antiviral Res.*, 2020, **182**, 104900.

84 D. Kumar, M. K. Meena, K. Kumari, R. V. Kumar, I. Bahadur, P. Jain and P. Singh, *J. Mol. Liq.*, 2021, **335**, 116164.

85 B. Nutho and T. Rungrotmongkol, *J. Mol. Graphics Modell.*, 2019, **92**, 227–235.

86 P. Yotmanee, T. Rungrotmongkol, K. Wichapong, S. B. Choi, H. A. Wahab, N. Kungwan and S. Hannongbua, *J. Mol. Graphics Modell.*, 2015, **60**, 24–33.

87 M. Mottin, R. C. Braga, R. A. da Silva, J. H. M. da Silva, A. L. Perryman, S. Ekins and C. H. Andrade, *Biochem. Biophys. Res. Commun.*, 2017, **492**, 643–651.

88 L. Lim, G. Gupta, A. Roy, J. Kang, S. Srivastava, J. Shi and J. Song, *Prog. Biophys. Mol. Biol.*, 2019, **143**, 52–66.
89 S. B. Mirza, R. E. Salmas, M. Q. Fatmi and S. Durdagi, *J. Mol. Graphics Modell.*, 2016, **66**, 99–107.
90 A. K. Timiri, S. Selvarasu, M. Kesherwani, V. Vijayan, B. N. Sinha, V. Devadasan and V. Jayaprakash, *Bioorg. Chem.*, 2015, **62**, 74–82.
91 M. K. Gilson and B. Honig, *Proteins: Struct., Funct., Genet.*, 1988, **4**, 7–18.
92 R. C. Rizzo, T. Aynechi, D. A. Case and I. D. Kuntz, *J. Chem. Theory Comput.*, 2006, **2**, 128–139.
93 D. Sitkoff, K. A. Sharp and B. Honig, *J. Phys. Chem.*, 1994, **98**, 1978–1988.
94 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.
95 R. Kumari, R. Kumar and A. Lynn, *J. Chem. Inf. Model.*, 2014, **54**, 1951–1962.
96 N. Homeyer and H. Gohlke, *Mol. Inf.*, 2012, **31**, 114–122.
97 J. Wang, R. Dixon and P. A. Kollman, *Proteins*, 1999, **34**, 69–81.
98 N. M. Henriksen, H. S. Hayatshahi, D. R. Davis and T. E. Cheathan III, *J. Chem. Inf. Model.*, 2014, **54**, 1758–1772.
99 X.-Y. Meng, H.-X. Zhang, M. Mezei and M. Cui, *Curr. Comput.-Aided Drug Des.*, 2011, **7**, 146–157.
100 F. Stanzione, I. Giangreco and J. C. Cole, in *Progress in Medicinal Chemistry*, ed. D. R. Witty and B. Cox, Elsevier B.V., 2021, pp. 273–343.
101 G. F. S. Passos, M. G. M. Gomes, T. M. de Aquino, J. X. de Araújo-Júnior, S. J. M. de Souza, J. P. M. Cavalcante, E. C. dos Santos, Ê. J. Bassi and E. F. da Silva-Júnior, *Pharmaceuticals*, 2020, **13**, 141.
102 L. R. Silva, A. S. Guimarães, J. do Nascimento, I. J. do Santos Nascimento, E. B. da Silva, J. H. McKerrow, S. H. Cardoso and E. F. da Silva-Júnior, *Bioorg. Med. Chem.*, 2021, **41**, 116213.
103 M. Bassetto, T. De Burghgraeve, L. Delang, A. Massarotti, A. Coluccia, N. Zonta, V. Gatti, G. Colombano, G. Sorba, R. Silvestri, G. C. Tron, J. Neyts, P. Leyssen and A. Brancale, *Antiviral Res.*, 2013, **98**, 12–18.
104 G. M. Morris and M. Lim-Wilby, in *Molecular Modeling of Proteins. Methods Molecular Biology*, ed. A. Kukol, Humana Press Inc., 2008, pp. 365–382.
105 A. Tiwari and S. Singh, *Bioinformatics*, Elsevier, 2022, pp. 207–217.
106 M. T. ul Qamar, S. Kiran, U. A. Ashfaq, M. R. Javed, F. Anwar, M. A. Ali and A. ul H. Gilani, *Int. J. Pharmacol.*, 2016, **12**, 621–632.
107 U. S. F. Tambunan, I. Kurniawan and A. A. Parikesit, *Afr. J. Biotechnol.*, 2013, **12**, 4419–4431.
108 F. Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister and U. A. Hellmich, *FEBS Lett.*, 2019, **593**, 2204–2213.
109 R. J. Leatherbarrow, *GraFit version 6*, Erithacus Software Ltd., Horley, UK, 2007.
110 S. Ludewig, M. Kossner, M. Schiller, K. Baumann and T. Schirmeister, *Curr. Top. Med. Chem.*, 2010, **10**, 368–382.
111 D. A. Craig, *Trends Pharmacol. Sci.*, 1993, **14**, 89–91.
112 M. Lew, *Trends Pharmacol. Sci.*, 1995, **16**, 328–337.
113 H. C. Cheng, *J. Pharmacol. Toxicol. Methods*, 2001, **46**, 61–71.
114 G. Amendola, R. Ettari, S. Previti, C. Di Chio, A. Messere, S. Di Maro, S. J. Hammerschmidt, C. Zimmer, R. A. Zimmermann, T. Schirmeister, M. Zappalà and S. Cosconati, *J. Chem. Inf. Model.*, 2021, **61**, 2062–2073.
115 G. P. Maier, C. M. Bernt and A. Butler, *Biomater. Sci.*, 2018, **6**, 332–339.
116 A. Angelini, L. Cendron, S. Chen, J. Touati, G. Winter, G. Zanotti and C. Heinis, *ACS Chem. Biol.*, 2012, **7**, 817–821.
117 M. Poreba, A. Szalek, W. Rut, P. Kasperkiewicz, I. Rutkowska-Wlodarczyk, S. J. Snipas, Y. Itoh, D. Turk, B. Turk, C. M. Overall, L. Kaczmarek, G. S. Salvesen and M. Drag, *Sci. Rep.*, 2017, **7**, 43135.
118 T. Steinmetzer, A. Schweinitz, A. Stürzebecher, D. Dönnecke, K. Uhland, O. Schuster, P. Steinmetzer, F. Müller, R. Friedrich, M. E. Than, W. Bode and J. Stürzebecher, *J. Med. Chem.*, 2006, **49**, 4116–4126.
119 R. Ettari, E. Nizi, M. E. Di Francesco, M.-A. Dude, G. Pradel, R. Vičík, T. Schirmeister, N. Micale, S. Grasso and M. Zappalà, *J. Med. Chem.*, 2008, **51**, 988–996.
120 A. Welker, C. Kersten, C. Müller, R. Madhugiri, C. Zimmer, P. Müller, R. Zimmermann, S. Hammerschmidt, H. Maus, J. Ziebuhr, C. Sotriffer and T. Schirmeister, *ChemMedChem*, 2021, **16**, 340–354.
121 P. Kirsch, A. M. Hartman, A. K. H. Hirsch and M. Empting, *Molecules*, 2019, **24**, 4309.
122 P. J. Hajduk and J. Greer, in *Ernst Schering Foundation Symposium Proceedings*, 2007, pp. 211–219.
123 M. Thompson, The 228th ACS National Meeting, Planaria Software LLC, Philadelphia, 2004.
124 A. Gibbs, B. Tounge and R. Steele, NMR structure of an inhibitor bound dengue NS3 protease, **https://www.rcsb.org/structure/2M9P** (accessed 4 November 2021).
125 G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–748.
126 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
127 W. L. DeLano, *The PyMOL Molecular Graphics System*, Schrödinger LLC, USA, 2006.
128 K. Lozano Untiveros, E. G. da Silva, F. C. de Abreu, E. F. da Silva-Júnior, J. X. de Araújo-Junior, T. Mendoça de Aquino, S. M. Armas, R. O. de Moura, F. J. B. Mendonça-Junior, V. L. Serafim and K. Chumbimuni-Torres, *Biosens. Bioelectron.*, 2019, **133**, 160–168.
129 C. C. Santana, E. F. Silva-Júnior, J. C. N. Santos, É. E. da, S. Rodrigues, I. M. da Silva, J. X. Araújo-Júnior, T. G. do Nascimento, L. A. Oliveira Barbosa, C. B. Dornelas, I. M. Figueiredo, J. C. C. Santos and L. A. M. Grillo, *Bioorg. Chem.*, 2019, **87**, 169–180.
130 K. M. Roque Marques, M. R. do Desterro, S. M. de Arruda, L. N. de Araújo Neto, M. do Carmo Alves de Lima, S. M. V. de Almeida, E. C. D. da Silva, T. M. de Aquino,

E. F. da Silva-Júnior, J. X. de Araújo-Júnior, M. de, M. Silva, M. D. de, A. Dantas, J. C. C. Santos, I. M. Figueiredo, M.-A. Bazin, P. Marchand, T. G. da Silva and F. J. B. Mendonça Junior, *Curr. Top. Med. Chem.*, 2019, **19**, 1075–1091.

131 R. A. Marques, A. O. C. V. Gomes, M. V. de Brito, A. L. P. dos Santos, G. S. da Silva, L. B. de Lima, F. M. Nunes, M. C. de Mattos, F. C. E. de Oliveira, C. do, Ó. Pessoa, M. O. de Moraes, Â. de Fátima, L. L. Franco, M. de M. Silva, M. D. de, A. Dantas, J. C. C. Santos, I. M. Figueiredo, E. F. da Silva-Júnior, T. M. de Aquino, J. X. de Araújo-Júnior, M. C. F. de Oliveira and A. A. Leslie Gunatilaka, *J. Photochem. Photobiol., B*, 2018, **179**, 156–166.

132 M. de M. Silva, T. S. Macedo, H. M. P. Teixeira, D. R. M. Moreira, M. B. P. Soares, A. L. Ana, V. de, L. Serafim, F. J. B. Mendonça-Júnior, M. do Carmo, A. de Lima, R. O. de Moura, E. F. da Silva-Júnior, J. X. de Araújo-Júnior, M. D. Maria, E. de, O. O. Nascimento, T. M. S. Maciel, T. M. de Aquino, I. M. Figueiredo and J. C. C. Santos, *J. Photochem. Photobiol., B*, 2018, **189**, 165–175.

133 E. F. da Silva-Junior, P. H. Barcellos Franca, F. F. Ribeiro, F. J. Bezerra Mendonca-Junior, L. Scotti, M. T. Scotti, T. M. de Aquino and J. X. de Araujo-Junior, *Curr. Comput.-Aided Drug Des.*, 2018, **14**, 68–78.

134 E. Silva-Júnior, P. França, L. Quintas-Júnior, F. Mendonça-Junior, L. Scotti, M. Scotti, T. de Aquino and J. Araújo-Júnior, *Curr. Comput.-Aided Drug Des.*, 2017, **13**, 1–9.

135 D. C. Rees, M. Congreve, C. W. Murray and R. Carr, *Nat. Rev. Drug Discovery*, 2004, **3**, 660–672.

136 H. Jhoti, G. Williams, D. C. Rees and C. W. Murray, *Nat. Rev. Drug Discovery*, 2013, **12**, 644.

137 A. Daina, O. Michielin and V. Zoete, *Sci. Rep.*, 2017, **7**, 42717.

138 T. D. Goddard, C. C. Huang and T. E. Ferrin, *Structure*, 2005, **13**, 473–482.

139 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

140 V. Zoete, M. A. Cuendet, A. Grosdidier and O. Michielin, *J. Comput. Chem.*, 2011, **32**, 2359–2368.

141 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.

142 P. J. Turner, Xmgrace, **https://cmg.soton.ac.uk/research/categories/visualisation-and-data-handling-software/xmgrace/** (accessed 30 July 2022).

143 H. Gohlke, C. Kiel and D. A. Case, *J. Mol. Biol.*, 2003, **330**, 891–913.

144 B. Kuhn and P. A. Kollman, *J. Med. Chem.*, 2000, **43**, 3786–3791.

145 W. Wang and P. A. Kollman, *J. Mol. Biol.*, 2000, **303**, 567–582.

146 F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, *J. Comput. Chem.*, 1995, **16**, 273–284.

147 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306–317.

148 K. M. Roque Marques, M. R. do Desterro, S. M. de Arruda, L. N. de Araújo Neto, M. do Carmo Alves de Lima, S. M. V. de Almeida, E. C. D. da Silva, T. M. de Aquino, E. F. da Silva-Júnior, J. X. de Araújo-Júnior, M. de M. Silva, M. D. de A. Dantas, J. C. C. Santos, I. M. Figueiredo, M.-A. Bazin, P. Marchand, T. G. da Silva and F. J. B. Mendonça Junior, *Curr. Top. Med. Chem.*, 2019, **19**, 1075–1091.

149 E. F. Silva-Junior, P. H. Barcellos Franca, L. J. Quintans-Junior, F. J. B. Mendonca-Junior, L. Scotti, M. T. Scotti, T. M. de Aquino and J. X. de Araujo-Junior, *Curr. Comput.-Aided Drug Des.*, 2017, **13**, 266–274.

150 T. C. Braga, T. F. Silva, T. M. S. Maciel, E. C. D. da Silva, E. F. da Silva-Júnior, L. V. Modolo, I. M. Figueiredo, J. C. C. Santos, T. M. de Aquino and Â. de Fátima, *New J. Chem.*, 2019, **43**, 15187–15200.

151 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.

152 A. M. Ferrari, G. Degliesposti, M. Sgobba and G. Rastelli, *Bioorg. Med. Chem.*, 2007, **15**, 7865–7877.

153 G. Rastelli, A. Del Rio, G. Degliesposti and M. Sgobba, *J. Comput. Chem.*, 2009, 797–810.

154 S. P. Brown and S. W. Muchmore, *J. Med. Chem.*, 2009, **52**, 3159–3165.

155 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 10037–10041.

156 G. Degliesposti, C. Portioli, M. D. Parenti and G. Rastelli, *J. Biomol. Screening*, 2011, **16**, 129–133.

157 A. Lindström, L. Edvinsson, A. Johansson, C. D. Andersson, I. E. Andersson, F. Raubacher and A. Linusson, *J. Chem. Inf. Model.*, 2011, **51**, 267–282.

158 G. Rastelli, G. Degliesposti, A. Del Rio and M. Sgobba, *Chem. Biol. Drug Des.*, 2009, **73**, 283–286.

159 K. H. Barakat, L. P. Jordheim, R. Perez-Pineiro, D. Wishart, C. Dumontet and J. A. Tuszynski, *PLoS One*, 2012, **7**, e51329.

160 T. Yang, J. C. Wu, C. Yan, Y. Wang, R. Luo, M. B. Gonzales, K. N. Dalby and P. Ren, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 1940–1951.

161 V. Kasam, J. Salzemann, M. Botha, A. Dacosta, G. Degliesposti, R. Isea, D. Kim, A. Maass, C. Kenyon, G. Rastelli, M. Hofmann-Apitius and V. Breton, *Malar. J.*, 2009, **8**, 88.

162 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712–725.

163 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 1950–1958.

164 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

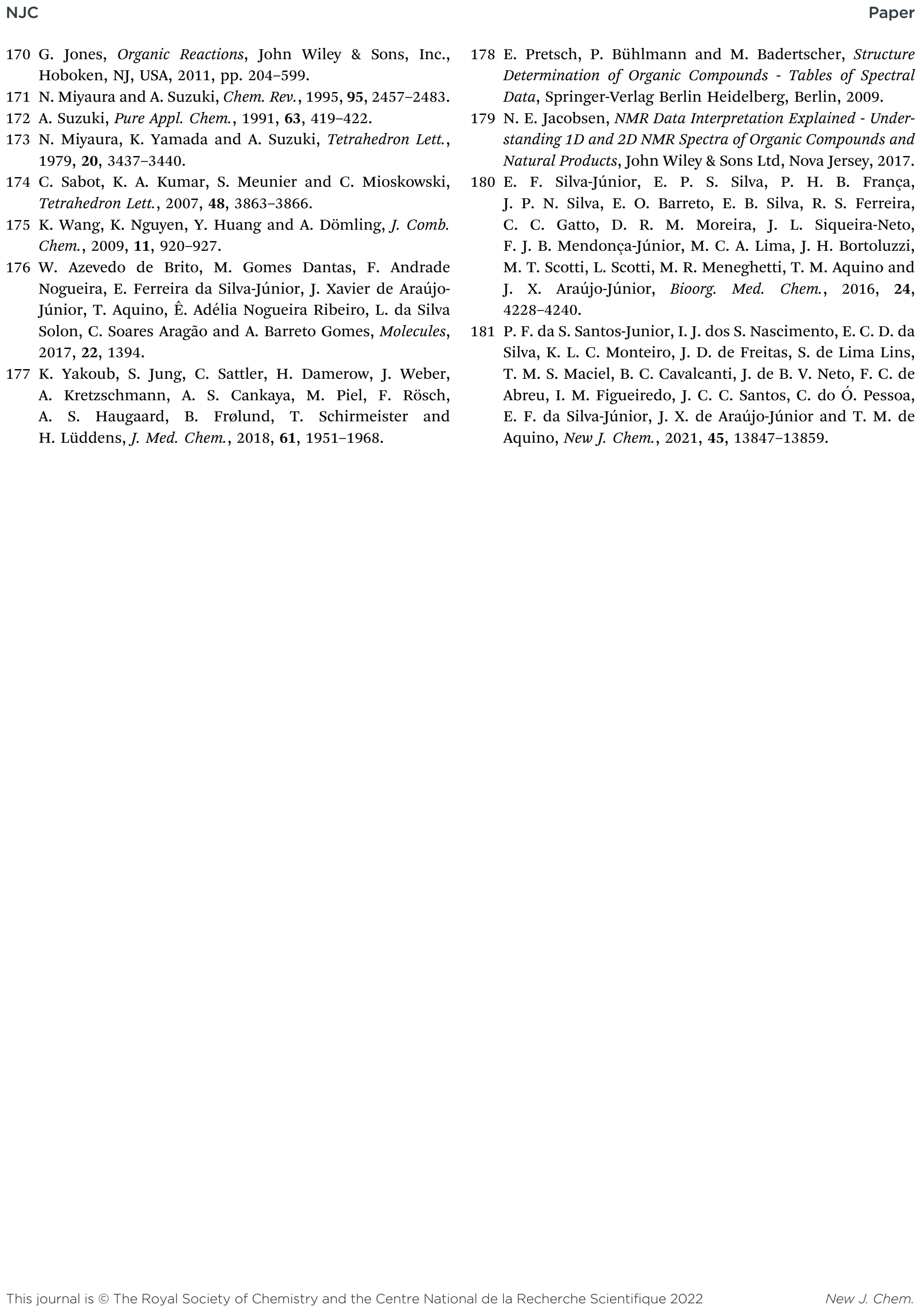
165 B. Honig and A. Nicholls, *Science*, 1995, **268**, 1144–1149.

166 L. T. Chong, J. W. Pitera, W. C. Swope and V. S. Pande, *J. Mol. Graphics Modell.*, 2009, **27**, 978–982.

167 R. M. Levy, L. Y. Zhang, E. Gallicchio and A. K. Felts, *J. Am. Chem. Soc.*, 2003, **125**, 9523–9530.

168 C. Tan, Y.-H. Tan and R. Luo, *J. Phys. Chem. B*, 2007, **111**, 12263–12274.

169 J. A. Wagoner and N. A. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 8331–8336.



170 G. Jones, *Organic Reactions*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011, pp. 204–599.

171 N. Miyaura and A. Suzuki, *Chem. Rev.*, 1995, **95**, 2457–2483.

172 A. Suzuki, *Pure Appl. Chem.*, 1991, **63**, 419–422.

173 N. Miyaura, K. Yamada and A. Suzuki, *Tetrahedron Lett.*, 1979, **20**, 3437–3440.

174 C. Sabot, K. A. Kumar, S. Meunier and C. Mioskowski, *Tetrahedron Lett.*, 2007, **48**, 3863–3866.

175 K. Wang, K. Nguyen, Y. Huang and A. Dömling, *J. Comb. Chem.*, 2009, **11**, 920–927.

176 W. Azevedo de Brito, M. Gomes Dantas, F. Andrade Nogueira, E. Ferreira da Silva-Júnior, J. Xavier de Araújo-Júnior, T. Aquino, Ê. Adélia Nogueira Ribeiro, L. da Silva Solon, C. Soares Aragão and A. Barreto Gomes, *Molecules*, 2017, **22**, 1394.

177 K. Yakoub, S. Jung, C. Sattler, H. Damerow, J. Weber, A. Kretzschmann, A. S. Cankaya, M. Piel, F. Rösch, A. S. Haugaard, B. Frølund, T. Schirmeister and H. Lüddens, *J. Med. Chem.*, 2018, **61**, 1951–1968.

178 E. Pretsch, P. Bühlmann and M. Badertscher, *Structure Determination of Organic Compounds - Tables of Spectral Data*, Springer-Verlag Berlin Heidelberg, Berlin, 2009.

179 N. E. Jacobsen, *NMR Data Interpretation Explained - Understanding 1D and 2D NMR Spectra of Organic Compounds and Natural Products*, John Wiley & Sons Ltd, Nova Jersey, 2017.

180 E. F. Silva-Júnior, E. P. S. Silva, P. H. B. França, J. P. N. Silva, E. O. Barreto, E. B. Silva, R. S. Ferreira, C. C. Gatto, D. R. M. Moreira, J. L. Siqueira-Neto, F. J. B. Mendonça-Júnior, M. C. A. Lima, J. H. Bortoluzzi, M. T. Scotti, L. Scotti, M. R. Meneghetti, T. M. Aquino and J. X. Araújo-Júnior, *Bioorg. Med. Chem.*, 2016, **24**, 4228–4240.

181 P. F. da S. Santos-Junior, I. J. dos S. Nascimento, E. C. D. da Silva, K. L. C. Monteiro, J. D. de Freitas, S. de Lima Lins, T. M. S. Maciel, B. C. Cavalcanti, J. de B. V. Neto, F. C. de Abreu, I. M. Figueiredo, J. C. C. Santos, C. do Ó. Pessoa, E. F. da Silva-Júnior, J. X. de Araújo-Júnior and T. M. de Aquino, *New J. Chem.*, 2021, **45**, 13847–13859.

4.3 Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates.

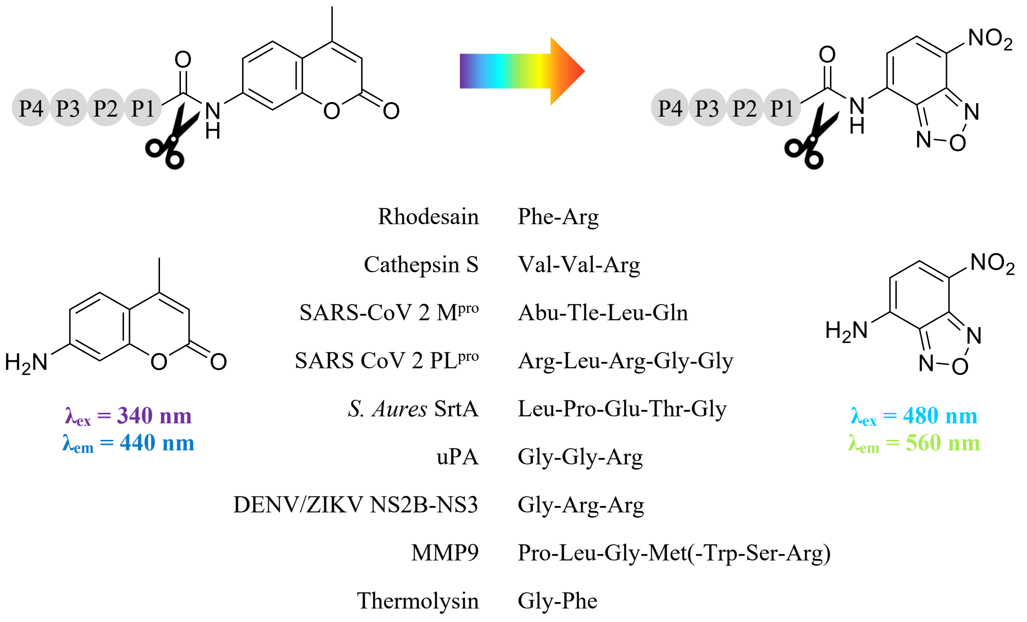
4.3.1 Context, Project Summary, and Own Contributions

Fluorometric assays are widely used in medicinal chemistry to analyze protease-inhibiting drug candidates due to their time efficiency and availability of fluorogenic protease substrates.[146,147] Fluorometric protease assays work by cleaving an amide bond in the substrate, changing fluorescence properties. Increasing fluorescence is measured, providing information on the reaction rate.[151] Inhibitors attenuate the enzyme's reaction rate, giving information about their inhibitory activity. However, it is crucial to address non-specific interference effects that can lead to false positive or negative results.[264–267] Shifting the assay wavelengths to the lower-energy range of light can mitigate these interferences.[268–270] The 7-nitrobenz-2-oxa-1,3-diazol-4-amine (NBD) fluorophore offers excitation and emission maxima in the fluorescein wavelength range, potentially avoiding common interferences associated with AMC and EDANS fluorophores.[148,150,271–273]

Within this study, we synthesized peptide-based NBD-amides and designed FRET substrate alternatives based on the observed Smiles rearrangement mechanism of selected NBD derivatives. The study describes the synthesis of substrates for ten different proteases from the cysteine, serine, and metalloprotease classes, including carboxypeptidases and endopeptidases (Figure 21). To validate the suitability of the NBD substrates, we determined enzyme- and substrate-specific kinetic parameters (K_M and k_{cat}) and compared them to the parent AMC and FRET substrates. In addition, the inhibition constants (K_i and k_{inact}) of known inhibitors determined with the new NBD substrates further confirmed the efficacy of the NBD substrates for various applications.

One notable advantage of using NBD substrates is their ability to avoid typical assay interferences encountered in medicinal chemistry-related assays, such as the inner filter effect, fluorophore quenching, and autofluorescence decay.[267,269] This was shown for an exemplarily NBD substrate compared to the corresponding AMC substrate.

In summary, we demonstrated that NBD substrates offer a suitable alternative to previously used fluorophores and have the potential to serve as the foundation for a new generation of fluorometric assays. Their successful synthesis and characterization, along with the favorable results obtained in enzyme kinetics and inhibitor studies, highlight the efficacy of NBD-based substrates for protease analysis in drug discovery applications.



Rhodesain	Phe-Arg
Cathepsin S	Val-Val-Arg
SARS-CoV 2 M^{pro}	Abu-Tle-Leu-Gln
SARS CoV 2 PL^{pro}	Arg-Leu-Arg-Gly-Gly
S. Aures SrtA	Leu-Pro-Glu-Thr-Gly
uPA	Gly-Gly-Arg
DENV/ZIKV NS2B-NS3	Gly-Arg-Arg
MMP9	Pro-Leu-Gly-Met(-Trp-Ser-Arg)
Thermolysin	Gly-Phe

Figure 21. Development of NBD-based substrates. Excitation and detection wavelengths are shifted to higher wavelengths compared to AMC, which allows for avoiding typical assay interferences. Substrates for ten different proteases were synthesized, and their suitability for use in the fluorometric assay was confirmed.

Own contributions: kinetic characterization of substrates, protease inhibition assays, investigations on assay interferences, writing of the first draft (kinetic characterization of NBD-based substrates, autocleavage of a DENV and ZIKV NS2B/NS3 protease substrate, harnessing NBD autocleavage for FRET substrate replacement, applications for protease inhibitor investigation, mitigation of typical protease assay interferences, conclusion, fluorometric assays) & editing of the manuscript.

Contribution from other authors: substrate synthesis, inhibitor synthesis, protein expression & purification, substrate stability assays, molecular docking, and writing parts of the first draft (introduction, synthesis, smartphone-based fluorometer for usage in classroom applications, chemistry) & editing of the manuscript.

This work has been published in Chemistry – A European Journal (impact factor: 5.02).

Article reprinted with permission of *Chemistry – A European Journal* **2023** “Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates.” © 2023 WILEY-VCH Verlag GmbH & Co. KGaA (Germany).

The appended Supporting Information represents an abridged version. The full version can be accessed online at doi: 10.1002/chem.202301855.

4.3.2 Publication

The following publication quoted within “” from page 102 to page 130 is the same as the manuscript cited on page 101.

“

Chemistry
A European
Journal

Chemistry
Europe
European Chemical
Societies Publishing

Accepted Article

Title: Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates

Authors: Hannah Maus, Patrick Müller, Mergim Meta, Sabrina N. Hoba, Stefan J. Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Tanja Schirmeister, and Fabian Barthels

To be cited as: *Chem. Eur. J.* **2023**, e202301855

Link to VoR: https://doi.org/10.1002/chem.202301855

WILEY-VCH

Chemistry - A European Journal 10.1002/chem.202301855

RESEARCH ARTICLE

Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates

Hannah Maus+[a], Patrick Müller+[a], Mergim Meta[a], Sabrina N. Hoba[a], Stefan J. Hammerschmidt[a], Robert A. Zimmermann[a], Collin Zimmer[a], Natalie Fuchs[a], Tanja Schirmeister[a], Fabian Barthels*[a]

[a] Hannah Maus, Patrick Müller, Mergim Meta, Sabrina N. Hoba, Stefan J. Hammerschmidt, Robert A. Zimmermann, Collin Zimmer, Natalie Fuchs, Prof. Dr. Tanja Schirmeister, Dr. Fabian Barthels
Institute of Pharmaceutical and Biomedical Sciences
Johannes Gutenberg-University
Staudingerweg 5, 55128 Mainz, Germany
E-mail: barthels@uni-mainz.de
Homepage: https://ak-barthels.pharmazie.uni-mainz.de/
[+] authors contributed equally

Supporting information for this article is given via a link at the end of the document.

Abstract: Fluorometric assays are one of the most frequently used methods in medicinal chemistry. Over the last 50 years, the reporter molecules for the detection of protease activity have evolved from first-generation colorimetric *p*-nitroanilides, through FRET substrates, and 7-amino-4-methyl coumarin (AMC)-based substrates. The aim of further substrate development is to increase sensitivity and reduce vulnerability to assay interferences. Here, we describe a new generation of substrates for protease assays based on 7-nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-amides). In this study, we synthesized and tested substrates for 10 different proteases from the serine-, cysteine-, and metalloprotease classes. Enzyme- and substrate-specific parameters as well as the inhibitory activity of literature-known inhibitors confirmed their suitability for application in fluorometric assays. Hence, we were able to present NBD-based alternatives for common protease substrates. In conclusion, these NBD substrates are not only less susceptible to common assay interference, but they are also able to replace FRET-based substrates with the requirement of a prime site amino acid residue.

Introduction

Fluorescence is one manifestation of the interaction of electromagnetic radiation with matter which is ubiquitously used for the analysis of all kinds of parameters in the life sciences.[1] It is widely used for the characterization of protein structures and protein/ligand interactions.[2] This comprises the analysis of protein structural changes, localization of proteins in cells, organisms, and the determination of binding affinities between proteins and their ligands.[3–6] Proteolytic enzymes represent a current field of research, because of their mechanistic involvement in many diseases from virus infections to cancer progression and immunological disorders.[7–9] Since 1964, more than 64 protease inhibitors were approved by the Food and Drug Administration (FDA) for application in human therapies[10,11]; exemplarily, the recently approved cysteine protease inhibitor Nirmatrelvir is being used to treat the SARS-CoV 2 infection.[12] This implies, that proteases are still important targets, and the scientific community strives for discovering new protease-targeting drugs.[13,14]

Due to their time efficiency and steadily growing supply of commercially available fluorogenic protease substrates, fluorometric assays are one of the most frequently used methods in medicinal chemistry to analyze protease-inhibiting drug candidates.[15,16] The first, in 1973 introduced fluorogenic substrates used for the measurement of protease activities were FRET-based (Förster resonance energy transfer) substrates which harbor a quencher and fluorophore molecular pair (Figure 1).[17] These replaced the previously utilized colorimetric *p*-nitroanilide (pNA) substrates, because of their higher detection sensitivity.[18] Only a few years later, in 1976, the corresponding amides of a fluorescent coumarin derivative (7-amino-4-methyl coumarin, AMC) were implemented as a new fluorogenic protease substrate which decreased the vulnerability to assay interferences compared to the FRET-based and colorimetric pNA substrates.[19]

In general, the principle of fluorometric protease assays is described as follows: The target enzyme cleaves an amide bond of the substrate, which leads to a change in fluorescence properties of the fluorogenic reporter, for example by separation of the quencher from the fluorophore (FRET substrate) or cleavage of the internally quenched fluorogenic residue from the substrate (AMC substrate).[20] The fluorophore is then excited within a cuvette fluorometer or a microplate reader at a specific excitation wavelength and the resulting fluorescence can be detected at its corresponding emission wavelength. The increasing fluorescence intensity over time describes a scale of enzymatic activity. By the addition of an inhibitor, the reaction rate of the enzyme is attenuated which gives information about the inhibitory activity by different evaluation methods. However, not less important is the resilience of a fluorometric assay system

Chemistry - A European Journal 10.1002/chem.202301855

WILEY VCH

regarding non-specific interference effects, that originate from various sources of errors like assay components, including the analyzed drug compounds themselves.[21–23] Assay interferences lead to both false positive and false negative results, which makes it necessary to validate potential hits with an additional, orthogonal methodology.[24]

A known strategy for mitigation of intrinsic assay interferences is based on the fact that many of the interfering physicochemical transitions do occur in the near UV-light spectrum, due to the molecular properties of mostly aromatic drug-like inhibitors.[25] One option to avoid inhibitor-induced interferences is to shift the assay wavelengths to the lower-energy range of light, and hence, there are existing biochemical methods using fluorescent dyes (BODIPY, Cy5, etc.) with their excitation and emission maxima in the red wavelength region (>580 nm), leading to minimized interferences.[26,27] Over the last 30 years, FRET substrates were optimized by an overall shift to longer wavelengths like Dabcyl/EDANS resulting in optimized sensitivity, photochemical, and low-interference properties.[28] In contrast to the most frequently used AMC- ($\lambda_{\mathrm{ex}} = 380\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 460\ \mathrm{nm}$) and FRET-based (Dabcyl/EDANS; $\lambda_{\mathrm{ex}} = 340\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 490\ \mathrm{nm}$) fluorogenic substrates, the 4-amino-7-nitro-benzofurazane (NBD) fluorophore displays excitation and emission maxima in the fluorescein wavelength magnitude ($\lambda_{\mathrm{ex}} = 485\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 535\ \mathrm{nm}$) and might consequently avoid common interferences associated with AMC and EDANS fluorophores.[17,19,29–31]

Applications for NBD-based chemosensors have already been described in the literature, leading to several hundreds of publications and the development of numerous commercial probes which have already been summarized in reviews.[32] For the analysis of enzymatic turnover, however, the use of NBD-based derivates has been only described for the study of histone acetyltransferases, deacetylases, and esterases.[33–37] To expand the scope, herein, we report the synthesis and characterization of peptide-based NBD-amides as substrates for various protease targets. Additionally, due to the observed Smiles rearrangement mechanism of selected NBD derivates, we were able to design substrates suited as corresponding FRET substrate alternatives (Figure 1).[38,39]

To evaluate the substrate affinity and turnover rate of the NBD-based substrates, we determined K_{M} and k_{cat} values. Both substrate benchmarks were compared with the corresponding literature known, AMC- or FRET-derived parent substrates. To demonstrate that this new assay protocol provides comparable results to the literature-described assays, K_{i} values were determined for model inhibitors known from the literature. Afterward, the mitigation of common assay interferences of conventional fluorogenic substrates was investigated by comparing a NBD-based substrate with its AMC-based counterpart.

Results and Discussion

Synthesis

All NBD substrates were prepared in a two-step synthetic procedure. First, the variable peptide sequences **2a–j** spanning from the P2-site to the *N*-capped terminus were synthesized following a standard fluorenylmethoxycarbonyl (Fmoc)-solid phase peptide synthesis (SPPS) protocol using a 2-chlorotrityl resin.[40]

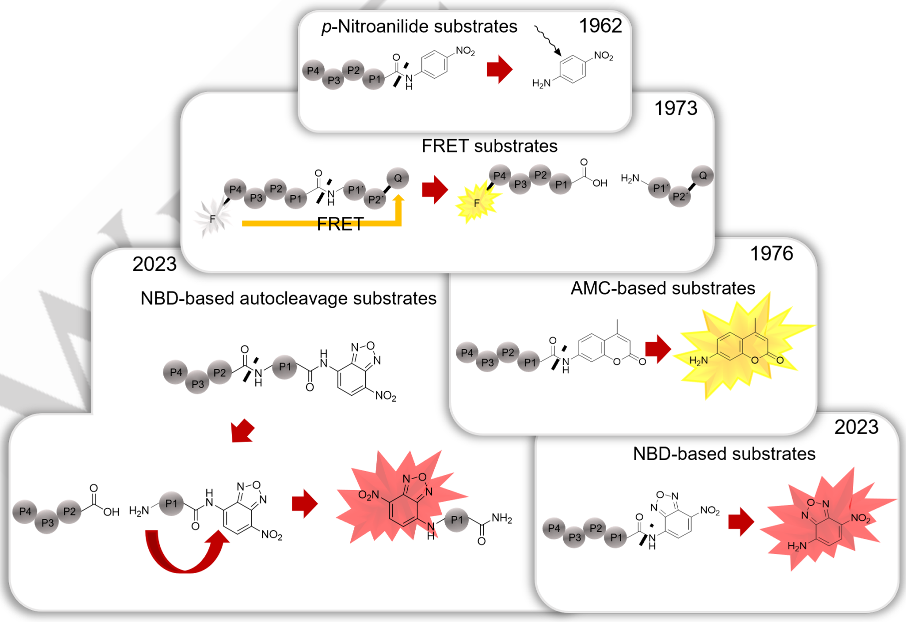


Figure 1. Timeline of protease substrates development with selected example substrates including *p*-nitroanilides (first described in 1962), FRET- (1973), and AMC-based (1976) fluorogenic substrates. In this study, we report the development of newly designed NBD-based protease substrates for biochemical and medicinal chemistry assays.

Chemistry - A European Journal 10.1002/chem.202301855
WILEY VCH
RESEARCH ARTICLE

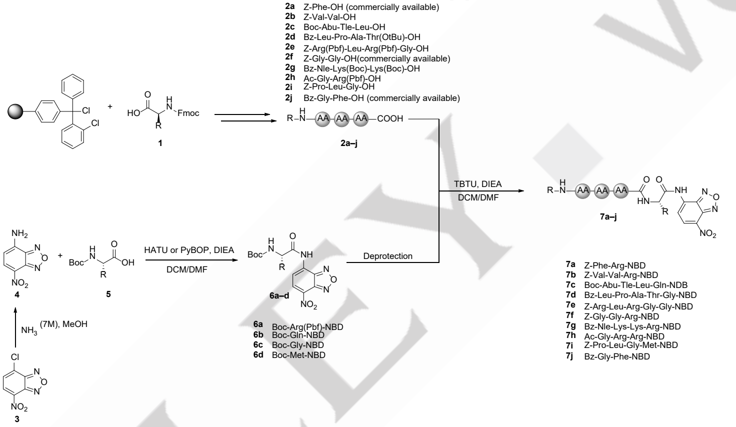
The second step in the synthetic sequence was the preparation of the fluorogenic reporter group through a nucleophilic aromatic substitution of precursor **3** in a methanolic ammonia solution.[41] Subsequently, the NBD amine **4** was coupled with the respective *tert*-butoxycarbonyl (Boc)-protected P1 amino acid **5**, using the *in situ* activating agents *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium-hexafluorophosphate (HATU) or benzotriazole-1-yl-oxytripyrrolidinophosphonium-hexafluorophosphate (PyBOP), and the organic base *N*,*N*-diisopropylethylamine (DIEA). The amino group of the intermediate **6a–d** was deprotected by trifluoroacetic acid, and subsequently, coupled with the remaining peptide sequence **2** to yield the desired fluorogenic substrates **7a–j** (Scheme 1, Table 1).

Kinetic Characterization of NDB-based Substrates

To verify and analyze the suitability of the proposed NBD-based substrates for fluorometric protease assays, the enzyme kinetic parameters K_m, v_{max}, k_{cat}, and the catalytic efficiency k_{cat}/K_M were determined for both the NBD substrates and their corresponding parent substrates. Proteolytic cleavage of the substrate release was detected in 96-well plate format (typically 200 µL). NBD substrates were excited at 485 nm, and the emission was detected at 535 nm.[42,43] The increase in fluorescence intensity over time (reaction rate) is a measure of the activity of the protease, and thus, the fluorescent progress traces follow the Michaelis Menten equation (Eq.1) which was used for enzyme kinetic analysis:

$$v_0 = \frac{v_{\max}\cdot[S]}{K_M+[S]} \qquad (1)$$

where v_0 is the initial reaction rate, at a given substrate concentration [*S*]. v_{max} is the maximum reaction rate, while K_M indicates [*S*] at which the conversion rate is half-maximal.



Scheme 1. SPPS/solution phase synthesis of the NBD-based substrates suited for cysteine-, serine-, and metalloproteases.

Table 1. Synthesized NBD-based substrates and their corresponding AMC-/FRET-based parent substrates.

Protease	Protease class	NBD-based substrate	Reference/parent substrate
T. brucei rhodesain	Cysteine	Z-Phe-Arg-NBD (**7a**)	Z-Phe-Arg-AMC[44]
Cathepsin S (human)	Cysteine	Z-Val-Val-Arg-NBD (**7b**)	Z-Val-Val-Arg-AMC[45]
SARS-CoV 2 M^{pro}	Cysteine	Boc-Abu-Tle-Leu-Gln-NBD (**7c**)	Boc-Abu-Tle-Leu-Gln-AMC[46]
S. aureus SrtA	Cysteine	Bz-Leu-Pro-Ala-Thr-Gly-NBD (**7d**)	Abz-Leu-Pro-Glu-Thr-Gly-Dap(Dnp)–OH[47]
SARS-CoV 2 PL^{pro}	Cysteine	Z-Arg-Leu-Arg-Gly-Gly-NBD (**7e**)	Z-Arg-Leu-Arg-Gly-Gly-AMC[48]
uPA (human)	Serine	Z-Gly-Gly-Arg-NBD (**7f**)	Z-Gly-Gly-Arg-AMC[49]
DENV NS2B/NS3 ZIKV	Serine	Bz-Nle-Lys-Lys-Arg-NBD (**7g**) Ac-Gly-Arg-Arg-NBD (**7h**)	Bz-Nle-Lys-Lys-Arg-AMC[50] Ac-Gly-Arg-Arg-AMC[51]
MMP9 (human)	Metallo	Z-Pro-Leu-Gly-Met-NBD (**7i**)	Dnp-Pro-Leu-Gly-Met-Trp-Ser-Arg-NH_2[52]
Thermolysin	Metallo	Bz-Gly-Phe-NBD (**7j**)	*N*-[3-(2-furyl)acryloyl]-glycyl-L-leucine amide (FAGLA)[53,54]

Chemistry - A European Journal 10.1002/chem.202301855
WILEY-VCH

The fluorescence progress curves were acquired for at least six different substrate concentrations for the corresponding substrate with at least a technical triplicate. K_M and v_{max} values were then calculated from the respective Michaelis-Menten curves (Table 2 and Figures S55–S63). Exemplarily, fluorescence progress curves for different substrate concentrations and the resulting Michaelis-Menten plot are shown for the cysteine protease rhodesain in Figure 2. The signal-to-noise ratio, as described by the detector-independent reaction rate slope ([au·s^{-1}] vs. [S]), was ~3 times higher for NBD than for AMC-based substrates (calculated by the gain factor-independent initial slope of Figure 2C vs. 2D). Consequently, the NBD-based fluorescence assays can be performed at lower enzyme and/or substrate concentrations than with the corresponding parent substrate. Especially for highly affine ligands, this might provide an advantage to avoid tight-binding inhibition ([E]>K_i).[55]

In general, the K_M values for the NBD-based substrates were found to be in the same order of magnitude or better than the K_M values determined for the AMC-based parent substrates, indicating similar or increased affinity compared to their AMC counterpart. While the K_M value of the cysteine proteases rhodesain, CatS, and M^{pro} substrates is virtually unchanged by the exchange of the fluorophore from AMC to NBD, the affinity of uPA, DENV/ZIKV NS2B/NS3, and PL^{pro} substrates is increased by up to a factor of ~40. The affinity increase was reasoned by analyzing the binding poses predicted by molecular docking. Comparing the docking poses of the AMC and the NBD substrate, it is noticeable that the positioning of the substrate peptide is mostly not affected by the exchange of the fluorophore residue. However, the positioning of the fluorophore itself might differ depending on the individual topology of the binding pocket. By its chemical nature, the NBD substructure can form more hydrophilic interactions compared to AMC as highlighted by the interaction analysis (Figure 3 and Figure S73). The formation of productive interactions with the charge-polarized catalytic dyad of a protease could be the reason for the improved affinity of NBD-based substrates. Some examples in which the NDB residue mediates increased affinity for protein and lipid surfaces have been previously documented.[32]

Considering the kinetics of catalytic cleavage, for the rhodesain, uPA, and PL^{pro} substrates, k_{cat} is lower for these NBD-based substrates than for the AMC-based substrates (10–100×). For the other enzymes, the k_{cat} values are in the same magnitude or slightly higher for the NBD-based substrates. Thus, no clear trend is apparent regarding the substrate cleavage rate, which seems to be an individual characteristic for each protease probably depending on the positioning of the susceptible amide bond towards the nucleophilic enzyme residue. Except for the rhodesain substrate, the catalytic efficiency k_{cat}/K_M of the NBD-based substrates was found to be as high as or higher than the one of the AMC-based substrates. In the case of the rhodesain NBD substrate (**7a**) the reduction of the k_{cat} value actually represents an advantage over the AMC substrate, since the fluorescence progress curves remain linear for a longer time in the steady-state regiment (Figure 2A) and there is no early saturation of the fluorescence by reaching an equilibrium, as is the case with the AMC substrate (Figure 2B).

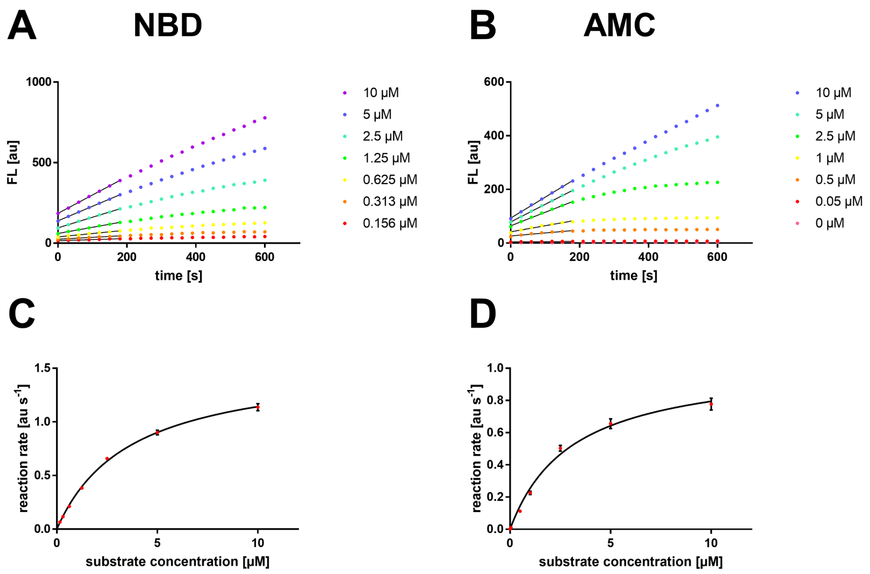


Figure 2. Exemplary fluorometric enzyme assay graphs for rhodesain. (**A**) Fluorescence progress curves for the conversion of different concentrations of the Z-Phe-Arg-NBD (**7a**) substrate by rhodesain. (**B**) Fluorescence progress curves for Z-Phe-Arg-AMC by rhodesain. (**C**) Michaelis-Menten plot for Z-Phe-Arg-NBD (**7a**). (**D**) Michaelis-Menten plot for Z-Phe-Arg-AMC.

Table 2. Enzyme- and substrate-specific kinetic constants for NBD-based and corresponding AMC-based substrates. Michaelis-Menten constant (K_M), maximum turnover rate (v_{max}), turnover number (k_{cat}), catalytic efficiency (k_{cat}/K_M).

protease	NBD substrate				AMC substrate			
	K_M [μM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·μmol^{-1} min^{-1}]	K_M [μM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·μmol^{-1} min^{-1}]
Rhodesain	2.84 ± 0.22	73.8 ± 2.6	59.0 ± 2.1	20.8 ± 1.8	3.08 ± 0.41	679 ± 43	1056 ± 66	343 ± 50
uPA	10.2 ± 0.63	174 ± 6	[a]	[b]	347 ± 35	1847 ± 90	[c]	[d]
CatS	28.0 ± 2.68	70.0 ± 3.1	7.00 ± 0.31	0.25 ± 0.03	34.4 ± 3.06	111 ± 23	0.45 ± 0.09	0.013 ± 0.003
DENV NS2B/NS3	88.1 ± 11.7	762 ± 37	3.05 ± 0.15	0.035 ± 0.005	869 ± 67	41.8 ± 1.9	0.17 ± 0.01	(1.92 ± 0.17) 10^{-4}
ZIKV NS2B/NS3	33.6 ± 1.51	252 ± 5	10.1 ± 0.2	0.30 ± 0.01	1381 ± 371	161 ± 34	1.29 ± 0.27	(9.3 ± 3.3) 10^{-4}
PL^{pro}	40.6 ± 4.11	151 ± 8	1.51 ± 0.08	0.037 ± 0.004	1332 ± 161	7049 ± 427	70.5 ± 4.3	0.053 ± 0.007
M^{pro}	18.9 ± 1.38	10.9 ± 0.4	2.62 ± 0.09	0.14 ± 0.01	56.7 ± 11.6	2.06 ± 0.16	0.50 ± 0.04	0.009 ± 0.002

The concentration of the uPA protein in mol·L^{-1} is unknown but given in units (U) by the vendor: [a] ($1.74\cdot10^{-5}$ ± $1.20\cdot10^{-7}$) mol·min^{-1}·U^{-1}; [b] (1.71 ± 0.12) L·min^{-1}·U^{-1}; [c] ($1.85\cdot10^{-4}$ ± $9.00\cdot10^{-6}$) mol·min^{-1}·U^{-1}; [d] (0.53 ± 0.06) L·min^{-1}·U^{-1}.

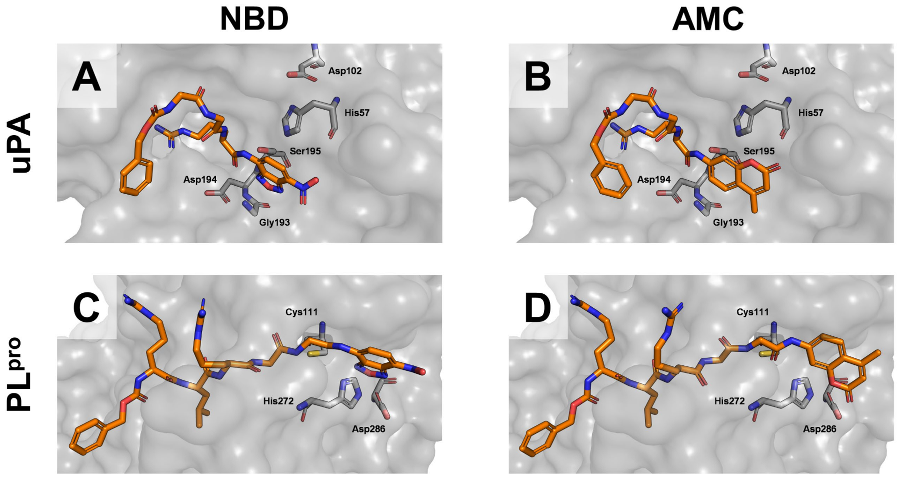


Figure 3. Predicted binding poses for the NBD- and AMC-based substrates for uPA and PL^{pro}. (**A**) Z-Gly-Gly-Arg-NBD (**7f**) in the active site of the crystal structure of uPA (pdb: 1LMW). (**B**) Z-Gly-Gly-Arg-AMC in the crystal structure of uPA (pdb: 1LMW). (**C**) Z-Arg-Leu-Arg-Gly-Gly-NBD (**7e**) in the crystal structure of PL^{pro} (pdb: 7RBS). (**D**) Z-Arg-Leu-Arg-Gly-Gly-AMC in the crystal structure of PL^{pro} (pdb: 7RBS).

Autocleavage of a DENV and ZIKV NS2B/NS3 Protease Substrate

Remarkably, in contrast to all other substrates presented in this study, our first DENV and ZIKV NS2B/NS3 protease substrate **7g** with the sequence Bz-Nle-Lys-Lys-Arg-NBD (Table 1) showed a fluorescence increase even in the absence of any enzyme. In an alkaline buffered solution, this fluorescence increase was found to be higher than at neutral pH values (Figure 4A and B). To analyze the molecular mechanism of this non-enzymatic fluorescence activation, the reaction mixture of the NS2B/NS3 substrate conversion was analyzed by HPLC/MS. By this, two LC peaks of the same mass (405.70 Da) but different fluorescence properties were detected, one corresponding to the substrate educt **7g** and the other to a fluorescent product **8**, suggesting an intramolecular rearrangement mechanism for fluorescence auto-activation (Figure S49). In the literature, an intramolecular Smiles rearrangement for NBD derivatives incorporating primary amine groups has been described during a histone deacetylase assay development.[37,39]

From these and the literature results, we hypothesized that substrate auto-cleavage is enabled by the presence of free amine groups in the substrate sequence (Figure 4C). Appropriately, we found that substrates without lysine residues are stable in their respective assay buffer. This was confirmed both by mass spectroscopy and by fluorescence spectroscopy (Figures S50–S52). Nonetheless, we aimed to develop a functional substrate for the NS2B/NS3 proteases, and thus, we replaced the lysine residues with arginine residues, resulting in a stable and functional substrate for both proteases (Ac-Gly-Arg-Arg-NBD **7h**, Table 2). Noteworthy, even high concentrations of buffers containing free amines (e.g., Tris 100 mM) do not lead to instability of the NBD substrates, however, the application scope of NBD-based substrates should be limited to lysine-free protease substrates. However, this does not represent a practical disadvantage in the development of protease substrates since there is no known protease that only tolerates lysine and not arginine.

Harnessing NBD-Amide Autocleavage for FRET Substrate Replacement

We developed the idea of harnessing the characteristic NBD autocleavage behavior for a FRET substrate replacement (Figure 5A). By this strategy, endopeptidases with a requirement for a specific amino acid in their prime site (S1') might also become accessible to the repertoire of NBD substrates. Such proteases that have specificity for a S1' amino acid usually must be assayed with a matching FRET substrate because AMC substrates usually do not fulfill the S1' specificity requirement either. To test the hypothesis if a free *N*-terminus of an H_2N-Xaa-NBD derivative is able to drive the Smiles rearrangement, a minimal substrate (H_2N-Arg(Pbf)-NBD **21a**) was analyzed fluorometrically and by LC/MS with enzyme-free assay conditions (100 mM Tris, 100 mM NaCl, variable pH) which highlighted that fluorescence auto-activation proceeds as expected (Figure S53), while at pH value between 5 and 8 the chemical turnover rate was moderately efficient (k = 0.01 min^{-1}). Outside this pH range the Smiles rearrangement might be less efficient but at very low or high pH values most proteases do not show enzymatic activity either.

We designed several NBD substrates from commonly known FRET-based parent substrates where the *N*-terminus of the prime site (S1') amino acid can induce a Smiles rearrangement, resulting in fluorescence increase after proteolytic cleavage and subsequent rearrangement (Figure 5A). Exemplary, such substrates were designed for the proteases SrtA (Figure 5B–C), MMP9, and thermolysin. LC/MS spectra of the enzymatic conversion can be found in the SI (Figure S54). Similar to the above-described non-prime site substrates, the NBD-based substrates for thermolysin, SrtA, and MMP9 showed similar or better affinities (K_M) compared to their parent substrates (Table 3). Both the turnover rates k_{cat} and catalytic efficiency k_{cat}/K_M of these NBD substrates were lower than those of the parent substrate. This might be explained by the two-step reaction mechanism which is limited by the second step of the chemical rearrangement with moderate efficiency. However, the disadvantage of a lower processivity will be outweighed by the higher detection sensitivity and decreased susceptibility to assay interferences (see below).

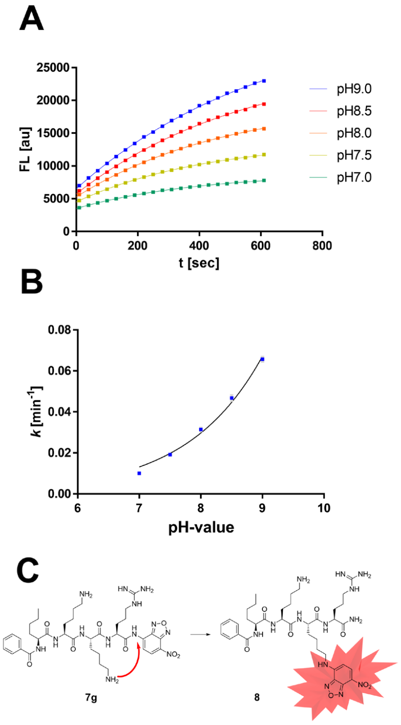


Figure 4. Autocleavage of the Bz-Nle-Lys-Lys-Arg-NBD (**7g**) substrate. (**A**) Enzyme-independent fluorescence increase at different pH values (100 µM Bz-Nle-Lys-Lys-Arg-NBD in DENV NS2B/NS3 buffer). (**B**) pH-dependence of the non-enzymatic turnover. (**C**) Putative Smiles rearrangement of lysine-containing substrates.

Applications for Protease Inhibitor Investigation

To confirm the applicability of the NBD-based substrates in drug discovery-relevant fluorometric assays, the inhibition constants K_i of literature-known inhibitors were determined for each protease using NBD-based and the matching parent substrates (Table S2 and Figures S64–S72). Inhibition constants (K_i) were determined from the fluorescence progress curves using the Cheng-Prusoff equation (Eq. 2).[56]

$$K_i = \frac{IC_{50}}{1+\left(\frac{[S]}{K_M}\right)} \qquad (2)$$

By this, inhibition values were corrected to the zero-substrate concentration, which 1.) allowed the comparison between different substrate affinities and their concentrations; and 2.) showed that the determined K_M values for NBD substrates (Tables 2 and 3) are valid compared to their reported parent substrates. Here, we could show that K_i values of NBD-based and parent substrates differ usually by less than 10% and in some cases by max. a factor of three. It can therefore be assumed that the NBD-based substrates are a well-suitable alternative for fluorometric assays in drug discovery.

Chemistry - A European Journal 10.1002/chem.202301855

WILEY-VCH

RESEARCH ARTICLE

Mitigation of Typical Protease Assay Interferences

The excitation and detection wavelengths of the NBD fluorophore ($\lambda_{\mathrm{ex}} = 485\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 535\ \mathrm{nm}$) are shifted towards the lower energy range compared to the commonly used AMC substrates ($\lambda_{\mathrm{ex}} = 380\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 460\ \mathrm{nm}$). In this regard, these assay conditions might be able to suppress the most typical assay interferences, because interfering physicochemical transitions do occur in the near UV-light spectrum.[24,25,27]

Using rhodesain as an example protease with a drug discovery context,[57] we investigated some compounds from our in-house library that are known to lead to false positive results while using the AMC-based rhodesain substrate. By using the newly developed NBD substrate, we aim to reduce these recurring assay interferences (Figure 6 and Figure S48). In this sense, we could show that the apparent inhibition in the AMC-based assay (attenuation of fluorescence increase) is due to common assay interferences, and the NBD-based substrate is not affected. Detailed molecular explanations of the interference effects can be found in the SI.

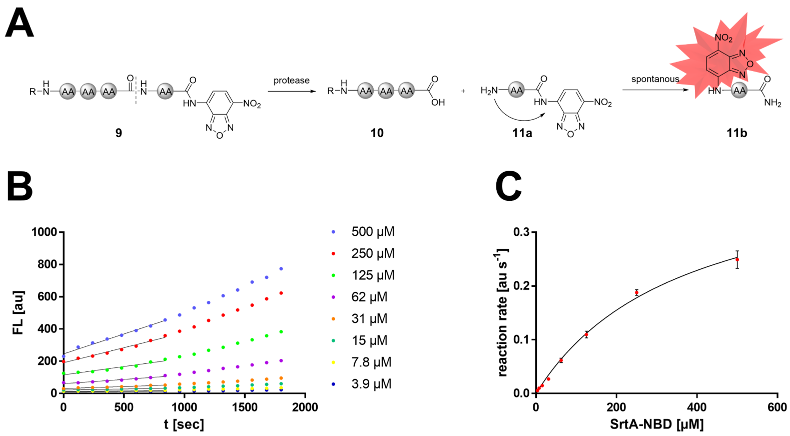


Figure 5. Prime site NBD auto-activating substrates as FRET substrate replacement. (**A**) Schematic depiction of the substrate activation mechanism. Rearrangement of the substrate after cleavage by the protease yields fluorophore activation. (**B**) Fluorescence progress curves for the conversion of the Bz-Leu-Pro-Ala-Thr-Gly-NBD (**7d**) substrate by SrtA. (**C**) Michaelis-Menten plot for the SrtA substrate.

Table 3. Enzyme- and substrate-specific kinetic constants for NBD-based and corresponding parent substrates. Michaelis-Menten constant (K_M), maximum turnover rate (v_{max}), turnover number (k_{cat}), catalytic efficiency (k_{cat}/K_M).

protease	NBD substrate				Parent substrate			
	K_M [µM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·µmol^{-1} min^{-1}]	K_M [µM]	v_{max} [nM·min^{-1}]	k_{cat} [min^{-1}]	k_{cat}/K_M [L·µmol^{-1} min^{-1}]
Thermolysin	56.4 ± 4.22	52.8 ± 2.3	73.1 ± 3.1	1.30 ± 0.11	703 ± 110	(3.0 ± 0.3) 10^4	(4.2 ± 0.4) 10^4	59.1 ± 10.7
SrtA	365 ± 35[a]	156 ± 7	0.12 ± 0.01	(3.28 ± 0.35) 10^{-4}	92.5 ± 9.62	459 ± 22	0.35 ± 0.02	0.0038 ± 0.0004
MMP9	26.7 ± 7.70	10.3 ± 1.03	4.12 ± 0.40	0.15 ± 0.05	24.7 ± 4.8	241 ± 27	96.5 ± 10.7	3.91 ± 0.88

[a] The apparent K_M value of the SrtA FRET substrate is in accordance with the literature value which is distorted by the inner filter effect (Dnp internal quenching). An HPLC-derived K_M value was found to be in the magnitude of K_M > 1 mM.[58]

Chemistry - A European Journal

10.1002/chem.202301855

WILEY-VCH

RESEARCH ARTICLE

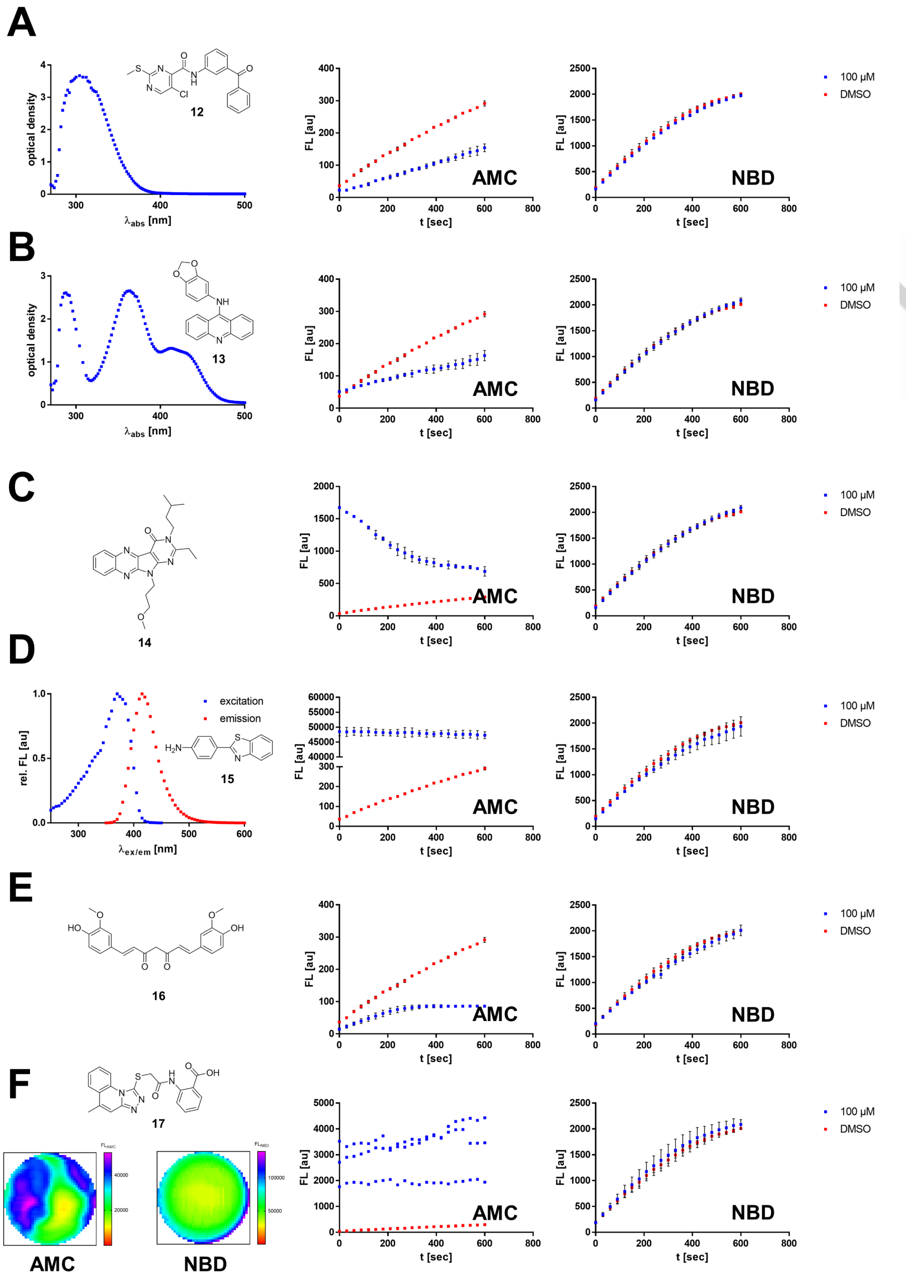


Figure 6. Mitigation of typical AMC assay interferences by an NBD substrate. Each interference example was performed for rhodesain assay conditions with the respective AMC and NBD substrate. (**A**) **Inner filter effect** of compound **12** (100 µM): compound absorbance spectrum; AMC-based assay; NBD-based assay. (**B**) **Fluorophore quenching** by compound **13** (100 µM): compound absorbance spectrum; AMC- and NBD-based assay. (**C**) **Autofluorescence decay** of compound **14** (100 µM) (**D**) **Detector non-linearity** by compound **15** (100 µM): compound fluorescence spectrum; AMC- and NBD-based assay. (**E**) **Photosensitization** of radical oxygen species by compound **16** (100 µM). (**F**) **Reflecting inhibitor aggregates** of compound **17** (100 µM): fluorometric well-homogeneity scan with AMC resp. NBD assay wavelengths; AMC- and NBD-based assay.

Smartphone-Based Fluorometer for Usage in Classroom Applications

The fluorescence of NBD-based dyes is relatively strong in quantum yield and brightness with the maximum sensitivity in the visible region (*vide supra*).[59,60] Thus, NBD substrates are potentially suitable for use with DIY or low-cost Vis-fluorometers to perform kinetic protease studies in classroom-like experiments.[61–63] Previous fluorogenic substrates do not allow for such conceptualizations, as monitoring substrate turnover of AMC and FRET substrates requires an excitation light source in the UV range, which is rarely available in the context of educational activities.

In fact, we were able to show that substrate turnover of Z-Phe-Arg-NBD (500 µM) using the cysteine protease rhodesain (1 µM) is possible by using only a conventional smartphone (Xiaomi Redmi 10) as a detection device (Figure 7). For this purpose, the camera lens and the built-in lamp were covered with LeeFilter films (cost < 10 ct): orange film (2×2 cm, LeeFilters 158 Deep Orange) and blue film (2×2 cm, LeeFilters 120 Deep Blue) as excitation resp. emission filters (Figure 7A). Recording a video of the reaction mixture in a 96-well plate with the built-in camera and analyzing the brightness in the well over time enabled the recording of a substrate turnover curve (Figure 7C and SI Movie 1).

Conclusion

In this study, protease substrates with NBD as a fluorogenic reporter group were designed, synthesized, and tested for utilization in drug discovery applications. Compared to the previous substrate gold standards, the absorption and emission characteristics of NBD are red-shifted to the visible spectrum. In total, we described the synthesis of substrates for 10 different proteases from the cysteine-, serine-, and metalloprotease classes. By this, we were able to obtain substrates for proteases with carboxypeptidase activity (XXX|NBD) as well as for proteases with endopeptidase activity and specificity requirements in the prime site (XX|**X**-NBD). The suitability of the new substrates was confirmed by determining enzyme- and substrate-specific kinetic parameters. The K_M values determined for the NBD-based substrates were in the same order of magnitude as the parent AMC- resp. FRET substrates. Determination of inhibition constants (K_i and k_{inact}) of known inhibitors confirmed that NBD substrates are well-suitable for various applications. In this regard, typical medicinal chemistry-related assay interferences such as the inner filter effect, fluorophore quenching, and autofluorescence decay could be avoided by using the NBD substrates. Thus, in summary, NBD substrates may not only provide a suitable alternative to previously used fluorophores but may form the basis of a new generation of fluorometric assays.

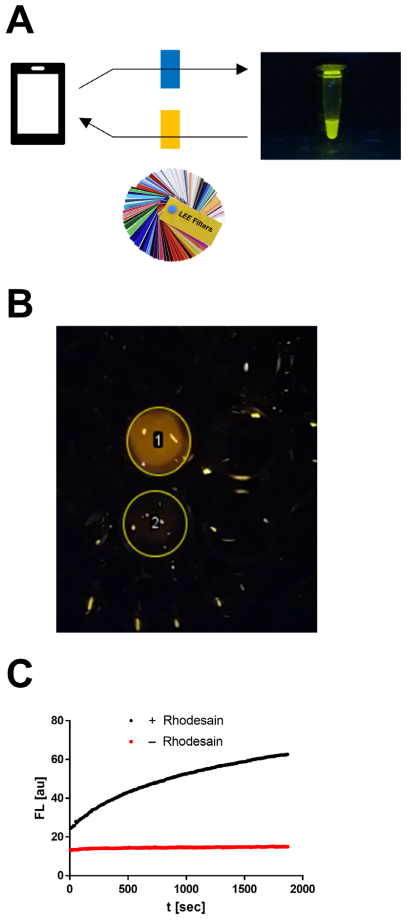


Figure 7. Smartphone-based assay for rhodesain activity. (**A**) Schematic representation of the experimental setup with a fluorescence-enabled photograph of an NBD-NH_2 solution (1 mM) in rhodesain assay buffer. (**B**) Screenshot of a reaction mixture containing rhodesain (1 µM) and the Z-Phe-Arg-NBD substrate (500 µM) in a 96-well plate. Circle 1 highlights a well after the addition of rhodesain enzyme, whereas circle 2 shows only NBD substrate in rhodesain assay buffer. (**C**) Fluorescence progress curves by analysis of the brightness in the wells (ImageJ 1.53).

Experimental Section

Fluorometric assays

Fluorometric assays were performed with a Tecan Spark 10M plate reader using white flat-bottom 96-well microtiter plates from Greiner Bio-One. Measurements were performed in at least three independent technical replicates. Typically, each well contained 180 µL buffer, 5 µL enzyme stock solution, 10 µL inhibitor in DMSO or pure DMSO for mock treatment, and 5 µL solution of the corresponding protease substrate in DMSO, resulting in a total volume of 200 µL (for MMP9: total volume 100 µL, 1 µL substrate, 98 µL buffer, 1 µL enzyme). The fluorescence was measured for 10 min or for 30 min every 30 s with the following excitation and emission wavelengths.

Emission and Excitation wavelengths. AMC: λ_{ex} = 380 nm, λ_{em} = 460 nm; FAGLA: λ_{abs} = 322 nm; Abz/Dap(dnp): λ_{ex} = 320 nm, λ_{em} = 430 nm; NBD: λ_{ex} = 485 nm, λ_{em} = 535 nm; Dnp/Trp: λ_{ex} = 280 nm, λ_{em} = 360 nm.

Enzyme Buffers and Substrates. **Rhodesain** (enzyme: 1 nM, 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, 5 mM DTT, 10 µM Z-Phe-Arg-AMC/10 µM Z-Phe-Arg-NBD); **SARS-CoV 2 M^{pro}** (enzyme: 250 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 60 µM Boc-Abu-Tle-Leu-Gln-AMC/30 µM Boc-Abu-Tle-Leu-Gln-NBD); **DENV2 and ZIKV NS2B/NS3** (enzyme: 250 nM (DENV) & 25 nM (ZIKV), 50 mM Tris pH 7.5, 100 µM Boc-Gly-Arg-Arg-AMC/50 µM Z-Gly-Arg-Arg-NBD); **Urokinase plasminogen activator (uPA)** (enzyme: 10 U, 50 mM Tris pH 7.4, 50 mM NaCl, 0.5 mM EDTA, 240 µM Z-Gly-Gly-Arg-AMC/10 µM Z-Gly-Gly-Arg-NBD); **Sortase A (SrtA)** (enzyme: 1.3 µM, 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM $CaCl_2$, 0.5 mM Gly_4, 25 µM Abz-LPETG-Dap(dnp)–OH/100 µM Bz-Leu-Pro-Ala-Thr-Gly-NBD); **SARS-CoV 2 PL^{pro}** (enzyme: 100 nM, 20 mM Tris pH 7.5, 0.1 mM EDTA, 200 mM NaCl, 1 mM DTT, 50 µM Z-Arg-Leu-Arg-Gly-Gly-AMC/40 µM Z-Arg-Leu-Arg-Gly-Gly-NBD); **Cathepsin S (CatS)** (enzyme: 10 nM, enzyme buffer: 35 mM K_3PO_4 pH 6.5, 35 mM sodium acetate, 2 mM DTT, 2 mM EDTA, assay buffer: 50 mM K_3PO_4 pH 6.5, 2.5 mM DTT, 2.5 mM EDTA, 10 µM Z-Val-Val-Arg-AMC/10 µM Z-Val-Val-Arg-NBD); **Matrixmetalloprotease 9 (MMP9)** (enzyme: 2.5 nM, 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 20 µM $ZnCl_2$, 1 mM $CaCl_2$, Dnp-Pro-Leu-Gly-Met-Trp-Ser-Arg-NH_2/40 µM Z-Pro-Leu-Gly-Met-NBD); **Thermolysin** (enzyme: 0.7 nM, 100 mM Tris, 100 mM NaBr, 2.5 mM $CaCl_2$, 100 µM FAGLA/20 µM Bz-Gly-Phe-NBD).

Chemistry

All reagents and solvents were of analytical grade quality and purchased from Sigma-Aldrich, Carbolution, BLDpharm, or FisherScientific. Chemicals were used without further purification. ^1H and ^{13}C spectra were recorded on a Bruker Fourier 300 or Bruker Avance III 600 using DMSO-d_6 or $CDCl_3$, as solvent. Chemical shifts δ are given in parts per million (ppm) using residual proton peaks of the solvent as an internal standard. UV-chromatograms and mass spectra were obtained by LC-MS consisting of an Agilent 1100 series HPLC system using an Agilent Poroshell 120 EC-C18 150 × 2.10 mm, 4 µm column or an Agilent Zorbax SB Aq. 150 × 4.6 mm, 5 µm column. Detection wavelengths were 210, 254, and 396 nm. The tested substrates and inhibitors displayed a purity $\geq$ 95% in all cases. The molecular mass was confirmed by an Agilent 1100 series LC/MSD Trap with electron spray ionization (ESI) in positive ionization mode. Melting points (uncorrected) were determined in open capillaries using a Schorpp Gerätetechnik MPM-H3 melting point device. Specific rotations $[\alpha]_D^{20}$ were measured on a P3000 polarimeter from Krüss and are reported in $cm^3\ g^{-1}\ dm^{-1}$. Detailed information on the synthesis and analytical data for characterization of NBD-substrates can be found in the Supporting Information.

Supporting Information

The authors have cited additional references within the Supporting Information.[27,41,47,52–54,56,64–89] Additional experimental Procedures (Protein Constructs and Expression, Fluorometric Assay, Molecular Docking, Chemistry) and Results and Discussion (Autocleavage of NS2B/NS3 Protease Substrate, Stability of Substrates, Fluorometric Assay Data).

Keywords: Assay Interferences • Fluorescent Probes • Medicinal Chemistry • Nitrobenzofurazane • Proteases

Acknowledgment

T.S. acknowledges financial support from the German Research Foundation (DFG) Collaborative Research Center (CRC) SFB 1066 (projects Q5). F.B. acknowledges financial support from the Johannes Gutenberg University of Mainz.

References

[1] S. Udenfriend, in *Fluoresc. Assay Biol. Med.*, Elsevier, **1969**, pp. 1–41.
[2] F. Ciruela, *Curr. Opin. Biotechnol.* **2008**, *19*, 338–343.
[3] T. D. Pollard, *Mol. Biol. Cell* **2010**, *21*, 4061–4067.
[4] M. Leopoldo, E. Lacivita, F. Berardi, R. Perrone, *Drug Discov. Today* **2009**, *14*, 706–712.
[5] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2023**, *32*.
[6] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *J. Phys. Chem. B* **2021**, *125*, 6837–6846.
[7] A. Eatemadi, H. T. Aiyelabegan, B. Negahdari, M. A. Mazlomi, H. Daraee, N. Daraee, R. Eatemadi, E. Sadroddiny, *Biomed. Pharmacother.* **2017**, *86*, 221–231.
[8] P. Müller, H. Maus, S. J. Hammerschmidt, P. M. Knaff, V. Mailänder, T. Schirmeister, C. Kersten, *Curr. Med. Chem.* **2022**, *29*, 635–665.
[9] C. López-Otín, J. S. Bond, *J. Biol. Chem.* **2008**, *283*, 30433–30437.
[10] G. Abbenante, D. Fairlie, *Med. Chem. (Los. Angeles).* **2005**, *1*, 71–104.
[11] P. D. Leeson, A. P. Bento, A. Gaulton, A. Hersey, E. J. Manners, C. J. Radoux, A. R. Leach, *J. Med. Chem.* **2021**, *64*, 7210–7230.
[12] Y. N. Lamb, *Drugs* **2022**, *82*, 585–591.
[13] M. Drag, G. S. Salvesen, *Nat. Rev. Drug Discov.* **2010**, *9*, 690–701.
[14] B. Turk, *Nat. Rev. Drug Discov.* **2006**, *5*, 785–799.
[15] V. Canbay, U. auf dem Keller, *Curr. Opin. Chem. Biol.* **2021**, *60*, 89–96.
[16] I. L. H. Ong, K.-L. Yang, *Analyst* **2017**, *142*, 1867–1881.
[17] A. Carmel, M. Zur, A. Yaron, E. Katchalski, *FEBS Lett.* **1973**, *30*, 11–14.
[18] H. Tuppy, U. Wiesbauer, E. Wintersberger, *Hoppe-Seyler´s Zeitschrift für Physiol. Chemie* **1962**, *329*, 278–288.
[19] M. Zimmerman, E. Yurewicz, G. Patel, *Anal. Biochem.* **1976**, *70*, 258–262.

[20] S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. Chung, N. P. Coussens, J. L. Dahlin, *Assay Guidance Manual*, Eli Lilly & Company And The National Center For Advancing Translational Sciences, **2004**.

[21] B. C. Chenna, L. Li, D. M. Mellott, X. Zhai, J. L. Siqueira-Neto, C. Calvet Alvarez, J. A. Bernatchez, E. Desormeaux, E. Alvarez Hernandez, J. Gomez, J. H. McKerrow, J. Cruz-Reyes, T. D. Meek, *J. Med. Chem.* **2020**, *63*, 3298–3316.

[22] T. T. Baird, C. S. Craik, *Handb. Proteolytic Enzym.* **2013**, *3*, 2594–2600.

[23] A. Jadhav, R. S. Ferreira, C. Klumpp, B. T. Mott, C. P. Austin, J. Inglese, C. J. Thomas, D. J. Maloney, B. K. Shoichet, A. Simeonov, *J. Med. Chem.* **2010**, *53*, 37–51.

[24] J. Comley, *Drug Discov. World Summer 2003* **2003**, 91–98.

[25] S. K. Grant, J. G. Sklar, R. T. Cummings, *SLAS Discov.* **2002**, *7*, 531–540.

[26] Y. Liu, W. Kati, C.-M. Chen, R. Tripathi, A. Molla, W. Kohlbrenner, *Anal. Biochem.* **1999**, *267*, 331–335.

[27] N. Thorne, D. S. Auld, J. Inglese, *Curr. Opin. Chem. Biol.* **2010**, *14*, 315–324.

[28] F. E. Jernigan, D. S. Lawrence, *Chem. Commun.* **2013**, *49*, 6728.

[29] J.-M. Delaissé, Y. Eeckhout, G. Vaes, *Biochem. Biophys. Res. Commun.* **1984**, *125*, 441–447.

[30] S. G. Deeks, M. Smith, M. Holodniy, J. O. Kahn, *J. Am. Med. Assoc.* **1997**, *277*, 145–153.

[31] S. Jo, H. Kim, S. Kim, D. H. Shin, M. Kim, *Chem. Biol. Drug Des.* **2019**, *94*, 2023–2030.

[32] C. Jiang, H. Huang, X. Kang, L. Yang, Z. Xi, H. Sun, M. D. Pluth, L. Yi, *Chem. Soc. Rev.* **2021**, *50*, 7436–7495.

[33] K. Okada, T. Yamaguchi, K. Dodo, M. Sodeoka, S. Obika, *Bioorg. Med. Chem.* **2019**, *27*, 1444–1448.

[34] Y. Xie, L. Chen, R. Wang, J. Wang, J. Li, W. Xu, Y. Li, S. Q. Yao, L. Zhang, Q. Hao, H. Sun, *J. Am. Chem. Soc.* **2019**, *141*, 18428–18436.

[35] M. He, Z. Han, J. Qiao, L. Ngo, M. P. Xiong, Y. G. Zheng, *Chem. Commun.* **2018**, *54*, 5594–5597.

[36] I. N. Gober, M. L. Waters, *J. Am. Chem. Soc.* **2016**, *138*, 9452–9459.

[37] Y. Xie, J. Ge, H. Lei, B. Peng, H. Zhang, D. Wang, S. Pan, G. Chen, L. Chen, Y. Wang, Q. Hao, S. Q. Yao, H. Sun, *J. Am. Chem. Soc.* **2016**, *138*, 15596–15604.

[38] J. M. An, S. Kang, E. Huh, Y. Kim, D. Lee, H. Jo, J. F. Joung, V. J. Kim, J. Y. Lee, Y. S. Dho, Y. Jung, J. K. Hur, C. Park, J. Jung, Y. Huh, J.-L. Ku, S. Kim, T. Chowdhury, S. Park, J. S. Kang, M. S. Oh, C.-K. Park, D. Kim, *Chem. Sci.* **2020**, *11*, 5658–5668.

[39] D. Castagnolo, M. Pagano, M. Bernardini, M. Botta, *Tetrahedron Lett.* **2012**, *53*, 5008–5011.

[40] P. M. Knaff, P. Müller, C. Kersten, L. Wettstein, J. Münch, K. Landfester, V. Mailänder, *Biomacromolecules* **2022**, *23*, 2236–2242.

[41] Y. Yang, D. Zhang, M. Xu, J. Wang, J. Chen, L. Wang, *Monatshefte für Chemie - Chem. Mon.* **2018**, *149*, 1003–1008.

[42] L. F. Bernal-Perez, L. Prokai, Y. Ryu, *Anal. Biochem.* **2012**, *428*, 13–15.

[43] P. B. Ghosh, M. W. Whitehouse, *Biochem. J.* **1968**, *108*, 155–156.

[44] S. Jung, N. Fuchs, C. Grathwol, U. A. Hellmich, A. Wagner, E. Diehl, T. Willmes, C. Sotriffer, T. Schirmeister, *Eur. J. Med. Chem.* **2022**, *238*, 114460.

[45] D. Dana, A. R. Davalos, S. De, P. Rathod, R. K. Gamage, J. Huestis, N. Afzal, Y. Zavlanov, S. S. Paroly, S. A. Rotenberg, G. Subramaniam, K. J. Mark, E. J. Chang, S. Kumar, *Bioorg. Med. Chem.* **2013**, *21*, 2975–2987.

[46] W. Rut, K. Groborz, L. Zhang, X. Sun, M. Zmudzinski, B. Pawlik, X. Wang, D. Jochmans, J. Neyts, W. Młynarski, R. Hilgenfeld, M. Drag, *Nat. Chem. Biol.* **2021**, *17*, 222–228.

[47] F. Barthels, G. Marincola, T. Marciniak, M. Konhäuser, S. Hammerschmidt, J. Bierlmeier, U. Distler, P. R. Wich, S. Tenzer, D. Schwarzer, W. Ziebuhr, T. Schirmeister, *ChemMedChem* **2020**, *15*, 839–850.

[48] A. Narayanan, M. Narwal, S. A. Majowicz, C. Varricchio, S. A. Toner, C. Ballatore, A. Brancale, K. S. Murakami, J. Jose, *Commun. Biol.* **2022**, *5*, 169.

[49] A. Dal Corso, M. Catalano, A. Schmid, J. Scheuermann, D. Neri, *Angew. Chemie* **2018**, *130*, 17424–17428.

[50] H. Lai, G. Sridhar Prasad, R. Padmanabhan, *Antiviral Res.* **2013**, *97*, 74–80.

[51] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.

[52] S. Netzel-Arnett, S. K. Mallya, H. Nagase, H. Birkedal-Hansen, H. E. Van Wart, *Anal. Biochem.* **1991**, *195*, 86–92.

[53] P. A. Bartlett, C. K. Marlowe, *Biochemistry* **1987**, *26*, 8553–8561.

[54] S. M. Khan, D. W. Darnall, *Anal. Biochem.* **1978**, *86*, 332–336.

[55] D. J. Murphy, *Anal. Biochem.* **2004**, *327*, 61–67.

[56] T. Schirmeister, J. Kesselring, S. Jung, T. H. Schneider, A. Weickert, J. Becker, W. Lee, D. Bamberger, P. R. Wich, U. Distler, S. Tenzer, P. Johé, U. A. Hellmich, B. Engels, *J. Am. Chem. Soc.* **2016**, *138*, 8332–8335.

[57] R. Ettari, S. Previti, L. Tamborini, G. Cullia, S. Grasso, M. Zappalà, *Mini-Reviews Med. Chem.* **2016**, *16*, 1374–1391.

[58] R. G. Kruger, P. Dostal, D. G. McCafferty, *Anal. Biochem.* **2004**, *326*, 42–48.

[59] S. Uchiyama, T. Santa, T. Fukushima, H. Homma, K. Imai, *J. Chem. Soc. Perkin Trans. 2* **1998**, 2165–2174.

[60] B. E. Tebikachew, F. Edhborg, N. Kann, B. Albinsson, K. Moth-Poulsen, *Phys. Chem. Chem. Phys.* **2018**, *20*, 23195–23201.

[61] M. A. Hossain, J. Canning, S. Ast, T. L. Yen, P. Rutledge, A. Jamalipour, in *Adv. Photonics*, OSA, Washington, D.C., **2014**, p. SeTh2C.1.

[62] L. M. Azad, H. Ehtesabi, A. Rezaei, *Nano-Structures & Nano-Objects* **2021**, *26*, 100722.

[63] A. Jamalipour, M. A. Hossain, **2019**, pp. 15–36.

[64] G. Amendola, R. Ettari, S. Previti, C. Di Chio, A. Messere, S. Di Maro, S. J. Hammerschmidt, C. Zimmer, R. A. Zimmermann, T. Schirmeister, M. Zappalà, S. Cosconati, *J. Chem. Inf. Model.* **2021**, *61*, 2062–2073.

[65] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *J. Med. Chem.* **2019**, *62*, 11359–11382.

[66] L. Schmohl, J. Bierlmeier, N. von Kügelgen, L. Kurz, P. Reis, F. Barthels, P. Mach, M. Schutkowski, C. Freund, D. Schwarzer, *Bioorg. Med. Chem.* **2017**, *25*, 5002–5007.

[67] Y. M. Báez-Santos, S. J. Barraza, M. W. Wilson, M. P. Agius, A. M. Mielech, N. M. Davis, S. C. Baker, S. D. Larsen, A. D. Mesecar, *J.*

Med. Chem. **2014**, *57*, 2393–2412.

[68] J. Joossens, P. Van der Veken, G. Surpateanu, A.-M. Lambeir, I. El-Sayed, O. M. Ali, K. Augustyns, A. Haemers, *J. Med. Chem.* **2006**, *49*, 5785–5793.

[69] B. J. Buckley, A. Aboelela, E. Minaei, L. X. Jiang, Z. Xu, U. Ali, K. Fildes, C.-Y. Cheung, S. M. Cook, D. C. Johnson, D. A. Bachovchin, G. M. Cook, M. Apte, M. Huang, M. Ranson, M. J. Kelso, *J. Med. Chem.* **2018**, *61*, 8299–8320.

[70] A. Welker, C. Kersten, C. Müller, R. Madhugiri, C. Zimmer, P. Müller, R. Zimmermann, S. Hammerschmidt, H. Maus, J. Ziebuhr, C. Sotriffer, T. Schirmeister, *ChemMedChem* **2021**, *16*, 340–354.

[71] D. Brömme, *Curr. Protoc. Protein Sci.* **2000**, *21*.

[72] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.

[73] G. N. Lewis, M. Calvin, *Chem. Rev.* **1939**, *25*, 273–328.

[74] L. D. Dias, K. C. Blanco, I. S. Mfouo-Tynga, N. M. Inada, V. S. Bagnato, *J. Photochem. Photobiol. C Photochem. Rev.* **2020**, *45*, 100384.

[75] I. Ahmad, S. Ahmed, Z. Anwar, M. A. Sheraz, M. Sikorski, *Int. J. Photoenergy* **2016**, *2016*, 1–19.

[76] K. T. Kazantzis, K. Koutsonikoli, B. Mavroidi, M. Zachariadis, P. Alexiou, M. Pelecanou, K. Politopoulos, E. Alexandratou, M. Sagnou, *Photochem. Photobiol. Sci.* **2020**, *19*, 193–206.

[77] K. M. Nelson, J. L. Dahlin, J. Bisson, J. Graham, G. F. Pauli, M. A. Walters, *J. Med. Chem.* **2017**, *60*, 1620–1637.

[78] D. Steverding, *Molecules* **2019**, *25*, 143.

[79] S. Ludewig, M. Kossner, M. Schiller, K. Baumann, T. Schirmeister, *Curr. Top. Med. Chem.* **2010**, *10*, 368–382.

[80] Lord Rayleigh, *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **1881**, *12*, 81–101.

[81] S. Jung, N. Fuchs, P. Johe, A. Wagner, E. Diehl, T. Yuliani, C. Zimmer, F. Barthels, R. A. Zimmermann, P. Klein, W. Waigel, J. Meyr, T. Opatz, S. Tenzer, U. Distler, H.-J. Räder, C. Kersten, B. Engels, U. A. Hellmich, J. Klein, T. Schirmeister, *J. Med. Chem.* **2021**, *64*, 12322–12358.

[82] R. D. A. Wilkinson, A. Young, R. E. Burden, R. Williams, C. J. Scott, *Mol. Cancer* **2016**, *15*, 29.

[83] L. Fu, S. Shao, Y. Feng, F. Ye, X. Sun, Q. Wang, F. Yu, Q. Wang, B. Huang, P. Niu, X. Li, C. C. L. Wong, J. Qi, W. Tan, G. F. Gao, *MBio* **2021**, *12*.

[84] S. Previti, R. Ettari, E. Calcaterra, C. Di Chio, R. Ravichandran, C. Zimmer, S. Hammerschmidt, A. Wagner, M. Bogacz, S. Cosconati, T. Schirmeister, M. Zappalà, *ACS Med. Chem. Lett.* **2022**, *13*, 1083–1090.

[85] M. A. Abduraman, M. Hariono, R. Yusof, N. A. Rahman, H. A. Wahab, M. L. Tan, *Heliyon* **2018**, *4*, e01023.

[86] K. H. Nam, *J. Inorg. Biochem.* **2021**, *215*, 111319.

[87] M. Johansson, A. J. Brooks, D. A. Jans, S. G. Vasudevan, *J. Gen. Virol.* **2001**, *82*, 735–745.

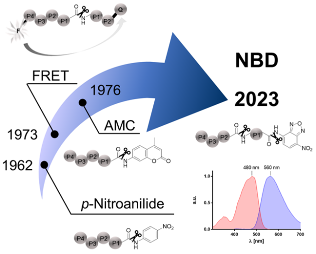
[88] F. Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister, U. A. Hellmich, *FEBS Lett.* **2019**, *593*, 2204–2213.

[89] C. R. Caffrey, E. Hansell, K. D. Lucas, L. S. Brinen, A. Alvarez Hernandez, J. Cheng, S. L. Gwaltney, W. R. Roush, Y.-D. Stierhof, M. Bogyo, D. Steverding, J. H. McKerrow, *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73.

Chemistry - A European Journal 10.1002/chem.202301855

RESEARCH ARTICLE

Entry for the Table of Contents



New Protease Substrates Light Up Medicinal Chemistry: Say Hello to NBD! Maus and Müller *et al.* have developed a game-changing generation of new protease assay substrates using the nitrobenzofurazane (NBD) fluorophore. With applications across different protease classes, this breakthrough promises to decisively improve fluorometric assays in protease drug discovery.

Results and Discussion

Discussion of AMC Assay Interference Effects and Mitigation Strategies by Using NBD Substrates

Inner Filter Effect. This effect describes the absorption of excitation UV light by the compound under investigation.[19] The chosen model compound **12** *N*-(3-benzoylphenyl)-5-chloro-2-(methylthio)pyrimidine-4-carboxamide (100 µM) has its absorption maximum at approx. 300 nm, while AMC is similarly excited at 280 nm (Figure 6A). The spectral overlap between the excitation light and the absorption of the compound reduced the detected fluorescence, and hence, led to an apparent but false-positive inhibition result. In contrast, the NBD fluorophore is excited at 485 nm, and thus, it was not affected by compound absorption. In this sense, the NBD fluorescence progress curves indicated that the apparent inhibition in the AMC-based assay is exclusively due to the absorption of the excitation light. Compounds that absorb excitation light of AMC (λ_{ex} = 380 nm) usually show a yellowish coloration (complementary color to UV/blue light). Compounds that absorb in the NBD excitation range (485 nm), on the other hand, are reddish to black. While yellowish compounds are frequently observed among organic drug candidate compounds, reddish compounds are rather rare.[20]

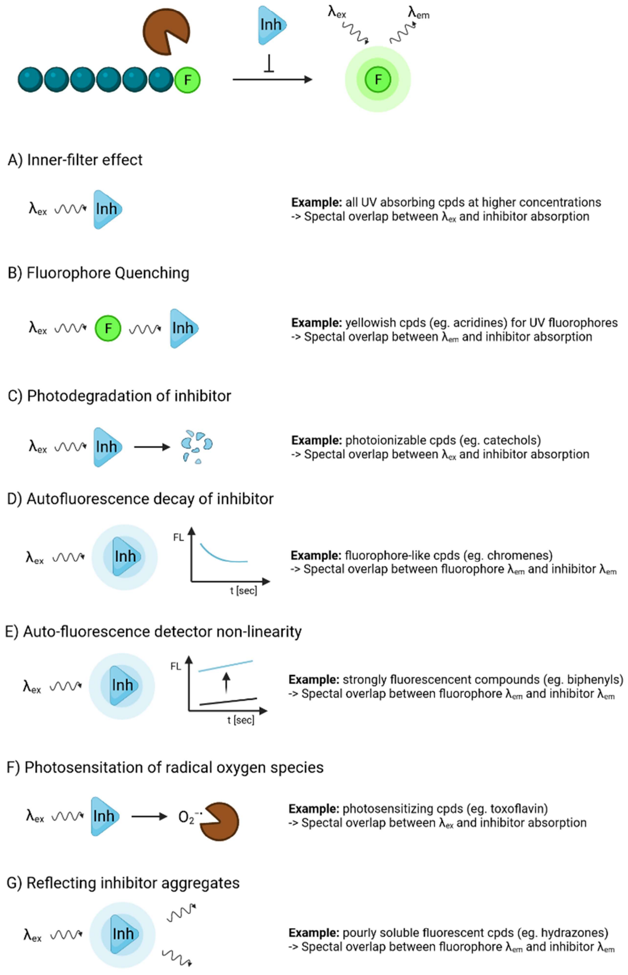
Fluorophore Quenching. Similar to the inner filter effect, fluorophore quenching is based on the absorption of light by the compound under investigation. However, in this case, the fluorescence emitted by the fluorophore is absorbed by an artificial quencher molecule.[19] Detection of the fluorescence emitted by AMC occurs usually at 460 nm. As an example compound to interfere with the assay systems, we used the strongly yellowish-colored compound **13** *N*-(benzo[*d*][1,3]dioxol-5-yl)acridine-9-amine (100 µM), which absorbs in a wavelength range of 350–460 nm (Figure 6B), and thus the AMC fluorescence was non-specifically attenuated by the absorption of this compound. The excitation and detection wavelengths of the NBD fluorescence ($\lambda_{\mathrm{ex}} = 485\ \mathrm{nm}$, $\lambda_{\mathrm{em}} = 535\ \mathrm{nm}$) are beyond this quenching absorbance, and hence, NBD-based assays were not affected by this model quencher compound.

Autofluorescence Decay. If the compound under investigation can be excited at the excitation wavelength of the fluorophore and at the same time emits in the range of the assay detection wavelength, this leads to an overall increased fluorescence level (autofluorescence). However, autofluorescence may decrease due to degradation or photobleaching of the compound under investigation. The overall fluorescence thus decreases over time, generating a false-positive inhibition appearance. Since not many organic small molecules show fluorescence in the wavelength range of NBD, these substrates are less prone to this effect (Figure 6C). We were able to demonstrate this behavior by using the model compound **14** 2-ethyl-3-isopentyl-11-(3-methoxypropyl)-3,11-dihydro-4*H*-pyrimido[5',4':4,5]pyrrolo[2,3-*b*]quinoxalin-4-one (100 µM).

Detector Nonlinearity. If compound-induced autofluorescence levels are very high, relatively large concentrations of the compound under investigation will yield apparent attenuation of fluorescence increase due to the nonlinearity of the fluorescence detector setup. This is problematic for compounds that have similar fluorescence properties to the fluorophore utilized. An example of a compound with fluorescent properties similar to AMC is compound **15** 4-(benzo[*d*]thiazol-2-yl)aniline (100 µM). The use of NBD, with different fluorescence properties prevents this model interference (Figure 6D).

Photosenzitation of Radical Oxygen Species. Irradiation can cause certain compounds to produce reactive oxygen species by radical or singlet oxygen pathways.[21–23] These may cause damage to the investigated enzyme. Consequently, the activity of the protease can be non-specifically attenuated. The formation of reactive oxygen species is caused predominantly by short-wavelength UV radiation required to excite the AMC fluorophore (Figure 6E). We demonstrated the formation of reactive oxygen species by irradiation at 360 nm for the known UV-sensitizer and pan-assay interference compound **16** curcumin (100 µM) which shows, in accordance with the literature, irreversible inhibition of rhodesain with an AMC substrate but no inhibition with the NBD substrate.[21,24,25]

Light Scattering at Inhibitor Aggregates. Compound aggregates (low nanometer diameter) caused by the compound under investigation can scatter the excitation light leading to fluctuating and non-predictable fluorescence traces in enzyme assays (Figure 6F).[26] Using the known aggregator model compound **17** 2-(2-((5-methyl-[1,2,4]triazolo[4,3-*a*]quinolin-1-yl)thio)acetamido)benzoic acid (100 µM), we showed that the NBD excitation light (λ_{ex} = 485 nm) was significantly less scattered than the AMC excitation (λ_{ex} = 380 nm), and hence, led to the generation of significantly more steady fluorescence progress curves. This might be reasoned by the effective cross-section of elastic light scattering by matter which depends on the fourth power on the wavelength, and thus, AMC excitation light is potentially scattered 3.5 times stronger than NBD excitation light by aggregates in the fluorometric assay.[27]



A) Inner-filter effect

Example: all UV absorbing cpds at higher concentrations
-> Spectal overlap between λ_{ex} and inhibitor absorption

B) Fluorophore Quenching

Example: yellowish cpds (eg. acridines) for UV fluorophores
-> Spectal overlap between λ_{em} and inhibitor absorption

C) Photodegradation of inhibitor

Example: photoionizable cpds (eg. catechols)
-> Spectal overlap between λ_{ex} and inhibitor absorption

D) Autofluorescence decay of inhibitor

Example: fluorophore-like cpds (eg. chromenes)
-> Spectal overlap between fluorophore λ_{em} and inhibitor λ_{em}

E) Auto-fluorescence detector non-linearity

Example: strongly fluorescent compounds (eg. biphenyls)
-> Spectal overlap between fluorophore λ_{em} and inhibitor λ_{em}

F) Photosensitation of radical oxygen species

Example: photosensitizing cpds (eg. toxoflavin)
-> Spectal overlap between λ_{ex} and inhibitor absorption

G) Reflecting inhibitor aggregates

Example: pourly soluble fluorescent cpds (eg. hydrazones)
-> Spectal overlap between fluorophore λ_{em} and inhibitor λ_{em}

Figure S48. Summary of common fluorometric assay interference effects caused by inhibitor compounds (cpds) occurring in protease assays and their molecular explanation. Interference examples of the model protease rhodesain can be found in Figure 6 in the manuscript.

Autocleavage of a DENV/ZIKV NS2B/NS3 Protease Substrate

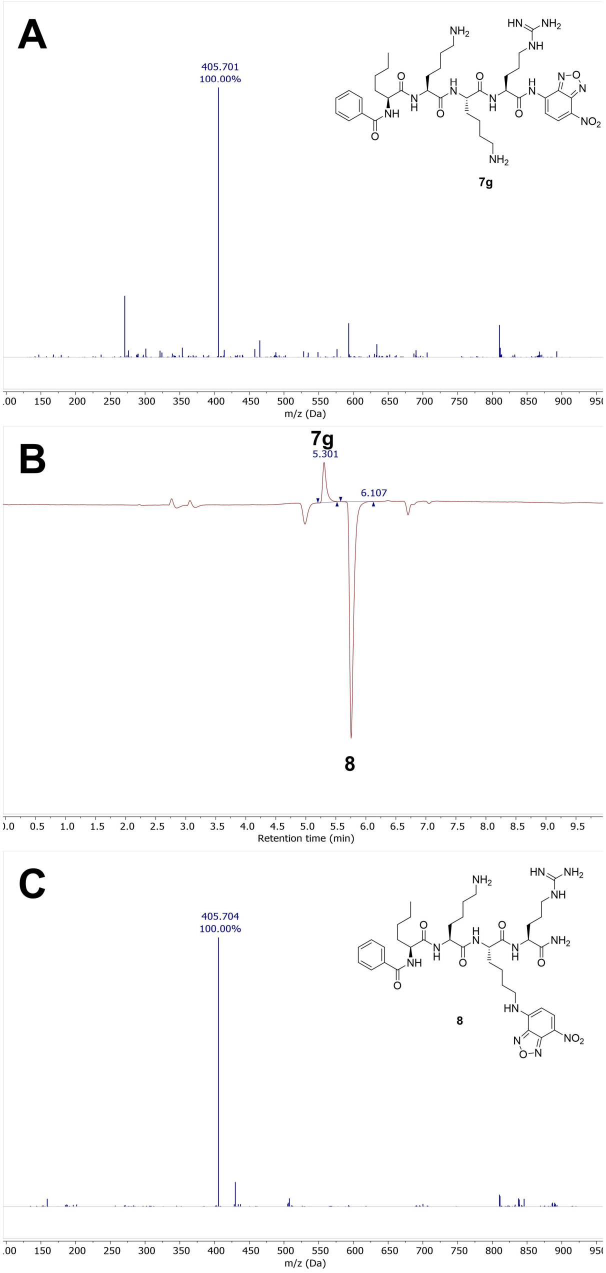


Figure S49. LC/MS analysis of NS2B/NS3 substrate (Bz-Nle-Lys-Lys-Arg-NBD) auto-activation. (**A**) Mass spectrum of the NS2B/NS3 substrate educt **7g**. (**B**) HPLC chromatogram at 396 nm highlighting two peaks with an identical mass of (405.70 Da). (**C**) Mass spectrum of the fluorescent product **8** after Smiles rearrangement.

Stability of Substrates and Autocleavage Investigations

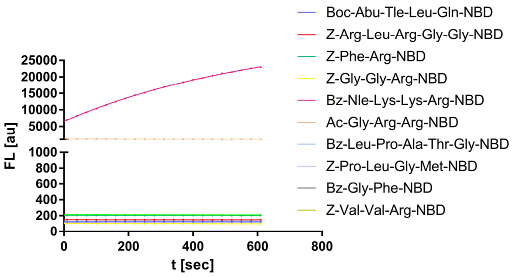


Figure S50. Substrate stability enzyme assay buffer. Fluorescence traces were recorded for 10 min at a substrate concentration of 100 µM in the respective assay buffer (without the addition of enzymes). No fluorescence increase was observed for any substrate except for Bz-Nle-Lys-Lys-Arg-NBD. The substrates with a basic residue in the P2 position (Ac-Gly-Arg-Arg-AMC and Bz-Nle-Lys-Lys-Arg-AMC) show a higher intrinsic fluorescence than all other substrates. We hypothesize that a positive charge in the P2 position might form a polar interaction with the nitro group of the NBD fluorophore. This enhances the electron-withdrawing effect of the nitro group on the NBD heteroaromatic ring, and thus, such an interaction could contribute to the push-pull substituent system on the NBD fluorophore and be responsible for partially reversing the self-quenching effect of the NBD amide

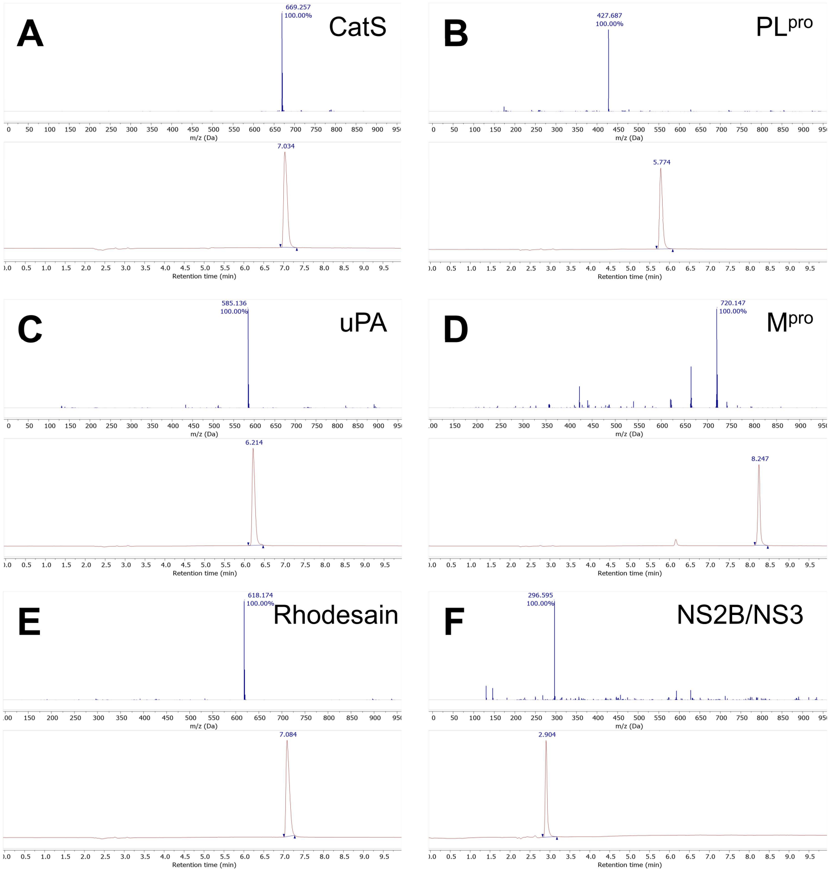


Figure S51. Substrate stability of NBD-based substrates analyzed by LC/MS. Substrates (500 µM) were diluted in their respective assay buffer and incubated for 60 min at room temperature. The reaction mixtures were diluted with acetonitrile (1:1) and subjected to LC/MS analysis. (**A**) Cathepsin S (CatS). (**B**) SARS-CoV 2 PL^{pro}. (**C**) Urokinase plasminogen activator (uPA). (**D**) SARS-CoV M^{pro}. (**E**) Rhodesain. (**F**) ZIKV/DENV NS2B/NS3 (Ac-Gly-Arg-Arg-NBD).

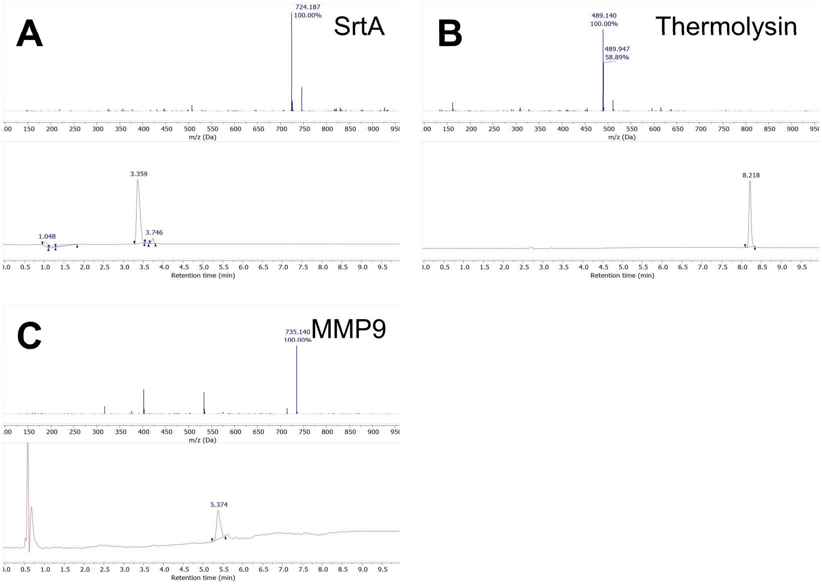


Figure S52. Substrate stability of auto-activation NBD-substrates analyzed by LC/MS. Substrates (500 µM) were diluted in their respective assay buffer and incubated for 60 min at room temperature. The reaction mixtures were diluted with acetonitrile (1:1) and subjected to LC/MS analysis. (**A**) Sortase (SrtA). (**B**) Thermolysin. (**C**) Matrix metalloprotease (MMP9).

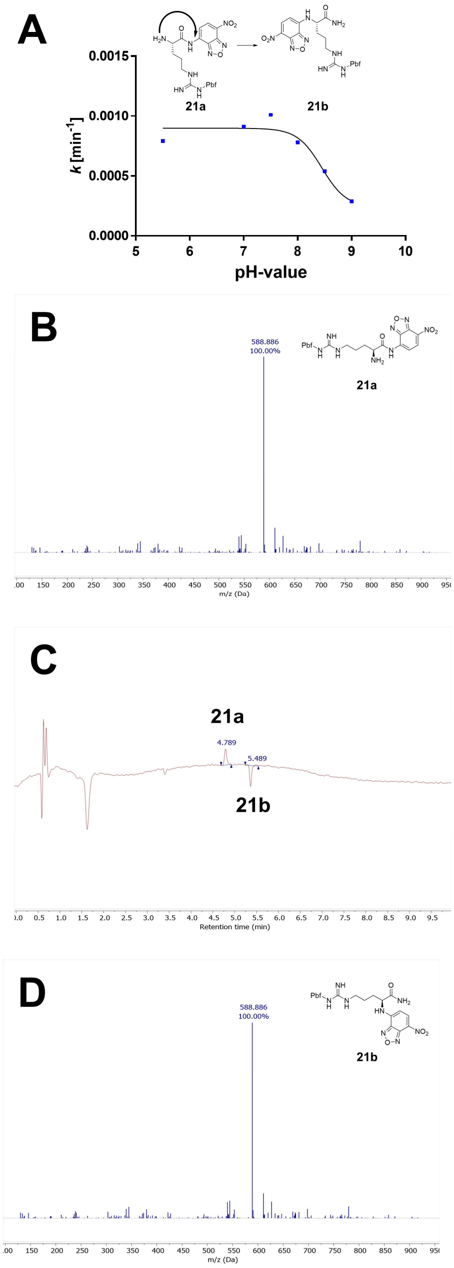


Figure S53. Investigation of a minimal substratefro fluorescence auto-activation. (**A**) pH dependence of the fluorescence autoactivation. For this, H-Arg(Pbf)-NBD (100 µM) was diluted in rhodesain assay buffer (pH 5.5) or NS2B/NS3 assay buffer (pH 7.0–9.0). Turnover rates were determined from the initial slopes of fluorescence progress curves. (**B**) Mass spectrum of the minimal substrate educt **21a**. (**C**) The reaction mixture at pH 7.5 was diluted with acetonitrile (1:1) and subjected to LC/MS analysis. HPLC chromatogram at 396 nm highlighting two peaks with an identical mass of (588.88 Da). (**D**) Mass spectrum of the fluorescent product **21b** after Smiles rearrangement.

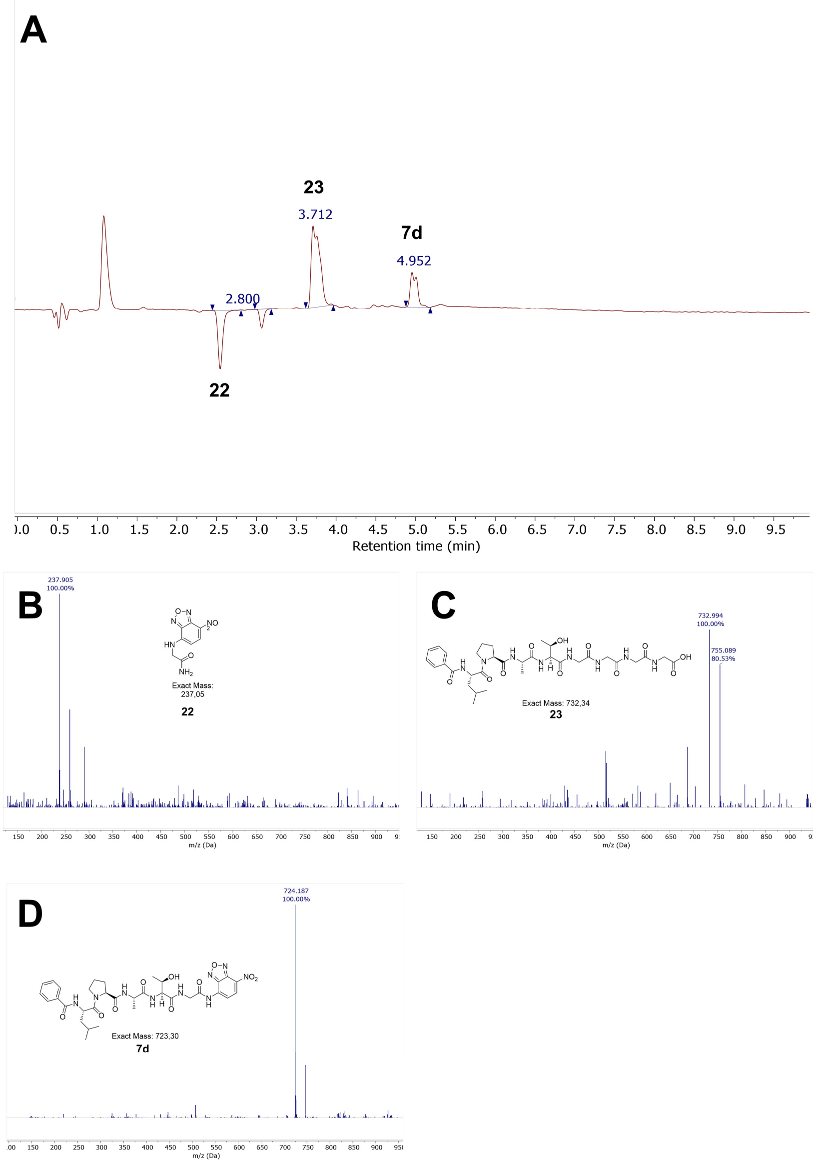


Figure S54. Investigation of the NBD-based SrtA auto-activation substrate by LC/MS. (**A**) HPLC chromatogram of the enzyme reaction mixture at 210 nm. Bz-Leu-Pro-Ala-Thr-Gly-NBD (500 µM) was diluted in SrtA assay buffer and incubated with SrtA (20 µg/mL) for 1 h at room temperature. The reaction mixture was diluted with acetonitrile (1:1) and subjected to LC/MS analysis. (**B**) Mass spectrum of the fluorescent product **22** after substrate cleavage and Smiles rearrangement. (**C**) Mass spectrum of the enzymatic co-product after transpeptidation with tetraglycine. (**D**) Mass spectrum of the Bz-Leu-Pro-Ala-Thr-Gly-NBD educt **7d**.

Fluorometric Assay Data

Michaelis-Menten plots

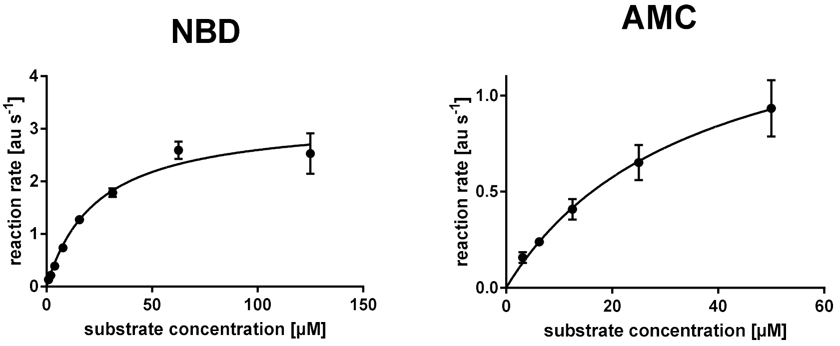


Figure S55. K_M of Cathepsin S (CatS) substrates.

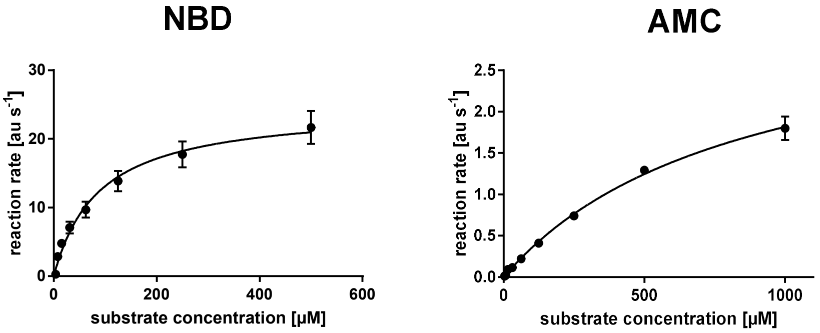


Figure S56. K_M of DENV NS2B/NS3 substrates.

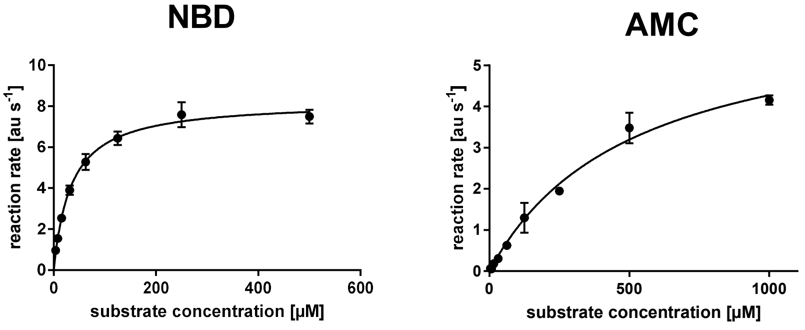


Figure S57. K_M of ZIKV NS2B/NS3 substrates.

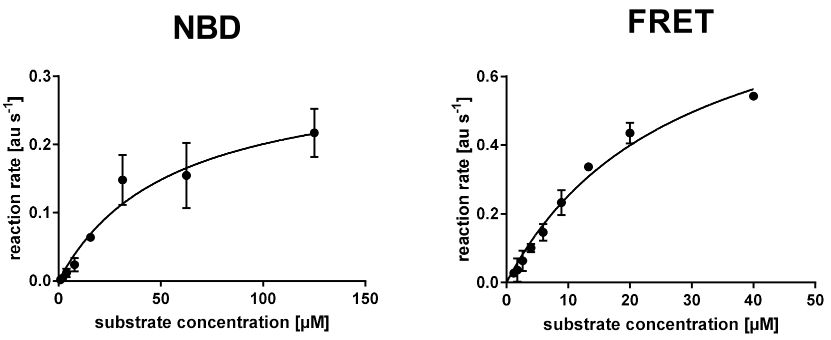


Figure S58. K_M of matrix metalloprotease 9 (MMP9) substrates.

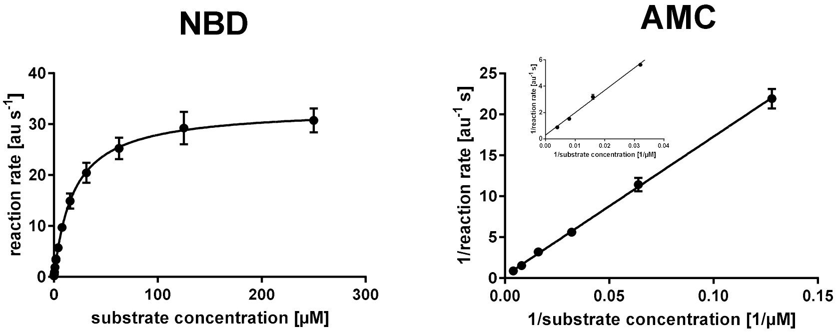


Figure S59. K_M of SARS-CoV 2 M^{pro} substrates. For the AMC substrate no reaction rate saturation was observed, thus, a Lineweaver-Burk linearization ($\frac{1}{k_{react}} = \frac{K_M}{k_{cat}}\frac{1}{[S]} + \frac{1}{k_{cat}}$) was employed to obtain K_M and k_{cat} values.

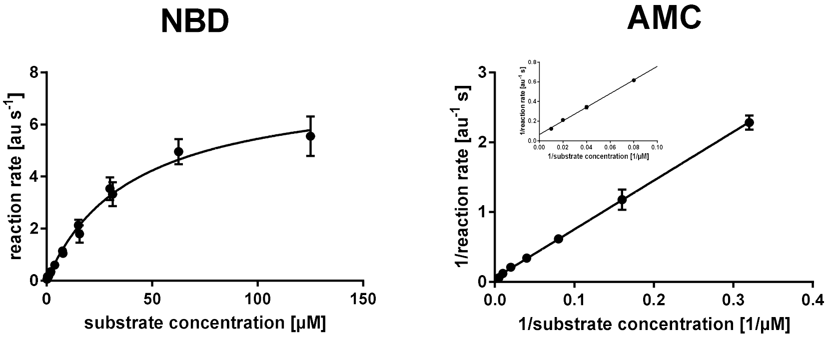


Figure S60. K_M of SARS-CoV 2 PL^{pro} substrates. For the AMC substrate no reaction rate saturation was observed, thus, a Lineweaver-Burk linearization ($\frac{1}{k_{react}} = \frac{K_M}{k_{cat}}\frac{1}{[S]} + \frac{1}{k_{cat}}$) was employed to obtain K_M and k_{cat} values.

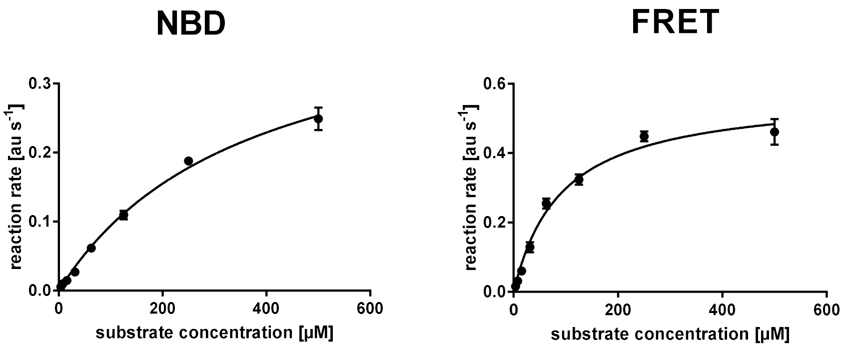


Figure S61. K_M of Sortase A (SrtA) substrates.

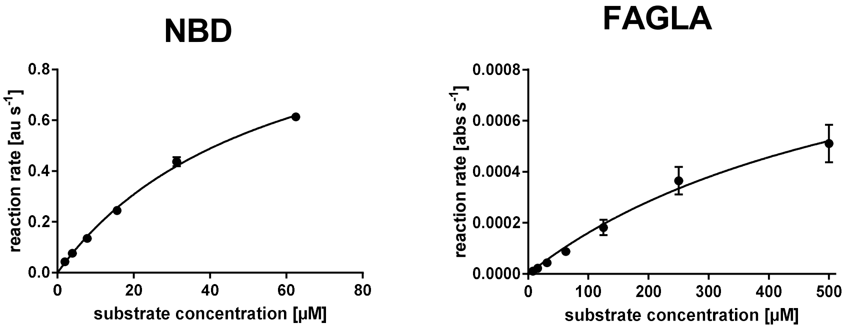


Figure S62. K_M of thermolysin substrates.

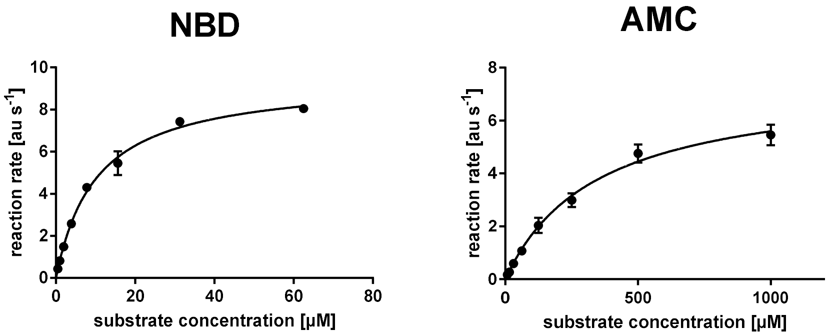
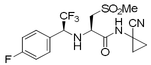
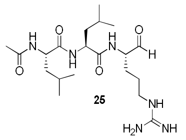
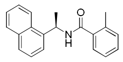
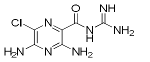
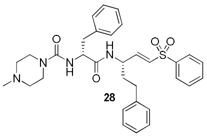
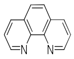
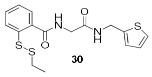


Figure S63. K_M of urokinase plasminogen activator (uPA) substrates.

Comparison of Literature Inhibitors by AMC and NBD Assays

Table S2. Inhibition constants for literature-known ligands were determined using the NBD-based and parent substrates. For covalent inhibitors (rhodesain & SrtA) K_i and k_{inact} values are reported according to the analysis procedure in the literature.[3,28]

		AMC substrate		NBD substrate	
protease	inhibitor	IC$_{50}$/k_{inact}	K_i	IC$_{50}$/k_{inact}	K_i
CatS	 **24** Wilkinson *et al.*[29]	1.64 ± 0.29 nM	1.27 ± 0.23 nM	1.63 ± 0.22 nM	1.20 ± 0.16 nM
M^{pro}	 **25** Fu *et al.*[30]	46.9 ± 4.78 µM	22.8 ± 3.37 µM	119 ± 13 µM	46.1 ± 5.6 µM
PL^{pro}	 **26** Welker *et al.*[12]	14.2 ± 0.96 µM	13.7 ± 0.93 µM	6.99 ± 0.64 µM	3.52 ± 0.37 µM
uPA	 **27** Buckley *et al.*[11]	8.87 ± 1.67 µM	5.24 ± 1.01 µM	10.0 ± 0.8 µM	5.06 ± 0.43 µM
Rhodesain	 **28** Previti *et al.*[31]	0.0062 ± 0.0003 s^{-1}	17.8 ± 3.5 nM	0.0056 ± 0.0002 s^{-1}	15.7 ± 2.3 nM
DENV NS2B/NS3	Aprotinin Abduraman *et al.*[32]	0.16 ± 0.01 µM	0.14 ± 0.01 µM	1.14 ± 0.08 µM	0.73 ± 0.06 µM
ZIKV NS2B/NS3	Aprotinin Abduraman *et al* [32]	1.11 ± 0.13 µM	1.04 ± 0.12 µM	2.35 ± 0.14 µM	0.94 ± 0.06 µM
Thermolysin	 **29** Nam *et al.*[33]	144 ± 9 µM	126 ± 8 µM	136 ± 6 µM	100 ± 5 µM
SrtA	 **30** Barthels *et al.*[3]	0.0068 ± 0.0004 s^{-1}	11.0 ± 1.9 µM	0.0066 ± 0.0004 s^{-1}	10.1 ± 1.81 µM

Dose-response curves for inhibitor evaluation

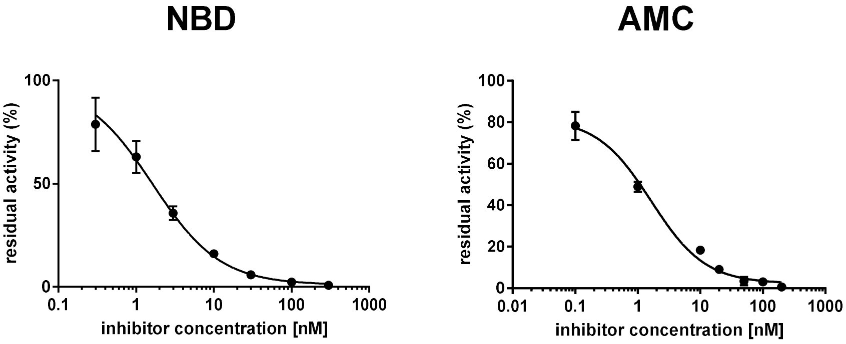


Figure S64. Dose-response curves of **24**[29] for CatS.

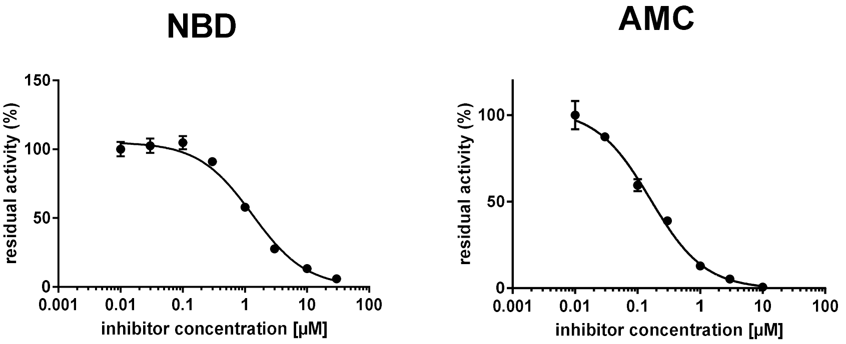


Figure S65. Dose-response curves of aprotinin[32] for DENV NS2B/NS3.

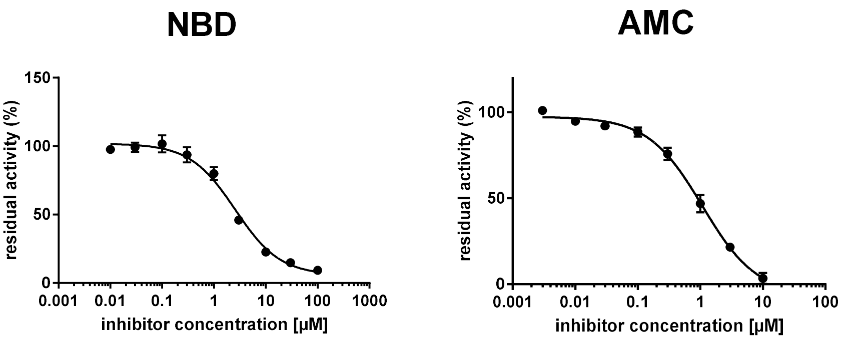


Figure S66. Dose-response curves of aprotinin[32] for ZIKV NS2B/NS3.

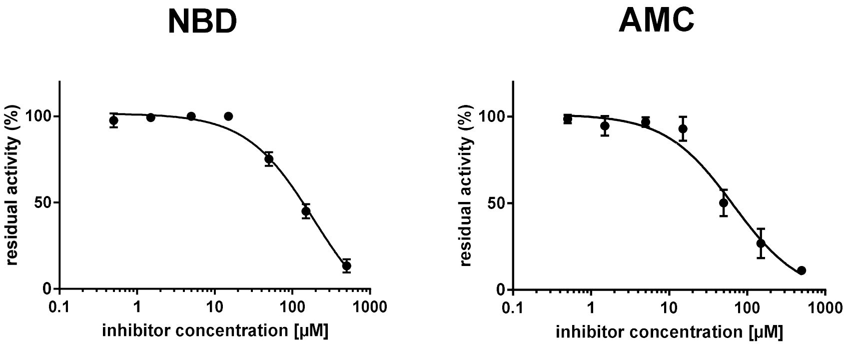


Figure S67. Dose-response curves of **25**[30] for SARS-CoV2 M^{pro}.

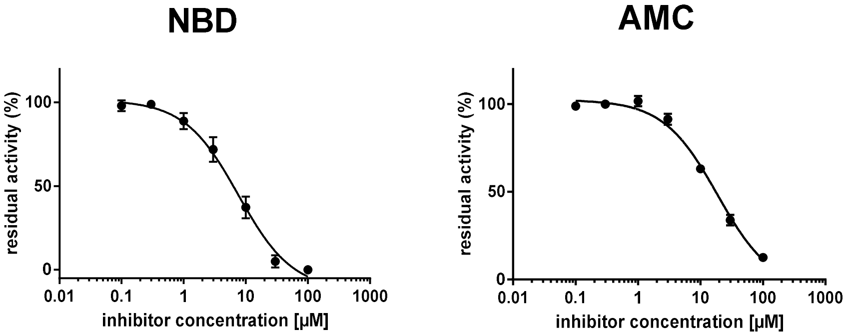


Figure S68. Dose-response curves of **26**[12] for SARS-CoV2 PL^{pro}.

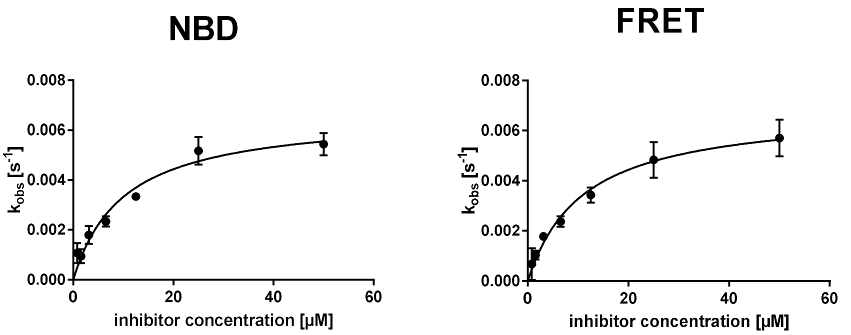


Figure S69. Dose-response curves of **30**[3] for SrtA.

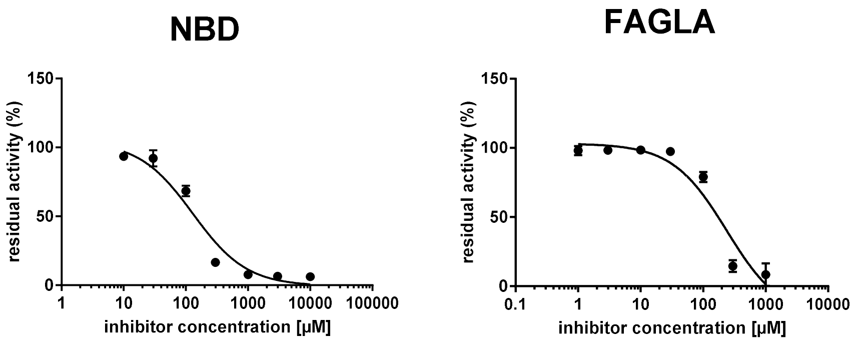


Figure S70. Dose-response curves of **29**[33] for thermolysin.

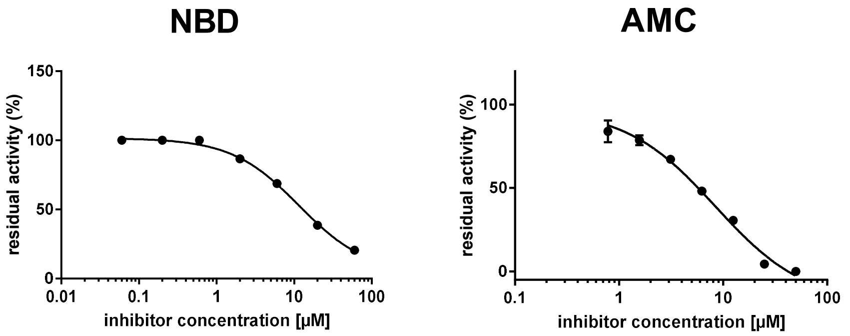


Figure S71. Dose-response curves of **27**[11] for uPA.

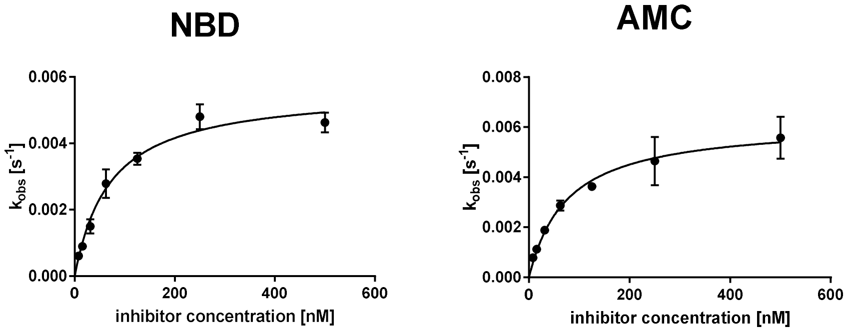


Figure S72. Dose-response curves of **28**[31] for rhodesain.

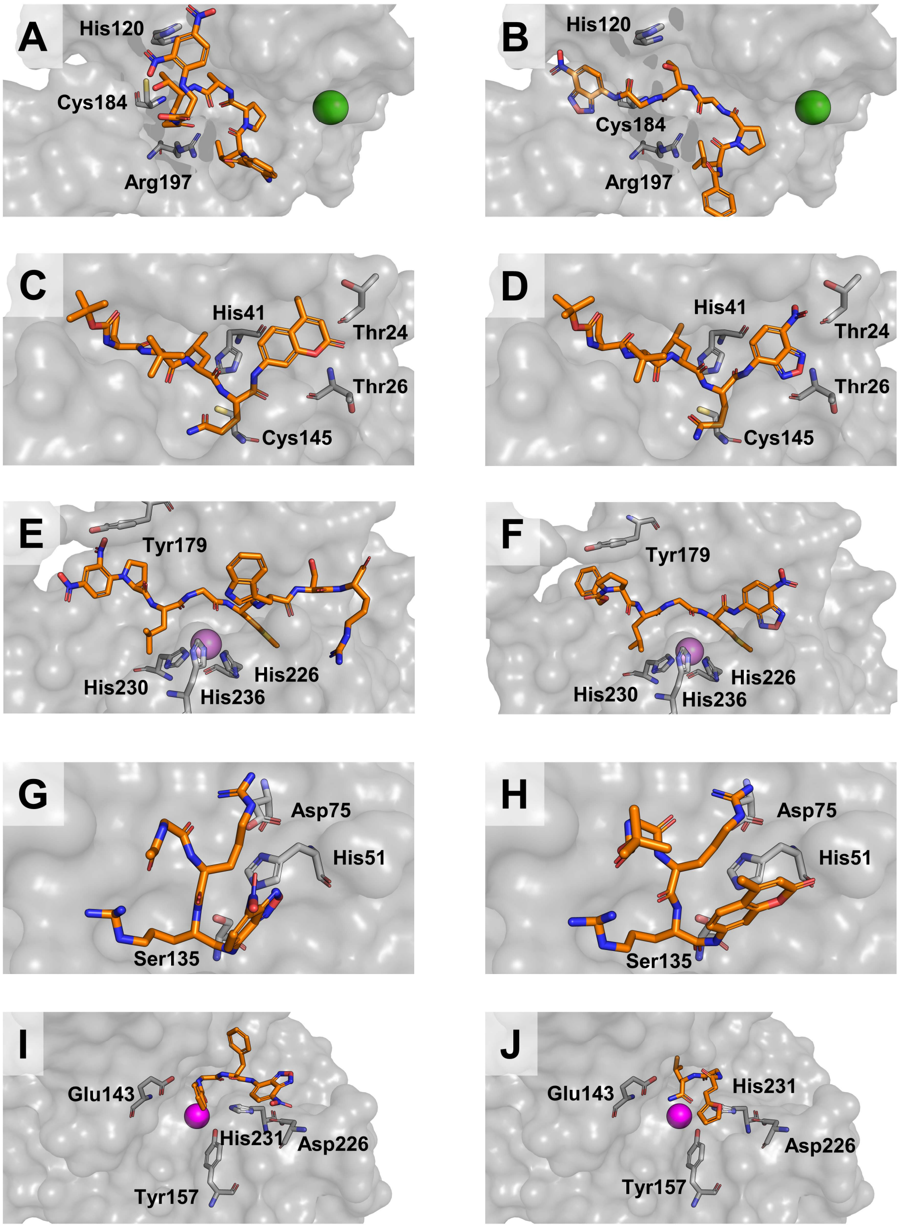


Figure S73. Predicted binding poses (Glide docking) of the NBD- and AMC-based substrates for (**A,B**) Sortase A (pdb: 2KID). (**C,D**) SARS-CoV M^{pro} (pdb: 6Z2E). (**E,F**) Matrix metalloprotease 9 (pdb: 4JQG). (**G,H**) DENV NS2B/NS3 (pdb: 3U1I). (**I,J**) Thermolysin (pdb: 1TMN).

‘‘

5 Project 2: NS2B-NS3 Protease-Ligand-Interactions

5.1 Conformational Dynamics of the Dengue Virus Protease Revealed by Fluorescence Correlation and Single-Molecule FRET Studies.

5.1.1 Context, Project Summary, and Own Contributions

It is known from crystal structures that the DENV protease adopts at least two different conformations, named open and closed conformation.[32,84,94] Although crystal structures provide valuable insights into the three-dimensional structure of proteins, they only capture a single energy-minimized state of the protein's conformation and do not provide information about conformational dynamics. Therefore, understanding the biological significance of observed conformations solely based on crystal structures can be challenging. This is particularly evident for a protein, such as the NS2B-NS3 protease, that exhibits conformational flexibility in solution.[106,113] Despite the common belief that the DENV protease's closed conformation represents the enzyme's catalytically active state[32,78], the overall conformational dynamics of the protein have yet to be extensively studied.

SmFRET has proven effective in uncovering static and dynamic heterogeneities in proteins[274,275] and was used in this study to enhance the understanding of the conformational dynamics of the protein in solution and investigate the influence of a competitive inhibitor on the conformational equilibrium. The protein was labeled with different dye-pairs (ATTO 488/ ATTO 647N or ATTO 488/ Cy5) at specific positions by introducing cysteine residues via site-directed mutagenesis (S79*C-S158C or S79*C-G14C), taking advantage of the fact that the native protein does not contain cysteines (Figure 22). These labeled mutants were then examined using confocal microscopy to study their behavior in solution. Auto- and cross-correlations were calculated to assess the existence of a dynamic equilibrium between different conformations and its modulation by the competitive inhibitor. Fluorescence burst analysis at the single-molecule level was performed to identify subpopulations and evaluate the effect of the inhibitor on conformational stability.

The results indicated the presence of a dynamic equilibrium between at least two conformations. Notably, the addition of the competitive inhibitor I_c^{DENV} significantly favored the closed state of the protein. The findings were consistent across different FRET pairs and labeling positions in the protein. Preliminary estimates suggested a conformational exchange timescale in the range of 10 ms.

The promising results obtained in this study lay the groundwork for investigations, including studying the response of the conformational equilibrium to allosteric inhibitors and immobilizing DENV protease for more in-depth analysis of individual molecules. These experiments hold the potential to provide quantitative information about conformational dynamics and contribute to improved antiviral drug design for DENV protease.

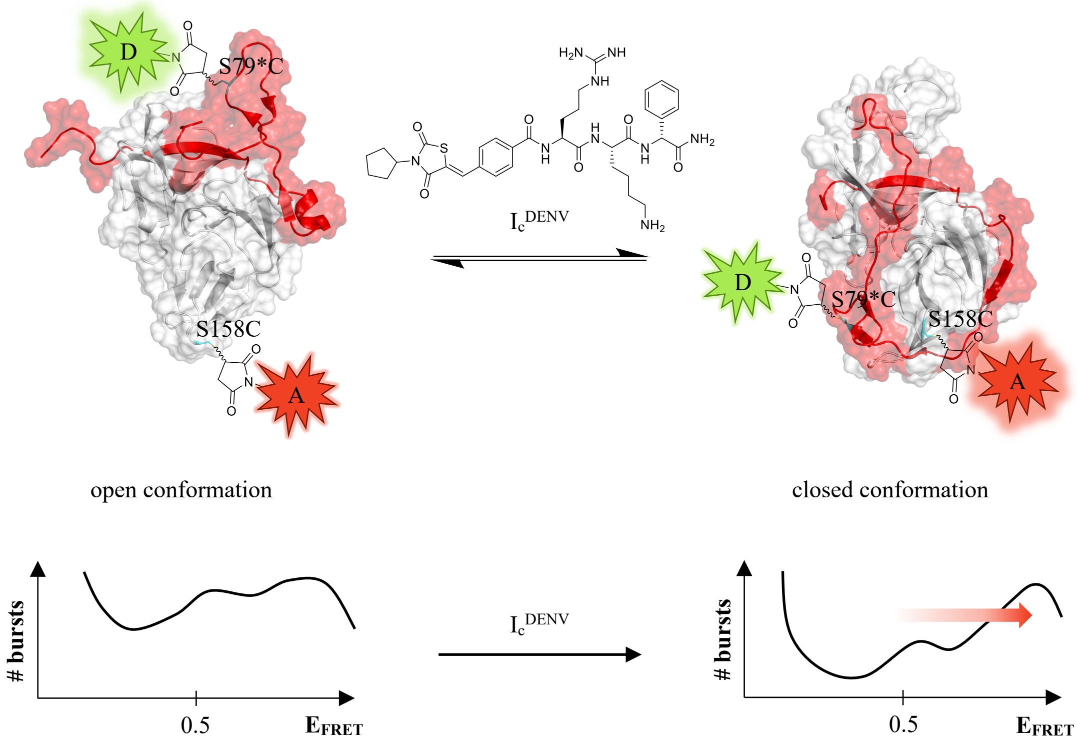


Figure 22. SmFRET experiments indicated competitive inhibitor' induced shift toward closed conformation of the previous existing conformational equilibrium. The distance of the mutated positions (S79*C and S158C) is smaller in the closed conformation of the DENV protease resulting in higher E_{FRET}. Open conformation (PDB-ID: 2fom), closed conformation (PDB-ID: 2m9p), NS2B is sown in red, NS3 in white. Mutated cysteine residues are shown with cyan sticks.

Own contributions: protein labeling of S79*C-S158C with ATTO 488 and Cy5 (together with Andrea Gellert), protein expression and purification (as part of my internship), fluorometric enzyme assay (activity of dye-labeled S79*C-S158C and control experiments not described in the publication), and manuscript editing.

Contributions from other authors: protein expression & purification, protein labeling, fluorometric enzyme assay, smFRET data curation, data analysis, and writing of the first draft & editing of the manuscript.

This work has been published in The Journal of Physical Chemistry B (impact factor: 3.47).

Article reprinted with permission of *The Journal of Physical Chemistry B* **2021**, 125 (25): 6837–6846 “Conformational dynamics of the dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies.” © 2021 American Chemical Society (United States).

5.1.2 Publication

The following publication quoted within “” from page 133 to page 147 is the same as the manuscript cited on page 132.

“

THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB Article

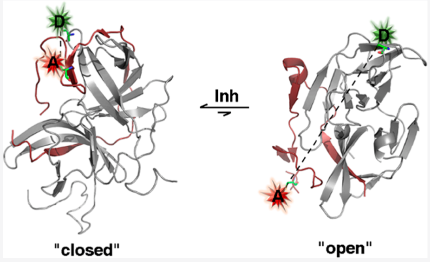
Conformational Dynamics of the Dengue Virus Protease Revealed by Fluorescence Correlation and Single-Molecule FRET Studies

Published as part of The Journal of Physical Chemistry virtual special issue “W. E. Moerner Festschrift”.

Christian Götz, Gerald Hinze, Andrea Gellert, Hannah Maus, Franziska von Hammerstein,
Stefan J. Hammerschmidt, Luca M. Lauth, Ute A. Hellmich, Tanja Schirmeister, and Thomas Basché*

Cite This: *J. Phys. Chem. B* 2021, 125, 6837−6846 **Read Online**

ACCESS | Metrics & More | Article Recommendations | Supporting Information



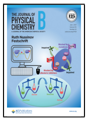
ABSTRACT: The dengue virus protease (DENV-PR) represents an attractive target for counteracting DENV infections. It is generally assumed that DENV-PR can exist in an open and a closed conformation and that active site directed ligands stabilize the closed state. While crystal structures of both the open and the closed conformation were successfully resolved, information about the prevalence of these conformations in solution remains elusive. Herein, we address the question of whether there is an equilibrium between different conformations in solution which can be influenced by addition of a competitive inhibitor. To this end, DENV-PR was statistically labeled by two dye molecules constituting a FRET (fluorescence resonance energy transfer) couple. Fluorescence correlation spectroscopy and photon-burst detection were employed to examine FRET pair labeled DENV-PRs freely diffusing in solution. The measurements were performed with two double mutants and with two dye couples. The data provide strong evidence that an equilibrium of at least two conformations of DENV-PR exists in solution. The competitive inhibitor stabilizes the closed state. Because the open and closed conformations appear to coexist in solution, our results support the picture of a conformational selection rather than that of an induced fit mechanism with respect to the inhibitor-induced formation of the closed state.

■ INTRODUCTION

The protease (PR) of the dengue virus (DENV) plays a crucial role in the replication cycle of the virus. It is involved in the cleavage of the viral polyprotein, resulting from the translation of the viral RNA, into the seven nonstructural proteins (NS) necessary for viral replication and maturation and three structural capsid, membrane precursor, and envelope proteins[1] that contribute to the formation of newly assembled virus particles.[2] Because of its essential function in virus replication, the DENV-PR is an attractive target for countering DENV infections. To create a basis for rational drug design, studies that provide structural insights into the dynamics of the DENV-PR are essential.

The DENV-PR is a heterodimeric trypsin-like serine protease consisting of the NS3 protease domain ($NS3_{pro}$) and the cofactor NS2B ($NS2B_{cf}$). Without the cofactor, the $NS3_{pro}$ domain shows only residual catalytic activity.[3] While the $NS3_{pro}$ domain adopts a similar spherical fold in all published structures, at least two different conformations were observed for the $NS2B_{cf}$ domain that differ mainly in the relative positions of the $NS2B_{cf}$ to the active center of the $NS3_{pro}$ domain.[4−8] As shown by cocrystal structures with ligands bound to the active site, in the closed conformation, the cofactor wraps around the $NS3_{pro}$ domain and contributes to the formation of the S2 and S3 pockets. Therefore, it is assumed that this conformation represents the catalytically active state.[4] In the open state, the $NS2B_{cf}$ domain is disordered and does not contribute to the formation of the substrate binding pocket.[5,6,9] Crystal structures provide substantial information about the three-dimensional fold of proteins; however, they represent only snapshots of energy-minimized states. Hence, it is difficult to evaluate the biological relevance of the two observed conformations only based on the available crystal structures. This is underscored by the fact that studies in solution provide evidence of conformational flexibility for this protein.[7,10] The formation of the closed state is often interpreted as a result of ligand binding to the

Received: February 27, 2021
Revised: May 21, 2021
Published: June 17, 2021



https://doi.org/10.1021/acs.jpcb.1c01797
J. Phys. Chem. B 2021, 125, 6837−6846

active site of the DENV-PR causing either an induced fit[7,8] or a conformational selection.[11,12] Most of the known competitive inhibitors of DENV-PR have a peptidic or peptidomimetic structure.[13,14] However, because of the preference of the S1 and S2 binding pockets for cationic residues, effective competitive inhibitors exhibit weak membrane permeabilities, thus lowering their bioavailability.[15−18] Furthermore, the development of selective competitive inhibitors is hampered by the shallow structure of the active site.[7,19] A putative allosteric binding site therefore represents an attractive alternative for effectively inhibiting the protease activity.[7,15,20] Interestingly, inhibitors targeting this allosteric site have been reported to stabilize the open state.[7]

Despite the widespread assumption that the closed conformation resembles the catalytically active state of the enzyme, the overall conformational dynamics of the protein have not yet been sufficiently explored. NMR studies employing different types of spin-labels have found at least two species in solution. These were interpreted as two conformations whose ratio was shifted to the closed conformation after addition of a competitive inhibitor.[21,22] In these studies, a higher flexibility of the open conformation and the possibility of intermediate states have been mentioned. Consequently, the open conformation may result from a superposition of structurally similar states. Later on, a luminescence assay was presented to monitor conformational transitions of the NS2B cofactor upon ligand binding to characterize potential allosteric inhibitors.[7] By use of a split luciferase complementation assay, it was found that binding of a competitive inhibitor to the PR induced the formation of the closed conformation, whereas binding of an allosteric inhibitor prevented the formation of the catalytically active closed state. Because only relative changes in luminescence could be detected with this method, no conclusions could be drawn regarding the occupancy and precise conformational states of subpopulations of the DENV-PR and whether a dynamic equilibrium between open and closed conformations in solution exists in the absence of ligands.

Förster resonance energy transfer (FRET),[23] implemented at the single-molecule level (smFRET),[24] has proven to uncover static and dynamic heterogeneities, which typically are averaged out in the ensemble. Thus, conformational subpopulations, conformational changes, and temporary fluctuations can be revealed.[25] In recent years, smFRET has become a common technique to analyze, for example, conformational dynamics of proteins including proteases,[26−28] the folding dynamics of RNA,[29] or the influence of ligands on the conformation of biomolecules.[30,31] In principle, smFRET experiments can be performed on immobilized or freely diffusing molecules.[32] In the latter case, individual photon bursts from molecules traversing the excitation volume are detected and analyzed to extract the unique single molecule information. The burst analysis allows to build histograms for the parameters of interest and to identify subpopulations in a heterogeneous ensemble.[32,33]

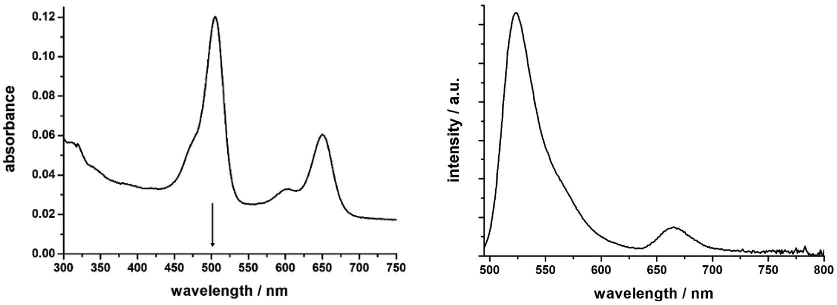
Since their first implementation, burst measurements have been continuously improved and refined. Experimentally accessible parameters comprise fluorescence wavelength,[34] intensity and lifetime,[33] and the polarization state of the excitation and emission.[35] Ideally, all intrinsic parameters of a fluorophore are detected simultaneously for each burst.[36,37] Recently, an analysis procedure has been presented for FRET pair labeled systems, which allowed for the visual detection of structural dynamics.[38] Briefly, FRET efficiencies and donor fluorescence lifetimes were determined for every burst and plotted against each other in a 2D histogram. It has been recognized that interconversion of conformational states which differ in their FRET efficiencies—during the molecule's transit time through the observation volume—will result in characteristic deviations from the so-called static FRET line.[38]

In a different approach it had been shown that conformational dynamics and the underlying kinetics of a FRET pair labeled system diffusing freely in solution can be addressed by comparing the auto- and cross-correlation functions of the donor and acceptor intensities.[39]

The present study aims to contribute to an improved understanding of the conformational dynamics of DENV-PR and to study the influence of a competitive inhibitor on the conformational equilibrium. Because the native DENV-PR does not contain any cysteines, two cysteine residues were introduced in each PR by site-directed mutagenesis. Importantly, one cysteine was placed in the NS3 core protease domain (G14C or S158C) and one within the NS2B cofactor (S79C). The double mutants were labeled by FRET pairs and investigated in a confocal microscope while freely diffusing in solution. The labeling positions were chosen to lead to significant changes of the FRET efficiencies after a conformational transition occurred. The stream of fluorescence photons, arising from individual PR molecules traversing the excitation volume, was subjected to two analyses based on the approaches discussed before. For an initial evaluation of whether a dynamic equilibrium between different conformations of the DENV-PR exists and whether this equilibrium is influenced by the addition of a competitive inhibitor, we calculated auto- and cross-correlations for each data set. Afterward, the data were analyzed at the single-molecule level by implementing a fluorescence burst analysis. For this purpose, individual bursts, attributed to individual DENV-PR molecules, were examined and sorted in a 2D histogram according to lifetime and FRET efficiency.[38] This plotting in a 2D histogram enabled the visual identification of the existing subpopulations. In addition, the histograms showed whether a ligand stabilizes a conformation and if so, which one. Furthermore, by comparing the positions of existing subpopulations with respect to the static FRET line, we had a second qualitative indicator for the presence of exchange dynamics. Our results point to a dynamic equilibrium of at least two conformations. Notably, the addition of a competitive inhibitor clearly favored the closed state of the enzyme. However, it has to be kept in mind that due to its flexibility the open conformation may actually be a superposition of structurally similar states. The results were found to be qualitatively independent of the selected FRET pair as well as of the position of the FRET labels in DENV-PR. Preliminary estimates indicate a time scale of the conformational exchange in the range of 10 ms.

■ METHODS

Protein Constructs and Cloning. A pET15b vector encoding the DENV2 NS2B$_{cf}$-NS3$_{pro}$ gene (GenBank ID: AY037116.1), with two-point mutations (I30A and L31A in NS3) for better protein solubility,[40] was used for introducing cysteine point-mutation pairs S79C/G14C and S79C/S158C in NS2B and NS3, respectively, through site-directed mutagenesis using the Kapa HiFi PCR kit (KapaBiosystems) and the following primers: (i) G14C: 5′-ccgccggtgtgcaaagcagaa-



(a) (b)

Figure 1. Ensemble spectra of the DENV-PR S79C−G14C double mutant labeled with the ATTO 488/ATTO 647N FRET pair in assay buffer with 10 vol % of DMSO. (a) Absorption spectrum (the arrow indicates the excitation wavelength (502 nm) for the correlation and burst experiments) and (b) emission spectrum. Following excitation of the donor (488 nm) emission from both donor and acceptor is observed.

ctggaagac-3′ and 5′-ctgctttgcacaccggcggcggagacgg-3′; (ii) S79C: 5′-gatcctgtcaattacgatctgcgaagatg-3′ and 5′-gacatgctaccatcttcgcagatcgtaattg-3′; (iii) S158C: 5′-ggatcgggcagctaccgctgatttctgcttg-3′ and 5′-cgtaggcaccacaacgggtcacgacaccattacc-3′.

Protein Expression and Purification. *E. coli* BL21-Gold (DE3) cells (Agilent Technologies, Santa Clara, CA) were transformed with a pET15b plasmid harboring the DENV2 $NS2B_{cf}$-$NS3_{pro}$ with an N-terminal hexa-histidine tag linked to $NS2B_{cf}$ via a thrombin cleavage site. Cells were grown in LB medium containing 100 μM ampicillin at 37 °C to an OD_{600} of ∼0.7. Expression was induced with 1 mM isopropyl-D-thiogalactoside (IPTG) for 16 h at 20 °C. After harvesting, cells were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.1 vol % Triton X-100, 20 mM imidazole, RNase, DNase, lysozyme, 1 mM PMSF, 1 mM DTT, 1 mM benzamidine) and lysed by sonication (Sonoplus, Bandelin, Berlin, Germany). The lysate was cleared by centrifugation (45 min at 15 krpm), and the protein was purified from the supernatant by gravity flow chromatography with Ni-NTA Agarose (Qiagen, Venlo, Netherlands). Eluted proteins were subsequently subjected to a gel-filtration step (HiLoad 16/600 Superdex 200 column, GE Healthcare) and eluted in the storage buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl). Purified proteins were concentrated, shock frozen in liquid nitrogen and stored at −80 °C until further use. Throughout all steps, protein concentrations were measured via absorbance at 280 nm, and sample purity was assessed via SDS-PAGE.

Fluorescent Labeling of the DENV-PR Double Mutants S79C−G14C and S79C−S158C. Purified protein was rebuffered by using Vivaspin protein concentrators (MWCO 10.0 kDa, GE Healthcare) in labeling buffer (20 mM Tris-HCl pH 6.8−7.0, 150 mM NaCl, 0.5 mM TCEP) and incubated for 30 min at 4 °C. DENV-PR S79C−G14C: A 4.7-fold excess of the maleimide derivative of ATTO 488 and a 5.8-fold excess of the maleimide derivative of ATTO 647N (both ATTO dyes were purchased from ATTO-TEC GmbH, Germany) were mixed with DENV-PR S79C−G14C (40 μM). DENV-PR S79C−S158C: TCEP was removed with a Vivaspin protein concentrator in the same buffer without TCEP. A 2.1-fold excess of ATTO 488 and a 2.6-fold excess of ATTO 647N or a 2.6-fold excess of ATTO 488 and 2.4-fold excess of Cy5 (GE Healthcare, United States) were used (24 μM protein for ATTO 488/ATTO 647N or 75 μM protein when labeled with ATTO 488/Cy5). Samples were incubated by gently shaking in the dark for 2 h at room temperature. This enables a Michael addition of the thiol group of cysteine to the β-position of the α,β-unsaturated carbonyl moiety (maleimide anchor) of the dyes, forming a stable thioether. Labeled protease was then dialyzed overnight at 4 °C (DENV-PR S79C−G14C: 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5 mM TCEP; DENV-PR S79C−S158C: 20 mM Tris-HCl pH 6.8, 150 mM NaCl, 1 mM CHAPS) and purified by gel filtration (HiLoad 16/60 Superdex 75 column, GE Healthcare) in the storage buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl). The activity of the labeled enzyme was confirmed in a fluorometric assay by using a fluorogenic peptide substrate. Additionally, the inhibitory activity of known competitive and noncompetitive inhibitors was examined.[20] For a more detailed description of the test of the enzyme activity, please refer to the Supporting Information. To determine the dye-to-protein ratio, the absorbance was measured at 280, 500, and 646 nm for DENV-PR S79C−G14C, ATTO 488, and ATTO 647N, respectively. Absorption and fluorescence emission spectra of the ATTO 488/ATTO 647 N FRET pair labeled DENV-PR S79C−G14C double mutant recorded in assay buffer (50 mM TRIS-HCl pH 9.0, 1 mM CHAPS, 20 vol % of glycerol) with 10 vol % of DMSO are shown in Figure 1. The spectra indicate a pronounced donor-only population, which may be attributed to remaining free dye or a large excess of singly or doubly donor-labeled proteases. The interference from these species can be minimized by interrogating donor- and acceptor-labeled proteases.

FCS-FRET and smFRET Experiments. The measurements on freely diffusing DENV-PR molecules were performed in self-made sample cells, consisting of a glass coverslip coated with poly(ethylene glycol) (PEG_01 & 02, MicroSurfaces Inc., United States)—to prevent nonspecific adsorption of proteins to the surface—and a glued-on cylinder, made of protein-repellent plastic (Protein LoBind Tube, Eppendorf AG, Germany). 150 μL of the respective solution of FRET pair labeled double mutant DENV-PR ($c \sim 100$ pM) in assay buffer (the assay buffer for the measurements on the DENV-PR S79C−G14C double mutant additionally contained 20 vol % of glycerol) was placed in the sample cells, and fluorescence photons were collected over a period of 1800 s in each case by using a custom-built confocal microscope. To ensure the solubility of the ligand, 10 vol % of DMSO was added to the solutions. After the measurements in the absence of the ligand, 150 μL of a solution of the competitive inhibitor ($c \sim 20$ μM) in assay buffer (with 10 vol % of DMSO) was added to the sample cells, and again fluorescence photons were collected. The recordings of both data sets—with and without ligand—were performed in direct succession.

The fluorophores were excited with the spectrally filtered output from a pulsed white light fiber laser (10 MHz, SC-OEM, YSL Photonics, China). An acousto-optical tunable filter (AOTF-VIS-DR, Fianium, United Kingdom) was used to obtain excitation pulses centered around 502 nm (FWHM ~ 1.5 nm). After passing a single-mode fiber, the laser light was recollimated and then focused into the solution by using a microscope objective (plan apochromat, 100×, NA = 1.4, oil immersion, Zeiss, Germany). The excitation power was measured in front of the objective and set to 10 μW. The same objective collected the fluorescence emitted by dye labeled DENV-PR while diffusing through the focus. Excitation and emission were separated by a dichroic mirror (ZT491 rdcxrxt-UF1, CHROMA, United States). The emitted fluorescence light finally passed a dichroic mirror (ZT640rdc-UF1, CHROMA, United States) in the detection beam path, which allowed red light to be transmitted, while light of higher energy was reflected. These two spectrally separated beams were then focused onto two APDs (acceptor channel A: SPCM-AQRH-15, PerkinElmer, United States; donor channel D: PDM 50ct, MPD, Italy) after appropriate spectral filtering (acceptor channel A: ET655lp, CHROMA, United States; donor channel D: FF01-550/88-25, Semrock, United States). This filter combination allowed the separate detection of both donor and acceptor fluorescence. The two detectors were connected to a HydraHarp 400 module (PicoQuant, Germany), which records the absolute and relative arrival time (relative to the excitation pulse) of the individual photons. This enabled the subsequent calculation of correlation functions and fluorescence intensity time traces for the burst analysis. Such time traces were calculated for both colors, and the fluorescence lifetime was determined for each burst within a trace. Therefore, we address these traces as color- and time-resolved fluorescence intensity time traces.

Analysis of FCS-FRET and smFRET Data. For a first tentative evaluation whether a dynamic equilibrium between different conformations of the DENV-PR exists and whether this equilibrium is influenced by the addition of a competitive inhibitor, correlation functions were calculated (FCS-FRET) from the macro arrival times of the individual photons recorded by using the HydraHarp 400 module. The fluorescence intensities on both APDs were auto- and cross-correlated (autocorrelations $G_{AA}(\tau)$, $G_{DD}(\tau)$, and cross-correlations $G_{AD}(\tau)$).

It had been shown that the cross-correlation can drop more slowly to zero than both autocorrelations if there is an exchange between two conformations that differ in their energy transfer efficiency provided that this process is not much slower than the diffusional time through the focus.[39,41] Exchange manifests itself in additional terms in the correlation functions; that is, in addition to the diffusional term without exchange, autocorrelations decay faster with an exchange time τ_{ex}, $G_{ii}(\tau) \propto (1 + a_{ii} \exp(-\tau/\tau_{ex}))$, while the cross-correlation function is extended by an increasing term $G_{AD}(\tau) \propto (1 - a_{AD} \exp(-\tau/\tau_{ex}))$. Without exchange ($\tau_{ex} \to \infty$), $G_{AD}(\tau)$ is expected to decay in between $G_{AA}(\tau)$ and $G_{DD}(\tau)$.[42] $G_{AA}(\tau)$ and $G_{DD}(\tau)$ could decay differently due to different sizes of the observation volumes. Notably, it is expected that the cross-correlation does no longer drop significantly slower to zero compared to the autocorrelations if an added ligand stabilizes one of the conformations of the protein, and therefore no or a significantly reduced conformational exchange takes place. Thus, by comparing the auto- and cross-correlation functions of the donor and acceptor intensities, we can draw conclusions about the existence of conformational dynamics.

To explore the presence of conformational subpopulations, data must be viewed at the single-molecule level (smFRET). For this purpose, individual bursts extracted from the fluorescence intensity time traces (bin time = 1 ms) were analyzed. For further evaluation, only bursts were taken into account which showed a total of at least 10 counts on both APDs. An average arrival time τ_{DA} of the donor photons in the presence of the acceptor relative to the excitation pulse (without taking the instrumental response function (IRF) into account) was calculated for each individual burst. Furthermore, the FRET efficiency E_{FRET} was calculated for each burst:

$$E_{FRET} = \frac{f_c \times \text{counts(acceptor channel A)}}{f_c \times \text{counts(acceptor channel A)} + \text{counts(donor channel D)}} \tag{1}$$

We did not account for spectral crosstalk or direct excitation of the acceptor. But we took into account a detection correction factor f_c, which compensates for the different detection efficiencies of the two APDs toward donor and acceptor fluorescence and the different quantum yields of both dyes, when calculating E_{FRET}. Note that the correction factor was calculated from the known characteristics of the optical components of the microscope, without further experimental validation. We found that in the case of the FRET pair ATTO 488/ATTO 647N the detection correction factor is close to 1. For the second FRET pair ATTO 488/Cy5, a correction factor $f_c = 2.3$, resulting from the lower quantum yield of Cy5 compared to ATTO 647N, was used.

The occurrences of the individual bursts (individual molecules) were then plotted in a 2D histogram, separated according to donor lifetime and FRET efficiency, which enables the visual identification of existing subpopulations. In addition, the 2D histograms show whether a ligand stabilizes a conformation and, if so, which one. Moreover, we exploited the linear relationship between FRET efficiency and donor lifetime in the presence (τ_{DA}) of the acceptor for molecules that remain in the same conformational state during the single particle's transit time through the focus, which is given by the following expression:

$$E_{FRET} = 1 - \frac{\tau_{DA}}{\tau_D} \tag{2}$$

τ_D represents the donor lifetime in the absence of the acceptor. Importantly, eq 2 gives rise to the so-called static FRET line because it only remains valid in the static case. If the studied system switches between various conformational states within a burst, which differ in their FRET efficiencies, there will be a characteristic deviation of the data from the static FRET line toward higher donor lifetimes.[38] Such deviations represent a second qualitative indicator for the presence of exchange dynamics.

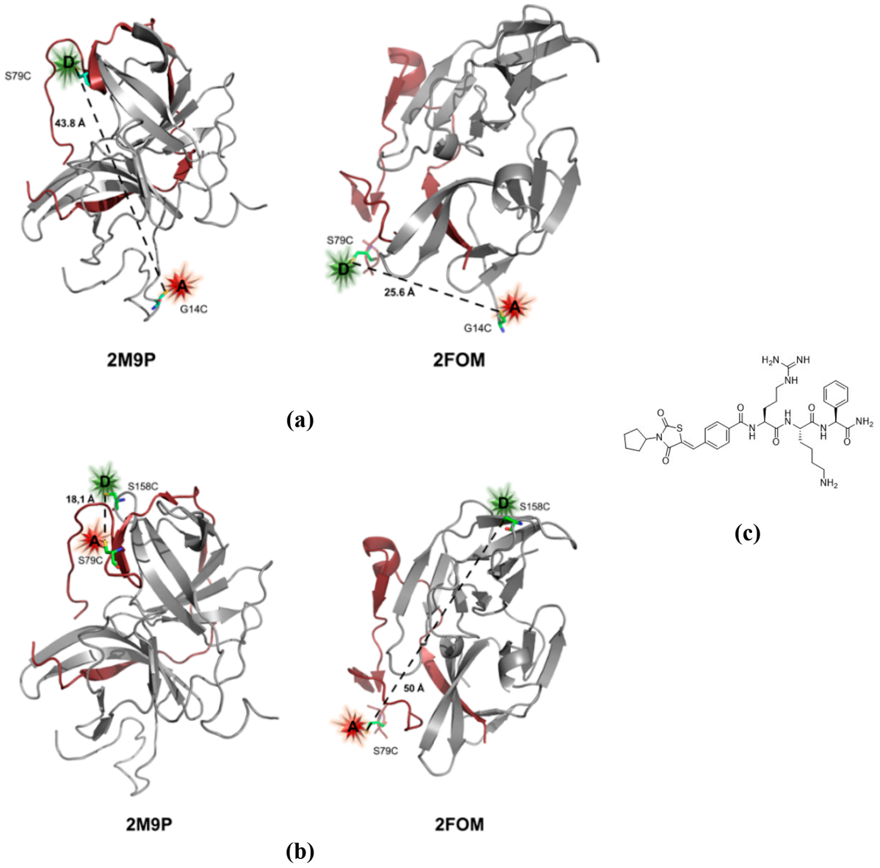


Figure 2. (a, b) Crystal structures of DENV-PR (PDB left 2M9P, right 2FOM) including the different labeling sites and the inter-residue distances. NS2B is colored in red and NS3 in gray. (a) Inter-residue distances for the DENV-PR S79C−G14C double mutant in the closed (left) and open (right) conformation. (b) Inter-residue distances for the DENV-PR S79C−S158C double mutant in the closed (left) and open (right) conformation. (c) Structure of the competitive inhibitor **1**.[43]

■ RESULTS AND DISCUSSION

FRET Pair Labeled DENV-PR Mutants. Two double mutants of DENV-PR were prepared by introducing cysteines to the sites of S79 in NS2B with either G14 (termed S79C-G14C) or S158 in NS3 (termed S79C-S158C) by site-directed mutagenesis. Importantly, the double mutants were selected to show significant distance changes between the mutated positions upon changes of conformation (Figure 2a,b). Moreover, in at least one of the conformations the inter-residue distance r was estimated to be close to the Förster radius R_0 of the respective dye pair. Finally, it was checked that the enzyme activity was not affected by the point mutations. The two cysteines of each double mutant were statistically labeled with the respective dyes equipped with maleimide linkers. Accordingly, a mixture of donor-only-, acceptor-only-, and donor−acceptor-labeled proteases is obtained. Informative molecules are those that are labeled with one donor and one acceptor dye.

Table 1 provides an overview of the different FRET pair labeled DENV-PR double mutants.

Fluorescence Measurements. *DENV-PR S79C−G14C before and after Addition of the Competitive Inhibitor **1**.* First, auto- and cross-correlations were calculated from the stream of emitted photons of the ATTO 488/ATTO 647N FRET pair labeled double mutant S79C−G14C of DENV-PR. The samples were excited at 502 nm, and in Figure 3 the correlation functions are shown before and after the addition of the competitive inhibitor **1** (Figure 2c). In both cases the cross-correlation $G_{AD}(\tau)$ drops to zero on the same time scale as $G_{AA}(\tau)$. Note that in the presence of conformational exchange $G_{AD}(\tau)$ should decay slower than both autocorrelations, while in the absence of any exchange $G_{AD}(\tau)$ is expected

Table 1. Overview of the Relevant Parameters of the FRET Pair Labeled Double Mutants of DENV-PR[a]

			r/nm		E_{FRET}/%	
sample	FRET pair	R_0/ nm	open	closed	open	closed
S79C−G14C	ATTO 488/ ATTO 647 N	4.9[44]	2.6	4.4	98.0	66.2
S79C−S158C A	ATTO 488/ ATTO 647N	4.9[44]	5.0	1.8	47.0	99.7
S79C−S158C B	ATTO 488/ Cy5	5.1[45]	5.0	1.8	53.0	99.8

[a] R_0: Förster radius; r: distances between the two labeling positions in the open and closed conformation of the DENV-PR; E_{FRET}: FRET efficiency, calculated according to $E_{FRET} = R_0^6/(R_0^6 + r^6)$. The distances r are based on the position of the cysteine residues according to crystallographic data. The actual distances between donor and acceptor dyes of the two FRET pairs may vary due to linking groups and the spatial dimensions of the dyes.

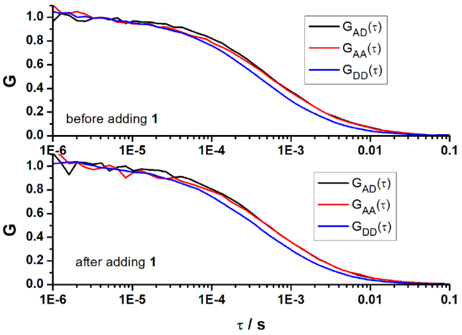


Figure 3. Comparison of the normalized cross-correlation (black curve) with the two normalized autocorrelations (red and blue), each before (upper graph) and after (lower graph) addition of the competitive inhibitor **1** to the ATTO 488/ATTO 647 N FRET pair labeled DENV-PR S79C−G14C double mutant.

to decay in between $G_{AA}(\tau)$ and $G_{DD}(\tau)$. Obviously, in this particular case the correlation data did not allow for a straightforward conclusion.

To prepare the data for the burst analysis, fluorescence intensity time traces were calculated with a bin time of 1 ms. A section of an intensity time trace before the addition of **1** is shown in Figure 4. It is seen that significantly more bursts are observed in the donor than in the acceptor channel. This is in accordance with the bulk solution spectra (see Figure 1) and probably due to the presence of a pronounced donor-only population. Note that the acceptor-only species is not observed with an excitation wavelength of 502 nm.

While the correlation data did not allow for conclusions on conformational dynamics, the burst analysis of the same data provides some evidence for the presence of different conformations and an impact of ligand **1** on the equilibrium. The lower part of Figure 5 shows the 2D histograms obtained from the burst analysis of the intensity time traces before and after addition of **1**.

First of all, it should be noted that the populations with a FRET efficiency of zero are due to DENV-PR molecules only

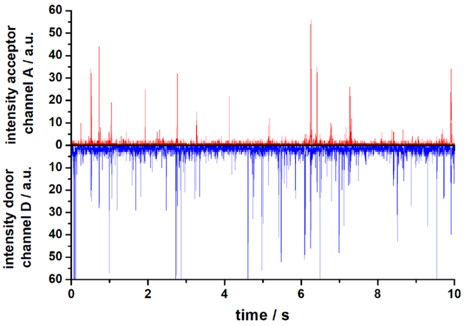


Figure 4. Fluorescence intensity time trace of the ATTO 488/ATTO 647 N FRET pair labeled DENV-PR S79C−G14C double mutant before adding **1**. The sample was excited with excitation pulses at 502 nm. The binning time was 1 ms.

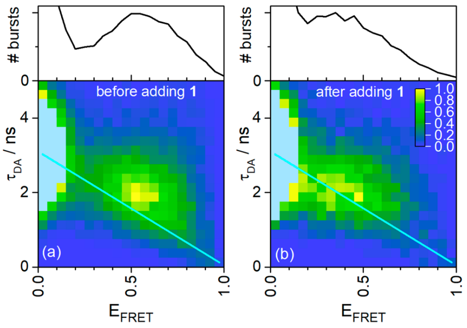


Figure 5. Plot of the normalized occurrences of the individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 647 N FRET pair labeled DENV-PR S79C−G14C double mutant before (a) and after (b) adding **1** in a 2D histogram, separated according to donor lifetime τ_{DA} and FRET efficiency E_{FRET}. For an easier visual comparison of the FRET populations before and after addition of the inhibitor, the respective maxima were normalized to one. The maxima of the donor-only species (for further explanation see main text) were not taken into account in the normalization, which is why values greater than one occur (shown in gray). The cyan lines represent the static FRET lines. The respective normalized 1D histograms of FRET efficiencies are shown as a projection on the top of the 2D histograms.

labeled with the donor dye ATTO 488, so that no energy transfer takes place. Therefore, these populations show a significantly longer fluorescence lifetime of the donor. Furthermore, no distinct FRET subpopulations are observed, neither before nor after adding **1**. Rather, broad point clouds are visible in both cases. These observations could be either due to the fact that the DENV-PR does not show significantly different conformations or due to the flexibility of the G14C point mutant, which could result in “smearing” of possibly existing conformational subpopulations. Another explanation could be that the conformations convert much faster into each other than the temporal resolution of the experiment (transit time through the focus ∼0.2 ms); thus, only an average of different conformations can be observed. Interestingly, before the addition of **1**, the observed FRET population is shifted significantly toward longer donor fluorescence lifetimes with respect to the static FRET line, which indicates conformational transitions during the transit time of the enzyme through the focal volume. Note that the beginning of the static FRET line (E_{FRET} = 0) is given by the mean lifetime of the donor-only species. After the addition of **1** (Figure 5b), the FRET population point cloud vertically shifts down closer toward the static FRET line, being compatible with reduced conformational exchange. At the same time, the point cloud shifts horizontally to smaller FRET efficiencies (see the projection in the upper part of Figure 5), indicating an increasing population of the closed conformation of DENV-PR. Before the inhibitor is added, the centroid of the FRET efficiency distribution of the DENV-PR molecules is located at ∼55%, whereas the addition of the inhibitor **1** leads to a shift of the centroid to values of 20%−35%.

Thus, although no distinct subpopulations were visible, the burst analysis supports the assumption of dynamical exchange between different conformations in the absence of an inhibitor. By the addition of **1**, the dynamics appears to be reduced and the equilibrium shifts to the closed conformation, as expected for binding of competitive inhibitors.[7,8]

DENV-PR S79C−S158C A before and after Addition of the Competitive Inhibitor 1. To prevent potential conformational flexibility, the G14C point mutation within a flexible linker of NS3 was replaced by the S158C point mutation in the same protease domain but on the opposite end within a β-turn (Figure 2). It was thus presumed that this residue should display less flexibility than G14. The two labeling positions of the DENV-PR S79C−S158C double mutant give rise to an inverted distance change compared to DENV-PR S79C−G14C, as it is now the closed rather than the open conformation where an increase in FRET efficiency is expected. Figure 6 shows the auto- and cross-correlations of the DENV-PR S79C−S158C double mutant before and after adding **1**. In the absence of the ligand, the cross-correlation drops to zero significantly slower than both autocorrelations. This fact can be understood as an indication that there is a significant exchange between two conformations that differ in their FRET efficiencies.[39,41] Interestingly, after addition of the competitive inhibitor, the cross-correlation decays in between both autocorrelations. This can be interpreted as a stabilization of one conformation upon inhibitor binding, thereby blocking or effectively reducing the conformational transition.

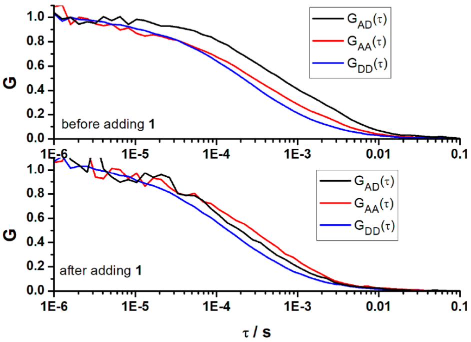


Figure 6. Comparison of the normalized cross-correlation (black curve) with the two normalized autocorrelations (red and blue), each before (upper graph) and after (lower graph) addition of the competitive inhibitor **1** to the ATTO 488/ATTO 647N FRET pair labeled DENV-PR S79C−S158C double mutant.

In Figure 7a, the 2D histogram obtained from the burst analysis of the intensity time trace is shown before addition of **1**.

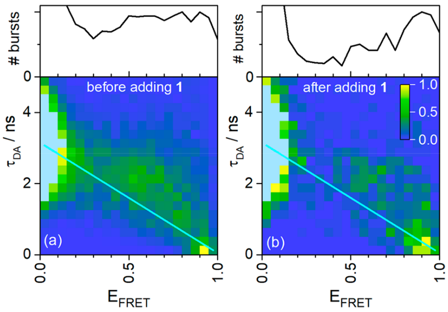


Figure 7. Plot of the normalized occurrences of the individual bursts within the intensity time trace of the ATTO 488/ATTO 647 N FRET pair labeled DENV-PR S79C−S158C double mutant before (a) and after (b) the addition of **1** in a 2D histogram, separated according to donor lifetime τ_{DA} and FRET efficiency E_{FRET}. For an easier visual comparison of the FRET populations before and after addition of the inhibitor, the respective maxima were normalized to one (see the inset in (b)). The maxima of the donor-only species were not taken into account in the normalization, which is why values greater than one occur (shown in gray). The cyan lines represent the static FRET lines. The respective normalized 1D histograms of FRET efficiencies are shown as a projection on the top of the 2D histograms.

The shape of the point cloud has changed significantly compared to the DENV-PR S79C−G14C double mutant (Figure 5). The point cloud is significantly elongated along the abscissa, indicating a much larger range of FRET efficiencies. With respect to the static FRET line, the data are shifted vertically toward longer donor fluorescence lifetimes, which indicates conformational transitions during the transit time of DENV-PR within a burst. After addition of the competitive inhibitor **1**, the populations at intermediate FRET efficiencies are diminished, while the high FRET efficiency population stabilizes (see Figure 7b). Because this population is located at the static FRET line, reduced conformational exchange takes place.

The impact of the inhibitor **1** on the subpopulations is seen even more clearly in the projections of the 2D histograms in the upper part of Figure 7. In the presence of **1** (Figure 7b), a clear preference for populations with high FRET efficiencies is noted. Indeed, a distinct maximum at FRET efficiencies of 90% is observed. Without inhibitor the distribution of FRET efficiencies is appreciably broadened. By fitting two Gaussian functions with similar weight (not shown) to the projection data in Figure 7a, we found maxima at 59% and 91% FRET efficiency. Theoretically, 47% is expected in the open state and ∼100% in the closed state.

To summarize, with respect to conformational exchange the burst analysis results in the same conclusion as the correlation analysis. With addition of the competitive inhibitor exchange between conformations is significantly reduced. Consequently, it can be assumed that in solution a dynamic equilibrium between different conformations of the DENV-PR does exist. After the addition of **1**, the burst analysis indicates the stabilization of the closed conformation yielding high FRET efficiencies. Accompanying time-resolved ensemble experiments confirm these results (see section S4 of the Supporting Information).

The reasonable agreement of our findings for two different double mutants of DENV-PR led us to check whether they might depend on the selected FRET pair. Therefore, the S79C−S158C double mutant of the DENV-PR was additionally labeled with the FRET pair ATTO 488/Cy5 (S79C−S158C B). The results of the corresponding measurements are described in full detail in the Supporting Information. Briefly, the correlation data show the same behavior as has been observed for DENV-PR S79C−S158C A. Without inhibitor, there appears to be an exchange between different conformations, while the presence of **1** stabilizes one of the conformations. This view is supported by the burst analysis, which in the absence of **1** indicates conformational exchange, whereas addition of **1** leads to the preference of the closed conformation. Although the burst data for DENV-PR S79C−S158C labeled with two different FRET pairs are quantitatively slightly different, they clearly follow the same trends. We conclude that the results are largely reproducible throughout all experiments, independent of the selected dye pair (ATTO 488/ATTO 647N vs ATTO 488/Cy5) and independent of the double mutant used (DENV-PR S79C−G14C vs DENV-PR S79C−S158C): In the absence of the competitive inhibitor **1** there is conformational exchange, while its presence leads to a stabilization of the closed conformation.

Finally, we would like to briefly describe a first attempt to address the time scale of the conformational exchange. Being aware of its lack of quantitative accuracy, we have performed a rough evaluation assuming a simple two-state model[39] with distributions of transition rates. Details of the calculations can be found in section S3 of the Supporting Information. The experimental data shown in Figure 6 (DENV-PR S79C−S158C A) are compatible with transitions between an open and a closed state on a mean time scale of $\tau_{\mathrm{ex}} \approx 10$ ms without addition of **1**, while exchange seems to be negligible after the addition of **1**. Note that only exchange on time scales in the micro- to millisecond regime significantly alters autocompared to cross-correlation functions as we have observed. Similar time scales have been suggested by NMR experiments on NS2B-NS3 PR.[22]

■ CONCLUSIONS AND OUTLOOK

In the literature, an open and a closed conformation of DENV-PR have been described. The aim of this study was to examine whether these conformations do exist in solution and the conformational equilibrium is shifted by the addition of a competitive inhibitor.

To address this issue, we performed FCS-FRET and smFRET measurements on two freely diffusing FRET pair labeled double mutants of DENV-PR. To exclude dependencies of the experimental results on labeling position and the FRET pair, the measurements were performed on two different double mutants and with two different FRET pairs. The double mutants were designed in such a way that distance changes between the labeling positions during conformational transitions would lead to significant changes of the FRET efficiency.

As a main result, we have found that the data from all three samples are compatible with a conformational exchange equilibrium of DENV-PR in solution. While the correlation data for the more flexible DENV-PR mutant S79C−G14C could not be properly analyzed, a comparison of the fluorescence intensity cross-correlation and autocorrelation functions of the DENV-PR S79C−S158C double mutants labeled with different FRET pairs gave strong indications that there exists an equilibrium between open and closed conformations before addition of the competitive inhibitor **1**. From the FCS-FRET experiments we conclude that exchange between conformations takes place on the micro- to millisecond time scale.

For all three systems examined, this view was supported by the location of the distribution of FRET efficiencies with respect to the static FRET line in the corresponding 2D lifetime−FRET efficiency histograms. Furthermore, we observed that in all cases the addition of inhibitor **1** induced a shift of the conformational equilibrium, indicating that the active site directed ligand stabilizes the closed conformation. In the case of the DENV-PR S79C−S158C double mutants a stabilization of one conformation was already visible in the correlation functions. Moreover, the preference for the closed state was evident from the fact that the subpopulations observed in the 2D histograms were in significantly better agreement with the static FRET line after addition of the inhibitor and from a shift of the related FRET population. The general trends were reproducible throughout all experiments and thus independent of the chosen dye pair and double mutant. This leads us to conclude that we have addressed intrinsic protein dynamics. In summary, our measurements provide strong evidence that there exists an equilibrium of at least two conformations of DENV-PR in solution. The addition of the competitive inhibitor leads to a stabilization of the closed state and thus causes a reduction of the conformational exchange. Based on the notion of a conformational equilibrium in solution, our results support the model of conformational selection rather than an induced fit description with respect to ligand binding. The open and closed conformations are already present in solution, and the inhibitor preferentially binds to the closed state, thereby shifting the equilibrium.

The promising results obtained in this first study provide the basis for future investigations. Employing the same methodology, it will be studied how the conformational equilibrium responds to the presence of an allosteric inhibitor.[20] In contrast to competitive inhibitors, we assume that an open conformation is stabilized by allosteric ligands. In a different approach we will immobilize DENV-PR and study individual molecules for extended periods of time. Preliminary experiments have shown that single immobilized FRET pair labeled proteases can be detected with sufficient signal-to-noise-ratio. Conformational changes will manifest themselves in an anticorrelated behavior in the donor and acceptor intensity time traces. The underlying conformational transitions and their kinetics will be extracted by means of a step analysis promising more quantitative information about the conformational dynamics. Moreover, such experiments promise to directly measure the residence time of a given inhibitor at its binding site in DENV-PR. We are confident that this study and future experiments will contribute to a better understanding of the conformational dynamics of DENV-PR, thereby delivering some input for improved antiviral drug design.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c01797.

> Fluorometric assay for enzymatic activity; absorption and emission spectra of FRET pair labeled mutants; conformational exchange model; ensemble fluorescence decays; correlation functions, burst measurements and data analysis for the DENV-PR S79C−S158C B mutant (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Thomas Basché − *Department of Chemistry, Johannes Gutenberg-University Mainz, Mainz, Germany;* orcid.org/0000-0003-0502-7088; Phone: +49 (0)6131 3922707; Email: thomas.basche@uni-mainz.de

Authors

Christian Götz − *Department of Chemistry, Johannes Gutenberg-University Mainz, Mainz, Germany*

Gerald Hinze − *Department of Chemistry, Johannes Gutenberg-University Mainz, Mainz, Germany;* orcid.org/0000-0003-4355-9325

Andrea Gellert − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany*

Hannah Maus − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany*

Franziska von Hammerstein − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany*

Stefan J. Hammerschmidt − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany*

Luca M. Lauth − *Department of Chemistry, Johannes Gutenberg-University Mainz, Mainz, Germany*

Ute A. Hellmich − *Department of Chemistry, Johannes Gutenberg-University Mainz, Mainz, Germany; Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe-University Frankfurt, Frankfurt, Germany*

Tanja Schirmeister − *Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c01797

Author Contributions

C.G. conducted experiments, analyzed data, interpreted results, and wrote the paper. G.H. designed research, analyzed data, interpreted results, and wrote the paper. A.G. performed protein labeling, designed experiments, conducted control experiments, and wrote the paper. H.M. performed protein labeling and cloning of protein constructs, conducted control experiments, and wrote the paper. F.v.H. performed cloning of protein constructs and protein expression. S.H. performed protein expression and purification and wrote the paper. L.M.L. performed cloning of protein constructs and protein expression. U.A.H. designed experiments and organized the project. T.S. designed experiments, organized the project, and wrote the paper. T.B. designed research, organized the project, interpreted the results, and wrote the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the group of Prof. Dr. W. Diederich, University of Marburg, Germany, for providing us with the DENV2 NS2B$_c$-NS3$_{pro}$ gene (GenBank ID: AY037116.1), with the two I30A and L31A point mutations in NS3$_{pro}$. Furthermore, we thank Prof. Dr. C. Klein, Heidelberg University, for providing us with the competitive inhibitor **1**. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2051 - Project-ID 390713860 (U.A.H.).

■ REFERENCES

(1) Behnam, M. A. M.; Nitsche, C.; Boldescu, V.; Klein, C. D. The medicinal chemistry of dengue virus. *J. Med. Chem.* **2016**, *59*, 5622−5649.

(2) Mukhopadhyay, S.; Kuhn, R. J.; Rossmann, M. G. A structural perspective of the flavivirus life cycle. *Nat. Rev. Microbiol.* **2005**, *3*, 13−22.

(3) Gupta, G.; Lim, L.; Song, J. NMR and MD studies reveal that the isolated dengue NS3 protease is an intrinsically disordered chymotrypsin fold which absolutely requests NS2B for correct folding and functional dynamics. *PLoS One* **2015**, *10*, No. e0134823.

(4) Noble, C. G.; Seh, C. C.; Chao, A. T.; Shi, P. Y. Ligand-bound structures of the dengue virus protease reveal the active conformation. *J. Virol.* **2012**, *86*, 438−446.

(5) Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **2006**, *13*, 372−373.

(6) Nitsche, C.; Holloway, S.; Schirmeister, T.; Klein, C. D. Biochemistry and medicinal chemistry of the dengue virus protease. *Chem. Rev.* **2014**, *114*, 11348−11381.

(7) Brecher, M.; Li, Z.; Liu, B.; Zhang, J.; Koetzner, C. A.; Alifarag, A.; Jones, S. A.; Lin, Q.; Kramer, L. D.; Li, H. A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease. *PLoS Pathog.* **2017**, *13*, No. e1006411.

(8) Chen, W.-N.; Loscha, K. V.; Nitsche, C.; Graham, B.; Otting, G. The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI. *FEBS Lett.* **2014**, *588*, 2206−2211.

(9) Phoo, W. W.; El Sahili, A.; Zhang, Z.; Chen, M. W.; Liew, C. W.; Lescar, J.; Vasudevan, S. G.; Luo, D. Crystal structures of full length DENV4 NS2B-NS3 reveal the dynamic interaction between NS2B and NS3. *Antiviral Res.* **2020**, *182*, 104900.

(10) Hill, M. E.; Yildiz, M.; Hardy, J. A. Cysteine disulfide traps reveal distinct conformational ensembles in dengue virus NS2B-NS3 protease. *Biochemistry* **2019**, *58*, 776−787.

(11) Paul, F.; Weikl, T. R. How to distinguish conformational selection and induced fit based on chemical relaxation rates. *PLoS Comput. Biol.* **2016**, *12*, No. e1005067.

(12) Behnam, M. A. M.; Klein, C. D. P. Conformational selection in the flaviviral NS2B-NS3 protease. *Biochimie* **2020**, *174*, 117−125.

(13) Sandner, A.; Hüfner-Wulsdorf, T.; Heine, A.; Steinmetzer, T.; Klebe, G. Strategies for late-stage optimization: Profiling thermodynamics by preorganization and salt bridge shielding. *J. Med. Chem.* **2019**, *62*, 9753−9771.

(14) Voss, S.; Nitsche, C. Inhibitors of the Zika virus protease NS2B-NS3. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 126965.

(15) Yildiz, M.; Ghosh, S.; Bell, J. A.; Sherman, W.; Hardy, J. A. Allosteric inhibition of the NS2B-NS3 protease from dengue virus. *ACS Chem. Biol.* **2013**, *8*, 2744−2752.

(16) Cregar-Hernandez, L.; Jiao, G.-S.; Johnson, A. T.; Lehrer, A. T.; Wong, T. A. S.; Margosiak, S. A. Small molecule pan-dengue and West Nile virus NS3 protease inhibitors. *Antiviral chemistry & chemotherapy* **2011**, *21*, 209−217.

(17) Lim, S. P.; Wang, Q.-Y.; Noble, C. G.; Chen, Y.-L.; Dong, H.; Zou, B.; Yokokawa, F.; Nilar, S.; Smith, P.; Beer, D.; et al. Ten years of dengue drug discovery: Progress and prospects. *Antiviral Res.* **2013**, *100*, 500−519.

(18) Weigel, L. F.; Nitsche, C.; Graf, D.; Bartenschlager, R.; Klein, C. D. Phenylalanine and phenylglycine analogues as arginine mimetics in dengue protease inhibitors. *J. Med. Chem.* **2015**, *58*, 7719−7733.

(19) Brecher, M.; Zhang, J.; Li, H. The flavivirus protease as a target for drug discovery. *Virol. Sin.* **2013**, *28*, 326−336.

(20) Millies, B.; von Hammerstein, F.; Gellert, A.; Hammerschmidt, S.; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; et al. Proline-based allosteric inhibitors of Zika and dengue virus NS2B/NS3 proteases. *J. Med. Chem.* **2019**, *62*, 11359−11382.

(21) Zhu, L.; Yang, J.; Li, H.; Sun, H.; Liu, J.; Wang, J. Conformational change study of dengue virus NS2B-NS3 protease using ^{19}F NMR spectroscopy. *Biochem. Biophys. Res. Commun.* **2015**, *461*, 677−680.

(22) de La Cruz, L.; Nguyen, T. H. D.; Ozawa, K.; Shin, J.; Graham, B.; Huber, T.; Otting, G. Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: Fold analysis by pseudocontact shifts. *J. Am. Chem. Soc.* **2011**, *133*, 19205−19215.

(23) Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **1948**, *437*, 55−75.

(24) Ha, T.; Enderle, T.; Ogletree, D. F.; Chemla, D. S.; Selvin, P. R.; Weiss, S. Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 6264−6268.

(25) *Single Particle Tracking and Single Molecule Energy Transfer*; Bräuchle, C., Lamb, D. C., Michaelis, J., Eds.; Wiley: 2009.

(26) Ha, T.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 893−898.

(27) Santoso, Y.; Joyce, C. M.; Potapova, O.; Le Reste, L.; Hohlbein, J.; Torella, J. P.; Grindley, N. D. F.; Kapanidis, A. N. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 715−720.

(28) Ruer, M.; Krainer, G.; Gröger, P.; Schlierf, M. ATPase and protease domain movements in the bacterial AAA+ protease FtsH are driven by thermal fluctuations. *J. Mol. Biol.* **2018**, *430*, 4592−4602.

(29) Zhuang, X.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T.; Herschlag, D.; Chu, S. A single-molecule study of RNA catalysis and folding. *Science (Washington, DC, U. S.)* **2000**, *288*, 2048−2051.

(30) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien, R. Y. Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. *Nature* **1997**, *388*, 882−887.

(31) Gouridis, G.; Schuurman-Wolters, G. K.; Ploetz, E.; Husada, F.; Vietrov, R.; de Boer, M.; Cordes, T.; Poolman, B. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat. Struct. Mol. Biol.* **2015**, *22*, 57−64.

(32) Weiss, S. Fluorescence spectroscopy of single biomolecules. *Science (Washington, DC, U. S.)* **1999**, *283*, 1676−1683.

(33) Eggeling, C.; Fries, J. R.; Brand, L.; Günther, R.; Seidel, C. A. Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1556−1561.

(34) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3670−3675.

(35) Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M. Identification of single molecules in aqueous solution by time-resolved fluorescence anisotropy. *J. Phys. Chem. A* **1999**, *103*, 331−336.

(36) Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A. M. Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Anal. Chem.* **2006**, *78*, 2039−2050.

(37) Kühnemuth, R.; Seidel, C. A. M. Principles of single molecule multiparameter fluorescence spectroscopy. *Single Mol.* **2001**, *2*, 251−254.

(38) Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. M. Detection of structural dynamics by FRET: A photon distribution and fluorescence lifetime analysis of systems with multiple states. *J. Phys. Chem. B* **2010**, *114*, 7983−7995.

(39) Torres, T.; Levitus, M. Measuring conformational dynamics: A new FCS-FRET approach. *J. Phys. Chem. B* **2007**, *111*, 7392−7400.

(40) Johansson, M.; Brooks, A. J.; Jans, D. A.; Vasudevan, S. G. A small region of the dengue virus-encoded RNA-dependent RNA polymerase, NS5, confers interaction with both the nuclear transport receptor importin-beta and the viral helicase, NS3. *J. Gen. Virol.* **2001**, *82*, 735−745.

(41) Price, E. S.; DeVore, M. S.; Johnson, C. K. Detecting intramolecular dynamics and multiple Förster resonance energy transfer states by fluorescence correlation spectroscopy. *J. Phys. Chem. B* **2010**, *114*, 5895−5902.

(42) Schwille, P.; Meyer-Almes, F. J.; Rigler, R. Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys. J.* **1997**, *72*, 1878−1886.

(43) Behnam, M. A. M.; Nitsche, C.; Vechi, S. M.; Klein, C. D. C-terminal residue optimization and fragment merging: Discovery of a potent peptide-hybrid inhibitor of dengue protease. *ACS Med. Chem. Lett.* **2014**, *5*, 1037−1042.

(44) ATTO-TEC GmbH. Förster-radius R0 of selected ATTO-dye pairs in Å (accessed 2020-06-25).

(45) Koch, A.; Ma, Q.; Frohnapfel, M.; Spomer, L.; Keitel-Anselmino, V.; Gertzen, C.; Gohlke, H.; Seidel, C. A. M. High-precision FRET analysis of the G-protein coupled receptor TGR5 in live cells. *Eur. J. Med.Res.* **2014**, *19*, S12.

Supporting Information

Conformational Dynamics of the Dengue Virus Protease revealed by Fluorescence Correlation and Single-Molecule FRET Studies

Christian Götz[a], *Gerald Hinze*[a], *Andrea Gellert*[b], *Hannah Maus*[b], *Franziska von Hammerstein*[b], *Stefan J. Hammerschmidt*[b], *Luca M. Lauth*[a,§], *Ute A. Hellmich*[a,c,#], *Tanja Schirmeister*[b], *Thomas Basché*[a,*]

[a]Department of Chemistry, Johannes Gutenberg-University Mainz, Germany

[b]Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Germany

[c]Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe-University Frankfurt, Germany

S1 Fluorometric assay

The enzymatic activity of the dye labeled protease was verified with the fluorogenic substrate *boc*-GRR-AMC in presence and absence of competitive or non-competitive inhibitors. The buffer contained Tris (50 mM, pH 9), CHAPS (1 mM) and optionally glycerol (20 vol%). Each measurement was performed with a Tecan Infinite F2000 PRO fluorimeter in duplicates in flat-bottom 96-well microtiter plates from Greiner Bio-One. A total volume of 200 μL, containing 180 μL buffer, 10 μL DMSO as control or inhibitor in DMSO (final concentration 20 μM), 5 μL enzyme solution and 5 μL substrate (final concentration of 100 μM) was used for each well. The inhibitory activity was compared to the results of the wildtype protease.

S2 Spectroscopic characterization of the FRET pair labeled double mutants of DENV-PR

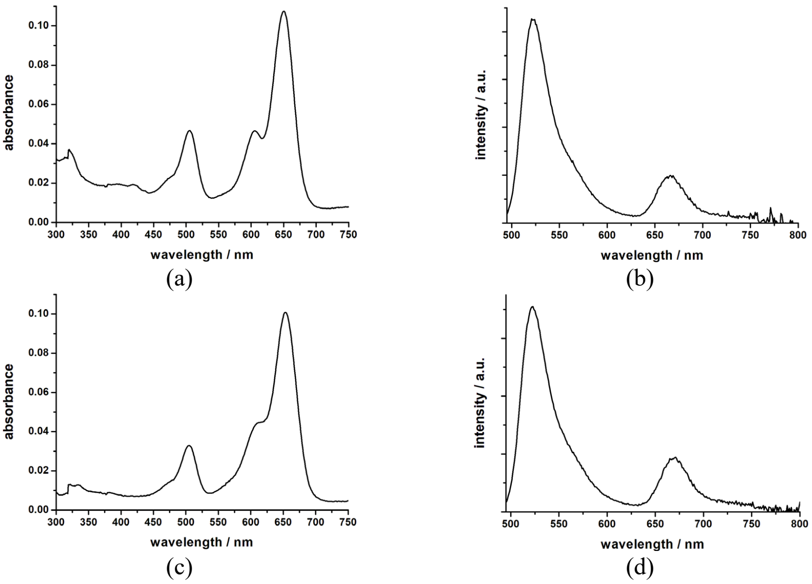
Absorption and fluorescence spectra of the FRET pair labeled double mutants of DENV-PR were recorded in assay buffer (50 mM TRIS-HCl pH 9.0, 1 mM CHAPS) with 10 vol% of DMSO. Notice that for historical reasons, the assay buffer for the measurements on the S79C-G14C double mutant contained 20 vol% of glycerol, while the assay buffer for the measurements on the other two samples contained no glycerol. Table S1 summarizes the concentrations used.

Table S1: Concentrations of the different samples used for ensemble characterization.

sample name	concentration / µM
S79C-G14C	0.82
S79C-S158C A	0.95
S79C-S158C B	0.72

To record the absorption spectra, 600 µL of the solutions were filled in a microcuvette. The absorption spectra were recorded at room temperature using a two-beam spectrophotometer (LAMBDA 850 UV/VIS, PerkinElmer, United States).

To record the fluorescence emission spectra, 50 µL of the solutions were filled in a precision cuvette. The spectra were measured at room temperature in a 90-degree geometry using a spectrofluorometer (Fluorolog3, Horiba Jobin-Yvon, Germany). For each measurement the spectral bandwidth of the excitation and emission monochromator was set to 1 nm. The absorption and emission spectra are shown in Figure S1.



(a) (b)

(c) (d)

Figure S1: Absorption and emission spectra, respectively, of the ATTO 488/ATTO 647 N labeled DENV-PR S79C-S158C A (a, b) and the ATTO 488/Cy5 labeled DENV-PR S79C-S158C B (c, d) double mutants, in assay buffer. Following excitation of the donor (488 nm) emission from both donor and acceptor is observed.

S3 DENV-PR S79C-S158C A: Conformational exchange without and with competitive inhibitor 1

Following a procedure proposed by Torres and Levitus [Torres, T; Levitus, M. J. Phys. Chem. B **2007**, 111, 7392-7400], in Figure S2 the ratios of the acceptor autocorrelation function G_{AA} and the crosscorrelation function G_{DA} are plotted before and after adding **1**. Within a simple two-state-model both correlation functions are composed of a diffusional part T_{ii} and a kinetic contribution X_{ii} depending on the kinetic rates k_{12}, k_{21} and the FRET efficiencies in both states E_1 and E_2, respectively. Assuming similar diffusion contributions for all correlation functions, $T_{AA} = T_{DA}$, the ratio solely depends on conformational exchange,

$$\frac{G_{AA}(\tau)}{G_{DA}(\tau)} = \frac{1 + a_{AA}\, exp(-\tau/\tau_{ex})}{1 - a_{DA}\, exp(-\tau/\tau_{ex})}$$

with $\tau_{ex} = (k_{12} + k_{21})^{-1}$, $a_{AA} = k_{12}k_{21}(E_1 - E_2)^2/(k_{21}E_1 + k_{12}E_2)^2$

and $a_{DA} = k_{12}k_{21}(E_1 - E_2)^2/\left[(k_{21}E_1 + k_{12}E_2)[k_{21}(1 - E_1) + k_{12}(1 - E_2)]\right]$. We have performed a fit to the ratio (in absence of the inhibitor) assuming $E_1 = 0.5$, $E_2 = 1.0$ and a log normal distribution of transition rates. Furthermore, we have accounted for an additional decay in the acceptor auto-correlation caused by triplet kinetics which is absent in the cross-correlation. We point out that our data evaluation is a first and rough approach, since fitting parameters strongly depend on the chosen model and the method relies on the assumption of $T_{AA} = T_{DA}$. Once the diffusion contribution for the donor decays faster than for the acceptor, the diffusion contribution for the cross-correlation decays faster than for the acceptor, too. As a consequence, besides conformational exchange the ratio shown in Figure S2 might contain some contribution from pure diffusion. This could explain the small increase in the calculated ratio in the presence of inhibitor, the amplitude of which, however, is close to the noise level. In principle, similar ratios could be calculated employing the donor autocorrelation function, however, the donor emission signal is strongly superimposed by donor only species. In addition, any scattering and fluorescing contributions from potential contaminations will be more easily seen in the donor channel, rather than in the acceptor channel.

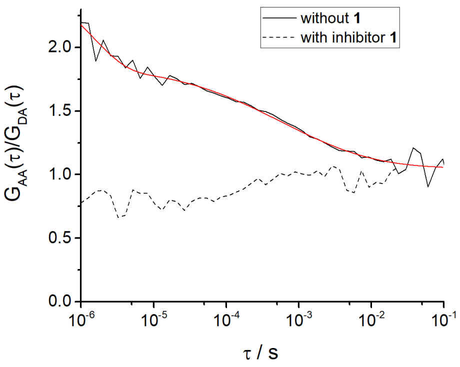


Figure S2: Calculated ratios of the acceptor auto-correlation and the cross-correlation functions without and with inhibitor **1**. The red line represents a fit to the data in the absence of inhibitor with a mean exchange time $\langle\tau_{ex}\rangle \approx 10ms$. The ratio of the mean transition rates $\langle k_{12}\rangle/\langle k_{21}\rangle \approx 4.5$ corresponds to the population ratios of the states. That is, the closed conformation appears to be more populated than the open conformation.

S4 DENV-PR S79C-S158C A: Ensemble fluorescence decays before and after adding the competitive inhibitor 1

Time correlated single photon counting (TCSPC)-experiments at the ensemble level have been performed in solution. In Figure S3 the rise/decay curves of the acceptor emission are plotted after pulsed excitation of the acceptor as well as the donor. In the latter case experiments were performed without and with inhibitor **1**. After donor excitation energy transfer to the acceptor leads to a delayed increase of the rise/decay curves compared to direct excitation of the acceptor. All curves were adjusted to the longtime tail solely determined by the acceptor fluorescence lifetime. As can be seen in Figure S3, without inhibitor the rise/decay curve significantly deviates at short times compared to the case of direct excitation, indicating slow (inefficient) energy transfer. The behavior significantly changes after adding **1**. Now the energy transfer to the acceptor is much more efficient, the longtime tail decay occurring much earlier.

Note that direct excitation resembles donor excitation with complete (instantaneous) energy transfer.

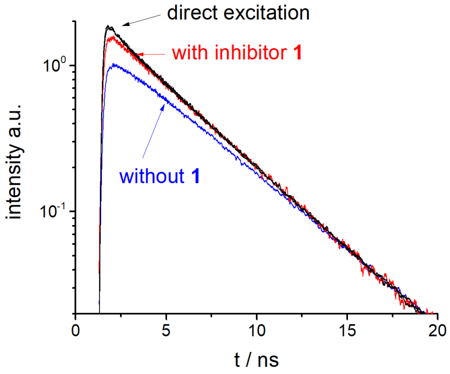


Figure S3: TCSPC-ensemble decay curves of the DENV-PR S79C-S158C A double mutant. For details see text.

S5 DENV-PR S79C-S158C B before and after adding the competitive inhibitor 1

Figure S4 shows the auto- and cross-correlations before and after addition of **1**. In the absence of the inhibitor, the cross-correlation drops to zero slower than the two autocorrelations. This again indicates that an exchange between different conformations, differing in their FRET efficiency, occurs.

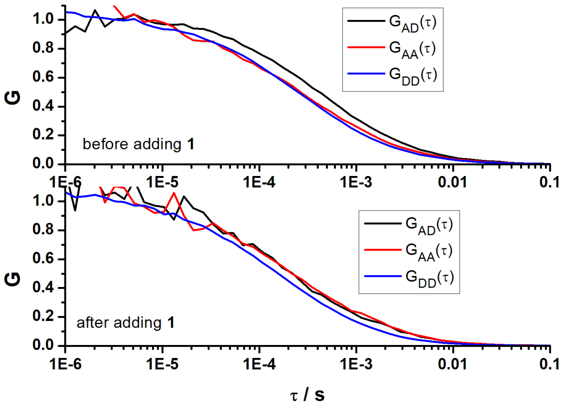


Figure S4: Comparison of the normalized cross-correlation (black curve) with the two normalized auto-correlations (red and blue), each before (upper graph) and after (lower graph) addition of the competitive inhibitor **1** to the ATTO 488/Cy5 FRET pair labeled DENV-PR S79C-S158C double mutant.

Similar to the ATTO 488/ATTO 647N labeled DENV-PR S79C-S158C double mutant (DENV-PR S79C-S158C A) described in the main text, the cross-correlation does not drop to zero slower than the acceptor auto-correlation after addition of **1**. Once again, this can be interpreted as an inhibitor induced stabilization of a conformation.

Figure S5 shows the 2D histograms obtained from the burst analysis of the intensity time traces before (a) and after (b) addition of **1**. Again, the population with a FRET efficiency of zero is due to DENV-PR molecules only labeled with the donor dye ATTO 488 and shows a significantly longer fluorescence lifetime of the donor. As with the previous studied ATTO 488/ATTO 647N FRET pair labeled DENV-PR S79C-S158C double mutant, the point cloud is significantly elongated along the abscissa before the addition of **1**. The data deviate characteristically from the static FRET line and are shifted towards longer donor lifetimes, in accordance with the observations described in the main text. Again, this indicates that the DENV-PR, during the transit time through the observation volume, fluctuates between different conformational substates, which differ in their FRET efficiencies. After addition of **1** (see, Figure S5 (b)), the bursts with medium FRET efficiencies which deviated characteristically from the static FRET line disappear almost completely. Solely the population with high FRET efficiencies survives, in good agreement with the static FRET line. This corresponds to the theoretical expectations, since for this double mutant, the distance between the two labeling positions (S79C and S158C) should decrease as a result of the transition from the open to the closed conformation (see Table 1 in the main text), leading to increased FRET efficiencies. The upper parts of Figure S5 (a) and (b) illustrate the corresponding normalized 1D histograms of FRET efficiencies before and after adding **1**. Almost exclusively one FRET population at high FRET efficiencies is observed after addition of the competitive inhibitor.

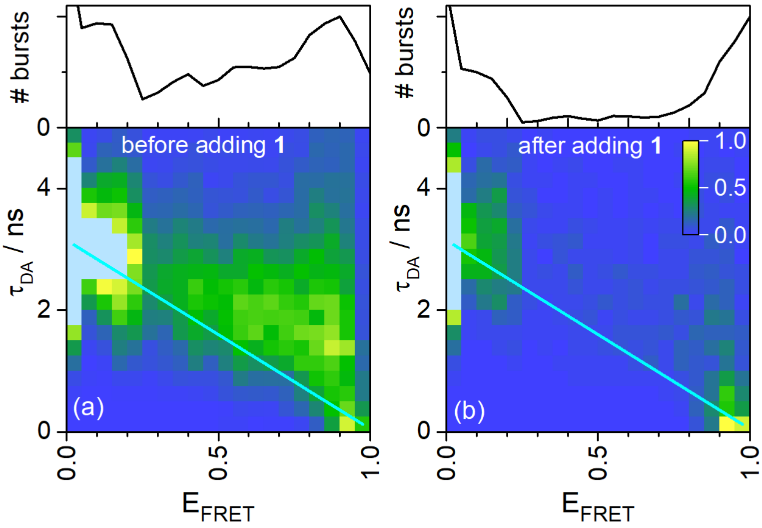


Figure S5: Plot of the normalized occurrences of the individual bursts (individual molecules) within the intensity time trace of the ATTO 488/Cy5 FRET pair labeled DENV-PR S79C-S158C double mutant before (a) and after (b) adding **1** in a 2D histogram, separated according to donor lifetime τ_{DA} and FRET efficiency E_{FRET}. For an easier visual comparison of the FRET populations before and after addition of the inhibitor, the respective maxima were normalized to one. The maxima of the donor-only species were not taken into account in the normalization, which is why values greater than one, shown in gray, occur. The cyan lines represent the static FRET lines. The respective normalized 1D histograms of FRET efficiencies are shown as a projection on the top of the 2D histograms.

“

5.2 A Competition smFRET Assay to Study Ligand-Induced Conformational Changes of the Dengue Virus Protease.

5.2.1 Context, Project Summary, and Own Contributions

NMR spectroscopy and split-luciferase assays have been used to study the conformational dynamics of the DENV protease in solution.[106,112,276] These methods detected multiple conformations, which competitive inhibitors and pH changes could influence.[106,112] However, previous studies had limitations in providing comprehensive information about the influence of allosteric inhibitors. NMR studies indicated that competitive inhibitors stabilize the closed conformation, while the split-luciferase assay showed that allosteric inhibitors do not stabilize this conformation.[106,112] The specific conformation induced by allosteric inhibitors and their effects on the DENV protease remained unclear.

To directly investigate the effect of an allosteric inhibitor on the protease conformation, we conducted smFRET experiments using a double mutant (S79*C-S158C) labeled with ATTO 488 and ATTO 643. Analysis of the fluorescence cross-correlation function in the presence of the allosteric inhibitor indicated stabilization of one conformation, but it was unclear which conformation was stabilized. However, the absence of a shift to higher FRET efficiencies after adding an allosteric inhibitor indicated that the closed conformation was not stabilized. Since the protease was predominantly in the open conformation under our experimental conditions, only stabilizations of the closed conformation could be directly visualized by shifts in E_{FRET}. A competition smFRET assay (Figure 23) addressed this point and visualized the allosteric inhibitor's induced shift to the open conformation. The assay followed a two-step process: first, the addition of a competitive inhibitor (I_c^{DENV}) shifted the equilibrium towards the closed conformation. Subsequently, the allosteric inhibitor (I_a) competed with I_c^{DENV}, causing the equilibrium to move toward the open conformation. The addition of increasing concentrations of I_a while keeping the concentration of I_c^{DENV} constant led to an increase in the fraction of the open conformation and a decrease in the fraction of the closed conformation. This demonstrated that I_a stabilized the open conformation. The binding affinity of I_a was determined by evaluating the ratio between open and closed conformations. The determined K_{d} values agreed with IC_{50} values obtained from a fluorometric assay, and the apparent K_{i} ($K_{\mathrm{i}}^{\mathrm{app}}$) constant determined by a Dixon plot. Kinetic modeling supported the assumption that the presence of the competitive inhibitor did not affect the binding affinity of the allosteric inhibitor. Additionally, inhibitor combination studies following the Chou-Talalay method[277] indicated antagonism between the two inhibitors. This supported our findings in the smFRET experiments, where we observed that the two inhibitors stabilize different conformations.

A comparable strategy has not yet been described and allowed, for the first time, the unambiguous assignment of the conformation of DENV protease stabilized by allosteric inhibitors. This finding is indispensable for rational drug design and the development of potential inhibitors.

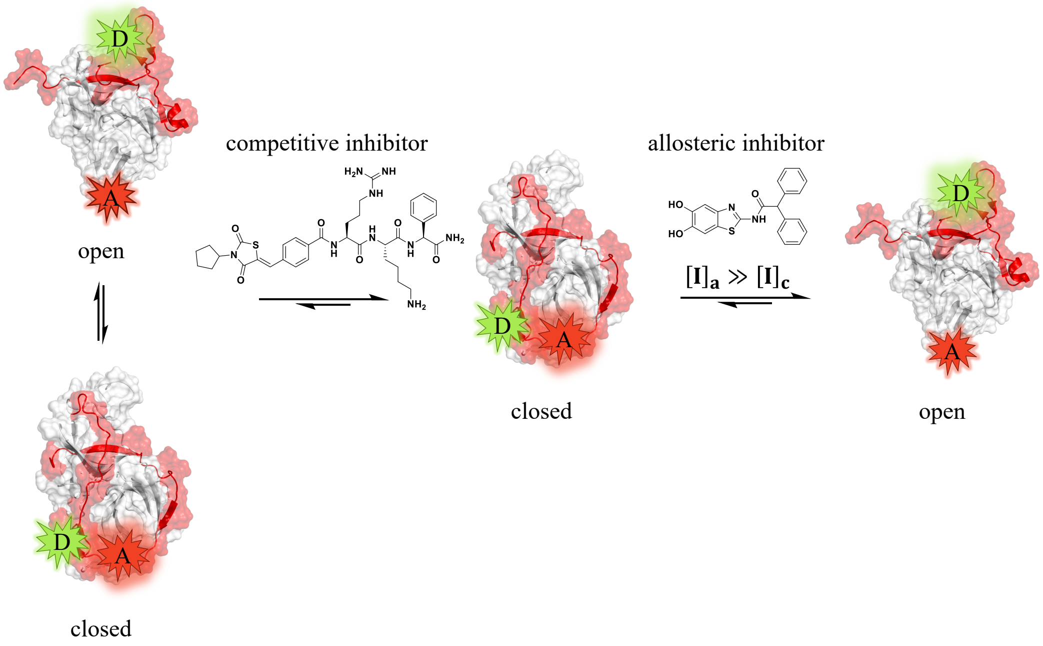


Figure 23. Principle of competition smFRET assay. The equilibrium between open and closed conformation is shifted to the closed conformation by adding a competitive inhibitor. The subsequent addition of an allosteric inhibitor in excess leads to reforming the open conformation. Open conformation (PDB-ID: 2fom), closed conformation (PDB-ID: 2m9p), NS2B is sown in red, NS3 in white.

Own contributions: protein labeling, fluorometric enzyme assays, smFRET data curation, and writing of the first draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, smFRET data analysis, kinetic modeling, and manuscript editing.

This work has been published in Protein Science (impact factor: 6.99).

Article reprinted with permission of *Protein Science* **2023**, 32 (1): e4526 “A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease.” © 2023 WILEY-VCH Verlag GmbH & Co. KGaA (Germany).

5.2.2 Publication

The following publication quoted within “” from page 150 to page 177 is the same as the manuscript cited on page 149.

“

Received: 23 June 2022 | Revised: 7 November 2022 | Accepted: 29 November 2022
DOI: 10.1002/pro.4526

FULL-LENGTH PAPER

THE PROTEIN SOCIETY WILEY

A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease

Hannah Maus[1] | **Gerald Hinze**[2] | **Stefan Josef Hammerschmidt**[1] | **Thomas Basché**[2] | **Tanja Schirmeister**[1]

[1]Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Mainz, Germany
[2]Department of Chemistry, Johannes Gutenberg-University, Mainz, Germany

Correspondence
Thomas Basché, Department of Chemistry, Johannes Gutenberg-University, Mainz, Germany.
Email: thomas.basche@uni-mainz.de

Tanja Schirmeister, Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Mainz, Germany.
Email: schirmei@uni-mainz.de

Funding information
University of Marburg, Grant/Award Number: AY037116

Review Editor: Aitziber L. Cortajarena

Abstract

Ligand binding to proteins often is accompanied by conformational transitions. Here, we describe a competition assay based on single molecule Förster resonance energy transfer (smFRET) to investigate the ligand-induced conformational changes of the dengue virus (DENV) NS2B-NS3 protease, which can adopt at least two different conformations. First, a competitive ligand was used to stabilize the closed conformation of the protease. Subsequent addition of the allosteric inhibitor reduced the fraction of the closed conformation and simultaneously increased the fraction of the open conformation, demonstrating that the allosteric inhibitor stabilizes the open conformation. In addition, the proportions of open and closed conformations at different concentrations of the allosteric inhibitor were used to determine its binding affinity to the protease. The K_D value observed is in accordance with the IC_{50} determined in the fluorometric assay. Our novel approach appears to be a valuable tool to study conformational transitions of other proteases and enzymes.

KEYWORDS
allosteric inhibition, competition assay, conformational change, flavivirus, NS2B-NS3 protease, smFRET

1 | INTRODUCTION

Understanding the operating mechanism of proteins is essential for the development of new effective inhibitors. In many cases, ligand binding leads to changes of the secondary and tertiary structure, as well as the dynamics of the proteins.[1] Various approaches already exist for tracking of ligand-induced conformational changes. For several proteins, known to undergo a conformational transition upon ligand binding (e.g., EF-hand proteins, maltose-binding protein, integrin I domains), residues have been mutated to clearly favor one of the states over the other, enabling investigation of their respective properties.[2–8]

Abbreviations: CI, combination index; DENV, dengue virus; FRET, Förster Resonance energy transfer; E^{*}_{ET}, FRET efficiency; I_a, allosteric inhibitor; IC_{50}, inhibitor concentration at half maximal inhibition; K_D, binding affinity; K_i^{app}, apparent K_i; NS, non-structural; $NS3_{pro}$, NS3 protease; R_0, Förster Radius; smFRET, single molecule FRET; τ, fluorescence lifetime.

However, additional tools are needed to obtain information about the relationship of protein sequence, dynamics, and function. NMR or single molecule Förster resonance energy transfer (smFRET) have proven to be suitable tools for studying population distributions of conformational substrates in solution.[9–16]

The dengue virus (DENV) NS2B-NS3 protease is essential for viral replication of the virus.[17] Besides host proteases, it is responsible for the cleavage of the viral precursor polyprotein, which is translated from the single-stranded RNA genome.[17] The protease thus represents an interesting drug target for countering DENV infections. The NS3 protease (NS3$_{pro}$) is a trypsin-like serine protease with the catalytic triad His51, Asp75, and Ser135.[18] Correct folding and catalytic activity of NS3$_{pro}$ strongly depends on the presence of the cofactor NS2B.[19,20] The latter can adopt at least two different conformations, which mainly differ in their relative position compared to the active center of NS3$_{pro}$. The two conformations are designated as "open" and "closed."[21–25] As already deduced from crystal structures, the binding of ligands to the active site leads to the stabilization of the closed conformation in which NS2B wraps around the NS3$_{pro}$ domain. Thus, the closed conformation is considered the active conformation.[21] In the open conformation, however, NS2B does not contribute to substrate recognition and is rather loosely bound to NS3$_{pro}$.[22,23,26] Although crystal structures provide important and fundamental information about proteins, they only represent energy minimized snapshots and, thus, do only provide very limited information about the state and dynamics of proteins in solution.

By using NMR spectroscopy, the conformational dynamics of the DENV protease in solution was previously investigated.[12,27] Hereby, at least two species were detected, and interpreted as different conformations. The ratio of the two species could be shifted by adding a competitive inhibitor or by changing the pH value of the buffer.[12] The use of a split-luciferase assay also enabled to observe conformational dynamics in solution. In this experiment, the addition of competitive inhibitors indicated the formation of the closed conformation, which was prevented by the addition of allosteric inhibitors.[24] However, the earlier studies have limitations in their information content. In the case of the NMR studies, it was reported that the competitive inhibitors did stabilize the closed conformation and in the case of the split luciferase assay, that the allosteric inhibitors did not stabilize this conformation. Hence, none of these studies unraveled which conformation is induced by allosteric inhibitors so that their effects on the DENV protease are not yet sufficiently understood. However, to design potential inhibitors, understanding the addressed protein is essential. This knowledge includes kinetics of the protein as well as the protein's mode of operation. For rational drug design in particular, the binding pocket with which a ligand interacts must be known. Accordingly, it is of great importance to know the conformation that is stabilized by the ligand. Since there is no direct proof that allosteric inhibitors do stabilize the open conformation, the impact of an allosteric inhibitor on the conformation of the protease will be in the focus of the present work. smFRET enables to observe conformational subpopulations, conformational transitions and temporal fluctuations that typically remain elusive in ensemble measurements.[28,29] Using this method, it was already shown that the DENV protease in solution is in an equilibrium between two conformations and that the presence of a competitive inhibitor stabilizes the closed conformation.[16]

Here, we provide evidence that the allosteric inhibitor stabilizes the open conformation of the protease by applying a competition smFRET assay. Our approach is based on the initial shift of the equilibrium between open and closed conformation toward the closed conformation by the addition of a competitive inhibitor I_c (Figure 1),[30] following work by Götz et al.[16] The subsequent addition of an allosteric inhibitor I_a (Figure 1)[31] leads to a competition between the two inhibitors in a way that the

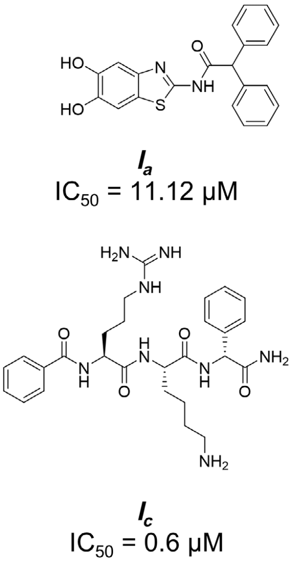


FIGURE 1 Structures and IC50 values of the allosteric (I_a)[31] and the competitive (I_C)[30] inhibitor

conformational equilibrium is shifted to the open conformation with increasing concentration of the inhibitor I_a. Under the experimental conditions, the concentration of I_c remains constant. For sufficiently high concentrations of I_a, an almost complete shift of the equilibrium towards the open conformation was achieved. Moreover, by evaluating the ratio between open and closed conformations the binding affinity of the allosteric inhibitor was determined. To the best of our knowledge, an assay as presented here has not been described in the literature before.

2 | RESULTS AND DISCUSSION

2.1 | FRET pair labeled DENV protease mutants

Since the DENV protease does not contain native cysteines, single cysteines can be introduced by site-directed mutagenesis and subsequently specifically labeled with dye molecules. As previously described,[16] the positions of the cysteines were chosen in a way that one of them is localized in the NS3$_{\text{pro}}$ domain (S158C) and the other one in the NS2B cofactor (S79C) (Figure 2).

The positions were selected to show significant distance changes between the mutated positions upon conformational changes.[16] Distances of the dyes ATTO 488 and ATTO 643 in the open and closed conformation were calculated using a toolkit from Seidel's laboratory.[32] Since the structure of ATTO 643 was not accessible, the structure of the related ATTO 647N was used for this calculation.[33] In each case, an ellipsoid was defined around the chromophore whose radii were determined using the python algorithm mol-ellipsize.[34] For this purpose, the lowest-energy conformers of the dyes including the linker were calculated using RDKit ETKDG (universal force field)[35] and the centers of mass of the chromophores as well as the linker lengths were computed using the custom PyMol center of mass plugin.[36] The resulting radii and linker lengths (Figure S1) were then used to estimate the mean distances between the dyes (Table 1). According to the data in Table 1, the distance in the open conformation is close to the Förster radius $R_0 = 5.2$ nm of the dye pair calculated by the same program.

Since the two cysteines were statistically labeled with ATTO 488 and ATTO 643, both equipped with maleimide linkers, a mixture of donor-only, acceptor-only, and donor-acceptor labeled proteases was obtained. However, in the smFRET experiments, exclusively molecules carrying one donor and one acceptor dye are informative and were considered for data analysis of fluorescence measurements. The activity of the labeled proteases was confirmed in a fluorometric assay by using a fluorogenic peptide substrate (Boc-GRR-AMC). The increase in fluorescence intensity as a function of time was used as a measure of activity. The proteolytic cleavage of the substrate releases AMC, whose fluorescence was measured at 460 nm. The unchanged turnover rate of the fluorogenic substrate of the S79C-S158C mutant and the dye-labeled mutant compared to the wild type shows that neither the mutations nor the dye labeling had a negative effect on the activity of the enzyme (Figure S3).

To determine the dye-to-protein ratio, the absorbance was measured at 280 nm (protein), 500 nm (ATTO 488), and 630 nm (ATTO 643). Degrees of labeling were determined to be 59% for ATTO 488 and 74% for ATTO 643. Binding of the dyes to the protease was examined using an SDS-PAGE gel. The scan at excitation wavelengths of 500 nm (iii, Figure 3b) and 630 nm (iv, Figure 3b), respectively, gave rise to a band at about 35 kDa in each

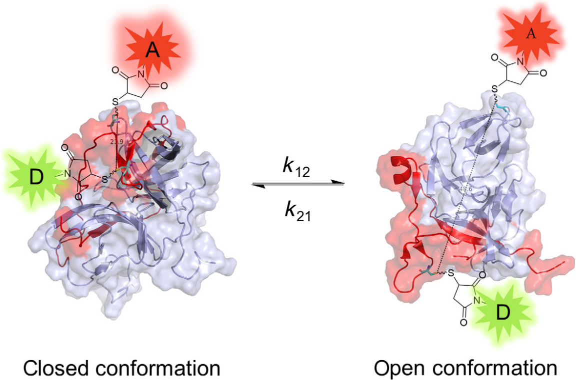


FIGURE 2 FRET pair labeled S79C-S158C double mutant of the DENV protease in the open and closed conformation. In the closed conformation (left), the distance between the dye labels is smaller than in the open conformation (right). DENV, dengue virus; FRET, Förster resonance energy transfer

TABLE 1 Mean distances of the dye pairs in the NS2B-NS3 protease in the open and closed conformation calculated by the toolkit from Seidel's laboratory[32]

	Dye label NS2B	Dye label NS3$_{pro}$	Mean distance (nm)
Open (2FOM)	ATTO 488-Maleimide	ATTO 647N-Maleimide	5.9 ± 0.5
	ATTO 647N-Maleimide	ATTO 488-Maleimide	5.8 ± 0.5
Closed (2M9P)	ATTO 488-Maleimide	ATTO 647 N-Maleimide	2.9 ± 0.7
	ATTO 647N-Maleimide	ATTO 488-Maleimide	3.0 ± 0.6

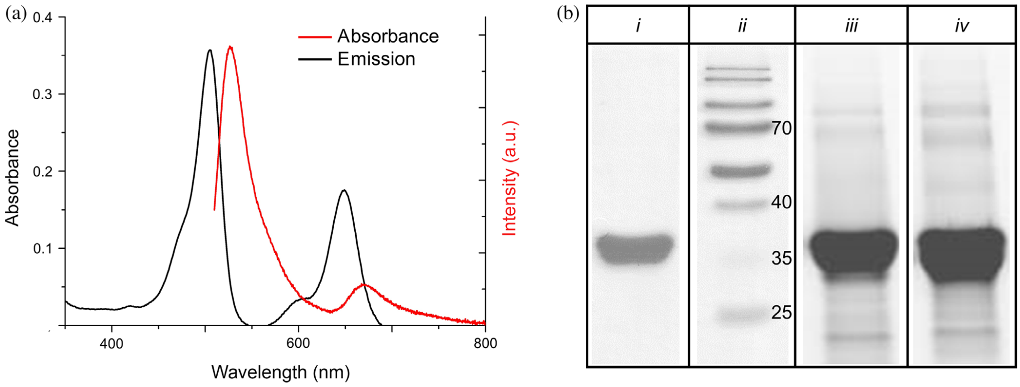


FIGURE 3 Ensemble spectra of the S79C-S158C DENV protease double mutant labeled with the ATTO 488/ATTO 643 FRET pair in buffer containing 10 vol% of DMSO and SDS PAGE gel of the labeled protease. (a) Absorption (black) and emission (red) spectra. The latter was measured at 500 nm excitation. (b) SDS-PAGE gel of ATTO 488/ATTO 643 FRET pair labeled DENV protease. (i) Coomassie stain, (ii) marker PageRuler™ Prestained protein ladder (in kDa), (iii) laser scan with excitation at 500 nm, (iv) laser scan with excitation at 630 nm. DENV, dengue virus

case, in accordance with the protein band visualized by coomassie staining (i, Figure 3b).

Absorption and fluorescence emission spectra of the ATTO 488/ATTO 643 FRET pair labeled S79C-S158C double mutant of the DENV protease recorded in buffer (50 mM TRIS–HCl pH 9.0, 1 mM CHAPS) with 10 vol% of DMSO are shown in Figure 3a. These conditions correspond to the conditions during smFRET measurements. The buffer was adapted to the assay conditions known from the literature,[31,37] including the DMSO added to ensure the solubility of the inhibitors. Under these conditions, the quantum yields of the dyes were also determined. High quantum yields of over 80% were obtained for both dyes, showing that the experimental conditions are favorable for both the protease and the dyes.

2.2 | S79C-S158C DENV NS2B-NS3 ATTO 488/ATTO 643 ± allosteric inhibitor I_a

In smFRET experiments in solution, individual photon bursts from dye labeled proteins diffusing through the excitation volume are monitored and analyzed to identify subpopulations of a typically heterogeneous ensemble.[38,39] Ideally, all intrinsic parameters of the fluorophore are detected simultaneously for each burst.[40,41] These parameters include the fluorescence wavelength,[42] intensity and lifetime.[39] Here, the emission wavelengths of individual bursts were not determined, however, a separation in terms of donor or acceptor fluorescence, respectively, was implemented.

To increase the stability of the allosteric inhibitor over the duration of the measurements, 1 mM TCEP was added to the buffer, in deviation from the previously used assay conditions.[31] The double mutant labeled with ATTO 488 and ATTO 643 was excited at 502 nm and the stream of emitted photons was recorded with and without the allosteric inhibitor at a concentration of 200 μM of the latter.[31]

For further evaluation, fluorescence intensity time traces with a binning time of 1 ms were calculated from the fluorescence bursts, considering only bursts which showed a total of at least 20 counts on both APDs. Sections of the intensity time traces of ATTO 488/ATTO

643 labeled S79C-S158C DENV protease with and without inhibitor are presented in Figure 4.

Auto- and cross-correlations were then calculated from the collected data. The cross-correlation G_{AD} after addition of the allosteric inhibitor (I_a) drops faster to zero than G_{AD} without I_a (Figure 5). This observation indicates the stabilization of one conformation by addition of I_a. It is noteworthy that the cross-correlation G_{AD} consists of a diffusion term and in case of fluctuating FRET dynamics an additional rise term, describing the exchange rates between different FRET states. Thus, fluctuating FRET dynamics effectively delay the decay of the cross-correlation G_{AD} compared to pure diffusion. Accordingly, the observation of a faster decay indicates less FRET dynamics or the stabilization of one conformation by addition of I_a. Neither triplet kinetics[43] nor quantum yields depending on I_a should affect the normalized cross-correlation.

As seen in Figure 5, the effect is rather moderate, and no conclusion can be drawn from the cross-correlation regarding which conformation is stabilized. Under the assumption that I_a stabilizes a conformation of the protease for at least the period of observation, a difference function (ΔG_{AD}) can be calculated from the correlation functions according to Torres et al.[44] The exchange correlation time ($(k_{12}+k_{21})^{-1}$) was estimated to be ~32 µs from the corresponding difference function (Appendix S1).

A burst analysis was performed, resulting in 2D histograms (Figure 6). Before (Figure 6a) and after addition (Figure 6b) of I_a, distinct point clouds can be seen in the 2D histograms. However, no shift of the E^*_{ET} frequency distributions was observed.

While from both, donor and acceptor, 2D histograms are available, point clouds at $E^*_{ET} > 0.5$ are better visible in the acceptor 2D histograms. Moreover, with increasing number of acceptor photons within a burst, the statistical significance of the calculated fluorescence lifetimes

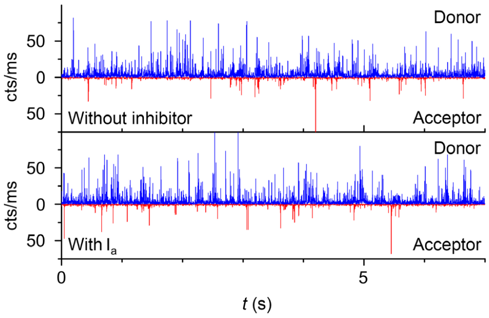


FIGURE 4 Section of the fluorescence intensity time trace of ATTO 488/ATTO 643 FRET pair labeled DENV protease with and without 200 µM of the allosteric inhibitor. The sample was excited with excitation pulses at 502 nm. The binning time was 1 ms. the donor channel is shown in blue, the acceptor channel is shown in red. DENV, dengue virus

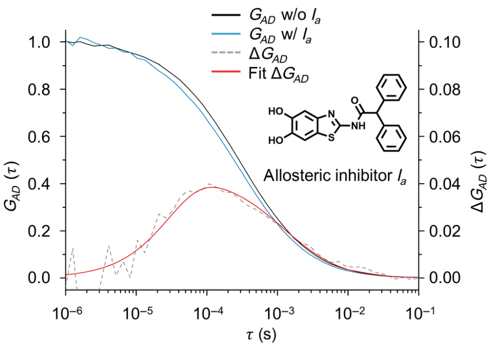


FIGURE 5 Cross-correlations (GAD) of the ATTO 488 and ATTO 643 labeled DENV protease S79C-S158C double mutant before (black) and after addition (blue) of 200 µM of the allosteric inhibitor (I_a).[31] The difference function ΔGAD of the cross-correlations is shown as a dashed line, the corresponding fit in red. DENV, dengue virus

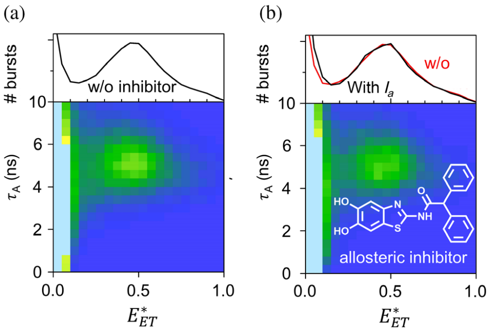


FIGURE 6 Plot of the normalized occurrences of the individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 643 FRET pair labeled DENV protease in a 2D histogram, separated according to acceptor lifetime τ_A and FRET efficiency E^*_{ET}. (a) without allosteric inhibitor. (b) Black: With allosteric inhibitor ($I_a = 200\,\mu\text{M}$), red: Without allosteric inhibitor. For an easier visual comparison of the FRET populations before and after addition of the inhibitor, the respective maxima were normalized to one. The respective normalized 1D histograms of FRET efficiencies are shown as a projection on the top of the 2D histograms. DENV, dengue virus; FRET, Förster resonance energy transfer

increases. Finally, counts in the acceptor channel mainly originate from FRET pair labeled protease, while in the donor channel donor–donor labeled protease and Raman scattering contribute as well. Therefore, we omit the data from the donor channel and will focus only on the acceptor 2D histograms.

In the 2D histogram, the acceptor lifetime τ_A, which was obtained as the mean delay time of the fluorescence of the acceptor dye without taking the instrumental response function (IRF) into account, is plotted against E^*_{ET} from the intensity ratios per burst. A change of the energy transfer efficiency leads to a shift of the point cloud, respectively the formation of an additional point cloud, which differs from the first one in its position of E^*_{ET} as well as of τ_A. To make even small changes visible, it is useful to examine the 2D diagrams in addition to the 1D histograms of the E^*_{ET} frequency distributions.

As seen in Figure 6, the point cloud is already located at intermediate FRET efficiencies before the addition of the allosteric inhibitor. In line with the observations of Zhu et al., who reported the prevalence of the open conformation of the protease at high pH values,[12] we conclude that the intermediate transfer efficiency of $E^*_{ET} \sim 0.5$ represents the open conformation. Accordingly, at this point the absence of a shift to higher FRET efficiencies in the presence of the allosteric inhibitor only indicates that the inhibitor does not stabilize the closed conformation. To reveal a potential shift of the conformational equilibrium of the protease in solution towards the open conformation by the addition of an allosteric inhibitor, the former approach had to be extended and a competition assay was designed.

2.3 | Competition assay

The conformation of the protease in solution depends on the pH value in the manner that the equilibrium of the two conformations is on the side of the open conformation at high pH values.[12] As a consequence, in our experiments performed at pH 9.0, only stimuli that lead to the stabilization of the closed conformation can be directly investigated. Note that the experimental conditions were adapted to the conditions during the fluorometric assay (50 mM Tris, 1 mM CHAPS, pH 9.0).[31,37]

It is assumed that the addition of allosteric inhibitors leads to the stabilization of the open conformation.[24] Therefore, to investigate the influence of an allosteric inhibitor on the conformation of the protease, a competition assay was designed.

To implement this approach, the stabilization of the closed conformation by a competitive inhibitor[30] was exploited to obtain a significant fraction of proteases in the closed conformation prior to addition of an allosteric inhibitor. Only then a shift of the equilibrium towards the open conformation could be observed. For this purpose, we first investigated whether different concentrations of the competitive inhibitor resulted in varying populations of the closed conformation. Indeed, as seen in Figure 7 the E^*_{ET} frequency distributions support the growth of the closed conformation with increasing concentration of the competitive inhibitor.

For the competition assay, a concentration of 2 μM of the competitive inhibitor was used. At this concentration, two distinct populations of protease can be seen in the frequency distribution of E^*_{ET} (Figure 7). Experiments were performed with different concentrations of the allosteric inhibitor (0.1–167 μM) while keeping the concentration of the competitive inhibitor constant. The collected fluorescence data were then subjected to burst analysis.

The E^*_{ET} frequency distributions show two distinct populations (Figure 8a,c) that can be interpreted as the open ($E^*_{ET} \sim 0.4-0.6$) and closed ($E^*_{ET} \sim 0.7-1.0$) conformation. As the concentration of I_a did increase, the fraction of the population of the closed conformation was reduced while the fraction of the open conformation was increased. This is a clear indication that the allosteric inhibitor stabilizes the open conformation of the protease. To determine the binding affinity, the ratio between open and closed conformations was determined using two different analysis methods. First, this ratio was derived from the cumulative integrals of the E^*_{ET} frequency distributions (Figure 8b). After normalization in the range of $E^*_{ET} \sim 0.4-0.6$, the cumulative integral P_i

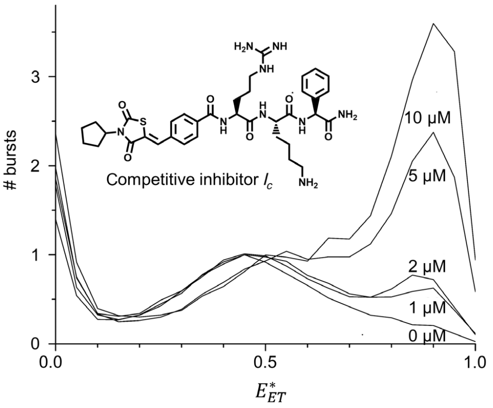


FIGURE 7 E^*_{ET} frequency distributions of ATTO 488/ATTO 643 FRET pair labeled DENV protease with different concentrations of the competitive inhibitor.[30] DENV, dengue virus; FRET, Förster resonance energy transfer

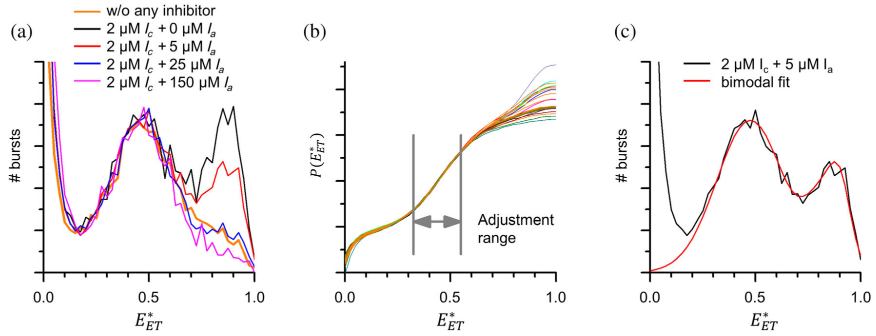


FIGURE 8 E^*_{ET} frequency distributions of the competition smFRET assay. (a) E^*_{ET} frequency distributions of ATTO 488/ATTO 643 FRET pair labeled DENV protease with different concentrations of the allosteric inhibitor (0–150 μM) in the presence of 2 μM of the competitive inhibitor. (b) Cumulative integrals of E^*_{ET} frequency distributions of ATTO 488/ATTO 643 FRET pair labeled DENV protease with different concentrations of the allosteric inhibitor (0–167 μM) in the presence of 2 μM of the competitive inhibitor normalized in the range of $E^*_{ET} \sim 0.4$–0.6. (c) Exemplary bimodal fit for the E^*_{ET} frequency distribution of ATTO 488/ATTO 643 FRET pair labeled DENV protease with 2 μM of the competitive inhibitor and 5 μM of the allosteric inhibitor. DENV, dengue virus; smFRET, single molecule Förster resonance energy transfer

was determined at $E^*_{ET} = 1.0$ and plotted against the concentration of the inhibitor resulting in a sigmoidal affinity curve (Figure 9). The binding affinity K_D = 15.9 ± 2.5 μM of the inhibitor was obtained from the inflection point of the sigmoidal curve. For the second analysis method, the E^*_{ET} frequency distributions were simply approximated by a bimodal fit composed of three Gaussians which turned out to satisfactorily reproduce the data. (Figure 8c) The ratio between open and closed conformation was then determined as the ratio of the corresponding amplitudes. The amplitude ratios were again plotted against the inhibitor concentration, yielding a K_D of 11.7 ± 2.2 μM (Figure S7).

The binding affinities obtained with these two methods are very similar and both are within the same range as the IC_{50} determined in the fluorometric assay under the same buffer conditions (IC_{50} = 12.8 ± 0.8 μM, Figure S4) and the apparent K_i (K_i^{app}) obtained from the Dixon plot (K_i^{app} = 14.7 ± 3.7 μM, Figure S8).

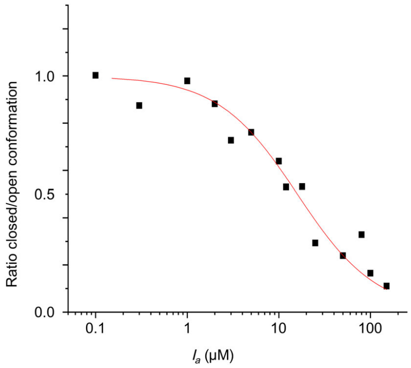


FIGURE 9 Binding affinity curve determined from the competitive smFRET assay based on the evaluation of the cumulative integrals of the E^*_{ET} frequency distributions. The sigmoidal fit provides a binding affinity of the allosteric inhibitor of K_D = 15.9 ± 2.5 μM. smFRET, single molecule Förster resonance energy transfer

2.4 | Inhibitor combination studies

Using the Chou-Talalay method, the impact of a combination of inhibitors was investigated.[45] IC_{50} values of the competitive and allosteric inhibitor were determined separately and in combination using a fluorometric assay (Figures S4–S6). Both compounds were subjected to assays at 7 concentrations at least, starting from the minimum dose required for the enzyme inhibition to that necessary to fully suppress protease activity. IC_{50} values have been obtained from dose response curves (Figures S4–S6). According to their IC_{50} values, I_c and I_a were used in a 1:2 ratio in the combination study. The combination index (CI) calculated using the formula for mutually non-exclusive inhibitors reads:

$$\mathrm{CI} = \left[\frac{(D)_a}{(D_i)_a}\right] + \left[\frac{(D)_c}{(D_i)_c}\right] + \frac{\left[(D)_a(D)_c\right]}{\left[(D_i)_a(D_i)_c\right]}$$

where $(D_i)_a$ and $(D_i)_c$ represent the concentrations of I_a and I_c alone, that are necessary to produce i% of protease

TABLE 2 CI for an inhibition of i% calculated with the Chou-Talalay method for mutually non-exclusive inhibitors

i	99%	90%	80%	70%	60%
CI	1.74	1.73	1.67	1.66	1.66
i	50%	40%	30%	20%	10%
CI	1.67	1.70	1.74	1.81	1.96

Abbreviation: CI, combination index.

inhibition, while $(D)_a$ and $(D)_c$ are the concentrations of both compounds able to produce i% of protease inhibition when they are used in combination. From the value of CI, we can infer how the combination of inhibitors affects their inhibition. When CI <1, synergism is indicated, for CI = 1 summation is indicated and CI >1 indicates antagonism. CI values were calculated for i = 10, 20, 30, 40, 50, 60, 70, 80, 90, and 99% inhibition of the protease (Table 2).

The calculated CI values, which are consistently greater than 1 for inhibitions between 10 and 99%, indicate antagonism of the two inhibitors. Since we cannot be sure whether the inhibitors are mutually exclusive or non-exclusive, the CI value was also calculated for exclusive inhibitors (data not shown). Here, the last summand of the above formula is not taken into account. The corresponding CI values are also higher than 1 and thus indicate antagonism. These results support our findings in the smFRET experiments, where we observed that the two inhibitors stabilize different conformations.

2.5 | Kinetic modeling

From the experiments described above, it was concluded that the open conformation is stabilized by binding of the allosteric inhibitor and that the inhibitors show antagonistic effects. To address the impact of the bound competitive inhibitor onto the binding of the allosteric inhibitor under the experimental setup, kinetic modeling was performed. Since no statement can be made about which conformation the inhibitor binds to, two different models were considered. In the following, model 1, (model 2, see Appendix S1) which assumes that the two inhibitors bind exclusively to the conformation that they stabilize (conformational selection), will be discussed in more detail. Based on this assumption, the following equilibria, with I_c as the competitive inhibitor, I_a as the allosteric inhibitor, E_c as the closed conformation of the protease, E_o as the open conformation of the protease, E_cI_c as the complex of closed conformation and competitive inhibitor, and E_oI_a as the complex of open conformation and allosteric inhibitor, can be established:

$$\begin{tikzcd}
\mathrm{E_cI_c} \arrow[r, harpoon, shift left, "K_{\mathrm{i,c}}"] & \mathrm{I_c} \arrow[l, harpoon, shift left] & + & \mathrm{E_c} \arrow[d, shift left, "K_{\mathrm{c}}"] \\
\mathrm{E_oI_a} \arrow[r, harpoon, shift left, "K_{\mathrm{i,a}}"] & \mathrm{E_o} \arrow[l, harpoon, shift left] & + & \mathrm{I_a} \arrow[u, shift left]
\end{tikzcd}$$

Thus, the following equilibrium constants can be formulated:

$$K_{i,c}=\frac{[E_c][I_c]}{[E_cI_c]},K_{i,a}=\frac{[E_o][I_a]}{[E_oI_a]},\text{and}\,K_c=\frac{[E_o]}{[E_c]}.$$

Using the equilibrium constants $K_{i,c}$, $K_{i,a}$, and K_c and the initial concentrations of the ligands $\left[I_a^0\right]$ and $\left[I_c^0\right]$, the total concentration of open $\left[E_o^*\right]$ and closed $\left[E_c^*\right]$ conformations of the protease present during the measurement can be described. Since the concentration of the protease used in the experiments is significantly smaller than the equilibrium constants $K_{i,c}$ and $K_{i,a}$, the following approximations can be made:

$$\left[E_c^*\right]\cong[E_c]\left(1+\frac{\left[I_c^0\right]}{K_{i,c}}\right)\text{ and }\left[E_o^*\right]\cong[E_o]\left(1+\frac{\left[I_a^0\right]}{K_{i,a}}\right).$$

The relationship between the closed and the open conformation can be described by the following term

$$\frac{\left[E_c^*\right]}{\left[E_o^*\right]}=F\frac{1}{\left(1+\frac{\left[I_a^0\right]}{K_{i,a}}\right)},$$

where $F=K_c^{-1}\left(1+\frac{\left[I_c^0\right]}{K_{i,c}}\right)$ remains constant during the addition of the allosteric inhibitor. The resulting titration curve $T\left(\left[I_a^0\right]\right)$ runs independently of the concentration of the competitive inhibitor from $T(0)=F$ over $T\left(\left[I_a^0\right]=K_{i,a}\right)=\frac{1}{2}F$ to $T(\infty)=0$. As such, it qualitatively reproduces the binding affinity curve displayed in Figure 9. The second binding model (induced fit) provides similar results which are discussed in detail in Appendix S1.

Remarkably, both models showed that the binding affinity of the allosteric inhibitor is not influenced by the presence of the competitive inhibitor under the experimental conditions, although antagonism of the inhibitors was indicated in the combination study. Which of the two models is more applicable to the binding of the ligand cannot be judged from the results described herein.

The agreement between the results from the fluorometric assay and the Dixon plot on one side and the smFRET competition assay on the other side indicates that the assay is suitable for the direct determination of binding affinities. It should also be mentioned that the inhibitory activity of I_a in the fluorometric assay is determined in the presence of a substrate, which is believed to stabilize the closed conformation.

3 | CONCLUSIONS

The existence of at least two different conformations of the DENV NS2B-NS3 protease, designated open and closed, has been reported by various groups. While the impact of competitive inhibitors on the protease has already been sufficiently investigated, the effect of allosteric inhibitors remains elusive. Various approaches showed that competitive inhibitors stabilize the closed conformation of the DENV protease. In contrast, the impact of allosteric inhibitors on the conformation of the protease has been inferred exclusively from the absence of signals, which was interpreted as stabilization of the open conformation.

To investigate the effect of an allosteric inhibitor on the conformation of DENV protease, we statistically labeled a double mutant (S79C-S158C) of the protease with ATTO 488 and ATTO 643 and then performed smFRET experiments. The positions of the point mutations were chosen in such a way that their distance changes significantly with the conformation. Analysis of the fluorescence cross-correlation function of the labeled double mutant in the absence and presence of the allosteric inhibitor indicated stabilization of one conformation, but it could not be decided which of the two conformations was stabilized. Since, under our experimental conditions, the protease is mostly present in the open conformation, only stabilizations of the closed conformation could be visualized by smFRET in the presence of a competitive inhibitor. The absence of a shift to higher FRET efficiencies after addition of an allosteric inhibitor thus only indicated that the closed conformation is not stabilized.

To directly visualize a shift to the open conformation induced by an allosteric inhibitor, a competition assay was designed. This assay is based on an initial shift of the conformational equilibrium toward the closed conformation by a competitive inhibitor. This shift was noticed in the E^*_{ET} frequency distributions by the appearance of a second peak ($E_{ET} \sim 0.6 - 1.0$), clearly separated from the first peak ($E_{ET} \sim 0.4 - 0.6$) originating from the open conformation. In the presence of a constant concentration of the competitive inhibitor, a titration with an allosteric inhibitor was then performed. It was found that with increasing concentration of the allosteric inhibitor, the fraction of the open conformation of the protease in solution increased, while the fraction of the closed conformation was reduced. Accordingly, we could demonstrate that the open conformation is stabilized by the allosteric inhibitor.

The relationship between open and closed conformation was accessed in two ways. First, the ratio of the peak areas of the E^*_{ET} frequency distributions was considered. For this purpose, cumulative integrals were calculated. Secondly, the areas of the two peaks were put into relation. In this case, the E^*_{ET} frequency distributions were approximated by appropriate functions. The outcomes of both methods, when plotted against the concentration of the allosteric inhibitor, allowed the determination of the binding affinity. The K_D values are in good agreement with the IC_{50} value determined by a fluorometric assay or the $K_\mathrm{i}^{\mathrm{app}}$, constant determined by a Dixon plot. The similarity of these values seems to indicate that equilibrium kinetics sufficiently describes the interplay of both inhibitors and the protease. This assumption was justified by kinetic modeling which showed that the presence of the competitive inhibitor has no effect on the binding affinity of the allosteric inhibitor.

The competition smFRET assay developed in this work appears as a promising tool to study ligand-protein interactions. As such, the methodology can be transferred to other enzymes that undergo conformational changes upon binding of ligands. An obvious case would be the structurally similar NS2B-NS3 protease of the Zika virus. Moreover, transfer to unrelated proteins is also conceivable.

4 | MATERIAL AND METHODS

4.1 | Protein constructs and cloning

Protein constructs were used and cloning was conducted as described by Millies et al.[37] In short, site directed mutagenesis was performed using the Kapa HiFi PCR kit (KapaBiosystems Inc., Woburn, MA) to generate the S79C and S158C mutations in a pET15b vector harboring the DENV *NS2B$_{cf}$-NS3$_{pro}$* gene (GenBank ID: AY037116.1). A detailed description of protein constructs and cloning can be found in Appendix S1.

4.2 | Protein expression and purification

Protein expression and purification were performed as described by Millies et al.[37] Briefly, competent *Escherichia coli* BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA) were grown in LB medium and expressed protein after induction with IPTG for 16 hr at 20°C. After harvesting and lysis of the cell pellets the resulting supernatant was subjected to an immobilized metal affinity chromatography (IMAC) on a HisTrap HP 5 ml column (GE Healthcare, Chicago, IL) and eluted in a linear gradient of buffer containing raising imidazole concentrations. Eluted fractions were further purified by size exclusion chromatography (SEC). A detailed description of protein expression and purification can be found in Appendix S1.

4.3 | Fluorescence labeling of the DENV-Protease double mutant S79C-S158C

The fluorescence labeling was performed analogously to Götz et al. using ATTO 643 instead of ATTO 647N.[16] In brief, a buffer containing TCEP was used to ensure free thiol groups within the cysteine double mutants. After an incubation time of 30 min at 4°C, TCEP was removed by rebuffering using spin concentrators. To the rebuffered protein solution were added a 2.1-fold excess of ATTO 488 and a 2.6-fold excess of ATTO 643, both dissolved in DMF. Dye labeling occurred during a 2 hr incubation time at room temperature. Labeled protease was then dialyzed and purified by SEC. For more details in Fluorescent Labeling, see Appendix S1.

4.4 | smFRET experiments

The smFRET experiments were performed analogously to the literature.[16] They were conducted in self-constructed sample cells made of a poly(ethylene glycol) coated glass coverslip and a glued-on plastic cylinder. Sample cells were loaded with 150 μl of FRET pair labeled double mutant of the protease ($c \sim 100$ pM) in buffer containing 10 vol% DMSO. Fluorescence photons were collected with a custom-built confocal microscope over a period of 1800 s. Measurements with inhibitor were performed after those without inhibitor, while the concentration of the protease and buffer conditions were kept constant. Excitation of the fluorophores was performed with a spectrally filtered output from a pulsed white light fiber laser (10 MHz, SC OEM, YSL Photonics, China). Excitation pulses were centered around 502 nm by using an acousto-optical tunable filter (AOTF-VIS-DR, Fianium, UK). Excitation and emission were separated by a dichroic mirror (ZT491 rdcxrxt-UF1, CHROMA, Bellows Falls, VT). Emitted fluorescence light was spectrally separated by a dichroic mirror (ZT640-rdc-UF1, CHROMA, Bellows Falls, VT) into red light and light of higher energy. Both beams were then focused onto two APDs (acceptor channel A: SPCM-AQRH-15, PerkinElmer, Waltham, MA; donor channel D: PDM 50ct, MPD, Italy). The absolute and relative arrival times (relative to excitation pulse) of the individual photons were detected by a HydraHarp 400 module (PicoQuant, Germany) which was connected to the two detector APDs. Further details on smFRET experiments can be found in Appendix S1.

4.5 | Analysis of smFRET data

The data were analyzed as described by Götz et al.[16] Cross-correlation functions $G_{\mathrm{AD}}(\tau)$ (FCS-FRET) were calculated from the arrival times of the individual photons. Individual bursts extracted from the fluorescence intensity time traces (bin time = 1 ms) were analyzed. Bursts with at least 20 counts on both APDs were considered for further evaluation. Average arrival times τ_A of the acceptor photons relative to the excitation pulse (without taking the IRF into account) were calculated for each individual burst. The IRF was recorded at the beginning of each measurement day. The full width at half-maximum height (FWHM) was about 770 ps. The FRET efficiency E^*_{ET} was calculated for each burst:

$$E^*_{ET} = \frac{\mathrm{counts}(A)}{\mathrm{counts}(A) + \mathrm{counts}(D)}.$$

The individual bursts were sorted with respect to their E^*_{ET} values and presented in frequency distributions. Details on the analysis of smFRET data can be found in Appendix S1.

4.6 | Analysis of competition assay

1. The cumulative integrals of the distributions of the FRET efficiencies $P(E_{ET})$ were calculated and normalized to each other in the range of the mean FRET efficiency ($E_{ET} \sim 0.4 - 0.6$). For this purpose, normalization factors f_i were determined so that the maxima of $V_{i,norm}(E_{ET}) = f_i \cdot V_i(E_{ET})$ overlap.

$$P_i(E_{ET}) = \int_0^{E_{ET}} V_i\left(E'_{ET}\right) dE'_{ET}$$

For a normalization of the cumulative distributions $P_i(E_{ET})$ in the corresponding range, an adjusted intercept b_i is necessary in addition to a stretching factor m_i, which considers the different contributions at small $E_{ET} < 0.4$, that is, $\int_0^{0.4} V_i(E'_{ET})dE'_{ET}$. The determination of the normalization pairs (m_i, b_i) was done via linear regression. From the values $P_i(E_{ET}=1.0)$, after subtracting an unknown offset, the proportions of the closed conformation for the individual measurements i could be determined. The offset B_{offset} was treated as a fit parameter. The $P_i\left(E_{ET}=1.0,\left[I_a^0\right]\right)$ obtained at different concentrations of the allosteric inhibitor were plotted as a function of the concentration of the allosteric inhibitor $\left[I_a^0\right]$. Titration curves were obtained by applying a global fit $P_i\left(E_{ET}=1.0,\left[I_a^0\right]\right)=A_j\frac{1}{\left(1+\frac{\left[I_a^0\right]}{K_{i,a}}\right)}+B_{offset}$.

The measurements for the competition assay were performed on three independent measuring days. Since each measurement differs due to the individual adjustment of the microscope and the exact protease concentrations used, a factor A_j was introduced to compensate these deviations.

2. For alternative evaluation, the E^*_{ET} frequency distributions were represented by a fit with a bimodal distribution composed of three Gaussians. For the range $E_{ET}\sim 0.4-0.6$ a Gaussian distribution with the parameters $E_{ET,1}=0.47560$ und $\sigma_1=0.15927$ (both determined) was assumed.

$$G_1(E_{ET})=M_1\,exp\left(-\frac{(E_{ET}-E_{ET,1})^2}{2\sigma_1^2}\right)$$

The distribution function for the range $E_{ET}\sim 0.7-1.0$ was represented by a composite function with $E_{ET,2a}=0.90448$, $\sigma_{2a}=0.05107$, $E_{ET,2b}=0.82242$, and $\sigma_{2b}=0.07923$.

$$G_2(E_{ET})=M_2\left[0.899\cdot exp\left(-\frac{(E_{ET}-E_{ET,2a})^2}{2\sigma_{2a}^2}\right)\right.$$
$$\left.+exp\left(-\frac{(E_{ET}-E_{ET,2b})^2}{2\sigma_{2b}^2}\right)\right]$$

For each distribution $V_i(E_{ET})$, the amplitude ratios $M_{2,i}/M_{1,i}$ were determined and plotted as a function of the concentration of the allosteric inhibitor. This titration curve was approximated by a global fit from which the binding affinity was determined as the inflection point.

4.7 | Fluorometric enzyme assay

The fluorometric assay was performed mainly as described in the literature.[31] Briefly, 250 nM of the protease in buffer (50 mM Tris–HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP) with 5 μl of the substrate in DMSO and 10 μl of the corresponding inhibitor in DMSO resulting in a total volume of 200 μl was added to a 96 well plate and the fluorescence of the released AMC was measured at 380 nm excitation and 460 nm emission. The percentage activity of the protease with the addition of the inhibitors was determined as the proportion of the slope with respect to the slope of the DMSO control. A detailed description can be found in Appendix S1.

AUTHOR CONTRIBUTIONS

Hannah Maus: Data curation (lead); validation (equal); writing – original draft (equal); writing – review and editing (equal). **Gerald Hinze:** Formal analysis (lead); software (lead); validation (equal); writing – original draft (equal); writing – review and editing (equal). **Stefan Josef Hammerschmidt:** Data curation (supporting); writing – original draft (equal); writing – review and editing (equal). **Thomas Basché:** Conceptualization (equal); funding acquisition (equal); project administration (equal); supervision (equal); writing – original draft (equal); writing – review and editing (equal). **Tanja Schirmeister:** Conceptualization (equal); funding acquisition (equal); project administration (equal); supervision (equal); writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

We thank the group of Prof. Dr. W. Diederich, University of Marburg, Germany, for providing us with the DENV2 NS2B$_{cf}$-NS3$_{pro}$ gene (GenBank ID: AY037116.1), with the two I30A and L31A point mutations in NS3$_{pro}$. Furthermore, we thank Prof. Dr. C. Klein, Heidelberg University, for providing us with the competitive inhibitor. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Data available on request from the authors

ORCID

Hannah Maus https://orcid.org/0000-0002-3512-5626
Stefan Josef Hammerschmidt https://orcid.org/0000-0002-0769-8435
Thomas Basché https://orcid.org/0000-0003-0502-7088
Tanja Schirmeister https://orcid.org/0000-0002-4587-5076

REFERENCES

1. Mizoue LS, Chazin WJ. Engineering and design of ligand-induced conformational change in proteins. Curr Opin Struct Biol. 2002;12(4):459–63. https://doi.org/10.1016/S0959-440X(02)00348-2

2. Ababou A, Shenvi RA, Desjarlais JR. Long-range effects on calcium binding and conformational change in the N-domain of Calmodulin. Biochemistry. 2001;40(42):12719–26. https://doi.org/10.1021/bi010405b
3. Nelson MR, Thulin E, Fagan PA, Forsén S, Chazin WJ. The EF-hand domain: a globally cooperative structural unit. Protein Sci. 2009;11(2):198–205. https://doi.org/10.1110/ps.33302
4. Ababou A. Solvation energetics and conformational change in EF-hand proteins. Protein Sci. 2001;10(2):301–12. https://doi.org/10.1110/ps.33601
5. Marvin JS, Hellinga HW. Manipulation of ligand binding affinity by exploitation of conformational coupling. Nat Struct Biol. 2001;8(9):795–8. https://doi.org/10.1038/nsb0901-795
6. Benson DE, Haddy AE, Hellinga HW. Converting a maltose receptor into a nascent binuclear copper oxygenase by computational design. Biochemistry. 2002;41(9):3262–9. https://doi.org/10.1021/bi011359i
7. Shimaoka M, Shifman JM, Jing H, Takagi J, Mayo SL, Springer TA. Computational design of an integrin I domain stabilized in the open high affinity conformation. Nat Struct Biol. 2000;7(8):674–8. https://doi.org/10.1038/77978
8. Dahiyat BI, Sarisky CA, De Mayo SL. Novo protein design: towards fully automated sequence selection. J Mol Biol. 1997; 273(4):789–96. https://doi.org/10.1006/jmbi.1997.1341
9. Evenäs J, Thulin E, Malmendal A, Forsén S, Carlström G. NMR studies of the E140Q mutant of the carboxy-terminal domain of Calmodulin reveal global conformational exchange in the Ca2+-saturated state. Biochemistry. 1997;36(12):3448–57. https://doi.org/10.1021/bi9628275
10. Wand AJ. Dynamic activation of protein function: a view emerging from NMR spectroscopy. Nat Struct Biol. 2001;8(11): 926–31. https://doi.org/10.1038/nsb1101-926
11. Kauk M, Hoffmann C. Intramolecular and intermolecular FRET sensors for GPCRs—monitoring conformational changes and beyond. Trends Pharmacol Sci. 2018;39(2):123–35. https://doi.org/10.1016/j.tips.2017.10.011
12. Zhu L, Yang J, Li H, Sun H, Liu J, Wang J. Conformational change study of dengue virus NS2B-NS3 protease using 19F NMR spectroscopy. Biochem Biophys Res Commun. 2015; 461(4):677–80. https://doi.org/10.1016/j.bbrc.2015.04.090
13. Ha T, Ting AY, Liang J, Caldwell WB, Deniz AA, Chemla DS, et al. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. Proc Natl Acad Sci U S A. 1999;96(3):893–8. https://doi.org/10.1073/pnas.96.3.893
14. Santoso Y, Joyce CM, Potapova O, le Reste L, Hohlbein J, Torella JP, et al. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. Proc Natl Acad Sci U S A. 2010;107(2):715–20. https://doi.org/10.1073/pnas.0910909107
15. Ruer M, Krainer G, Gröger P, Schlierf M. ATPase and protease domain movements in the bacterial AAA+ protease FtsH are driven by thermal fluctuations. J Mol Biol. 2018;430(22):4592–602. https://doi.org/10.1016/j.jmb.2018.07.023
16. Götz C, Hinze G, Gellert A, Maus H, von Hammerstein F, Hammerschmidt SJ, et al. Conformational dynamics of the dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies. J Phys Chem B. 2021;125(25): 6837–46. https://doi.org/10.1021/acs.jpcb.1c01797
17. Behnam MAM, Nitsche C, Boldescu V, Klein CD. The medicinal chemistry of dengue virus. J Med Chem. 2016;59(12):5622–49. https://doi.org/10.1021/acs.jmedchem.5b01653
18. Amberg SM, Nestorowicz A, McCourt DW, Rice CM. NS2B-3 Proteinase-Mediated Processing in the Yellow Fever Virus Structural Region: In Vitro and in Vivo Studies. J Virol. 1994; 68(6):3794–802.
19. Gupta G, Lim L, Song J. NMR and MD studies reveal that the isolated dengue NS3 protease is an intrinsically disordered chymotrypsin fold which absolutely requests NS2B for correct folding and functional dynamics. PLoS One. 2015;10(8):1–24. https://doi.org/10.1371/journal.pone.0134823
20. Yusof R, Clum S, Wetzel M, Murthy HMK, Padmanabhan R. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. J Biol Chem. 2000;275(14): 9963–9. https://doi.org/10.1074/jbc.275.14.9963
21. Noble CG, Seh CC, Chao AT, Shi PY. Ligand-bound structures of the dengue virus protease reveal the active conformation. J Virol. 2012;86(1):438–46. https://doi.org/10.1128/jvi.06225-11
22. Erbel P, Schiering N, Arcy AD, et al. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat Struct Mol Biol. 2006;13(4):2005–6. https://doi.org/10.1038/nsmb1073
23. Nitsche C, Holloway S, Schirmeister T, Klein CD. Biochemistry and medicinal chemistry of the dengue virus protease. Chem Rev. 2014;114(22):11348–81. https://doi.org/10.1021/cr500233q
24. Brecher M, Li Z, Liu B, Zhang J, Koetzner CA, Alifarag A, et al. A conformational switch high-throughput screening assay and allosteric inhibition of the Flavivirus NS2B-NS3 protease. PLoS Pathog. 2017;13(5):e1006411. https://doi.org/10.1371/journal.ppat.1006411
25. Chen WN, Loscha KV, Nitsche C, Graham B, Otting G. The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI. FEBS Lett. 2014;588(14):2206–11. https://doi.org/10.1016/j.febslet.2014.05.018
26. Phoo WW, El Sahili A, Zhang ZZ, et al. Crystal structures of full length DENV4 NS2B-NS3 reveal the dynamic interaction between NS2B and NS3. Antiviral Res. 2020;182(July):104900. https://doi.org/10.1016/j.antiviral.2020.104900
27. De La Cruz L, Nguyen THD, Ozawa K, et al. Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus Ns2b-Ns3 protease: fold analysis by pseudocontact shifts. J Am Chem Soc. 2011;133(47):19205–15. https://doi.org/10.1021/ja208435s
28. Bräuchle C, Lamb DC, Michaelis J. Single particle tracking and single molecule energy transfer. Wiley-VCH, Weinheim; 2009. https://doi.org/10.1002/9783527628360
29. Förster T. Zwischenmolekulare Energiewanderung Und Fluoreszenz. Ann Phys. 1948;6(2):55–75.
30. Behnam MAM, Nitsche C, Vechi M, et al. C-terminal residue optimization and fragment merging: discovery of a potent peptide-hybrid inhibitor of dengue protease. ACS Med Chem Lett. 2014;5:1037–42. https://doi.org/10.1021/ml500245v
31. Maus H, Barthels F, Hammerschmidt SJ, Kopp K, Millies B, Gellert A, et al. SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases. Bioorg Med Chem. 2021;47(May):116392. https://doi.org/10.1016/j.bmc.2021.116392

32. Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat Methods. 2012;9(12):1218–25. https://doi.org/10.1038/nmeth.2222
33. GmbH, A. T. ATTO 643 https://www.atto-tec.com/ATTO-643.html?language=en
34. Tarzia A. andrewtarzia/mol-ellipsize: v1.0.1 (v1.0.1). Zenodo. 2021. https://doi.org/10.5281/zenodo.5136958
35. Cheminformatics, O.-S. RDKit http://www.rdkit.org.
36. Henschel. Center of mass https://pymolwiki.org/index.php/Center_of_mass
37. Millies B, Von Hammerstein F, Gellert A, et al. Proline-based allosteric inhibitors of Zika and dengue virus NS2B/NS3 proteases. J Med Chem. 2019;62(24):11359–82. https://doi.org/10.1021/acs.jmedchem.9b01697
38. Weiss S. Fluorescence spectroscopy of single biomolecules. Science. 1999;283(5408):1676–83. https://doi.org/10.1126/science.283.5408.1676
39. Eggeling C, Fries JR, Brand L, Günther R, Seidel CAM. Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. Proc Natl Acad Sci U S A. 1998; 95(4):1556–61. https://doi.org/10.1073/pnas.95.4.1556
40. Widengren J, Kudryavtsev V, Antonik M, Berger S, Gerken M, Seidel CAM. Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. Anal Chem. 2006;78(6):2039–50. https://doi.org/10.1021/ac0522759
41. Kühnemuth R, Seidel CAM. Principles of single molecule multiparameter fluorescence spectroscopy. Single Mol. 2001;2(4):251–4. https://doi.org/10.1002/1438-5171(200112)2:4<251::AID-SIMO251>3.0.CO;2-T
42. Deniz AA, Dahan M, Grunwell JR, Ha T, Faulhaber AE, Chemla DS, et al. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. Proc Natl Acad Sci U S A. 1999;96(7):3670–5. https://doi.org/10.1073/pnas.96.7.3670
43. Margittai M, Widengren J, Schweinberger E, Schröder GF, Felekyan S, Haustein E, et al. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of Syntaxin 1. Proc Natl Acad Sci U S A. 2003;100(26):15516–21. https://doi.org/10.1073/pnas.2331232100
44. Torres T, Levitus M. Measuring conformational dynamics: a new FCS-FRET approach. J Phys Chem B. 2007;111(25):7392–400. https://doi.org/10.1021/jp070659s
45. Chou TC, Talalay P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. Adv Enzyme Regul. 1984;22:27–55. https://doi.org/10.1016/0065-2571(84)90007-4

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

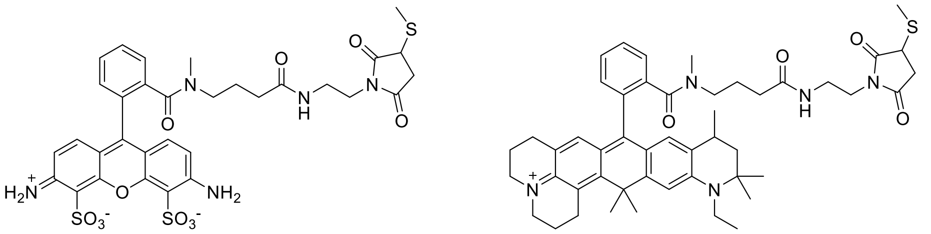
How to cite this article: Maus H, Hinze G, Hammerschmidt SJ, Basché T, Schirmeister T. A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease. Protein Science. 2023;32(1):e4526. https://doi.org/10.1002/pro.4526

Supplementary Material

Content

1 Calculation of the Dye Distances

Dye distances in the protein were calculated using the toolkit from Seidel's laboratory (Kalinin, S.; Peulen, T.; Sindbert, S.; et al. A Toolkit and Benchmark Study for FRET-Restrained High-Precision Structural Modeling. Nat. Methods **2012**, 9 (12), 1218–1225. https://doi.org/10.1038/nmeth.2222). Approximately, ATTO647N was used as acceptor dye for the calculation since no structure of ATTO643 is available. The radii describing the ellipsoid enclosing the chromophore and the corresponding linker lengths for the dye maleimides are given in Figure S1.



ATTO488
r_1 = 0.57 nm
r_2 = 0.68 nm
r_3 = 0.75 nm
linker length = 0.97 nm

ATTO647N
r_1 = 0.54 nm
r_2 = 0.75 nm
r_3 = 0.92 nm
linker length = 1.47 nm

Figure S1: Structures of the donor and acceptor dye ATTO488 and ATTO647N. The structure of ATTO647N was used as an approximation for ATTO643. The radii r_1, r_2 and r_3 define the ellipsoid enclosing the chromophores. The linker length indicates the distances between the center of mass of the chromophores and the sulfur atom attached to the maleimide.

The determined accessible volume of the dyes is shown in Figure S2.

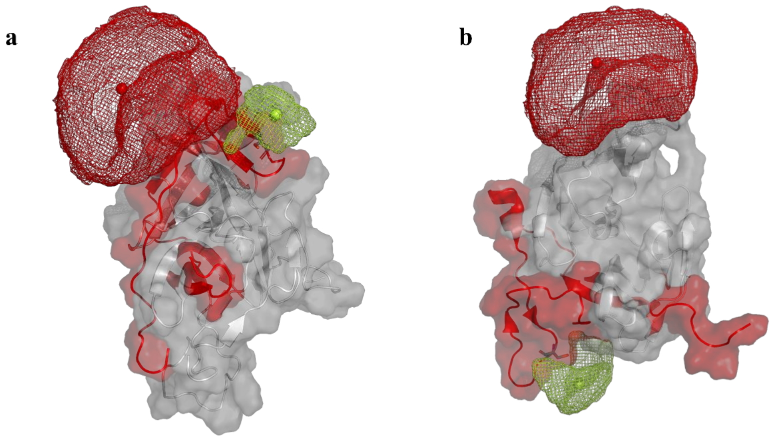


Figure S2: Accessible volume of the donor dye ATTO488 (green) and the acceptor dye ATTO647N (red) in the open (PDB: 2FOM) and closed (PDB: 2M9P) conformation of the NS2B-NS3 protease that were calculated using the toolkit from Seidel's laboratory (Kalinin, S.; Peulen, T.; Sindbert, S.; et al. A Toolkit and Benchmark Study for FRET-Restrained High-Precision Structural Modeling. Nat. Methods **2012**, 9 (12), 1218–1225. https://doi.org/10.1038/nmeth.2222).

2 Activity of the Labelled Protein

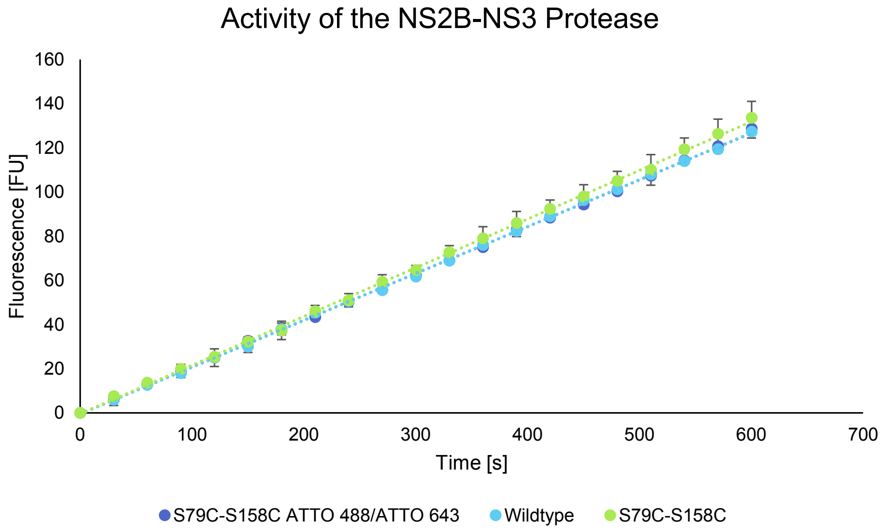


Figure S3: Progress curves of the wildtype NS2B-NS3 protease, the double mutant S79C-S158C and the dye labelled double mutant S79C-S158C ATTO488/ATTO643 at a protein concentration of 200 nM. Errors are indicated as bars.

3 Difference Function of the Cross-Correlation

The following notation bases on the publication from Torres, T. and Levitus, J. Phys. Chem. B 2007, 111 (25), 7392–7400 (https://doi.org/10.1021/jp070659s).

It is assumed, that fluorescence intensity fluctuations are caused by translational diffusion and by changes in energy transfer efficiencies between donor and acceptor moieties. Then the intensity cross-correlation function $G_{xy}(\tau)$ can be factorized into a diffusion term $T(\tau)$ and $X_{xy}(\tau)$, which solely depends on the intensity fluctuation due to switching between conformations with different energy transfer rates E_1 and E_2, $G_{xy}(\tau) = T(\tau)X_{xy}(\tau)$. k_{12} and k_{21} denote the transition rates between both conformations. With $k_{ex} = k_{12} + k_{21}$ and a constant A the exchange term simplifies to $X_{AD}(\tau) = 1 - Ae^{-K_{ex}\tau}$. Without conformational change $X_{AD}(\tau) = 1$ follows.

After addition of a ligand, there are two species of proteases in solution, those that are present as a ligand-protease complex and free protease. It is assumed that the free protease changes between two conformations, the protease in complex with the ligand, however, is fixed in one conformation (at least during the observation period). Thus, the following relationships can be formulated:

$$G_{AD}^{after}(\tau) = T(\tau)[A_1 X_{AD}^{free} + A_2 X_{AD}^{complex}(\tau)] = T(\tau)[A_1 + A_2 - A_1 e^{-K_{ex}\tau}]$$

The proportions A_1 and A_2 depend on the concentration of the ligand. Before addition of the ligand, the following equation applies

$$G_{AD}^{before}(\tau) = T(\tau)[1 - B_1 e^{-K_{ex}\tau}].$$

For further evaluation both cross-correlation functions must be normalized to 1,

$$G_{AD}^{before}(\tau = 0) = G_{AD}^{after}(\tau = 0) = 1.$$

By normalizing $T(\tau = 0) = 1$ as well, the following relationships result:

$$G_{AD}^{before}(\tau) = T(\tau)[C_1 - C_2 e^{-K_{ex}\tau}] = T(\tau)[C_1 - (C_1 - 1)e^{-K_{ex}\tau}]$$

$$G_{AD}^{after}(\tau) = T(\tau)[D_1 - D_2 e^{-K_{ex}\tau}] = T(\tau)[D_1 - (D_1 - 1)e^{-K_{ex}\tau}],$$

from which the difference can be calculated,

$$\Delta G(\tau) = G_{AD}^{before}(\tau) - G_{AD}^{after}(\tau) = T(\tau)[C_1 - D_1 - C_1 e^{-K_{ex}\tau} + D_1 e^{K_{ex}\tau}]$$

Since a conformation is stabilized after addition of the ligand, the following relation applies

$\frac{C_2}{C_1} > \frac{D_2}{D_1}$, i.e., $C_1 > D_1$. By defining $P = C_1 - D_1$ with $P > 0$, $\Delta G(\tau) = T(\tau)P[1 - e^{-K_{ex}\tau}]$ is obtained.

Assuming free diffusion and an FCS excitation volume in ellipsoid form with $a = b = c/3$, a function for a fit can be given:

$$\Delta G(\tau) = P\left(1 + \frac{\tau}{t_d}\right)^{-1}\left(1 + \frac{\tau}{9t_d}\right)^{-\frac{1}{2}}[1 - e^{-K_{ex}\tau}].$$

4 Binding Affinities of the Inhibitors

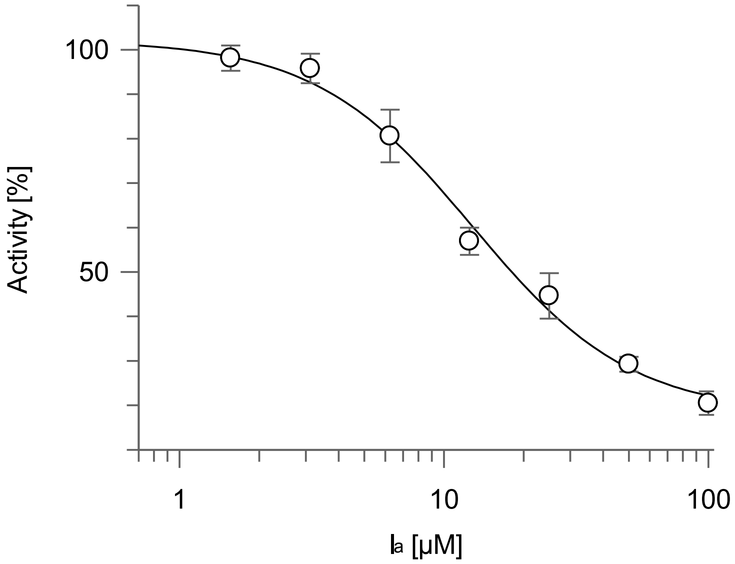


Figure S4: Binding affinity curve of the allosteric inhibitor I_a. The inflection point of the curve indicates the $IC_{50} = 12.8 \pm 0.8$ µM. Errors are indicated as bars.

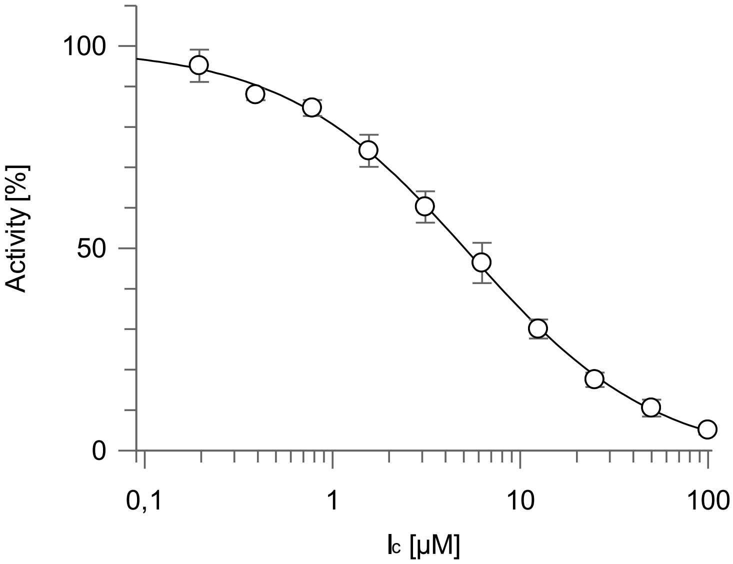


Figure S5: Binding affinity curve of the competitive inhibitor I_c. The inflection point of the curve indicates the IC_{50} = 5.4 ± 0.4 µM. Errors are indicated as bars.

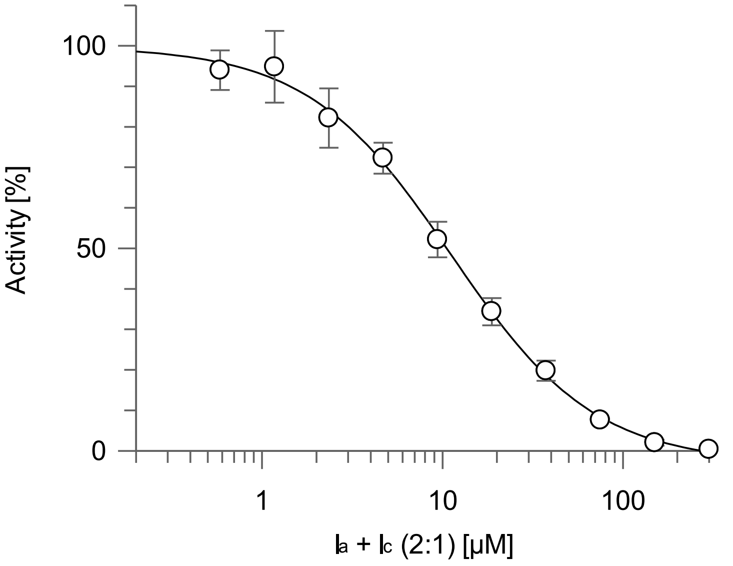


Figure S6: Binding affinity curve of the combination of allosteric inhibitor I_a and competitive inhibitor I_c. Ratio of I_a to I_c was set to 2:1. The inflection point of the curve indicates the IC_{50} = 11.2 ±0.5 µM. Errors are indicated as bars.

5 Binding Curve Resulting from the Bimodal Fit

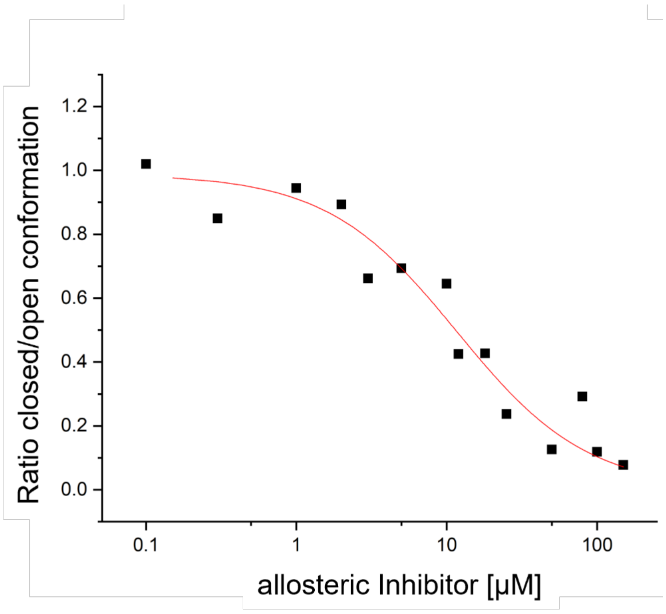
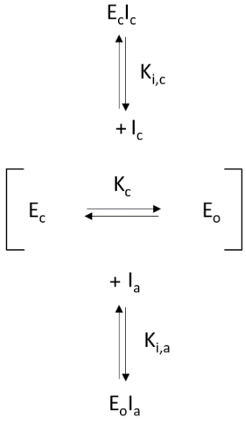


Figure S7. Binding curve of the allosteric inhibitor resulting from the bimodal fit. The K_D = 11.7 ± 2.2 µM was obtained from the inflection point of the sigmoidal curve.

6 Kinetic Modelling

Model 2 (induced fit): Both inhibitors can bind to both conformations of the protease independent from the conformation that is stabilized by the respective inhibitor. It is assumed that the allosteric inhibitor stabilizes the open and the competitive inhibitor stabilizes the closed conformation.

The following equilibrium reactions can be formulated, where E_o is the open conformation of the enzyme, E_c is the closed conformation of the enzyme, I_c is the competitive inhibitor and I_a is the allosteric inhibitor:



The corresponding equilibrium constants are $K_{i,c} = \frac{\left[E_c \overset{K_C}{\Leftrightarrow} E_o\right][I_c]}{[E_c I_c]}$, $K_{i,a} = \frac{\left[E_c \overset{K_C}{\Leftrightarrow} E_o\right][I_a]}{[E_o I_a]}$ and $K_c = \frac{[E_o]}{[E_c]}$.

Using the equilibrium constants $K_{i,c}$, $K_{i,a}$ and K_c and the initial concentrations of the ligands $[I_a^0]$ and $[I_c^0]$, the total concentration of open $[E_o^*]$ and closed $[E_c^*]$ protease present during the measurement can be described. Since the applied concentration of the protease is markedly smaller than the equilibrium constants $K_{i,c}$ and $K_{i,a}$, the following approximations can be made:

$$[E_c^*] = [E_c] + [E_c I_c] = [E_c] + \frac{[E_c + E_o][I_c^0]}{K_{i,c} + [E_c + E_o]} \cong [E_c] + \frac{[E_c + E_o][I_c^0]}{K_{i,c}}$$

$$[E_o^*] = [E_o] + [E_o I_a] = [E_o] + \frac{[E_c + E_o][I_a^0]}{K_{i,a} + [E_c + E_o]} \cong [E_o] + \frac{[E_c + E_o][I_a^0]}{K_{i,a}}$$

Taking into account $[E_o] = [E_c + E_o] - [E_c]$ and $[E_c] = \frac{[E_o]}{K_c}$, it follows $[E_0] = \frac{[E_c + E_o]}{1 + \frac{1}{K_c}}$ and thus

$$[E_o^*] = [E_c + E_o]\left(\frac{1}{1 + \frac{1}{K_c}} + \frac{[I_a^0]}{K_{i,a}}\right).$$

The concentration ratio of closed and open conformation follows

$$\frac{[E_c^*]}{[E_o^*]} = F \frac{1}{\left(\frac{K_c}{K_c + 1} + \frac{[I_a^0]}{K_{i,a}}\right)} = F \frac{1}{\left(\frac{K_c}{K_c + 1} + R\right)},$$

where $F = \left(\frac{1}{1+K_c} + \frac{[I_c^0]}{K_{i,c}}\right)$ remains constant during the titration. For a simplified representation $R = \frac{[I_a^0]}{K_{i,a}}$ is defined. The resulting titration curve $T([I_a^0])$ progresses with increasing concentration of the allosteric inhibitor $[I_a^0]$ from $T(0) = \frac{F}{\frac{K_c}{K_c+1}}$ to $T(\infty) = 0$.

The amplitude has dropped to half at:

$T(x) = \frac{1}{2}\frac{F}{\frac{K_c}{K_c+1}} = \frac{F}{\left(\frac{K_c}{K_c+1}+R\right)}$ with R=$\frac{K_c}{K_c+1} = \frac{x}{K_{i,a}}$,

i.e., $[I_a^0]_{\frac{1}{2}} = x = K_{i,a} \cdot \frac{K_c}{K_c+1}$

If, without inhibitors, the open conformation is significantly present, i.e., $K_c \gg 1$, this can be used to further approximate $T([I_a^0]) \cong F\frac{1}{(1+R)}$

In contrast to model 1 (conformational selection), in this model the titration curve depends on the equilibrium constant K_c. Only in the limiting case $K_c \gg 1$ it becomes independent of it.

7 Protein Constructs and Cloning

As a starting point for generating the S79C (NS2B) and the S158C ($NS3_{pro}$) mutations, a pET15b vector harboring the DENV2 $NS2B_{cf}$-$NS3_{pro}$ gene (GenBank ID: AY037116.1), with two-point mutations (I30A and L31A in NS3) for better protein solubility and an N-terminal hexa-histidine tag linked to $NS2B_{cf}$ *via* a thrombin cleavage site, was used. Site-directed mutagenesis was performed using the Kapa HiFi PCR kit (KapaBiosystems Inc, Woburn, MA, USA) and following primers:

(i) S79C: 5′-GATCCTGTCAATTACGATCTGCGAAGATG-3′ and 5′-GACATGCTACCATCTTCGCAGATCGTAATTG-3′

(ii) S158C: 5′-GGATCGGGCAGCTACCGCTGATTTCTGCTTG-3′ and 5′-CGTAGGCACCACAACGGGTCACGACACCATTACC-3′.

8 Protein Expression and Purification

Competent *E. coli* BL21-Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA) were used for expression of the protein. After transformation, cells were grown in LB medium containing 100 μM ampicillin at 37 °C until they reached an OD_{600} of ~0.7. Expression was then induced with 1 mM Isopropyl-β-D-thiogalactopyranosid (IPTG) for 16 h at 20 °C. After harvesting by centrifugation, cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.1 vol% $\text{Triton}_{\text{X-100}}$, 20 mM imidazole, DNase, lysozyme,) and lysed by sonication (Sonoplus, Bandelin, Berlin, Germany). The lysate was cleared *via* centrifugation (10 krpm, 45 min) and the supernatant was subjected to an immobilized metal affinity chromatography (IMAC) on a HisTrap HP 5 mL column (GE Healthcare, Chicago, IL, USA). After washing with 5 column volumes (CV) of buffer A (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole), the protein was eluted in a linear gradient of buffer B (20 mM Tris-HCl pH 8, 300 mM NaCl, 250 mM imidazole). Eluted fractions were further purified by size exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75 pg (GE Healthcare, Chicago, IL, USA) and eluted in SEC buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl). Prior to storage at –80 °C, proteins were concentrated, and flash-frozen with liquid nitrogen. Throughout all purification steps, protein concentrations were measured *via* absorbance at 280 nm, and sample purity was assessed *via* SDS-PAGE.

9 Fluorescent Labelling of the DENV-Protease double Mutant S79C-S158C

Purified protein was rebuffered by using Vivaspin protein concentrators (MWCO 10.0 kDa, GE Healthcare, Chicago, IL, USA) in labelling buffer (20 mM Tris-HCl pH 6.8, 150 mM NaCl, 0.5 mM TCEP) and incubated for 30 min at 4 °C. TCEP was removed using a Vivaspin protein concentrator in the same buffer without TCEP. A 2.1- fold excess of ATTO 488 and a 2.6-fold excess of ATTO 643 were added to the protein solution (26 µM protein) and samples were incubated by gently shaking in the dark for 2 h at room temperature. This enables a Michael addition of the thiol group of cysteine to the *β*-position of the *α,β*-unsaturated carbonyl moiety (maleimide anchor) of the dyes, forming a stable thioether. Labelled protease was then dialyzed overnight at 4 °C (20 mM Tris-HCl pH 6.8, 150 mM NaCl, 1 mM CHAPS) and purified by gel filtration (HiLoad 16/600 Superdex 75 pg column, GE Healthcare, Chicago, IL, USA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl).

10 smFRET Experiments

For smFRET measurements of the FRET-paired double mutant DENV protease, self-constructed sample cells consisting of a poly(ethylene glycol) (PEG_01 & 02, MicroSurfaces Inc., Minneapolis, MN, USA) coated glass coverslip and a glued-on cylinder of protein repellent plastic (Protein LoBind Tube, Eppendorf SE, Hamburg, Germany) were used. 150 µL of the respective solution of FRET pair labelled double mutant of the DENV protease (c ~ 100 pM) in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP) was placed in the sample cell, and fluorescence photons were collected over a period of 1800 s using a custom-built confocal microscope. The ligands were dissolved in DMSO and added to the protease solution, resulting in a 10 vol% DMSO content in the sample cell. First, in each case a measurement was performed without ligand (76 pM protease, 50 mM Tris-HCl, 1 mM CHAPS, 1 mM TCEP, 10 vol% DMSO). Then a solution of a ligand in buffer with protease was added. In this way, the concentration of protease and the DMSO content were kept constant during all measurements. The concentration of the competitive inhibitor was varied between 1 µM and 10 µM, and that of the allosteric between 0.1 µM and 150 µM. The recordings of both data sets with and without the ligand were performed in direct succession. The fluorophores were excited with the spectrally filtered output from a pulsed white light fiber laser (10 MHz, SC OEM, YSL Photonics, China). To obtain excitation pulses centered around 502 nm (FWHM ~ 1.5 nm), an acousto-optical tunable filter (AOTF-VIS-DR, Fianium, United Kingdom) was used. The laser light, after passing a single-mode fiber, was recollimated and then focused into the solution by using a microscope objective (plan apochromat, 100×, NA = 1.4, oil immersion, Zeiss, Germany). In front of the objective, the excitation power was measured and set to 10 µW. Fluorescence emitted by FRET pair labelled DENV protease diffusing through the focus was collected through the objective. Excitation and emission were separated by a dichroic mirror (ZT491 rdcxrxt-UF1, CHROMA, USA). The emitted fluorescence light passed a dichroic mirror (ZT640rdc-UF1, CHROMA, USA) in the detection beam path, which allowed red light to be transmitted, while light of higher energy was reflected. These two spectrally separated beams were then focused onto two APDs (acceptor channel A: SPCM-AQRH-15, PerkinElmer, United States; donor channel D: PDM 50ct, MPD, Italy) after appropriate spectral filtering (acceptor channel A: ET655lp, CHROMA, United States; donor channel D: FF01-550/88-25, Semrock, USA). The use of the described filter combination ensured separate detection of donor and acceptor fluorescence. The two detector APDs were connected to a HydraHarp 400 module (PicoQuant, Germany), which records the absolute and relative arrival time (relative to the excitation pulse) of the individual photons. This allowed the subsequent calculation of correlation functions, lifetime, and intensity time traces for the burst analysis. Such time traces were calculated for both channels.

11 Analysis of smFRET Data

First, for the initial assessment of the conformational influence of ligands on DENV protease, cross-correlation functions $G_{AD}(\tau)$ (FCS-FRET) were calculated from the macro arrival times of the individual photons recorded by using the HydraHarp 400 module (PicoQuant, Berlin, Germany). It had been shown that the cross-correlation can drop more slowly to zero if there is an exchange between two conformations that differ in their energy transfer efficiency provided that this process is not much slower than the diffusional time through the focus.[36,39] Exchange manifests itself as additional term in the correlation function; that is, in addition to the diffusional term without exchange, the cross-correlation function is extended by an increasing term $\mathrm{G_{AD}} \propto (1 - \mathrm{a_{AD}} \mathrm{e}^{-\frac{\tau}{\tau_{\mathrm{ex}}}})$. Notably, it is expected that the cross-correlation does no longer drop significantly slower to zero if an added ligand stabilizes one of the conformations of the protein, and therefore no or a significantly reduced conformational exchange takes place. It is consequently possible to draw conclusions about the conformational dynamics based on the cross-correlations of the donor and acceptor intensities with and without added ligands. To explore the presence of conformational subpopulations, data must be viewed at the single-molecule level (smFRET). For this purpose, individual bursts extracted from the fluorescence intensity time traces (bin time = 1 ms) were analyzed. For further evaluation, only bursts were considered which showed a total of at least 20 counts on both APDs. Average arrival times τ_A of the acceptor photons relative to the excitation pulse (without taking the instrumental response function (IRF) into account) were calculated for each individual burst. Furthermore, the FRET efficiency E_{ET} was calculated for each burst:

$$E_{ET} = \frac{counts(A)}{counts(A)+counts(D)}$$

We did not account for spectral crosstalk or direct excitation of the acceptor. Furthermore, we omitted a detection correction factor, which could compensate for the different detection efficiencies of the two APDs toward donor and acceptor fluorescence and the different quantum yields of both dyes.

The individual bursts were sorted with respect to their E_{ET}^{*} and presented in frequency distributions. These allow the identification of subpopulations that differ in their E_{ET}. The occurrences of the individual bursts (individual molecules) were then plotted in a 2D histogram, separated according to acceptor lifetime and FRET efficiency, which enables the visual identification of existing subpopulations.

12 Fluorometric Enzyme Assay

The inhibitory activity of the allosteric inhibitor towards DENV protease was determined using an assay based on the fluorogenic substrate Boc-GRR-AMC (Bachem, Bubendorf, Switzerland). Substrate and inhibitor were pipetted as DMSO stock solutions to the protease (250 nM) dissolved in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP). The measurements were carried out in flat-bottom 96-well microtiter plates (Greiner bio-one, Kremsmünster, Austria) with a Tecan Infinite F2000 PRO fluorimeter or a Tecan Spark 10M (Agilent Technologies, Santa Clara, USA) in three independent experiments. In each well, a total volume of 200 µL was used, consisting of 180 µL of buffer, 5 µL of the enzyme solution, resulting in a final concentration of 125 nM, 10 µL of the inhibitor in DMSO or pure DMSO as control, and 5 µL of a solution of the substrate with a final concentration of 100 µM. The fluorescence was measured every 30 s for 10 min at 25 °C with 380 nm excitation and 460 nm emission wavelengths. The Dixon plot was determined by repeating the measurement at different substrate concentrations (50 µM, 150 µM and 200 µM).

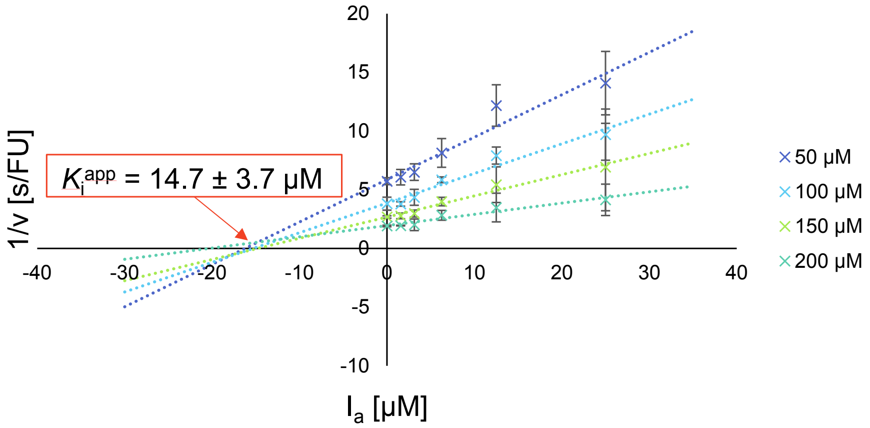


Figure S8: Dixon plot of the DENV NS2B-NS3 protease with the allosteric inhibitor. The intersection of the lines reveals the apparent K_i = 14.7 ± 3.7 µM. Errors are indicated as bars.

“

5.3 The Effects of Allosteric and Competitive Inhibitors on ZIKV Protease Conformational Dynamics Explored Through smFRET, nanoDSF, DSF, and ^{19}F-NMR.

5.3.1 Context, Project Summary, and Own Contributions

Only a few studies have examined the conformational dynamics of the ZIKV protease in solution, mainly using NMR spectroscopy.[124,278–282] These studies focused on the impact of competitive ligands, while the effects of allosteric ligands have yet to be thoroughly investigated. Research on the DENV protease using smFRET (see previous Chapters 5.1 and 5.2), a split-luciferase assay, and NMR spectroscopy revealed an equilibrium between two conformations, demonstrating that competitive ligands favor the closed conformation.[106,114,283,284] In contrast, allosteric ligands stabilize the open conformation.[284] To study the ZIKV protease, which has native cysteine residues, site-specific labeling for smFRET experiments requires mutagenesis and the introduction of cysteine residues.

We employed smFRET to investigate the conformational dynamics of a glycine-linked ZIKV protease construct (gZiPro, WT) with five cysteine point mutations (S84*C-C80S-C143S-S160C-C178S, ^{5}ZiPro, Figure 24A). This construct maintained enzymatic activity and structural integrity, making it suitable for studying ZIKV protease dynamics, shown by fluorometric assay and CD spectroscopy. Like the DENV protease, we found that a competitive inhibitor (I_c^{ZIKV}) stabilized the closed conformation, while an allosteric inhibitor (I_a) favored the open conformation.

Furthermore, we utilized more accessible methods such as ^{19}F-NMR spectroscopy, DSF, and nanoDSF to investigate ligand-protein interactions of the ZIKV protease in solution. Opposite effects of the competitive and the allosteric inhibitor were observed using thermal shift assays. Together with the smFRET results, it was speculated that the respective in- or decrease in melting temperature caused by the different ligands is a consequence of stabilizing two different protease conformations. For ^{19}F-NMR experiments, fluor-labeling was performed by expressing the protease in a defined media containing 7F-tryptophane to be incorporated into the protein (Figure 24B). In ^{19}F-NMR spectroscopy, the shift of W50 allowed tracking of the respective binding mode of the ligands. However, no shift was observed for the allosteric inhibitor, which was explained by the further stabilization of the conformation, predominantly present without a ligand, resulting in the lack of response in the NMR experiments.

Overall, this study represented the first comprehensive investigation of the conformational dynamics of the ZIKV protease in response to allosteric ligands using multiple biophysical techniques. The findings enhanced our understanding of inhibitor interactions with the ZIKV protease.

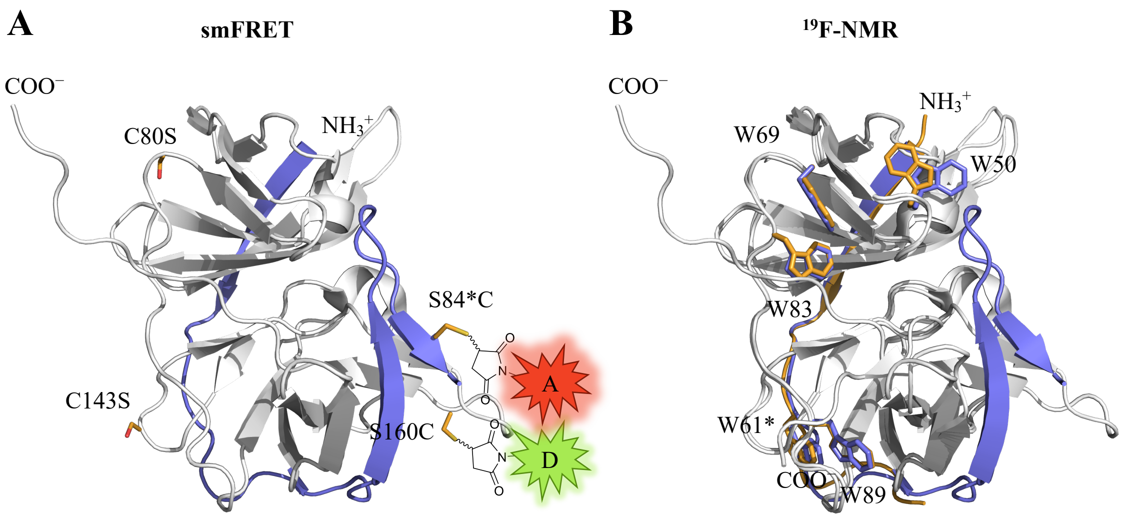


Figure 24. Structure of the ZIKV protease. **A** Quintuple mutant (S84*C-C80S-C143S-S160C-C178S) of the closed conformation (PDB-ID: 6y3b) of the ZIKV Protease (^{5}ZiPro). Mutated amino acids are labeled and highlighted with yellow sticks. C178S is not resolved in the crystal structure and consequently not shown. S84* and S160 are labeled with maleimide containing donor (D) and acceptor (A) dyes. **B** Positions of the native tryptophan residues in the open (PDB-ID: 5gxj) and closed (PDB-ID: 6y3b) conformation of the ZIKV protease. W5 is not resolved in the crystal structure and thus is not shown. NS3 is shown in white, NS2B in yellow (open conformation) or blue (closed conformation). Tryptophan residues are highlighted with sticks in yellow (open conformation) or blue (closed conformation).

Own contributions: protein labeling for smFRET experiments, smFRET data curation, DSF experiments, nanoDSF experiments, CD spectroscopy experiments, fluorometric enzyme assays, protein labeling for ^{19}F-NMR experiments (together with Stefan J. Hammerschmidt), ^{19}F-NMR data curation (together with Stefan J. Hammerschmidt and Victor H. Pérez), and writing of the original draft & editing of the manuscript.

Contributions of others: protein expression & purification, site-directed mutagenesis, smFRET data analysis, protein labeling for ^{19}F-NMR experiments, ^{19}F-NMR data curation, ^{19}F-NMR data analysis, and writing parts of the original draft & editing of the manuscript.

This work was published in European Journal of Medicinal Chemistry (impact factor:7.09).

Article reprinted with permission of *European Journal of Medicinal Chemistry* **2023**, 258 (February): 115573 “The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR.” © 2023 Elsevier B.V. (Netherlands).

5.3.2 Publication

The following publication quoted within “” from page 180 to page 200 is the same as the manuscript cited on page 179.

“

European Journal of Medicinal Chemistry 258 (2023) 115573

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: www.elsevier.com/locate/ejmech

Research paper

The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F NMR

Hannah Maus [a,1], Stefan J. Hammerschmidt [a,1], Gerald Hinze [b], Fabian Barthels [a], Victor H. Pérez Carrillo [c], Ute A. Hellmich [c], Thomas Basché [b], Tanja Schirmeister [a,*]

[a] Institute of Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg-University, Mainz, Germany
[b] Department of Chemistry, Johannes Gutenberg-University, Mainz, Germany
[c] Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University Jena, Jena, Germany

ARTICLE INFO

Keywords:
ZIKV NS2B-NS3 protease
Allosteric/competitive inhibition
Conformational dynamics
smFRET
(Nano)DSF
^{19}F NMR

ABSTRACT

ABSTRACT: Zika and dengue viruses cause mosquito-borne diseases of high epidemic relevance. The viral NS2B-NS3 proteases play crucial roles in the pathogen replication cycle and are validated drug targets. They can adopt at least two conformations depending on the position of the NS2B cofactor. Recently, we reported ligand-induced conformational changes of dengue virus NS2B-NS3 protease by single-molecule Förster resonance energy transfer (smFRET). Here, we investigated the conformational dynamics of the homologous Zika virus protease through an integrated methodological approach combining smFRET, thermal shift assays (DSF and nanoDSF) and ^{19}F NMR spectroscopy. Our results show that allosteric inhibitors favor the open conformation and competitive inhibitors stabilize the closed conformation of the Zika virus protease.

1. Introduction

Zika and dengue viruses (ZIKV/DENV) cause two of the most widespread and epidemically relevant mosquito-borne viral diseases of modern times [1–3]. Both pathogens require a viral NS2B-NS3 serine protease, which, together with host proteases, is responsible for the cleavage of the viral precursor polyprotein [4]. Thus, the enzyme constitutes an essential component of the viral replication process [5]. Accordingly, the structurally conserved viral NS2B-NS3 serine protease is a valid drug target to combat virus replication [5–7]. The N-terminal domain of ZIKV NS3 consists of a trypsin-like serine protease ($NS3^{pro}$), with the catalytic triad composed of His51, Asp75 and Ser135 [8]. Both the correct folding and the catalytic activity of the protease depend on the presence of the associated C-terminal cofactor region of $NS2B^{cf}$ [9–12]. Two conformations of NS2B-NS3 have been described that differ mainly with regard to the relative position of the C-terminus of $NS2B^{cf}$ [9,13,14]. In the so-called “closed” conformation, the $NS2B^{cf}$ is wrapped around the $NS3^{pro}$ core, while it is rather loosely bound to $NS3^{pro}$ in the “open” conformation [9,13–16]. For both conformations, crystal structures for ZIKV [17,18] (open: 5GXJ, closed: 5LC0) and DENV [14,16] (open: 2FOM, closed: 3U1I) protease are available. Despite the overall structural similarity and high sequence homology of the two proteases, the position of a loop within the $NS2B^{cf}$, i.e. residues 152–167 in the ZIKV proteas in the open conformation differs compared to the DENV protease [19]. Furthermore, the C-terminus of the ZIKV $NS2B^{cf}$ seems to exhibit increased flexibility compared to the DENV protein. Roy et al. attributed this to reduced interactions between $NS2B^{cf}$ and $NS3^{pro}$ in the ZIKV protein [19–21]. While extremely valuable, static structures obtained by X-ray crystallography need to be complemented with a dynamic view of a protein in solution through an additional set of optical and spectroscopic methods [22–30]. However, to date, only a few studies looked into the conformational dynamics of the ZIKV protease in solution, mostly by NMR spectroscopy [21,31–36]. These studies investigated the consequence of competitive ligand binding, but the dynamic effect of allosteric ligands on the ZIKV protease has not yet been investigated in detail [31]. For the DENV protease, it was shown by smFRET and NMR spectroscopy that the protein is in an equilibrium between two conformations [28,37]. Importantly, smFRET allows to observe conformational subpopulations, conformational transitions and temporal fluctuations that typically remain elusive in fluorescence

---

[*] Corresponding author. Institute for Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg-University, Mainz, Germany.
E-mail addresses: schirmei@uni-mainz.de, schirmei@uni-mainz.de (T. Schirmeister).

[1] authors contributed equally.

https://doi.org/10.1016/j.ejmech.2023.115573
Received 13 February 2023; Received in revised form 5 June 2023; Accepted 6 June 2023

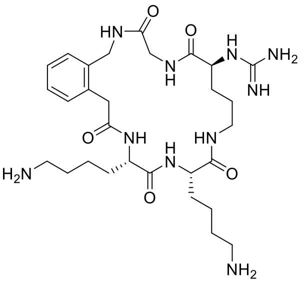
ensemble measurements [38,39]. To this end, a target protein is labeled with a donor and an acceptor dye, typically by taking advantage of a free thiol group of a cysteine residue *via* maleimide coupling. Unlike the cysteine-free DENV protease, ZIKV NS2B-NS3 has three native cysteine residues [40]. Site-specific labeling thus requires the mutagenesis of the native cysteine residues and introduction of cysteine residues at the desired labeling sites. Recently, we successfully used smFRET on DENV protease to show that while competitive ligands favor the closed conformation [28], allosteric ligands lead to a stabilization of the open conformation [30]. Herein, we describe the conformational dynamics of a glycine-linked ZIKV protease construct (gZiPro, WT) with five cysteine point mutations (^{5}ZiPro) by smFRET. This construct exhibits enzymatic activity comparable to the native protease and retains its secondary structure. It can thus be used as a suitable model to study ZIKV protease structural dynamics. Similar to DENV protease, we found that a competitive inhibitor [41] (**I$_c$**, Fig. 1) stabilized the closed conformation, whereas an allosteric inhibitor [42] (**I$_a$**, Fig. 1) stabilized the open conformation.

Importantly, using the results of our smFRET study as the basis for data interpretation, we show how more easily accessible approaches such as ^{19}F NMR spectroscopy or differential scanning fluorimetry (DSF) and nanoDSF can be used to study ligand-protein interactions of the ZIKV protease in solution.

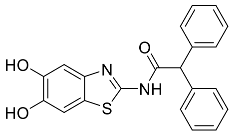
2. Results

Preparation and activity of a ZIKV protease quintuple mutant. To selectively label the ZIKV protease for smFRET at the desired positions, its three native cysteine residues had to be replaced and two new cysteine residues introduced, thus creating a ZIKV quintuple construct (C80S, C143S, C178S, S84*C, S160C, ^{5}ZiPro). To minimize disruption of the protein structure, replacement of cysteine residues with serine and *vice versa* was chosen. Similar to what was shown for DENV NS2B-NS3 [30], the reporters in ZIKV NS2B-NS3 should change their distance significantly when a conformational change occurs between open and closed conformations, reflected in changes in the FRET intensity (Fig. 2). To ensure the functional integrity of the mutant, the proteolytic activity (k_{cat}) and substrate affinity (K_M) of ^{5}ZiPro were compared with the WT protease. A fluorometric activity assay with a Boc-Gly-Arg-Arg-AMC substrate (Fig. 1) was used to detect the fluorescence increase upon proteolytic release of AMC by an active protease. The K_M values determined for ^{5}ZiPro (K_M = 794 ± 115 μM) and the WT enzyme (K_M = 641 ± 20 μM) under the same buffer conditions are very similar (Table 1). However, the k_{cat} of ^{5}ZiPro is lower compared to the WT (Table 1). Hence, to ascertain the structural integrity of ^{5}ZiPro, circular dichroism (CD) spectroscopy was used (Fig. 3). No significant differences in the secondary structure content of the protease variants were observed in the absence or presence of reducing agents, and the expected high β-sheet content of the protein could be confirmed (Fig. 3 b–e) [43].

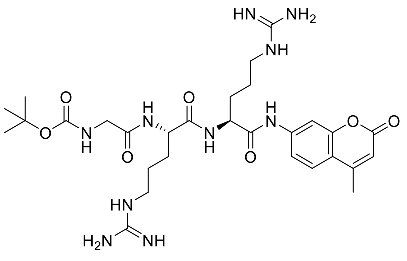
Labeling of ^{5}ZiPro for smFRET. The ^{5}ZiPro mutant was stochastically labeled with maleimide linker-containing ATTO 488 and ATTO 643 dyes. Similar to the unlabeled protein, the catalytic activity and the substrate binding affinity of the labeled protease agreed with the values determined for the WT protein (Table 1). The degree of labeling was determined by measuring the absorbance at 280 nm (protein), 500 nm



I$_c$
IC$_{50}$ (WT) = 0.98 ± 0.04 μM
IC$_{50}$ (^{5}ZiPro) = 3.14 ± 0.19 μM



I$_a$
IC$_{50}$ (WT) = 34.3 ± 4.0 μM
IC$_{50}$ (^{5}ZiPro) = (52.6 ± 0.92) % @ 100 μM



Substrate
Boc-Gly-Arg-Arg-AMC

Fig. 1. Structures of the competitive inhibitor **I$_c$** [41], and the allosteric inhibitor **I$_a$** [42] as well as the fluorogenic-substrate for enzyme activity assays used in this study.

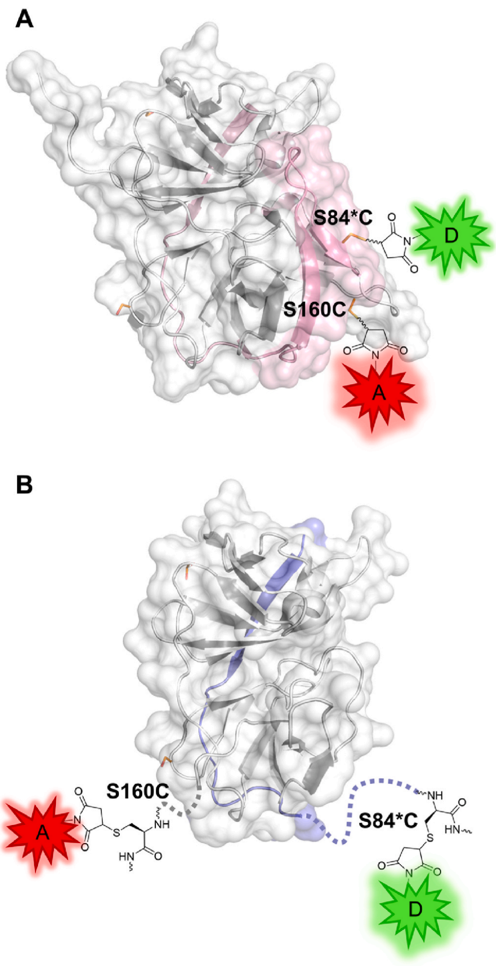


Fig. 2. Closed (**a**) and open (**b**) conformation of the ZIKV protease quintuple mutant and relative smFRET label positions. NS3^{pro} is shown in white; NS2B of the closed conformation is colored light red or light blue for the open conformation. Mutated positions are shown as orange sticks. Dyes are shown schematically as stars. (**a**) closed conformation (PDB: 5LC0). (**b**) open conformation (PDB: 5GXJ). Neither the positions of the introduced cysteine in NS3^{pro} nor the introduced cysteine in NS2B^{cf} are resolved in the used X-ray structure of the open conformation. Missing amino acids are indicated by a dashed line. Graphic created with pyMOL [44], ChemDraw and Microsoft PowerPoint.

Table 1
K_M and k_{cat} values as well as the catalytical efficiency of the wildtype protease measured by us for the current study (WT_{TCEP}) and reported previously (WT_{lit}), and of labeled variants used in this study. Given is the mean value from three independent measurements and its standard deviation.

Protease	K_M [µM]	k_{cat} [s^{-1}]	$k_{cat} \bullet K_M^{-1}$ [$M^{-1} \bullet s^{-1}$]
WT_{lit}^{43}	558 ± 62	$2.6 \bullet 10^{-2} \pm 1.4 \bullet 10^{-3}$	47.2 ± 5.8
WT_{TCEP}	641 ± 20	$2.4 \bullet 10^{-2} \pm 1.1 \bullet 10^{-3}$	37.5 ± 2.1
$^5ZiPro_{TCEP}$	794 ± 115	$1.4 \bullet 10^{-4} \pm 1.4 \bullet 10^{-5}$	0.18 ± 0.03
$^5ZiPro^{ATTO488/ATTO643}{}_{TCEP}$	261 ± 20	$1.8 \bullet 10^{-4} \pm 0.8 \bullet 10^{-5}$	0.70 ± 0.07
^{19}F-Trp $eZiPro_{TCEP}$	885 ± 56	$5.3 \bullet 10^{-2} \pm 3.2 \bullet 10^{-3}$	59.6 ± 5.2
$eZiPro_{TCEP}$	1202 ± 111	$2.0 \bullet 10^{-2} \pm 1.6 \bullet 10^{-3}$	16.6 ± 2.1

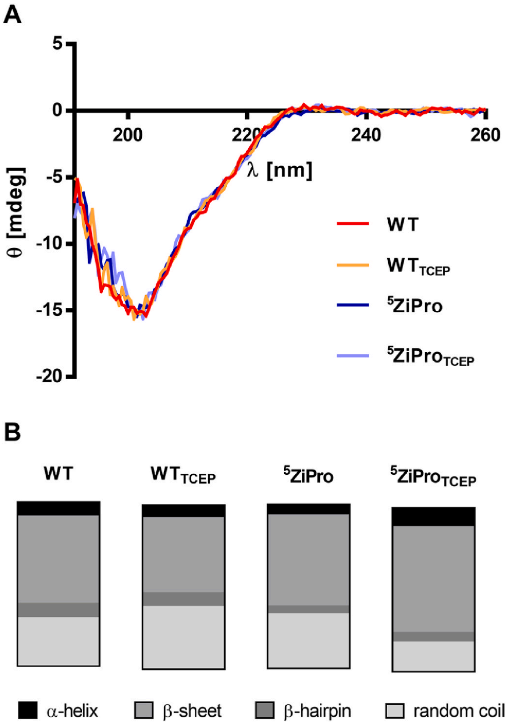


Fig. 3. CD spectra of the ZIKV NS2B-NS3 wildtype protease (WT) and its quintuple mutant (^{5}ZiPro) with and without the addition of a reducing agent (1 mM TCEP) at a protein concentration of 5 µM. (**a**) recorded CD spectra. (**b**) CD spectroscopy based secondary structure prediction of the WT and ^{5}ZiPro with and without reducing agent (1 mM TCEP) determined with BeStSel [45]. This figure was created using GraphPad Prism [46].

(ATTO 488) and 630 nm (ATTO 643) and could be confirmed by SDS-PAGE and an in-gel fluorescence scan at the respective excitation wavelengths (Fig. 4). Relatively low labeling efficiencies were found, with 14% for ATTO 488 and 20% for ATTO 643. Nonetheless these proved to be sufficient for smFRET experiments (see below).

smFRET. ^{5}ZiPro labeled with ATTO 488 and ATTO 643 was excited at 502 nm and the stream of emitted photons from individual proteases diffusing through the excitation volume was recorded. This allowed to identify subpopulations of different conformations of the protease. From the fluorescence bursts, fluorescence intensity time traces with a binning time of 1 ms were calculated. Sections of the intensity time traces for the donor and acceptor dye of the labeled ^{5}ZiPro are shown in Fig. 5.

smFRET of ^{5}ZiPro in the presence of a competitive inhibitor. To assess the influence of ligands on the conformation of the protease, auto- and cross-correlations were determined from the recorded data (Fig. 6).

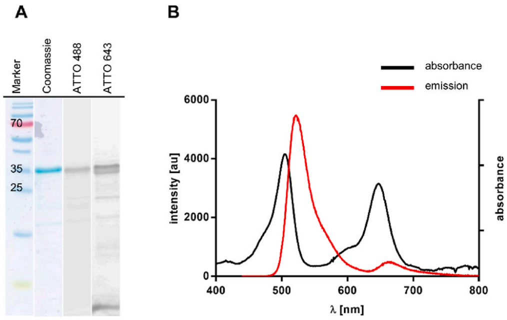


Fig. 4. Successful fluorophore labeling of ^{5}ZiPro. (**a**) SDS-PAGE of ATTO 488/ ATTO 643 FRET pair labeled ZIKV protease. The marker PageRuler™ Prestained Protein Ladder (in kDa), Coomassie stain, ATTO 488 Laser scan with excitation at 500 nm, ATTO 643 Laser scan with excitation at 630 nm. (**b**) Absorption (black) and emission (red) spectra. The latter was measured at 500 nm excitation. This figure was created using Microsoft PowerPoint and GraphPad Prism [46].

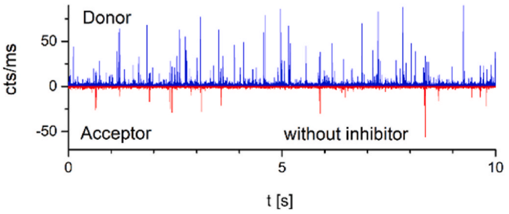


Fig. 5. Section of the fluorescence intensity time trace of ATTO 488/ATTO 643 FRET pair labeled ZIKV protease. The sample was excited with excitation pulses at 502 nm. The binning time was 1 ms. The donor channel is shown in blue, the acceptor channel is shown in red. This figure was created using Origin.

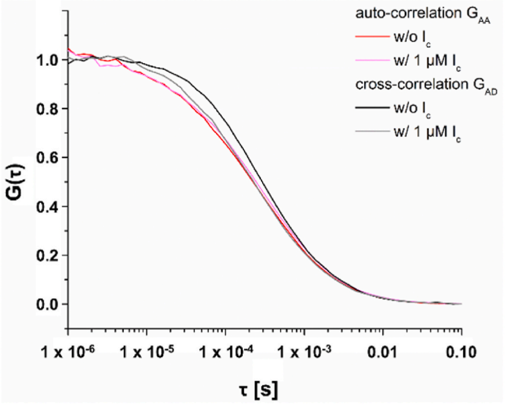


Fig. 6. Auto- (G_{AA}) and cross-correlations (G_{AD}) of the ATTO 488 and ATTO 643 labeled ^{5}ZiPro before and after adding 1 μM of $\mathbf{I_c}$.

After adding the competitive ligand, the cross-correlation function G_{AD} (gray in Fig. 6) approached zero faster than G_{AD} without the inhibitor (black in Fig. 6). Additionally, a burst analysis was performed, resulting in frequency distributions of FRET energy transfer efficiencies (E_{ET}) visualized in 2D histograms. Only bursts that showed a total of at least 20 counts on both avalanche photodiodes (APDs) were considered. The acceptor lifetime τ_A was obtained as the mean delay time of the fluorescence signal of the acceptor dye after excitation without taking the instrumental response function (IRF) into account. τ_A was plotted against E_{ET} from the intensity ratios per burst. Although 2D histograms are available from the donor and acceptor, we focused on the acceptor 2D histograms, since point clouds at $E_{ET} > 0.5$ were more clearly visible in the acceptor 2D histograms [47]. Additionally, counts in the acceptor channel mainly originate from FRET pair labeled protease, while in the donor channel donor-donor labeled protease and Raman scattering contribute as well. Moreover, with an increasing number of acceptor photons within a burst, the statistical significance of the calculated fluorescence lifetimes increases.

In the absence of ligands, a single distinct point cloud was seen in the 2D histogram (Fig. 7 a). The maximum of the corresponding frequency distribution of E_{ET} was around $E_{ET} \approx 0.5$. The addition of 5 μM of $\mathbf{I_c}$ led to the formation of a second point cloud at higher E_{ET} (Fig. 7 b). Accordingly, a second peak was observed in the frequency distributions of E_{ET}. Modifying the concentrations of $\mathbf{I_c}$, the proportion of the peak at higher E_{ET} increased with increasing concentration of the inhibitor (Fig. 8).

smFRET of ^{5}ZiPro in the presence of an allosteric inhibitor. In the absence of ligands, the observed point cloud was at medium E_{ET} (Fig. 7 a). As already described for DENV protease, direct observation of the impact of an allosteric ligand $\mathbf{I_a}$ in the smFRET measurements is not possible when it stabilizes the open state [47], since this is the conformation the protease predominantly populates in our experimental conditions in the absence of ligands. To study the effects of $\mathbf{I_a}$ on the ZIKV protease, a competition assay [47] with both $\mathbf{I_c}$ and $\mathbf{I_a}$ was carried out. First, by adding the competitive ligand, a subpopulation of the protease with high $E_{ET} > 0.7$ was generated (Fig. 9, red line). Afterwards, in the presence of a constant concentration of the competitive ligand, $\mathbf{I_a}$ was added in excess. The proportion of the peak at high $E_{ET} > 0.7$ formed by the addition of $\mathbf{I_c}$ was reduced by the addition of $\mathbf{I_a}$.

Thermal Shift Assay. The thermal shift assay or differential scanning fluorimetry (DSF) is a commonly used approach for detecting protein-ligand interactions [48,49]. By heating a protein of interest, its thermal unfolding transition can be monitored spectrophotometrically in the presence of a fluorescent dye as SYPRO Orange (DSF) [50] or based on intrinsic tryptophan and tyrosine fluorescence (nanoDSF) [51]. Ligands that interact with the protein usually influence the unfolding transition and shift the inflection point of the fluorescence time trace (i.e., the melting temperature T_m) [50]. The simplicity of the method as well as the low protein consumption make it an interesting method for the study of protein-ligand interactions [48,50].

We performed thermal shift assays with SYPRO Orange (DSF) and without dye addition (nanoDSF) with the wildtype ZIKV protease in the absence and presence of $\mathbf{I_c}$ and $\mathbf{I_a}$ (Figs. 10 and 11).

The DSF approach with SYPRO Orange yielded a T_m of 54.29 ± 0.50 °C for the ZIKV protease in the apo state (Fig. 10). The addition of the allosteric ligand increased this temperature to 55.45 ± 0.28 °C, while addition of the competitive ligand reduced the melting temperature to 51.48 ± 1.31 °C. When the first derivative of the measured fluorescence was plotted against the temperature, the differences between the two ligands became even clearer. While the presence of $\mathbf{I_a}$ led to the formation of a shoulder at higher temperatures (Fig. 10, blue arrow), the presence of $\mathbf{I_c}$ gave rise to a peak-shoulder at lower temperatures (Fig. 10, red arrow). Higher concentrations of the competitive inhibitor (100 μM, Fig. 10 g and h) enhanced the shoulder at lower temperatures but did not further decrease the resulting melting temperature (51.48 ± 1.31 °C) compared to what was observed at lower concentrations (10 μM) of $\mathbf{I_c}$.

In contrast to DSF, no dye is needed for nanoDSF. The fluorescence intensity ratio is measured at emission wavelengths of 330 nm and 350

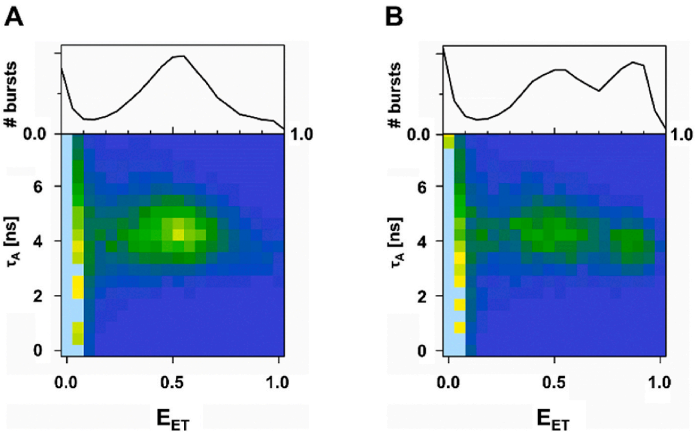


Fig. 7. A plot of the normalized occurrences of individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 643 FRET pair labeled ^{5}ZiPro plotted in a 2D histogram, depending on acceptor lifetime τ_A and FRET efficiency E_{ET}. (**a**) Without competitive inhibitor and (**b**) with competitive inhibitor ($\mathbf{I_c}$ = 5 μM). For an easier visual comparison of the FRET populations before and after the addition of the inhibitor. The respective normalized 1D histograms of FRET efficiencies are shown as a projection on the top of the 2D histograms.

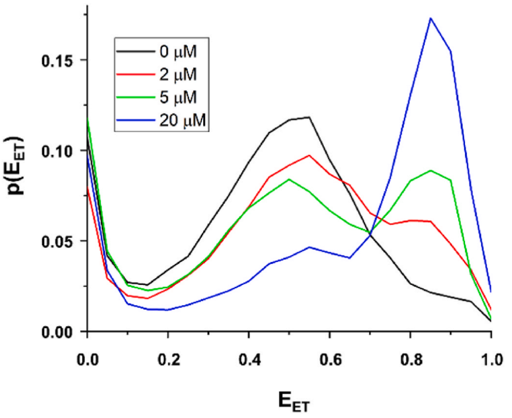


Fig. 8. E_{ET} frequency distributions of ATTO 488/ATTO 643 FRET pair labeled ^{5}ZiPro with different concentrations of the competitive inhibitor.

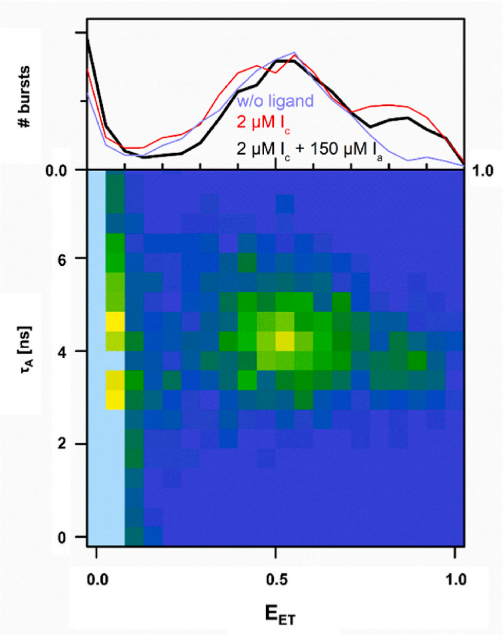


Fig. 9. Plot of the normalized occurrences of the individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 643 FRET pair labeled ^{5}ZiPro in a 2D histogram, separated according to acceptor lifetime τ_A and FRET efficiency E_{ET}. (top) blue: without ligand, red: with 2 μM $\mathbf{I_c}$, black: with 2 μM $\mathbf{I_c}$ and 150 μM $\mathbf{I_a}$. (bottom) 2D histogram: with $\mathbf{I_c}$ and $\mathbf{I_a}$.

nm and plotted against the temperature [52]. Analogously to DSF, the melting point can be derived from the inflection point of the resulting sigmoidal curve and yielded a value of 44.03 ± 0.01 °C for the apo protease (Fig. 11 a). The presence of $\mathbf{I_a}$ led to a lowering of the melting temperature to 43.35 ± 0.03 °C (Fig. 11 a and b, blue line). $\mathbf{I_c}$ (10 μM), on the other hand, led to a shift to a higher melting temperature of 44.86 ± 0.01 °C (Fig. 11 a and b, light red line), which is further increased to 46.40 ± 0.01 °C by higher concentrations (100 μM, Fig. 11 a and b, red line). Concentration dependency was also observed for $\mathbf{I_a}$. Plotting the melting temperatures against the concentration resulted in a saturation curve which allowed to estimate a K_D value for $\mathbf{I_a}$ of 42.1 ± 1.56 μM which is in the same range as the IC_{50} determined under identical buffer conditions (34.3 ± 4.0 μM).

^{19}F-NMR. Fluorine has become a valuable reporter to probe interactions between biomacromolecules and to investigate protein structural dynamics by NMR spectroscopy [53–55]. A fluorine reporter can be incorporated post-translationally through the labeling of cysteine residues [56,57] or the co-translational insertion of fluorinated amino acids such as fluoro-tryptophan [57,58]. The ZIKV protease contains six tryptophan residues; five within the core NS3 protease domain and one within the NS2B^{cf} (Fig. 11 b). Qianzhu et al. successfully used site-specific incorporation of the unnatural amino acid 7F-Trp into ZIKV NS2B-NS3 by genetic code expansion, i.e. the introduction of an amber stop codon recognized by a mutant aminoacyl-tRNA synthetase that site-specifically incorporates 7F-Trp for NMR studies [58]. Here, we

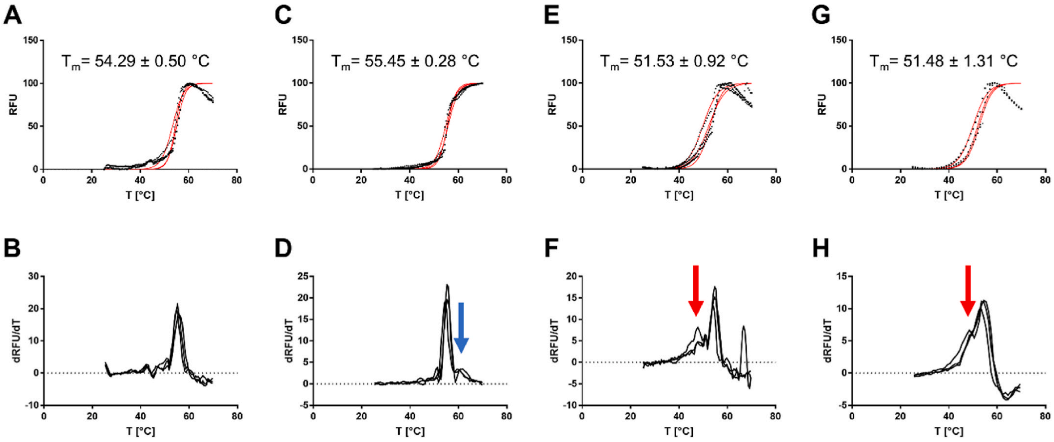


Fig. 10. Melting curves of ZIKV WT determined by DSF. Three independent measurements (black) and their corresponding sigmoidal fit curves (red) are shown. (**a, b**) measured fluorescence in the temperature range from 20 °C to 70 °C for the ZIKV protease without ligand and the corresponding 1st derivative (**c, d**) melting curve of the WT with 200 μM of $\mathbf{I_a}$ and 1st derivative. The blue arrow shows the shoulder that appears at a higher temperature with $\mathbf{I_a}$. (**e, f**) melting curve of the WT with 10 μM of the $\mathbf{I_c}$ and 1st derivative. (**g, h**) melting curve of the WT with 100 μM of the $\mathbf{I_c}$ and 1st derivative. The red arrows show the shoulder in the first derivative of the fluorescence signal at lower temperatures caused by $\mathbf{I_c}$.

aimed to simplify this approach by using uniform fluoro-tryptophan labeling thus circumventing mutagenesis, dedicated expression strains or the use of specific tRNA synthetases. Instead, uniform incorporation of the labeled amino acid can be achieved by including it in the bacterial growth medium [57]. Since it has been shown that the fluorine substitution position in the indole ring affects the chemical shift range and sensitivity to ligands in a site-specific manner [59], we incorporated 5F-Trp into ZIKV NS2B-NS3 thus complementing the available data with 7F-Trp (Fig. 11 a). Activity and binding affinity for the substrate (Fig. 1) of the labeled protease were shown to be similar to that of the wildtype under the same buffer conditions (Table 1).

In the ^{19}F NMR spectrum of the labeled protease, six peaks of differing linewidth were observed (Fig. 12 a), displaying an overall broader dispersion than previously seen for the 7F-Trp labeled samples [58]. Using single tryptophan point mutations, a subset of residues could be tentatively assigned (Fig. S3). As expected for a highly flexible residue, W5 at the N-terminus of NS3 gives rise to the narrowest peak (Fig. 12). This agrees with the fact that this residue could not be resolved in available crystal structures of NS2B-NS3 (Fig. 12 b). Of note, ZIKV NS2B-NS3 W83F and W89F could be expressed and purified, but were catalytically inactive, precipitated under the NMR conditions and yielded aberrant NMR spectra. The tentative position of the ^{19}F chemical shifts for these residues in the WT protease can thus only be inferred through the process of exclusion.

While the addition of the inhibitors caused only subtle effects on the spectra, the addition of $\mathbf{I_c}$ caused a chemical shift perturbation for the W50 residue. This tryptophan resides directly adjacent to the active site and interacts with the hairpin of NS2B in the closed conformation (Fig. 12 b). Notably, this residue's chemical shift was not affected by $\mathbf{I_a}$. This finding is also in line with the previous study using 7F-Trp labeled protease, where 4-nitrophenyl-4-guanidinobenzoate was used as a generic competitive inhibitor and found to strongly affect W50 [58]. This shows that in the absence of an elaborate, site-specific labeling scheme, uniform labeling allows distinguishing between the tryptophan resonances and that W50 can be exploited in a straight-forward manner to inform on the nature of a ligand bound to ZIKV NS2B-NS3.

3. Discussion

In this study a quintuple mutant of ZIKV protease (^{5}ZiPro) was successfully generated for smFRET measurements, A fluorometric peptide cleavage assay and CD spectroscopy confirmed its functional and structural integrity. Importantly, this shows that the ZIKV protease is relatively tolerant towards mutagenesis, an important requirement for many biophysical methods. We showed that easily accessible methods such as DSF and nanoDSF are valuable tools to characterize the protease behavior and ligand binding and that a simplified NMR spectroscopic approach using uniform ^{19}F–5F-Trp labeling can be used to derive site-specific information on inhibitor binding.

The time-resolved smFRET data allowed to paint a detailed picture of the structural consequences of the interaction of ZIKV NS2B-NS3 protease with competitive and allosteric inhibitors. Although the fluorescent dye labeling efficiency was rather low, presumably due to the formation of disulfide bridges and/or reduced solvent accessibility of the respective labeling sites, smFRET experiments could nonetheless be carried out with a sufficiently increased protein concentration. The faster decrease in the cross-correlation function G_{AD} after addition of $\mathbf{I_c}$ compared to G_{AD} without inhibitor showed that one conformation of the protease was stabilized. However, assigning the actual state was impossible without additional information, thus highlighting the importance of a multi-pronged methodological approach. Based on available data from a crystal structure with $\mathbf{I_c}$ in the closed conformation [41], it was concluded that the conformation stabilized by $\mathbf{I_c}$ in the smFRET experiment was the closed one, which was later confirmed with the help of 2D histograms as E_{ET} can be used as a nanoruler to measure distances [60]. Without a ligand, a single distinct point cloud at intermediate E_{ET} was seen in the 2D histogram (Fig. 7 a), showing that the protease is predominantly in the open conformation in the absence of a ligand. $\mathbf{I_c}$ stabilized the closed conformation, while the allosteric inhibitor favored the open conformation, a process that occurred in a concentration dependent manner.

Notably, the behavior of ZIKV NS2B-NS3 resembles that of DENV NS2B-NS3,[28,47] which shows that it is indeed possible to draw meaningful conclusions about one flaviviral protease based on the studies of another. While a similar dynamic behavior is usually inferred for

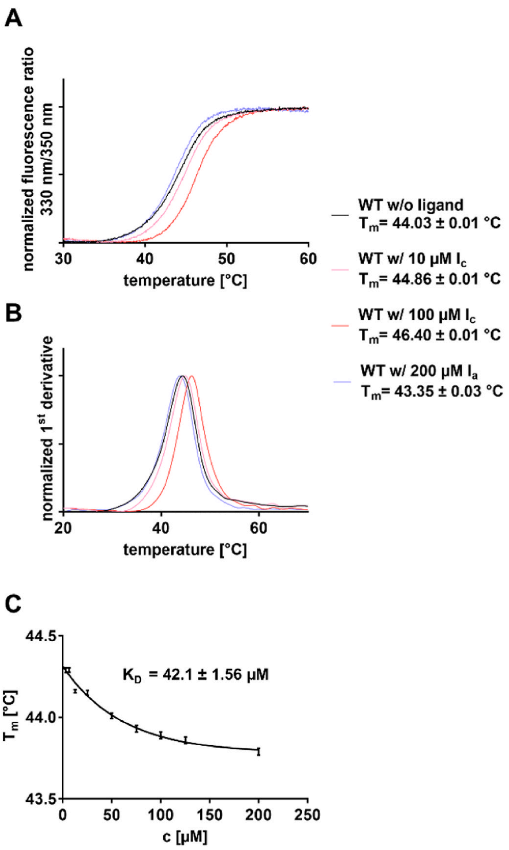


Fig. 11. Thermal shift assay performed by nanoDSF. Shown is the mean of three independent measurements. (**a**) Melting curves of the WT determined by nanoDSF with and without competitive and allosteric inhibitors. black: without ligand, pink: 10 µM of $\mathbf{I_c}$, red: 100 µM of $\mathbf{I_c}$ and blue: 200 µM of $\mathbf{I_a}$. (**b**) first derivative of the measured fluorescence ratio. black: without ligand, pink: 10 µM of $\mathbf{I_c}$, red: 100 µM of $\mathbf{I_c}$ and blue: 200 µM of $\mathbf{I_a}$. (**c**) Plot of melting temperatures versus concentration of $\mathbf{I_a}$. $K_D = 42.1 \pm 1.6$ µM.

structurally closely related proteins, this nonetheless needs to be shown unambiguously for each protein. Despite its high information content, sample preparation for smFRET and the subsequent data analysis are challenging. Furthermore, not all laboratories have easy access to the instrumental equipment. Thus, we applied other, more accessible methodologies to observe whether the allosteric and competitive inhibitor have conformational consequences. The results from the smFRET study were used to guide the respective interpretation.

Using thermal shift assays, opposite effects of the inhibitors were observed. Together with the smFRET results, it can be speculated that the respective in- or decrease in melting temperature caused by the different ligands is a consequence of stabilizing two different protease conformations. It should be noted that SYPRO Orange used for DSF may associate with hydrophobic regions as present in the allosteric binding site of NS2B-NS3 (Fig. S1), thereby possibly affecting the protein's conformational equilibrium [13,28,48,61]. Other disadvantages of SYPRO Orange are its susceptibility to non-protein induced fluorescence such as EDTA aggregates, plastic surfaces, lipids, glycerol, and detergents [48,62,63]. In these cases, nanoDSF, which is based on the intrinsic fluorescence of aromatic amino acids, does not require a dye reporter and may be the more suitable option.

In ^{19}F spectroscopy, the shift of W50 allowed pinpointing the respective binding mode of the ligands. In contrast to the open conformation, W50 was seen to adopt a flipped orientation in crystal structures of the closed conformation, thus directly interacting with the NS2B hairpin (Fig. S2) [41]. All other tryptophan residues are not or only marginally affected by the open/closed transition (Fig. S2). Hence, the W50 residue should be particularly well suited to inform on conformational transitions from the open to the closed state. Indeed, the resonance of W50 showed a clear shift upon addition of $\mathbf{I_c}$, but not $\mathbf{I_a}$. As demonstrated by smFRET, the apo protease is predominantly present in the open conformation. This conformation is further stabilized by $\mathbf{I_a}$, thereby providing an explanation for the lack of response in the NMR expermiments.

4. Conclusion

With this work, we were the first to investigate the effects of an allosteric ligand on the conformational dynamics of ZIKV protease using various biophysical methods. Together, smFRET, (nano)DSF and ^{19}F NMR, showed differential effects of allosteric and competitive inhibitors for the ZIKV protease conformational dynamics. This study contributes to a deeper understanding of the interactions of different inhibitor classes on the ZIKV protease, thereby providing a basis for developing of potent inhibitors.

5. Material and methods

5.1. Expression constructs and site-directed mutagenesis

For the generation of the ZIKV protease constructs, a pET11a vector harboring the sequence of the French Polynesia ZIKV strain (GenBank ID: KJ776791.2, NS2B^{cf}: aa 1423–1467 and NS3^{pro}: aa 1503–1688) [40] as used by us before [64] served as a starting point. Both domains are covalently linked by a glycine-rich linker (G_4SG_4) and an R95A mutation was inserted into NS2B to suppress autocatalytic *cis*-cleavage (glycine-linked Zika protease: gZiPro) [43]. The construct contains an N-terminal hexahistidine (His_6) tag followed by a tobacco etch virus protease (TEV)-cleavage site. The ZIKV NS2B/NS3 ^{5}ZiPro construct was obtained by a one-pot site-directed mutagenesis approach using the Kapa HiFi PCR kit (Kapa Biosystems Inc. Woburn, MA, USA) with gZiPro as template and the following ten primers (aa in NS2B are given with an *): (i) C80S: 5′–CTACTCCGGTCCGTGGAAACTGG–3′ and 5′–GGACCGGAGTAGCTAACCAGGTC–3′, (ii) C143S: 5′–GGACAAATCCGGCCGTGTTATTGG–3′ and 5′–CGGCCGGATTTGTCCAGAATCGG–3′, (iii) C178S: 5′–GTGGAGTCCTTTGAACCGAGCATG–3′ and 5′–GTTCAAAGGACTCCACCGGGGTTTC–3′, (iv) S84*C: 5′–GTGACTTCTGCCTGGTTGAAGACG–3′ and 5′–CCAGGCAGAAGTCACCGCTTTCG–3′, and (v) S160C: 5′–GAACGGTTGCTACGTTAGCGCGATC–3′ and 5′–CTAACGTAGCAACCGTTCTTGATCACAAC–3′. PCR was performed with annealing temperatures increasing about 0.33 °C each cycle ranging from 65 to 75 °C. Besides several variants not harboring all desired mutations, the correct sequence could be obtained from ~15 clones. For the enzymatically cleavable construct, the glycine linker is replaced by the native cleavage site of NS2B, namely a Glu (aa 1468) and the native cleavage site KTGKR of NS2B/NS3 (aa 1498–1502), similar to a previous eZiPro denoted construct [65]. Here, the following primers were used: (vi) 5′–GAAAAAACCGGCAAACGCAGCGGCGCGCTGTGGGACGTGCCG–3′ and 5′–GCGTTTGCCGGTTTTTTCACGCATCGGCGGACCATCGTCTTCAACC–3′. For the signal assignment of ^{19}F NMR experiments, the single

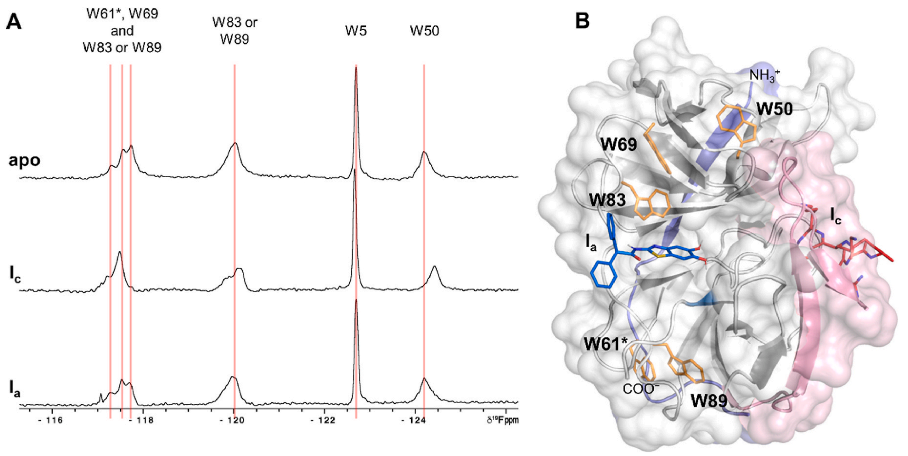


Fig. 12. Sensing of inhibitor binding by 5F-Trp and their respective positions. (**a**) ^{19}F NMR spectra of the 5F-Trp labeled ZIKV protease with and without ligands. from up to down: spectrum of the apo protease; spectrum of the protease with 100 μM of $\mathbf{I_c}$; spectrum of the protease with 100 μM of $\mathbf{I_a}$. The tryptophan residue within NS2B^{cf}, W61*, is marked with an asterisk. (**b**) Structure of the ZIKV protease (PDB: 5GXJ and 6Y3B) with the position of the native tryptophan residues highlighted in orange. The docking pose of $\mathbf{I_a}$ is shown in blue. $\mathbf{I_c}$ is shown in red. NS2B is shown in light red (closed conformation) and light blue (open conformation). For a clear view, only NS3 of the open conformation (PDB: 5GXJ) is shown in white. Residue W5 in the NS3 N-terminus was not resolved in these structures and is thus not shown.

W/F mutants of gZiPro as well as eZiPro were generated using the primers: (vii) W61$^{\#}$F 5′–GACATTACCTTTGAAAAGGATGCGGAAGTTACC–3′ and 5′–CTTTTCAAAGGTAATGTCGCCCGCACGC–3′, (vii) W5F: 5′–CGCTGTTTGACGTGCCGGCGCC–3′ and 5′–GCACGTCAAACAGCGCGCCGCTAC–3′, (ix) W50F: 5′–CACCATGTTTCACGTTACCAAGGGTAGCG–3′ and 5′–GTAACGTGAAACATGGTGTGGAAAACACCTTC–3′, (x) W69F: 5′–CCGTATTTTGGTGATGTTAAGCAGGACCTG–3′ and 5′–CATCACCAAAATACGGGTCCAGACGACC–3′, (xi) W83F: 5′–CGGTCCGTTTAAACTGGACGCGGCG–3′ and 5′–CCAGTTTAAACGGACCGCAGTAGCTAACC–3′, (xii) W89F: 5′–CGGCGTTTGATGGTCACAGCGAAGTG–3′ and 5′–GTGACCATCAAACGCCGCGTCCAG–3′.

5.2. Recombinant protein expression and preparation

The glycine-linked ZIKV NS2B^{cf}-NS3^{pro} (gZiPro, WT), the quintuple mutant (^{5}ZiPro) of gZiPro, which is suitable for smFRET experiments, and the enzymatically cleavable linked ZIKV NS2B-NS3 protease (eZiPro), with the respective W/F single mutants for the assignment of the tryptophan residues in ^{19}F NMR experiments, were expressed and purified as described previously [42,66]. In brief, the respective expression vectors (pET11a) were transformed into calcium chloride competent *Escherichia coli* (*E.* coli) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). After reaching an optical density (OD_{600}) of ~0.8 in LB medium containing 100 mg L^{-1} ampicillin at 37 °C, overexpression was induced by the addition of 1 mM isopropyl-*β*-ᴅ-thiogalactoside (IPTG) at 18 °C for ~20 h. Cell pellets were harvested by centrifugation at 9000 rpm at 4 °C for 15 min. For purification, cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.1 vol% Triton$_{X\text{-}100}$, RNase, DNase, lysozyme, 1 mM DTT) and lysed by 10 cycles à 45 s of sonication at 60% power under constant cooling with ice (Sonoplus HD 2200; Bandelin, Berlin, Germany). The crude lysate was cleared by centrifugation (20,000 rpm at 4 °C for 1 h) and purified by immobilized metal affinity chromatography (IMAC) using a HisTrap HP 5 ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany) on an ÄKTA start FPLC device (Cytiva). After a washing step with 5 column volumes (CV) of IMAC buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole), elution of the His$_6$ tagged protein was performed with a linear gradient in IMAC buffer B with 250 mM imidazole. Eluted NS2B-NS3 was concentrated using Vivaspin 10 MWCO spin concentrators (Sartorius AG, Gottingen, Germany) and subjected to a size exclusion chromatography step (HiLoad 16/600 Superdex 75 column, GE Healthcare) in SEC buffer A (20 mM Tris-HCl, pH 8.0, 150 mM NaCl) or SEC buffer B for the ^{5}ZiPro (with additional 1 mM DTT). Fractions containing NS2B/NS3 were concentrated, flash-frozen in liquid nitrogen, and stored at −80 °C. The purity and identity of NS2B/NS3 were confirmed by SDS-PAGE with Coomassie blue staining and subsequent fluorometric enzyme activity assays.

For ^{19}F NMR labeled proteins, *E. coli* BL21 Gold (DE3) cells (Agilent Technologies) were grown in LB media as described above. After an OD_{600} of ~0.6 was reached, cells were pelleted, washed and resuspended in a defined expression medium composed of all 20 aa, but containing only 10% of non-fluorinated tryptophane and 90% of the labeled amino acid (4-fluoro-tryptophane, 5-fluoro-tryptophane, 6-fluoro-tryptophane) [67]. Purification was performed as described above but with size exclusion chromatography directly with the NMR buffer (50 mM Tris-HCl pH 9.0, 20 mM NaCl, 1 mM Chaps) instead of the SEC buffer A. After a set of benchmark experiments consisting of expression, purification, and ^{19}F NMR studies with gZiPro and eZiPro, 5-fluoro-tryptophan and eZiPro were chosen for subsequent ^{19}F NMR experiments based on expression levels, the solubility of the protease, quality of the ^{19}F NMR spectra and the sensitivity to inhibitor treatment of the ^{19}F NMR spectra (data not shown).

5.3. Fluorometric assay

Fluorometric assays were performed as described [42]. Briefly, 25 nM–160 nM of the purified protease in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS) with 5 μL substrate (100 μM Boc-Gly-Arg-Arg-AMC) in DMSO and 10 μL of the corresponding inhibitor in DMSO resulting in a total volume of 200 μL was added to a 96 well plate. The fluorescence of the released AMC was measured at 380 nm excitation and 460 nm emission. The percentage activity of the protease with the addition of the inhibitors was determined as the proportion of the slope with respect to the slope of the DMSO control. 1 mM TCEP or 5 mM DTT were used as reducing agents.

5.4. CD spectroscopy

Circular dichroism measurements were performed on a JASCO-1500 spectrometer from JASCO (Tokyo, Japan) at 25 °C with an enzyme concentration of 5 μM in buffer (10 mM Tris pH 9.0). 1 mM TCEP was used as a reducing agent. The scan speed was 100 nm/min and the bandwidth and scan intervals were set to 5 nm and 1 nm, respectively. For each sample, 84 measurements were recorded in aquartz-cuvettes with 1 mm pathlength (Hellma GmbH & Co. KG, Müllheim, Germany) from which mean values were calculated. Proportions of different secondary structural features were determined from the spectra using BeStSel (https://bestsel.elte.hu/index.php) [45].

5.5. Labeling for smFRET

The fluorescence labeling was performed analogously to Götz et al. and Maus et al. using ATTO 488 as donor and ATTO 643 as acceptor dyes to label ^{5}ZiPro [28,47]. In brief, a buffer (20 mM Tris-HCl pH 6.8, 150 mM NaCl, 1 mM CHAPS) containing 0.5 mM TCEP as a reducing agent was used. TCEP was removed after an incubation time of 30 min at 4 °C by rebuffering using spin concentrators. A 2.1-fold excess of ATTO 488 and a 2.6-fold excess of ATTO 643, both dissolved in DMF, were then added to the protein solution, and incubated for 2 h at room temperature. The labeled protease was dialyzed and purified again by SEC to remove unreacted dye-labels.

5.6. smFRET experiments

The smFRET experiments were performed as described previously [28,47] using a self-made sample cell consisting of a polyethylene glycol-coated glass coverslip and a glued-on plastic cylinder. Samples contained 150 μL of FRET-pair labeled protease (c ~100 pM) in buffer with 10 vol% DMSO. Fluorescence photons were collected with a custom built confocal microscope over a period of 1800 s. Concentration of the protease and buffer conditions were kept constant for measurements in direct sequence. Fluorophores were excited with a spectrally filtered output from a pulsed white light fiber laser (10 MHz, SC OEM, YSL Photonics, Wuhan, China). Excitation pulses were centered around 502 nm by using an acousto-optical tunable filter (AOTF-VIS-DR, Fianium, Southampton, United Kingdom). Excitation and emission beams were separated by a dichroic mirror (ZT491 rdcxrxt-UF1, Chroma Systems Solutions Inc., Lake Forest, CA, USA). Emitted fluorescence light was spectrally separated by a dichroic mirror (ZT640-rdc-UF1, Chroma Systems Solutions Inc.) into red light and light of higher energy. Both beams were then focused onto two APDs (acceptor channel A: SPCM-AQRH-15, PerkinElmer Inc., Waltham, MA, USA; donor channel D: PDM 50 ct, Micro Photon Devices S.r.l., Bolzano, Italy). The absolute and relative arrival times (relative to the excitation pulse) of the individual photons were detected by a HydraHarp 400 module (PicoQuant GmbH, Berlin, Germany) which was connected to the two detector APDs. Further details on smFRET experiments can be found in the Supplementary Material.

5.7. Analysis of smFRET data

The data were analyzed as described before [28,47]. Cross-correlation functions $G_{AD}(\tau)$ (FCS-FRET) were calculated from the arrival times of the individual photons. Individual bursts extracted from the fluorescence intensity time traces (bin time = 1 ms) were analyzed. Bursts with at least 20 counts on both APDs were considered for further evaluation. Average arrival times τ_A of the acceptor photons relative to the excitation pulse (without taking the IRF into account) were calculated for each individual burst. The IRF was recorded at the beginning of each measurement day. The full width at half-maximum height (FWHM) was about 770 ps. The FRET efficiency E_{ET} was calculated for each burst:

$$E_{ET} = \frac{counts(A)}{counts(A) + counts(D)}$$

The individual bursts were sorted with respect to their E_{ET} values and presented in frequency distributions. Details on the analysis of smFRET data can be found in the Supplementary Material.

5.8. DSF experiments

The DSF experiments were performed in quadruplicate on the openDSF instrument [68]. Measurements were performed in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP) containing 10 vol% DMSO. The test solution contained 5 μM protein and 10 or 100 μM of the corresponding inhibitor in DMSO as well as 5x SYPRO Orange (10 μM). The samples were heated from 20 °C to 80 °C with a heating rate of 2 °C/min. The measured fluorescence was plotted as a function of the respective temperature and the melting temperature was calculated from the inflection point of the resulting sigmoidal fluorescence curve.

5.9. nanoDSF experiments

Thermoshift assays were carried out in triplicate on the Prometheus NT.48 nano DSF instrument (NanoTemper Technologies GmbH, München, Germany) using the manufacturer-designated capillaries for the instrument. Sample solutions contained 5 μM protein in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP) and 10 vol% DMSO. In the capillaries, the sample solutions were heated from 20 °C to 80 °C at a heating rate of 1.5 °C/min, and fluorescence was recorded at 330 nm and 350 nm. Excitation was carried out at 280 nm. The measured fluorescence ratio of the detected fluorescence at 330 nm and 350 nm was plotted as a function of temperature using the GraphPad Prism 7.04 program [46]. The melting temperature was calculated as the inflection point of the resulting sigmoidal curve.

5.10. ^{19}F NMR experiments

For ^{19}F NMR experiments, labeled proteins were concentrated with Vivaspin 10 MWCO spin concentrators (Sartorius) and diluted in NMR buffer (50 mM Tris-HCl pH 9.0, 20 mM NaCl, 1 mM Chaps) to the same stock concentration. In the NMR sample, a final concentration of 100 μM protein with 25 μL of DMSO-d_6 or inhibitors dissolved in DMSO-d_6 (final concentration: 100 μM), 10 vol% D_2O, and 0.01% TFA as an internal standard was used. ^{19}F NMR spectra were recorded on a Bruker AVANCE 3600 MHz spectrometer equipped with a Prodigy TCI cryoprobe (Bruker AXS GmbH, Karlsruhe, Germany). ^{1}H and ^{19}F are measured in the same coil (and thus without proton decoupling during ^{19}F data acquisition). All measurements were carried out at 298 K with 8192 scans and referenced to TFA. Spectra were processed using Bruker TopSpin 4.0.8.

Supporting information

The Supporting Information contains:

1. Tryptophan residues in Open and Closed Conformation of the ZIKV Protease; 2. Assignment of Tryptophan residues using W/F point mutants; 3. Fluorescence Labeling of the 5ZiPro; 4. smFRET Experiments; 5. Analysis of smFRET Data; 6. Fluorometric Enzyme Assay; 7. KM Curves.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Victor Hugo Perez Carrillo reports financial support was provided by German Research Foundation.

Data availability

Data will be made available on request.

Acknowledgments

We thank Prof. Torsten Steinmetzer for providing us with the competitive inhibitor.

We thank Prof Heermann's group for providing the nanoDSF. In particular, we thank Friederike Pisarz for the instruction on device usage.

We thank Luca M. Lauth for the generation of the eZiPro mutant.

We thank Dr. Nadja Hellmann for her support in recording the CD spectra.

VHPC acknowledges a DAAD-CONACYT PhD fellowship. Supported by the DFG (Deutsche Forschungsgemeinschaft) under Germany's Excellence Strategy - EXC 2051 - Project ID 390713860 (to UAH).

We thank our bachelor students Olivia R. Englert and Valerie M. L. Kadenbach (Johannes Gutenberg-University, Mainz) for their experimental support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmech.2023.115573.

References

[1] H. Harapan, A. Michie, M. Mudatsir, et al., Epidemiology of dengue hemorrhagic fever in Indonesia: analysis of five decades data from the national disease surveillance, BMC Res. Notes 12 (1) (2019) 350, https://doi.org/10.1186/s13104-019-4379-9.

[2] V. Sharma, M. Sharma, D. Dhull, et al., Zika virus: an emerging challenge to public health worldwide, Can. J. Microbiol. 66 (2) (2020) 87–98, https://doi.org/10.1139/cjm-2019-0331.

[3] W.H.O. Prevention, Control of Dengue and Dengue and Haemorrhagic Fever. *New Delhi World Heal. Organ. South East Asia Reg*, Off, 2011.

[4] M.A.M. Behnam, C. Nitsche, S.M. Vechi, et al., C-terminal residue optimization and fragment merging: discovery of a potent peptide-hybrid inhibitor of dengue protease, ACS Med. Chem. Lett. 5 (9) (2014) 1037–1042, https://doi.org/10.1021/ml500245v.

[5] C. Nitsche, Proteases from dengue, west nile and Zika viruses as drug targets, Biophys. Rev. 11 (2) (2019) 157–165, https://doi.org/10.1007/s12551-019-00508-3.

[6] M. Yildiz, S. Ghosh, J.A. Bell, et al., Allosteric inhibition of the NS2B-NS3 protease from dengue virus, ACS Chem. Biol. 8 (12) (2013) 2744–2752, https://doi.org/10.1021/cb400612h.

[7] H.A. Rothan, H.C. Han, T.S. Ramasamy, et al., Inhibition of dengue NS2B-NS3 protease and viral replication in vero cells by recombinant retrocyclin-1, BMC Infect. Dis. 12 (1) (2012) 314, https://doi.org/10.1186/1471-2334-12-314.

[8] S.M. Amberg, A. Nestorowicz, D.W. McCourt, et al., NS2B-3 proteinase-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies, J. Virol. 68 (6) (1994) 3794–3802, https://doi.org/10.1128/jvi.68.6.3794-3802.1994.

[9] C. Nitsche, S. Holloway, T. Schirmeister, et al., Biochemistry and medicinal chemistry of the dengue virus protease, Chem. Rev. 114 (22) (2014) 11348–11381, https://doi.org/10.1021/cr500233q.

[10] G. Gupta, L. Lim, J. Song, NMR and MD studies reveal that the isolated dengue NS3 protease is an intrinsically disordered chymotrypsin fold which absolutely requests NS2B for correct folding and functional dynamics, PLoS One 10 (8) (2015), e0134823, https://doi.org/10.1371/journal.pone.0134823.

[11] R. Yusof, S. Clum, M. Wetzel, et al., Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro, J. Biol. Chem. 275 (14) (2000) 9963–9969, https://doi.org/10.1074/jbc.275.14.9963.

[12] A.E. Shannon, K.J. Chappell, M.J. Stoermer, et al., Simultaneous uncoupled expression and purification of the dengue virus NS3 protease and NS2B Co-factor domain, Protein Expr. Purif. 119 (2016) 124–129, https://doi.org/10.1016/j.pep.2015.11.022.

[13] M. Brecher, Z. Li, B. Liu, et al., A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease, PLoS Pathog. 13 (5) (2017), e1006411, https://doi.org/10.1371/journal.ppat.1006411.

[14] C.G. Noble, C.C. Seh, A.T. Chao, et al., Ligand-bound structures of the dengue virus protease reveal the active conformation, J. Virol. 86 (1) (2012) 438–446, https://doi.org/10.1128/JVI.06225-11.

[15] W.-N. Chen, K.V. Loscha, C. Nitsche, et al., The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI, FEBS Lett. 588 (14) (2014) 2206–2211, https://doi.org/10.1016/j.febslet.2014.05.018.

[16] P. Erbel, N. Schiering, A.D. Arcy, et al., Structural basis for the activation of flaviviral NS3 proteases from dengue and west nile virus, Nat. Struct. Mol. Biol. 13 (4) (2006) 372–373, https://doi.org/10.1038/nsmb1073.

[17] H. Yang, X. Chen, X. Ji, Y. Xiong, K. Yang, Crystal Structure of Zika Virus NS2B-NS3 Protease in Apo-form, 2016, https://doi.org/10.2210/pdb5GXJ/pdb.

[18] J. Lei, G. Hansen, C. Nitsche, et al., Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor, Science 353 (6298) (2016) 503–505, https://doi.org/10.1126/science.aag2419.

[19] X. Chen, K. Yang, C. Wu, et al., Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease, Cell Res. 26 (11) (2016) 1260–1263, https://doi.org/10.1038/cr.2016.116.

[20] J. Lei, G. Hansen, C. Nitsche, et al., Crystal structure of Zika virus ns2b-ns3 protease in complex with a boronate inhibitor, Science 353 (6298) (2016) 503–505, https://doi.org/10.1126/science.aag2419.

[21] A. Roy, L. Lim, S. Srivastava, et al., Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants, PLoS One 12 (7) (2017), e0180632, https://doi.org/10.1371/journal.pone.0180632.

[22] J. Evenäs, E. Thulin, A. Malmendal, et al., NMR studies of the E140Q mutant of the carboxy-terminal domain of calmodulin reveal global conformational exchange in the Ca 2+ -saturated state, Biochemistry 36 (12) (1997) 3448–3457, https://doi.org/10.1021/bi9628275.

[23] A.J. Wand, Dynamic activation of protein function: a view emerging from NMR spectroscopy, Nat. Struct. Biol. 8 (11) (2001) 926–931, https://doi.org/10.1038/nsb1101-926.

[24] M. Kauk, C. Hoffmann, Intramolecular and intermolecular FRET sensors for GPCRs – monitoring conformational changes and beyond, Trends Pharmacol. Sci. 39 (2) (2018) 123–135, https://doi.org/10.1016/j.tips.2017.10.011.

[25] L. Zhu, J. Yang, H. Li, et al., Conformational change study of dengue virus NS2B-NS3 protease using 19F NMR spectroscopy, Biochem. Biophys. Res. Commun. 461 (4) (2015) 677–680, https://doi.org/10.1016/j.bbrc.2015.04.090.

[26] T. Ha, A.Y. Ting, J. Liang, et al., Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism, Proc. Natl. Acad. Sci. USA 96 (3) (1999) 893–898, https://doi.org/10.1073/pnas.96.3.893.

[27] Y. Santoso, C.M. Joyce, O. Potapova, et al., Conformational transitions in DNA polymerase I revealed by single-molecule FRET, Proc. Natl. Acad. Sci. USA 107 (2) (2010) 715–720, https://doi.org/10.1073/pnas.0910909107.

[28] C. Götz, G. Hinze, A. Gellert, et al., Conformational dynamics of the dengue virus protease revealed by fluorescence correlation and single-molecule FRET studies, J. Phys. Chem. B 125 (25) (2021) 6837–6846, https://doi.org/10.1021/acs.jpcb.1c01797.

[29] M. Ruer, G. Krainer, P. Gröger, et al., ATPase and protease domain movements in the bacterial AAA+ protease FtsH are driven by thermal fluctuations, J. Mol. Biol. 430 (22) (2018) 4592–4602, https://doi.org/10.1016/j.jmb.2018.07.023.

[30] H. Maus, G. Hinze, S.J. Hammerschmidt, et al., A competition <scp>smFRET</Scp> assay to study ligand-induced conformational changes of the dengue virus protease, Protein Sci. 32 (1) (2023) e4526, https://doi.org/10.1002/pro.4526.

[31] Q. Li, C. Kang, Insights into structures and dynamics of flavivirus proteases from NMR studies, Int. J. Mol. Sci. 21 (7) (2020) 2527, https://doi.org/10.3390/ijms21072527.

[32] Y. Li, Z. Zhang, W.W. Phoo, et al., Structural dynamics of Zika virus NS2B-NS3 protease binding to dipeptide inhibitors, Structure 25 (8) (2017) 1242–1250.e3, https://doi.org/10.1016/j.str.2017.06.006.

[33] Y. Li, Z. Zhang, W.W. Phoo, et al., Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor, Structure 26 (4) (2018) 555–564.e3, https://doi.org/10.1016/j.str.2018.02.005.

[34] M.C. Mahawaththa, B.J.G. Pearce, M. Szabo, et al., Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease, Antivir. Res. 142 (2017) 141–147, https://doi.org/10.1016/j.antiviral.2017.03.011.

[35] Y. Li, W.W. Phoo, Y.R. Loh, et al., Structural characterization of the linked <scp>NS</Scp> 2B- <scp>NS</Scp> 3 protease of Zika virus, FEBS Lett. 591 (15) (2017) 2338–2347, https://doi.org/10.1002/1873-3468.12741.

[36] C. Kang, T.H. Keller, D. Luo, Zika virus protease: an antiviral drug target, Trends Microbiol. 25 (10) (2017) 797–808, https://doi.org/10.1016/j.tim.2017.07.001.

[37] Y.M. Kim, S. Gayen, C. Kang, et al., NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex, J. Biol. Chem. 288 (18) (2013) 12891–12900, https://doi.org/10.1074/jbc.M112.442723.

[38] C. Bräuchle, D.C. Lamb, J. Michaelis, in: C. Bräuchle, D.C. Lamb, J. Michaelis (Eds.), Single Particle Tracking and Single Molecule Energy Transfer, Wiley, Weinheim, 2009, https://doi.org/10.1002/9783527628360.

[39] T. Förster, Zwischenmolekulare energiewanderung und fluoreszenz, Ann. Phys. 437 (1–2) (1948) 55–75, https://doi.org/10.1002/andp.19484370105.

[40] C. Baronti, G. Piorkowski, R.N. Charrel, et al., Complete coding sequence of Zika virus from a French Polynesia outbreak in 2013, Genome Announc. 2 (3) (2014), e00500-e00514, https://doi.org/10.1128/genomeA.00500-14.

[41] N.J. Braun, J.P. Quek, S. Huber, et al., Structure-based macrocyclization of substrate analogue NS2B-NS3 protease inhibitors of Zika, west nile and dengue viruses, ChemMedChem 15 (15) (2020) 1439–1452, https://doi.org/10.1002/cmdc.202000237.

[42] H. Maus, F. Barthels, S.J. Hammerschmidt, et al., SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases, Bioorg. Med. Chem. 47 (August) (2021), 116392, https://doi.org/10.1016/j.bmc.2021.116392.

[43] F. von Hammerstein, L.M. Lauth, S. Hammerschmidt, et al., Cis autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs, FEBS Lett. 593 (16) (2019) 2204–2213, https://doi.org/10.1002/1873-3468.13507.

[44] The PyMOL Molecular Graphics System, CCP4 newsletter on protein crystallography, Schrödinger (2002) 82–92.

[45] A. Micsonai, F. Wien, É. Bulyáki, et al., BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra, Nucleic Acids Res. 46 (W1) (2018) W315–W322, https://doi.org/10.1093/nar/gky497.

[46] GraphPad Prism, GraphPad Software, San Diego, California USA, 2014.

[47] H. Maus, G. Hinze, S.J. Hammerschmidt, et al., A competition SmFRET assay to study ligand-induced conformational changes of the dengue virus protease, Protein Sci. 32 (1) (2023), e4526, https://doi.org/10.1002/pro.4526.

[48] K. Gao, R. Oerlemans, M.R. Groves, Theory and applications of differential scanning fluorimetry in early-stage drug discovery, Biophys. Rev. 12 (1) (2020) 85–104, https://doi.org/10.1007/s12551-020-00619-2.

[49] M. Vivoli, H.R. Novak, J.A. Littlechild, et al., Determination of protein-ligand interactions using differential scanning fluorimetry, J. Vis. Exp. 91 (2014) 1–13, https://doi.org/10.3791/51809.

[50] F.H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, Nat. Protoc. 2 (9) (2007) 2212–2221, https://doi.org/10.1038/nprot.2007.321.

[51] Y. Ahmad, S.A. Shamsi, A novel direct synthesis of the N -oxides of quinoline derivatives, Bull. Chem. Soc. Jpn. 39 (1) (1966), https://doi.org/10.1246/bcsj.39.195, 195–195.

[52] A.O. Magnusson, A. Szekrenyi, H. Joosten, et al., NanoDSF as screening tool for enzyme libraries and biotechnology development, FEBS J. 286 (1) (2019) 184–204, https://doi.org/10.1111/febs.14696.

[53] S.L. Cobb, C.D. Murphy, 19F NMR applications in chemical biology, J. Fluor. Chem. 130 (2) (2009) 132–143, https://doi.org/10.1016/j.jfluchem.2008.11.003.

[54] C. Dalvit, Ligand- and substrate-based 19F NMR screening: principles and applications to drug discovery. *Progress in nuclear magnetic resonance spectroscopy*, November 30 (2007) 243–271, https://doi.org/10.1016/j.pnmrs.2007.07.002.

[55] D. Rose-Sperling, M.A. Tran, L.M. Lauth, et al., 19F NMR as a versatile tool to study membrane protein structure and dynamics, Biol. Chem. 400 (10) (2019) 1277–1288, https://doi.org/10.1515/hsz-2018-0473.

[56] U.A. Hellmich, N. Pfleger, C. Glaubitz, 19 F-MAS NMR on proteorhodopsin: enhanced protocol for site-specific labeling for general application to membrane proteins, Photochem. Photobiol. 85 (2) (2009) 535–539, https://doi.org/10.1111/j.1751-1097.2008.00498.x.

[57] K. Lang, J.W. Chin, Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins, Chem. Rev. 114 (9) (2014) 4764–4806, https://doi.org/10.1021/cr400355w.

[58] H. Qianzhu, E.H. Abdelkader, I.D. Herath, et al., Site-specific incorporation of 7-fluoro- <scp>L</Scp> -tryptophan into proteins by genetic encoding to monitor ligand binding by ^{19} F NMR spectroscopy, ACS Sens. 7 (1) (2022) 44–49, https://doi.org/10.1021/acssensors.1c02467.

[59] C. Kenward, K. Shin, J.K. Rainey, Mixed fluorotryptophan substitutions at the same residue expand the versatility of 19 F protein NMR spectroscopy, Chem. Eur J. 24 (14) (2018) 3391–3396, https://doi.org/10.1002/chem.201705638.

[60] J. Zheng, in: T. Jue (Ed.), Biomedical Applications of Biophysics, Humana Press, Totowa, NJ, 2010, https://doi.org/10.1007/978-1-60327-233-9.

[61] M.E. Hill, M. Yildiz, J.A. Hardy, Cysteine disulfide traps reveal distinct conformational ensembles in dengue virus NS2B-NS3 protease, Biochemistry 58 (6) (2019) 776–787, https://doi.org/10.1021/acs.biochem.8b00978.

[62] T. Kroeger, B. Frieg, T. Zhang, et al., EDTA aggregates induce SYPRO orange-based fluorescence in thermal shift assay, PLoS One 12 (5) (2017), e0177024, https://doi.org/10.1371/journal.pone.0177024.

[63] T. Wu, J. Yu, Z. Gale-Day, et al., Three essential resources to improve differential scanning fluorimetry (DSF) experiments, bioRxiv 2020 (2020), https://doi.org/10.1101/2020.03.22.002543, 03.22.002543.

[64] B. Millies, F. von Hammerstein, A. Gellert, et al., Proline-based allosteric inhibitors of Zika and dengue virus NS2B/NS3 proteases, J. Med. Chem. 62 (24) (2019) 11359–11382, https://doi.org/10.1021/acs.jmedchem.9b01697.

[65] W.W. Phoo, Y. Li, Z. Zhang, et al., Structure of the NS2B-NS3 protease from Zika virus after self-cleavage, Nat. Commun. 7 (2016), 13410, https://doi.org/10.1038/ncomms13410.

[66] S.J. Hammerschmidt, S. Huber, N.J. Braun, et al., Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs, Arch. Pharm. 356 (4) (2023), e2200518, https://doi.org/10.1002/ardp.202200518.

[67] P.B. Crowley, C. Kyne, W.B. Monteith, Simple and inexpensive incorporation of 19F-tryptophan for protein NMR spectroscopy, Chem. Commun. 48 (86) (2012), 10681, https://doi.org/10.1039/c2cc35347d.

[68] F. Barthels, S.J. Hammerschmidt, T.R. Fischer, et al., A low-cost 3D-printable differential scanning fluorometer for protein and RNA melting experiments, HardwareX 11 (2022), e00256, https://doi.org/10.1016/j.ohx.2022.e00256.

Supporting Information: The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR.

Hannah Maus [a#], Stefan J. Hammerschmidt [a#], Gerald Hinze [b], Fabian Barthels [a], Victor H. Pérez Carrillo [c], Ute A. Hellmich [c], Thomas Basché [b], Tanja Schirmeister [a*]

[#] authors contributed equally

* corresponding author

Tanja Schirmeister – Institute for Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg-University, Mainz, Germany; orcid.org/0000-0002-4587-5076; Phone: +49 (0)6131 3925742; Email: schirmei@uni-mainz.de

[a] Institute of Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg-University, Mainz, Germany

[b] Department of Chemistry, Johannes Gutenberg-University, Mainz, Germany

[c] Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University, Jena, Germany

KEYWORDS: ZIKV NS2B-NS3 protease, allosteric/competitive inhibition, conformational dynamics, smFRET, (nano)DSF, ^{19}F-NMR.

CONTENT

Hydrophobicity of the Surface of the ZIKV Protease

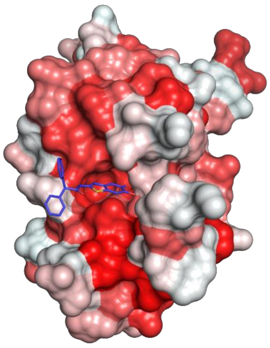


Figure S1. Representation of the surface of the ZIKV protease (PDB: 5GXJ) stained according to hydrophobicity. Deeper red color shades indicate less polarity of the surface.[1] The allosteric ligand $\mathbf{I_a}$ is shown in blue.

1. Tryptophan Residues in Open and Closed Conformation of the ZIKV Protease

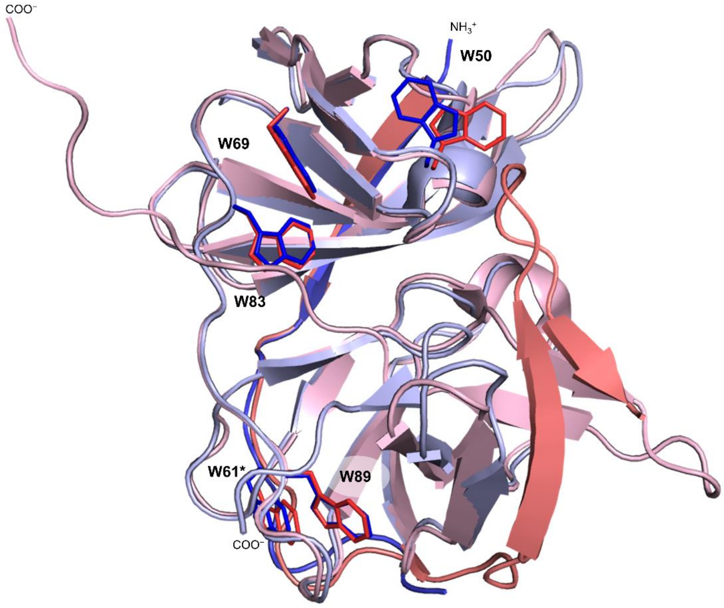


Figure S2. Tryptophan residues in the open (blue, PDB: 5GXJ) and closed conformation (red, PDB: 6Y3B) of the ZIKV protease. Tryptophan residues are highlighted as sticks. NS2B is shown in darker colors and NS3 in light colors. Residue W61 within NS2B is labeled by an asterisk. W5 at the N-terminus of NS3 is not resolved in the structures and thus not included here.

2. Assignment of Tryptophan Residues Using W/F Point Mutants

To assign the analyzed peak around 125 ppm to one of the six tryptophane residues, six different single W/F mutants were generated, each with one of the tryptophan residues replaced by a phenylalanine by site directed mutagenesis. Consequently, the mutated tryptophan is not labeled with ^{19}F upon expression. The absence of the peak allows a definite assignment. The asterisk indicates residues in the NS2B^{cf} region.

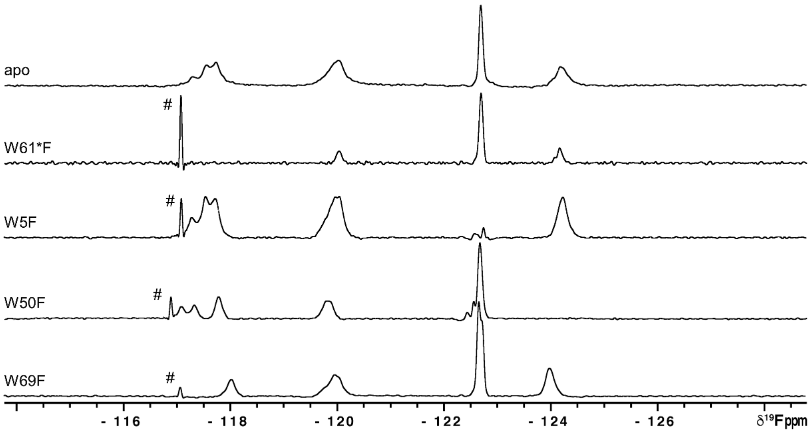


Figure S3. ^{19}F-NMR spectra of the 5F-5Trp labeled W/F mutants of ZIKV protease for assignment purposes. W83F and W89F did not show activity and aberrant shift patterns, so they were excluded from this analysis. # residual solvent peak.

3. Fluorescence Labeling of the ^{5}ZiPro

Purified protein was rebuffered by using Vivaspin 10 MWCO spin concentrators (Sartorius AG, Göttingen, Germany) in labeling buffer (20 mM Tris-HCl pH 6.8, 150 mM NaCl, 0.5 mM TCEP) and incubated for 30 min at 4 °C. TCEP was removed using a Vivaspin protein concentrator in the same buffer without TCEP. A 2.1-fold excess of ATTO 488 and a 2.6-fold excess of ATTO 643 were added to the protein solution (5 µM) and samples were incubated by gently shaking in the dark for 2 h at room temperature. This enables a Michael addition of the thiol group of cysteine to the *β*-position of the *α,β*-unsaturated carbonyl moiety (maleimide anchor) of the dyes, forming a stable thioether. Labeled protease was then dialyzed overnight at 4 °C (20 mM Tris-HCl pH 6.8, 150 mM NaCl, 1 mM CHAPS) and purified by gel filtration (HiLoad 16/600 Superdex 75 pg column, GE Healthcare, Chicago, IL, USA, 20 mM Tris-HCl pH 8.0, 150 mM NaCl).

4. smFRET Experiments[2,3]

For smFRET measurements of the FRET-paired quintuple mutant of the ZIKV protease (^{5}ZiPro), self-constructed sample cells consisting of a poly(ethylene glycol) (PEG_01 & 02, MicroSurfaces Inc., Minneapolis, MN, USA) coated glass coverslip and a glued-on cylinder of protein repellent plastic (Protein LoBind Tube, Eppendorf SE, Hamburg, Germany) were used. 150 µL of the respective solution of FRET pair labeled ^{5}ZiPro (c ~ 1000 pM) in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP) was placed in the sample cell, and fluorescence photons were collected throughout 1800 s using a custom-built confocal microscope. The ligands were dissolved in DMSO and added to the protease solution, resulting in a 10 vol% DMSO content in the sample cell. First, a measurement was performed for each experiment without ligand (900 pM protease, 50 mM Tris-HCl, 1 mM CHAPS, 1 mM TCEP, 10 vol% DMSO). Then a solution of a ligand in buffer with protease was added. In this way, the protease concentration and the DMSO content were kept constant during all measurements. The concentration of the competitive inhibitor (**I$_c$**) varied between 1 µM and 20 µM, and the allosteric inhibitor was used at 150 µM. The recordings of both data sets with and without the ligand were performed in direct succession. The fluorophores were excited with the spectrally filtered output from a pulsed white light fiber laser (10 MHz, SC OEM, YSL Photonics, China). To obtain excitation pulses centered around 502 nm (FWHM ~1.5 nm), an acousto-optical tunable filter (AOTF-VIS-DR, Fianium, Southampton, United Kingdom) was used. After passing a single-mode fiber, the laser light was recollimated and then focused into the solution using a microscope objective (plan apochromat, 100×, NA = 1.4, oil immersion, Carl Zeiss AG, Oberkochen, Germany). In front of the objective, the excitation power was measured and set to 10 µW. Fluorescence emitted by the FRET pair labeled ^{5}ZiPro diffusing through the focus was collected through the objective. Excitation and emission were separated by a dichroic mirror (ZT491 rdcxrxt-UF1, Chroma Systems Solutions Inc., Lake Forest, CA, USA). The emitted fluorescence light passed a dichroic mirror (ZT640rdc-UF1, Chroma Systems Solutions Inc.) in the detection beam path, which allowed red light to be transmitted, while the light of higher energy was reflected. These two spectrally separated beams were then focused onto two APDs (acceptor channel A: SPCM-AQRH-15, PerkinElmer Inc., Waltham, MA, USA; donor channel D: PDM 50ct, Micro Photon Devices S.r.l., Bolzano, Italy) after appropriate spectral filtering (acceptor channel A: ET655lp, Chroma Systems Solutions Inc.; donor channel D: FF01-550/88-25, Semrock Inc., Rochester, NY, USA). Using the described filter combination ensured separate detection of donor and acceptor fluorescence. The two detector APDs were connected to a HydraHarp 400 module (PicoQuant GmbH, Berlin, Germany), which records the absolute and relative arrival time (relative to the excitation pulse) of the individual photons. This allowed the subsequent calculation of correlation functions, lifetime, and intensity time traces for the burst analysis. Such time traces were calculated for both channels.

5. Analysis of smFRET Data[2,3]

First, for the initial assessment of the conformational influence of ligands on ^{5}ZiPro, cross-correlation functions $G_{AD}(\tau)$ (FCS-FRET) and auto-correlation functions $G_{AA}(\tau)$ were calculated from the macro arrival times of the individual photons recorded by using the HydraHarp 400 module (PicoQuant GmbH, Berlin, Germany). It has been shown that the cross-correlation can drop more slowly to zero if there is an exchange between two conformations that differ in their energy transfer efficiency provided that this process is not much slower than the diffusional time through the focus.[4,5] Exchange manifests as an additional term in the correlation function; that is, in addition to the diffusional term without exchange, extended by an increasing term:

$G_{AD} \propto (1 - a_{AD} e^{-\frac{\tau}{\tau_{ex}}})$.

Notably the cross-correlation is expected to no longer drop significantly slower to zero if an added ligand stabilizes one of the conformations of the protein. Therefore, no or a considerably reduced conformational exchange takes place. It is consequently possible conclude the conformational dynamics based on the cross-correlations of the donor and acceptor intensities with and without added ligands. To explore the presence of conformational subpopulations, data must be viewed at the single-molecule level (smFRET). For this purpose, individual bursts extracted from the fluorescence intensity time traces (bin time = 1 ms) were analyzed. For further evaluation, only bursts were considered which showed at least 20 counts on both APDs. Average arrival times (τ_A) of the acceptor photons relative to the excitation pulse (without taking the instrumental response function (IRF) into account) were calculated for each individual burst. Furthermore, the FRET efficiency E_{ET} was calculated for each burst:

$$E_{ET} = \frac{counts(A)}{counts(A) + counts(D)}$$

We did not account for spectral crosstalk or direct excitation of the acceptor. Furthermore, we omitted a detection correction factor, which could compensate for the different detection efficiencies of the two APDs toward donor and acceptor fluorescence and the different quantum yields of both dyes. The individual bursts were sorted with respect to their E_{ET} and presented in frequency distributions. These allow the identification of subpopulations that differ in their E_{ET}. The occurrences of the individual bursts (individual molecules) were then plotted in a 2D histogram, separated according to acceptor lifetime and FRET efficiency, which enables the visual identification of existing subpopulations.

6. Fluorometric Enzyme Assay

The inhibitory activity of the allosteric inhibitor towards ZIKV protease was determined using an assay based on the fluorogenic substrate Boc-GRR-AMC (Bachem, Bubendorf, Switzerland). Substrate and inhibitor were pipetted as DMSO stock solutions to the protease (25–125 nM) dissolved in buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS, 1 mM TCEP). The measurements were carried out in flat-bottom 96-well microtiter plates (Greiner bio-one, Kremsmünster, Austria) with a Tecan Infinite F2000 PRO fluorimeter or a Tecan Spark 10M (Agilent Technologies, Santa Clara, CA, USA) in three independent experiments. In each well, a total volume of 200 µL was used, consisting of 180 µL of buffer, 5 µL of the enzyme solution, resulting in a final concentration of 25–125 nM, 10 µL of the inhibitor in DMSO or pure DMSO as control, and 5 µL of a solution of the substrate with a final concentration of 100 µM. The fluorescence was measured in 30 s intervals for 10 min at 25 °C with 380 nm excitation and 460 nm emission wavelengths. K_M values were determined for substrate concentrations between 0 µM and 500 µM.

7. K_M Curves

K_M values indicate the affinity of proteases and mutants for the substrate. K_M values were determined for all proteases analyzed in this paper. They range between 361 µM and 1641 µM. The affinity of the single W/F mutants to the substrate is slightly lower than that of the wild type and ^{5}ZiPro.

Table S1. K_M, k_{cat} and catalytic efficiency of all proteases and mutants used in this study.

protein	K_M [µM]	v_{max} [nM s^{-1}]	k_{cat} [s^{-1}]	k_{cat} K_M^{-1} [L mol^{-1} s^{-1}]
WT$_{Lit}$[6]	558 ± 62	-	0.026 ± 0.001	40.6 ± 2.5
WT$_{TCEP}$	641 ± 20	0.600 ±0.028	0.024 ± 0.001	37.5 ± 2.1
^{5}ZiPro$_{TCEP}$	794 ± 115	0.018 ± 0.002	$(1.42 \pm 0.14) \cdot 10^{-4}$	0.179 ± 0.031
^{5}ZiPro$^{ATTO\ 488/\ ATTO\ 643}_{TCEP}$	261 ± 23	$(9.17 \pm 0.41) \cdot 10^{-3}$	$(1.83 \pm 0.08) \cdot 10^{-4}$	0.70 ±0.07
eZiPro$_{TCEP}$	1202 ± 111	0.500 ± 0.041	0.020 ± 0.002	16.6 ± 2.1
5F-W eZiPro$_{TCEP}$	885 ± 56	1.32 ± 0.08	0.053 ± 0.003	59.6 ± 5.2
eZiPro W61*F$_{TCEP}$	1120 ± 134	0.217 ± 0.020	0.009 ± 0.001	7.11 ± 1.03
eZiPro W5F$_{TCEP}$	1318 ± 150	0.377 ± 0.038	0.015 ± 0.001	11.5 ± 1.7
eZiPro W50F$_{TCEP}$	1642 ± 157	0.585 ± 0.052	0.023 ± 0.002	14.3 ± 1.9
eZiPro W69F$_{TCEP}$	1493 ± 168	0.449 ± 0.045	0.018 ± 0.002	12.0 ± 1.8
eZiPro W83F$_{TCEP}$	n.a.	n.a.	n.a.	n.a.
eZiPro W89F$_{TCEP}$	n.a.	n.a.	n.a.	n.a.

n.a. indicates not active

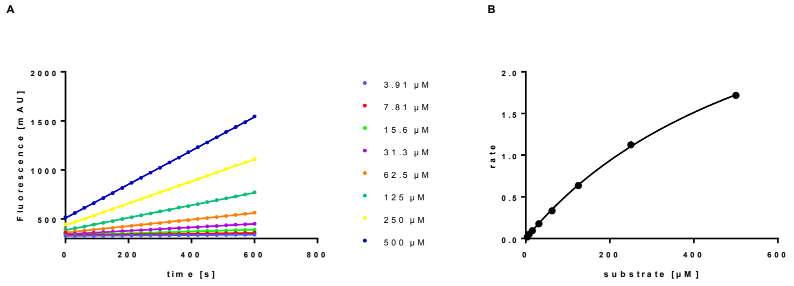


Figure S4. Fluorometric assay of the WT (gZiPro) of the ZIKV protease with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

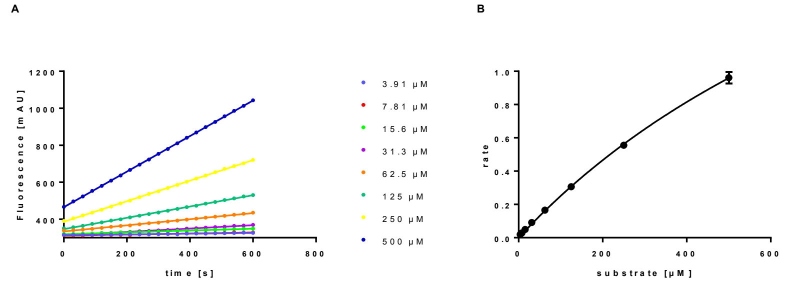


Figure S5. Fluorometric assay of the enzymatically cleavable linked ZIKV protease (eZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

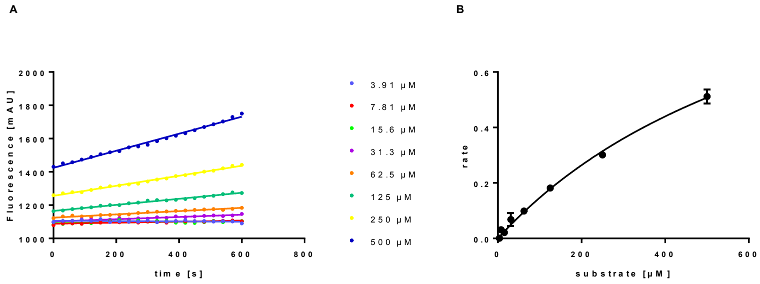


Figure S6. Fluorometric assay of the quintuple mutant of the wildtype of the ZIKV protease (^{5}ZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

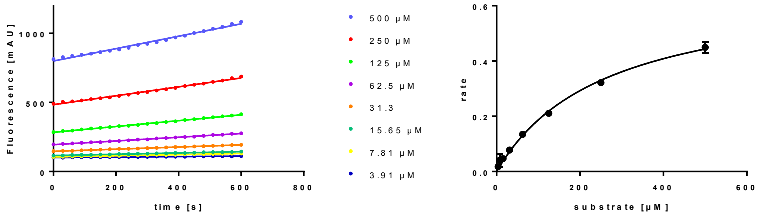


Figure S7. Fluorometric assay of the dye labeled quintuple mutant of the wildtype of the ZIKV protease (^{5}ZiPro$^{ATTO\ 488/\ ATTO\ 643}$) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

5F-tryptophan labeled eZiPro

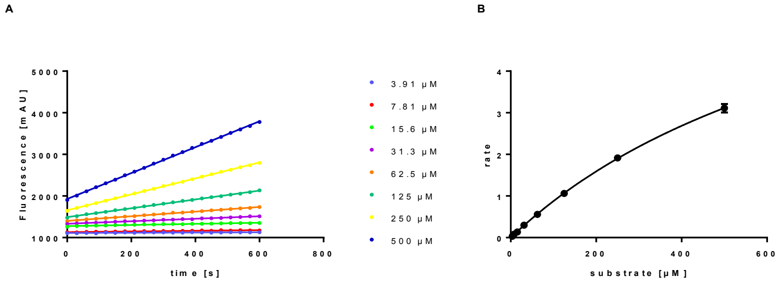


Figure S8. Fluorometric assay of the 5F-tryptophane labeled ZIKV protease (5F-W eZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

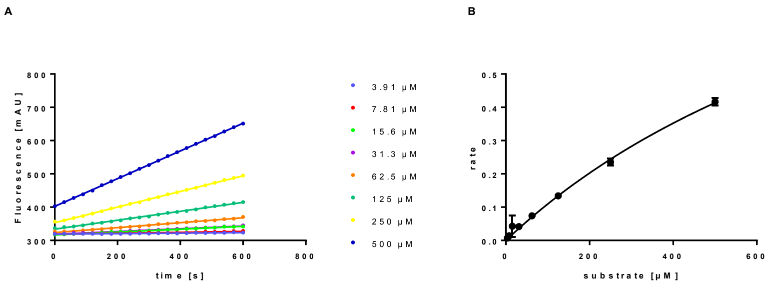


Figure S9. Fluorometric assay of the W61*F mutant of the ZIKV protease (eZiPro W61*F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

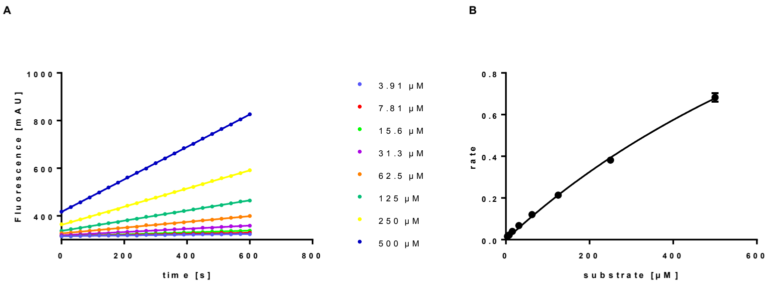


Figure S10. Fluorometric assay of the W5F mutant of the ZIKV protease (eZiPro W5F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

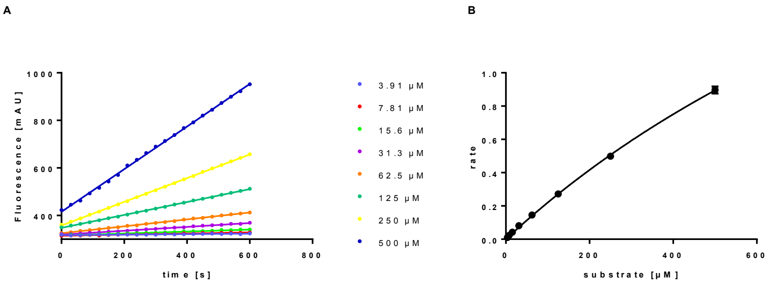


Figure S11. Fluorometric assay of the W50F mutant of the ZIKV protease (eZiPro W50F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

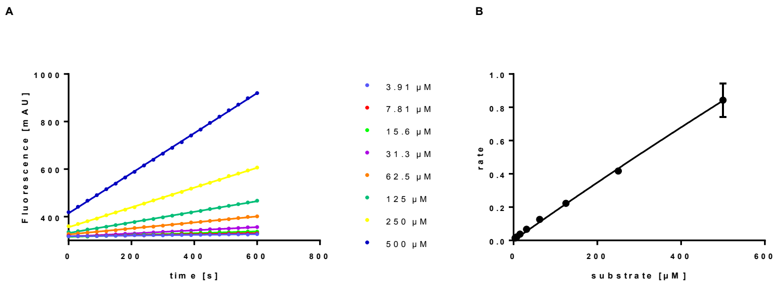


Figure S12. Fluorometric assay of the W69F mutant of the ZIKV protease (eZiPro W69F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

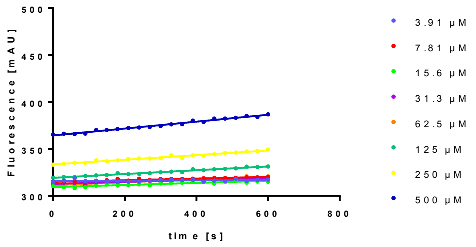


Figure S13. Fluorometric assay of the W83F mutant of the ZIKV protease (eZiPro W83F) with 1 mM TCEP.

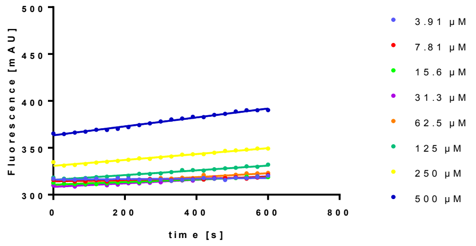


Figure S14. Fluorometric assay of the W89F mutant of the ZIKV protease (eZiPro W89F) with 1 mM TCEP.

(1) Eisenberg, D.; Schwarz, E.; Komarony, M.; et al. Amino Acid Scale: Normalized Consensus Hydrophobicity Scale. *J. Mol. Biol.* **1984**, *179*, 125–142.

(2) Maus, H.; Hinze, G.; Hammerschmidt, S. J.; et al. A Competition SmFRET Assay to Study Ligand-induced Conformational Changes of the Dengue Virus Protease. *Protein Sci.* **2022**, No. November 2022. https://doi.org/10.1002/pro.4526.

(3) Götz, C.; Hinze, G.; Gellert, A.; et al. Conformational Dynamics of the Dengue Virus Protease Revealed by Fluorescence Correlation and Single-Molecule FRET Studies. *J. Phys. Chem. B* **2021**, *125* (25), 6837–6846. https://doi.org/10.1021/acs.jpcb.1c01797.

(4) Torres, T.; Levitus, M. Measuring Conformational Dynamics: A New FCS-FRET Approach. *J. Phys. Chem. B* **2007**, *111* (25), 7392–7400. https://doi.org/10.1021/jp070659s.

(5) Price, E. S.; Devore, M. S.; Johnson, C. K. Detecting Intramolecular Dynamics and Multiple Fo¨rster Resonance Energy Transfer States by Fluorescence Correlation Spectroscopy. *J. Phys. Chem. B* **2010**, *114*, 5895–5902.

(6) von Hammerstein, F.; Lauth, L. M.; Hammerschmidt, S.; et al. Cis Autocatalytic Cleavage of Glycine-Linked Zika Virus NS2B-NS3 Protease Constructs. *FEBS Lett.* **2019**, *593* (16), 2204–2213. https://doi.org/10.1002/1873-3468.13507.

“

5.4 Conformational Selection and Induced Fit: The Behavior of Two Homologous Proteases.

5.4.1 Context, Project Summary, and Own Contributions

Trypsin-like serine proteases, specifically flaviviral NS2B-NS3 proteases, have been extensively studied concerning ligand recognition models.[122,184,285] Structural studies have revealed conformational changes in the NS2B-NS3 proteases during substrate recognition, small molecule binding, and catalysis.[32,84,106,278] Previous research has investigated the ligand binding mechanism of these proteases, but the appropriate model (conformational selection or induced fit) remains to be determined. In our earlier studies of the ZIKV and DENV2 NS2B-NS3 proteases (see previous Chapters 5.1–5.3), we utilized quintuple (^{5}ZiPro) and double mutants (^{2}DPro), respectively, to examine their conformational changes as a response to inhibitors.[283,284,286] Despite structural similarities and sequence homology, there are differences in the three-dimensional structure of ZIKV and DENV protease (see Chapter 1.2.4)[115,124], making the detailed study of the interactions of both proteases with ligands interesting.

We compared the conformational changes induced by small molecule inhibitors in these homologous proteases and employed smFRET analysis to study the binding mechanism. For this purpose, the labeled proteases were immobilized on a Ni-NTA-modified glass slide using their His_6-tags (Figure 25). This allowed the observation of individual proteases for the lifetime of the dyes. From the recorded time traces of the individual proteases, we calculated empirical cumulative distribution functions of the dwell times in the open and closed conformations for different concentrations of the competitive ligands. The results showed that, as shown before, both proteases favored the closed conformation in the presence of competitive ligands, but their conformational kinetics differed. The ^{5}ZiPro ligand reduced dwell times in the open state with increasing concentrations of I_c^{ZIKV}, while dwell times in the closed conformation remained independent of ligand concentration. Conversely, the ^{2}DPro exhibited dwell times independent from the ligand concentration in the open state but increased dwell times in the closed state with increasing concentrations of I_c^{DENV}. Following Weikl *et al.*, the reaction rate depending on the ligand concentration can be utilized to unravel the binding mechanism of a ligand (Figure 25).[250] Determination of the rates of conformational transitions in the presence of different concentrations of the competitive inhibitors allowed us to assign the interaction mechanisms unambiguously. While ^{5}ZiPro follows the IF model, the interaction of I_c^{DENV} with the ^{2}DPro follows the CS model. K_d values of the competitive inhibitors estimated using the kinetic parameters were in the same size range as the K_i values determined in the fluorometric assay, further supporting the reliability of our approach.

These findings demonstrated that specific binding mechanisms observed in one protease may not apply to an entire class of proteases, emphasizing the importance of detailed analysis for each protease. Additionally, we showed that smFRET is a powerful method to distinguish between different ligand binding models by analyzing conformational transitions.

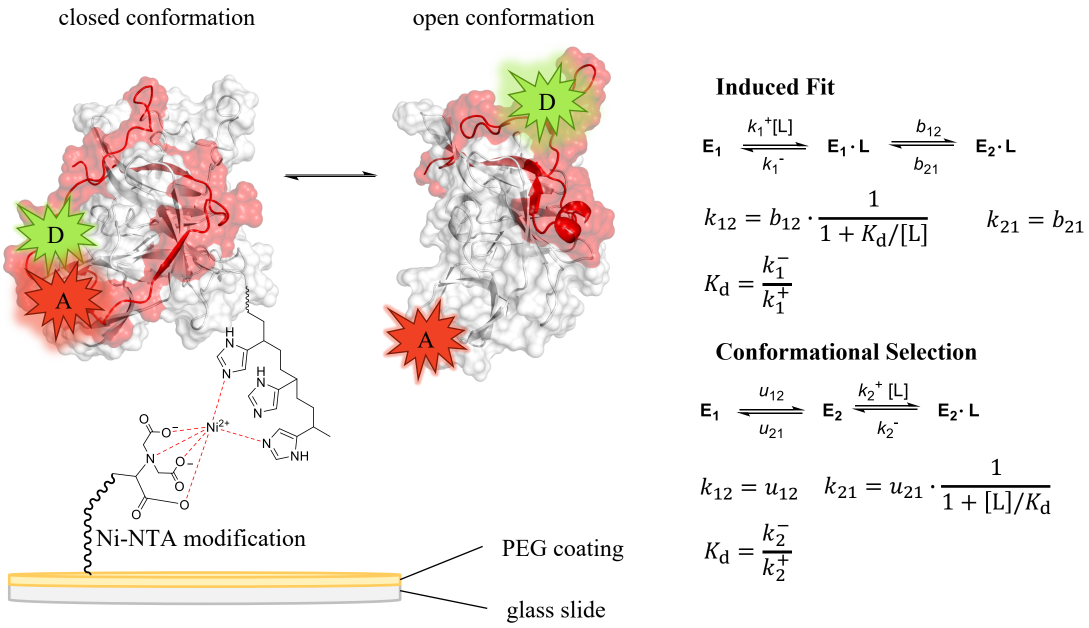


Figure 25. On a Ni-NTA modified glass slide, immobilized DENV protease. The protease is in an equilibrium between open (PDB-ID: 2fom) and closed (PDB-ID: 3u1i) conformation. Kinetic parameters obtained from the immobilized protease allow discrimination between induced fit and conformational selection models.

Own contributions: protein labeling, smFRET data curation, and writing of the original draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, smFRET data analysis, and manuscript editing.

Manuscript submitted to Angewandte Chemie (impact factor: 16.82).

5.4.2 Manuscript

The following manuscript from page 203 to page 217 contains unpublished data.

Conformational Selection and Induced Fit: The Behavior of Two Homologous Proteases

Hannah Maus[a], Gerald Hinze[b], Stefan J. Hammerschmidt[a], Tanja Schirmeister*[a], Thomas Basché*[b]

[a] Prof. Dr. Tanja Schirmeister
Institute of Pharmaceutical and Biomedical Sciences (IPBS)
Johannes Gutenberg-University
Staudingerweg 5, 55128 Mainz, Germany
E-mail: schirmei@uni-mainz.de

[b] Prof. Dr. Thomas Basché
Department of Chemistry
Johannes Gutenberg-University
Johann-Joachim-Becher Weg-18-20, 55128 Mainz, Germany
E-mail: basche@uni-mainz.de

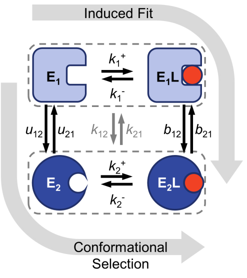
Supporting information for this article is given via a link at the end of the document.

Abstract. Protein-ligand interactions are crucial for many cellular processes, with details of the binding mechanism being discussed as essential for biological functions. Interestingly, protein binding often involves conformational changes between two or more states, whereby different binding mechanisms are possible even with a simple two-state description. Two models are widely used to portray protein-ligand interactions: Induced fit and conformational selection. However, distinguishing them experimentally is challenging. Single-molecule Förster resonance energy transfer (smFRET) has emerged as a powerful tool to resolve structural dynamics at the level of single proteins. Here, we investigated immobilized Zika virus (ZIKV) and dengue virus (DENV2) NS2B-NS3 proteases using smFRET to compare their conformational changes upon binding to competitive small molecule inhibitors. The analysis of the smFRET data allowed us to distinguish between induced fit and conformational selection models and assign the binding mechanism from the kinetic parameters obtained. Although DENV and ZIKV protease are proteins with high structural similarities, our results reveal that they have opposite binding mechanisms for competitive ligands. While the protein-ligand interaction in the ZIKV protease follows an induced fit mechanism, the DENV protease follows the conformational selection mechanism.

Introduction

Proteins are one of the fundamental macromolecules of life with their various functions being determined by their three-dimensional structures.[1–4] Adopting these structures is crucial for many cellular processes, including signal transduction, metabolism, and gene regulation.[5,6,15,16,7–14] In particular, conformational changes of proteins are of utmost importance in protein-ligand interactions, where small molecules bind to the protein to modulate its function.[14,17,18] The two most commonly used models to describe protein-ligand interactions accompanied by conformational changes are the induced fit (IF) and conformational selection (CS) models.[19,20]

The IF model proposes that the protein (E) undergoes a significant conformational change upon binding to the ligand (L). In contrast, the CS model suggests that the protein has a pre-existing conformational equilibrium that is shifted (and stabilized) upon ligand binding (Figure 1).[18,21–23] Thus, conformational changes can occur before (CS) or after (IF) ligand binding.



Induced Fit

$$\mathbf{E_1} \underset{k_1^-}{\overset{k_1^+[\mathrm{L}]}{\rightleftharpoons}} \mathbf{E_1 \cdot L} \underset{b_{21}}{\overset{b_{12}}{\rightleftharpoons}} \mathbf{E_2 \cdot L}$$

$$k_{12} = b_{12} \cdot \frac{1}{1 + K_\mathrm{d}/[\mathrm{L}]} \qquad k_{21} = b_{21}$$

$$K_\mathrm{d} = \frac{k_1^-}{k_1^+}$$

Conformational Selection

$$\mathbf{E_1} \underset{u_{21}}{\overset{u_{12}}{\rightleftharpoons}} \mathbf{E_2} \underset{k_2^-}{\overset{k_2^+[\mathrm{L}]}{\rightleftharpoons}} \mathbf{E_2 \cdot L}$$

$$k_{12} = u_{12} \qquad k_{21} = u_{21} \cdot \frac{1}{1 + [\mathrm{L}]/K_\mathrm{d}}$$

$$K_\mathrm{d} = \frac{k_2^-}{k_2^+}$$

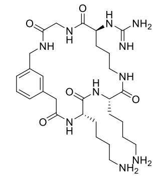
Figure 1. Induced fit and conformational selection model. Induced fit postulates a conformational transition between the complexes $E_1 \cdot L$ and $E_2 \cdot L$ that optimizes binding of the ligand. k_{12} and k_{21} denote the transition rates between the two conformations E_1 / $E_1 \cdot L$ and E_2 / $E_2 \cdot L$, depending on the ligand concentration. Conformational selection postulates a pre-existing equilibrium between the conformations E_1 and E_2 of the protein, of which only E_2 binds the ligand. [24]

Both models possess distinct differences in their underlying assumptions, and determining which model describes best a particular protein-ligand interaction is an active research area.[3,17,25,26] Besides these two borderline cases, combinations of both mechanisms are also discussed.[17,20,27] Providing experimental evidence to distinguish between these models is challenging. Traditional biochemical and biophysical techniques yield ensemble-averaged data that can mask the heterogeneity of protein conformations.[27–30]

Single-molecule Förster resonance energy transfer (smFRET) can overcome these limitations by studying conformational changes one protein at a time. Thereby, it presents a unique opportunity to follow the dynamics of protein-ligand interactions in real-time, thus providing information beyond that from ensemble measurements.[31–33] SmFRET allows for measuring distances and distance changes between fluorescent labels attached to different protein regions, yielding spatial information about conformational transitions in response to ligand binding.[34–39] By recording transitions between conformational states for a certain time, it is possible to extract their time scales or rates. The reaction rate depending on the ligand concentration, has often been utilized to gain more insights into ligand binding, thus being of paramount importance to unravel the underlying mechanistic details.[17,24]

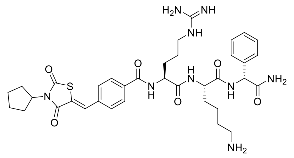
Trypsin-like serine proteases have often been used in the literature to discuss ligand recognition models.[3,26,40] Among these, the flaviviral NS2B-NS3 proteases are of special interest since they exist in at least two conformations. These proteases are key enzymes in the replication cycles of dengue (DENV), Zika (ZIKV), Japanese encephalitis, West Nile, tick-borne encephalitis, and Yellow fever viruses turning them into validated drug targets.[41–45] High-resolution structural studies have revealed key conformational changes of the NS2B-NS3 proteases that occur during the proteases' substrate recognition, small molecule binding, and catalysis.[43,46–48] Both conformations differ mainly in the relative positioning of the NS2B cofactor to the NS3 protease (NS3^{pro}). In the so-called "open conformation", the NS2B cofactor (NS2B$_{cf}$) region is disordered, and the active site is not fully formed, while in the closed conformation the NS2B cofactor is fully engaged with the NS3 protease and contributes to substrate recognition.[43,46,47,49,50] Up to now, the ligand-binding mechanism of these proteases has been investigated by different groups.[3,26,40,51] Yet, it has not yet been unambiguously clarified whether CS or IF is the appropriate model.

In our study of the Zika virus (ZIKV) and dengue virus 2 (DENV2) NS2B-NS3 proteases we have employed a quintuple mutant (C80S, C134S, C178S, S84C, S160C, ^{5}ZiPro) in case of the ZIKV protease, while for DENV protease a double mutant (S79C, S158C, ^{2}DPro) was used. Despite the overall structural similarity and high sequence homology of the two proteases, the position of a loop within the ZIKV NS2B$_{cf}$, i.e. residues 152–167, in the open conformation differs compared to the position of the respective DENV protease loop. The loop in the ZIKV protease kinks at Gly151 and runs in the opposite direction compared to the closed conformation. The stabilization of the loop in this position is achieved by hydrophobic interactions.[52] Furthermore, the C-terminus of the ZIKV NS2B$_{cf}$ seems to exhibit increased flexibility compared to the DENV2 protein. Roy *et al*. attributed this to reduced interactions between NS2B$_{cf}$ and NS3^{pro} in the ZIKV protein.[52–54] Herein, we compared the conformational changes of the homologous proteases upon binding small molecule inhibitors[55–57] (Figure 2). Analyzing the smFRET data of immobilized proteases provided substantial evidence to distinguish between induced fit and conformational selection and to assign the binding mechanism from the kinetic parameters obtained.



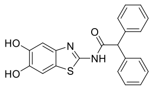
I$_c$^{ZIKV}

IC$_{50}$ (ZIKV NS2B-NS3) = 0.98 ± 0.04 µM
IC$_{50}$ (^{5}ZiPro) = 3.14 ± 0.19 µM
K_i (^{5}ZiPro) = 2.53 ± 0.15 µM



I$_c$^{DENV}

IC$_{50}$ (Behnam *et al.*) = 0.6 µM
IC$_{50}$ (^{2}DPro) = 5.4 ± 0.4 µM
K_i = 5.02 ± 0.37 µM



I$_a$

IC$_{50}$ (ZIKV NS2B-NS3) = 34.3 ± 4.0 µM
IC$_{50}$ (^{5}ZiPro) = (52.6 ± 0.92) % @ 100 µM

Figure 2. Structures and inhibition constants of the competitive inhibitors for ZIKV (I$_c$^{ZIKV})[57] and DENV2 protease (I$_C$^{DENV})[56] and the allosteric inhibitor I$_a$[55]. IC$_{50}$ values were determined for the corresponding mutants (^{5}ZiPro, ^{2}DPro) under the conditions used in the smFRET experiments.

Results and Discussion

smFRET of immobilized Proteases

The impact of competitive and allosteric inhibitors on the conformations of ZIKV and DENV2 NS2B-NS3 protease, respectively, has already been studied using smFRET, NMR, thermal shift assays, and a luciferase split assay.[38,39,46,58–60] In these studies, proteases diffusing freely in solution were examined, allowing to retrieve the ratio of open and closed conformations as a function of ligand concentrations. However, it could not be deduced from these data whether IF or CS caused the ligand-dependent conformational change. Nevertheless, based on the smFRET experiments in solution we had speculated that the binding of small molecules to DENV protease might follow the CS model.[38]

As described before, the proteases (^{5}ZiPro, ^{2}DPro) were labeled with a donor and an acceptor dye[38,39,59] Using their His$_6$-tags, they were immobilized at low coverage on a glass slide modified with Ni-NTA (Figure 3 A). Thus, individual dye-labeled proteases could be addressed. The positions of the doubly labeled proteases immobilized on the glass slides were determined from fluorescence images (Figure 3 B).

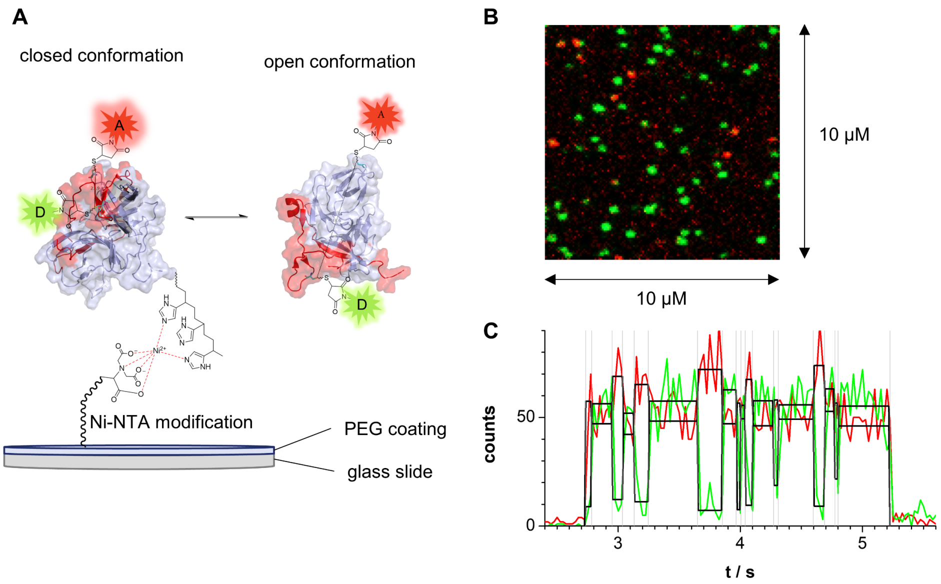


Figure 3. Immobilized proteases. (**A**) Schematic depiction of the ^{2}DPro protease immobilized on the Ni-NTA modified glass slide. The protease is in equilibrium between closed (left, PDB: 3U1I) and open (right, PDB: 2FOM) conformation. (**B**) Fluorescence image to localize dye-labeled ^{2}DPro. Proteases were excited at 502 nm and both the donor and acceptor dyes were detected. Donor-acceptor labeled proteases are seen as red spots, and donor-only labeled proteases as green spots. Acceptor-only and unlabeled proteases are not visible. (**C**) Exemplary time trace of ^{2}DPro with 2 µM of $\mathbf{I_c^{DENV}}$. red line: acceptor time trace, green line: donor time trace, black line: mean, gray lines: state boundaries.

After localization, the stream of emitted photons of individual proteases was recorded until photobleaching of one of the dyes occurred. In this way, time traces of the donor and acceptor emission intensity were recorded simultaneously for every single protease. As exemplarily seen in Figure 3 C, the donor and the acceptor intensities fluctuate in an anticorrelated manner. When the acceptor signal is high, the donor signal is low and vice versa. These fluctuations reflect the conformational transitions between open and closed states of the protease. In the closed conformation the donor and the acceptor dye are expected to be at a smaller distance as compared to the open conformation (Figure 3 A). Accordingly, the FRET efficiency is high, and the acceptor signal is larger than the donor signal. Upon a transition to the open state the situation is reversed. The fluctuating and anticorrelated signals are a unique characteristic of smFRET allowing to follow conformational transitions in real time. Thus, residence or dwell times of individual conformations are accessible.

In total, over 3,500 individual (donor and acceptor) time traces were recorded, of which about 2,000 could be evaluated. Time traces of the proteases were recorded in the absence and presence of different concentrations of competitive inhibitors. Without ligands, in 73% of the time traces only one conformational state was detected for the ^{5}ZiPro. For ^{2}DPro in the apo state, about half of the time traces did not show conformational transitions. For both proteases only time traces that showed at least two different states were used for further analysis. Time traces that did show only one intensity level (no conformational transitions within the limits given by the signal-to-noise-ratio) were not considered because the apparent dwell time in the conformational state was determined by photobleaching of one of the dyes. Thus, the observation time did not correspond to the actual residence time of the protease in this conformation. With the same argument, the very last state of each multi-state trajectory was omitted. The exclusion of these time traces and last states led to an underrepresentation of the more frequent conformation (Table S2) in the following analyses but was mandatory because the real dwell time could not be estimated in a reliable way.

After adding the competitive inhibitor $\mathbf{I_c^{ZIKV}}$ (Figure 2) the proportion of time traces with more than one conformational state increased in case of the ^{5}ZiPro (Table S1). Accordingly, the competitive inhibitor induced a more frequent switching between the different states of the proteases. On the other hand, the allosteric inhibitor $\mathbf{I_a}$ caused a reduction in transition frequency between the different conformations of the ^{5}ZiPro and was not investigated further, since only a fraction of 12% of the time traces eventually showed at least two different states. Similarly, for the ^{2}DPro, the proportion of time traces with multiple states was increased by the addition of $\mathbf{I_c^{DENV}}$ (Table S2). However, the effect was smaller than for the ^{5}ZiPro, probably due to the already higher proportion of multiple-state time traces without ligand and could not further be increased with growing concentration of $\mathbf{I_c^{DENV}}$.

Immobilized ^{5}ZiPro

Empirical cumulative distribution functions (eCDF) of the dwell times in the open and closed conformations were plotted for immobilized ^{5}ZiPro in the presence of different concentrations of $\mathbf{I_c^{ZIKV}}$ (Figure 4A, B). As seen in Figure 4A, higher concentrations of $\mathbf{I_c^{ZIKV}}$ led to a shift of the eCDF towards shorter times designating significantly reduced dwell times in the open conformation (*open-times*). A doubling of the inhibitor concentration caused a shortening of the average dwell time in the open conformation by almost one half. On the other hand, the dwell times of the closed conformation (*closed-times*) were found to be independent of the inhibitor concentration. Consequently, the probability of the protease to adopt the closed conformation was on average increased in presence of the inhibitor. This fits well with previous studies on the influence of competitive ligands on the conformation of ZIKV protease in solution.[59]

From the mean intensities of each state energy transfer efficiencies have been calculated without any additional assumptions (Fig. S1 A). The resulting distribution of energy transfer efficiencies are similar to results obtained from ZIKV protease freely diffusing in solution.[59] This agreement indicates that immobilization of the protease did not impact the conformational states involved in the smFRET experiments.

Immobilized ^{2}DPro

Remarkably, the data from ^{2}DPro reveal a completely different picture. In Figure 4C and D, the eCDFs of the dwell times in the open and closed conformation are presented for the immobilized ^{2}DPro. While the dwell times of the open conformation of the ^{5}ZiPro strongly decreased by adding $\mathbf{I}_{\mathbf{c}}^{\mathbf{ZIKV}}$, the *open-times* of the ^{2}DPro were independent from the inhibitor. Simultaneously, the dwell times of the ^{2}DPro in the closed conformation significantly increased with higher inhibitor concentration. Overall, $\mathbf{I}_{\mathbf{c}}^{\mathbf{DENV}}$ on average increased the probability of the ^{2}DPro being in the closed conformation, as described previously.[38,39] We emphasize that regarding ^{2}DPro, this was caused by an increase in the *closed-times*. In contrast, a decrease in the *open-times* was responsible for the higher probability of the closed state of the ^{5}ZiPro.

For the immobilized ^{2}DPro, the distributions of energy transfer efficiency (Figure S1 B) were also compatible to those of the protease freely diffusing in solution.[39] The relative underrepresentation of the open conformation in case of the immobilized protease without ligands is attributed to the discarding of time traces showing only a single state which more frequently was the open state (Table S2).

Conformational Selection or Induced Fit?

From crystal structures and experiments in solution is known that competitive inhibitors favor the closed conformation of flaviviral proteases.[38,39,57] The mechanism by which this preference is established can be understood using the results from the smFRET studies with the immobilized proteases described above. Under the condition that the applied ligand concentration [L] is significantly higher than the concentration [E] of the protein[17] and assuming that ligand binding and unbinding is substantially faster than the conformational changes of the protease, the equilibrium kinetics shown in Figure 1 can be simplified as follows.[24] For the IF model, the rate constants k_{12} (closing) and the k_{21} (opening) are $k_{12}^{\mathrm{IF}} = b_{12}\frac{1}{1+K_{\mathrm{d}}/[\mathrm{L}]}$ and $k_{21}^{\mathrm{IF}} = b_{21}$, respectively. For the model of CS, k_{12} becomes independent from [L], $k_{12}^{\mathrm{CS}} = u_{12}$, while k_{21} decreases with ligand concentration, $k_{21}^{\mathrm{CS}} = u_{21}\frac{1}{1+[\mathrm{L}]/K_{\mathrm{d}}}$. Similar expressions were derived from ensemble experiments for the case of first-order rate constants.[17]

The average open- and closed-times are the reciprocal of the rate constants of the conformational equilibrium, $\tau_{\mathrm{open}} = \frac{1}{k_{12}}$ and $\tau_{\mathrm{closed}} = \frac{1}{k_{21}}$. Thus, in the case of the IF model, the dwell times of the open state decrease with increasing ligand concentration, $\tau_{\mathrm{open}}^{\mathrm{IF}} = \frac{1+K_{\mathrm{d}}/[\mathrm{L}]}{b_{12}}$, while the dwell times of the closed states remain unaffected by [L], $\tau_{\mathrm{closed}}^{\mathrm{IF}} = \frac{1}{b_{21}}$. This is precisely what we observed for the ^{5}ZiPro. In the case of ^{2}DPro, a different mechanism was revealed. Here, the dwell times of open states remained unaffected by [L] while the dwell times of closed states increased with [L]. This is both predicted by the CS model, with $\tau_{\mathrm{open}}^{\mathrm{CS}} = \frac{1}{u_{12}}$ and $\tau_{\mathrm{closed}}^{\mathrm{CS}} = \frac{1+[\mathrm{L}]/K_{\mathrm{d}}}{u_{21}}$, respectively.

In the next step, it was checked whether the calculated concentration-dependent rate constants k_{12}^{IF} and k_{21}^{CS} lead to reasonable predictions regarding binding constants. Approximate binding affinities (K_{d} values) were obtained by fitting a model function (see SI) to the ligand concentration dependent rate constants (Figure 5). The K_{d} of $\mathbf{I}_{\mathbf{c}}^{\mathbf{ZIKV}}$ was estimated to be 8.0 µM. For the DENV protease, the K_{d} of $\mathbf{I}_{\mathbf{c}}^{\mathbf{DENV}}$ was determined to be 2.6 µM. Both K_{d} values lie in the same range as the inhibition constants obtained in fluorometric enzyme assays: $K_{\mathrm{i}}(\mathbf{I}_{\mathbf{c}}^{\mathbf{ZIKV}}) = 2.5 \pm 0.2$ µM, $K_{\mathrm{i}}(\mathbf{I}_{\mathbf{c}}^{\mathbf{DENV}}) = 5.0 \pm 0.37$ µM (see Figure 1). We interpret this satisfying agreement as further confirmation that kinetic details of protease-ligand interactions can be derived by our approach.

From our data, the ligand-binding of ZIKV and DENV2 proteases can be assigned to the IF and CS models, respectively. Despite the fact that ZIKV and DENV2 NS2B-NS3 proteases are structurally highly similar proteins, they follow the opposite limiting cases of protein-ligand interactions.

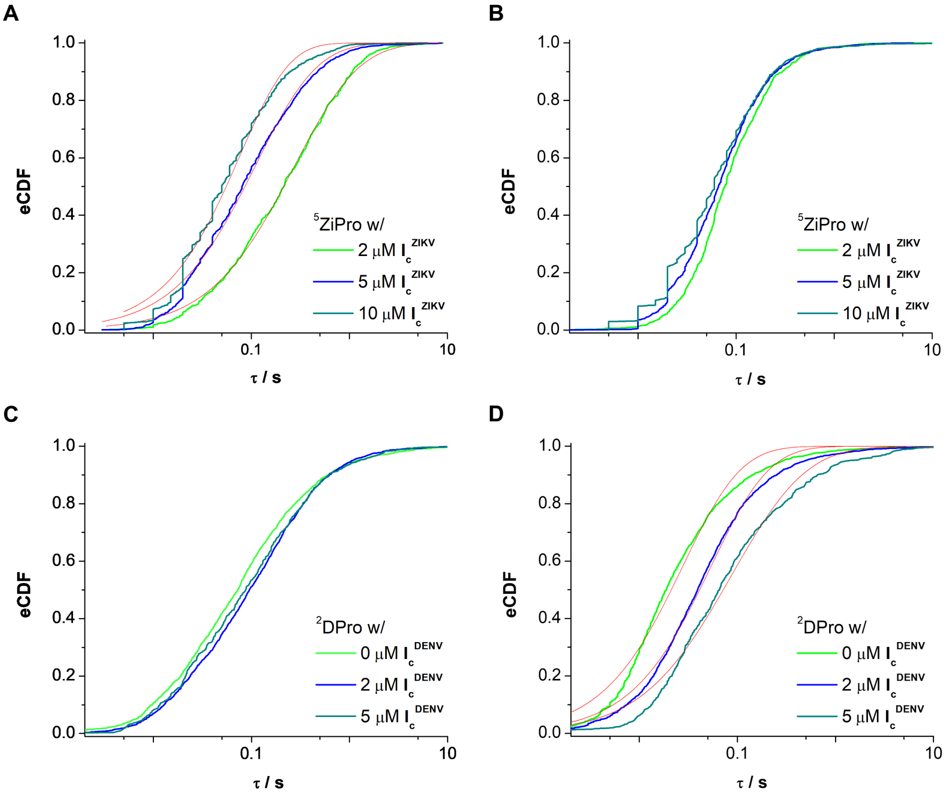


Figure 4. Empirical cumulative distribution functions (eCDF) of the dwell times in the open and closed conformation of ZIKV (**A**, **B**) and DENV2 (**C**, **D**) protease, respectively. (**A**) Open conformation of the ZIKV protease. (**B**) Closed conformation of the ZIKV protease. (**C**) Open conformation of the DENV2 protease. (**D**) Closed conformation of the DENV2 protease. The orange lines are fits to the experimental data; see SI.

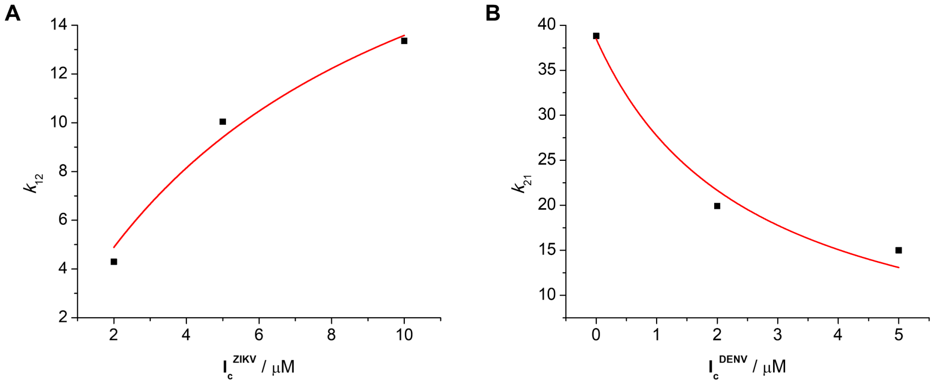


Figure 5. Calculated rate constants *ki* as a function of competitive inhibitor concentration. The red lines are fits to the data (see SI). (**A**) Induced fit model of ZIKV NS2B-NS3 protease. K_d = 8.0 µM. (**B**) Conformational selection modelDENV2 NS2B-NS3 protease. K_d = 2.6 µM.

Conclusion

Conformational changes that are triggered by protein-ligand interactions are essential for protein function. The present study utilized smFRET to study the binding mechanism of competitive small molecule inhibitors to individual immobilized ZIKV and DENV2 NS2B-NS3 proteases. Although for both proteases an increased preference for the closed conformation was observed in the presence of competitive ligands, as already shown in experiments in solution,[38,39,59] the kinetics of the conformational changes was different for the two proteases. The ^{5}ZiPro exhibited decreased dwell times in the open state with increasing ligand concentrations, whereas dwell times in the closed conformation were independent from the ligand concentration. This indicates that the binding mechanism follows the IF model. In contrast, in the case of the ^{2}DPro, the dwell times in the open conformation were found to be independent from the ligand concentration, while the dwell times in the closed state increased with increasing ligand concentrations. This implies that the binding mechanism follows the CS model. Our results highlight the power of smFRET to distinguish between different models of ligand binding by unravelling the kinetics of the conformational transitions involved. Moreover, they suggest in a more general way that a particular binding mechanism determined for a specific protease may not be representative for an entire class of proteases (e.g., trypsin-like proteases),[3] as has been assumed in the past, even for dissimilar proteins (flaviviral NS2B-NS3 proteases).[26] In contrast, it appears that each specific protease has to be analyzed in detail. Zika and dengue viruses are increasingly getting into the focus of current research due to their growing prevalence.[61] To date, there is no reliable treatment option nor a vaccine to prevent infection with these viruses.[62] In the search for potent drugs, the flaviviral proteases in particular represent an interesting target.[63–65] Our surprising finding of opposing binding mechanisms of competitive ligands to structurally similar proteases is expected to open new avenues for drug discovery and rational drug design.

Supporting Information

The authors have cited additional references within the Supporting Information.[38,39,55,66–69]

Experimental Procedures (Protein Constructs and Site-Directed Mutagenesis, Protein Expression and Purification, Fluorescent Dye Labeling of the NS2B-NS3 Proteases, smFRET Experiments); Data Analysis and Modeling (Analysis of smFRET Data, Data Interpretation and Modeling); Experimental Results (Lists of Time Traces of Immobilized ZIKV and DENV2 NS2B-NS3 Proteases, Distributions of Energy Transfer Efficiencies).

Acknowledgements

We thank Prof. Dr. Christian Klein and Prof. Dr. Torsten Steinmetzer for providing the competitive inhibitors and Christian Götz for experimental groundwork.

Keywords: Binding Mechanism of Competitive Inhibitors • Flaviviral Proteases • Molecular Dynamics • Protein Folding • smFRET

References

[1] N. Kitadai, S. Maruyama, *Geosci. Front.* **2018**, *9*, 1117–1153.

[2] N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, *Cell* **2009**, *138*, 774–786.

[3] N. Pozzi, A. D. Vogt, D. W. Gohara, E. Di Cera, *Curr. Opin. Struct. Biol.* **2012**, *22*, 421–431.

[4] B. Willenberg, K. Vollmar Hughes, *Home Econ. Guid.* **1983**, 1916.

[5] T. Pawson, P. Nash, *Genes Dev.* **2000**, *14*, 1027–1047.

[6] D. Wu, G. LaRosa, M. Simon, *Science* **1993**, *261*, 101–103.

[7] L. Birnbaumer, M. Birnbaumer, in *Biomembranes*, Wiley-VCH Verlag GmbH, Weinheim, Germany, **2008**, pp. 153–252.

[8] C. Londos, C. Sztalryd, J. T. Tansey, A. R. Kimmel, *Biochimie* **2005**, *87*, 45–49.

[9] M. Ptashne, *Nature* **1986**, *322*, 697–701.

[10] P. H. von Hippel, *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 79–105.

[11] R. K. Montange, R. T. Batey, *Annu. Rev. Biophys.* **2008**, *37*, 117–133.

[12] X. Deupi, B. Kobilka, **2007**, pp. 137–166.

[13] R. Weatherman, R. Fletterick, T. Scanlan, *Annu. Rev. Biochem.* **1999**, *68*, 559–581.

[14] G. G. Hammes, *Enzyme Catalysis and Regulation*, Elsevier, **1982**.

[15] C. H. Henkels, J. C. Kurz, C. A. Fierke, T. G. Oas, *Biochemistry* **2001**, *40*, 2777–2789.

[16] P. E. Wright, H. J. Dyson, *Curr. Opin. Struct. Biol.* **2009**, *19*, 31–38.

[17] G. G. Hammes, Y.-C. Chang, T. G. Oas, *Proc. Natl. Acad. Sci.* **2009**, *106*, 13737–13741.

[18] D. E. Koshland, **2006**, pp. 45–97.

[19] S. Kumar, B. Ma, C.-J. Tsai, N. Sinha, R. Nussinov, *Protein Sci.* **2008**, *9*, 10–19.

[20] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789–796.

[21] D. E. Koshland, *Angew. Chemie* **1994**, *106*, 2468–2472.

[22] D. E. Koshland, *Proc. Natl. Acad. Sci.* **1958**, *44*, 98–104.

[23] D. E. Koshland, G. Némethy, D. Filmer, *Biochemistry* **1966**, *5*, 365–385.

[24] T. R. Weikl, F. Paul, *Protein Sci.* **2014**, *23*, 1508–1518.

[25] A. D. Vogt, E. Di Cera, *Biochemistry* **2012**, *51*, 5894–5902.

[26] M. A. M. Behnam, C. D. P. Klein, *Biochimie* **2020**, *174*, 117–125.

[27] N. S. Hatzakis, *Biophys. Chem.* **2014**, *186*, 46–54.

[28] H. P. Lu, *Science* **1998**, *282*, 1877–1882.

[29] W. Tan, E. S. Yeung, *Anal. Chem.* **1997**, *69*, 4242–4248.

[30] D. M. Rissin, H. H. Gorris, D. R. Walt, *J. Am. Chem. Soc.* **2008**, *130*, 5349–5353.

[31] T. Förster, *Ann. Phys.* **1948**, *437*, 55–75.

[32] C. Bräuchle, D. C. Lamb, J. Michaelis, *Single Particle Tracking and Single Molecule Energy Transfer*, Wiley, Weinheim, **2009**.

[33] G. Agam, C. Gebhardt, M. Popara, R. Mächtel, J. Folz, B. Ambrose, N. Chamachi, S. Y. Chung, T. D. Craggs, M. de Boer, D. Grohmann, T. Ha, A. Hartmann, J. Hendrix, V. Hirschfeld, C. G. Hübner, T. Hugel, D. Kammerer, H.-S. Kang, A. N. Kapanidis, G. Krainer, K. Kramm, E. A. Lemke, E. Lerner, E. Margeat, K. Martens, J. Michaelis, J. Mitra, G. G. Moya Muñoz, R. B. Quast, N. C. Robb, M. Sattler, M. Schlierf, J. Schneider, T. Schröder, A. Sefer, P. S. Tan, J. Thurn, P. Tinnefeld, J. van Noort, S. Weiss, N. Wendler, N. Zijlstra, A. Barth, C. A. M. Seidel, D. C. Lamb, T. Cordes, *Nat. Methods* **2023**, *20*, 523–535.

[34] S. Kalinin, T. Peulen, S. Sindbert, P. J. Rothwell, S. Berger, T. Restle, R. S. Goody, H. Gohlke, C. A. M. Seidel, *Nat. Methods* **2012**, *9*, 1218–1225.

[35] Y. Santoso, C. M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. P. Torella, N. D. F. Grindley, A. N. Kapanidis, *Proc. Natl. Acad. Sci.* **2010**, *107*, 715–720.

[36] R. Roy, S. Hohng, T. Ha, *Nat. Methods* **2008**, *5*, 507–516.

[37] Y. Zhao, D. S. Terry, L. Shi, M. Quick, H. Weinstein, S. C. Blanchard, J. A. Javitch, *Nature* **2011**, *474*, 109–113.

[38] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *J. Phys. Chem. B* **2021**, *125*, 6837–6846.

[39] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2023**, *32*, e4526.

[40] J.-P. Changeux, S. Edelstein, *F1000 Biol. Rep.* **2011**, *3*, 1–15.

[41] V. Boldescu, M. A. M. Behnam, N. Vasilakis, C. D. Klein, *Nat. Rev. Drug Discov.* **2017**, *16*, 565–586.

[42] M. A. M. Behnam, C. Nitsche, V. Boldescu, C. D. Klein, *J. Med. Chem.* **2016**, *59*, 5622–5649.

[43] C. Nitsche, S. Holloway, T. Schirmeister, C. D. Klein, *Chem. Rev.* **2014**, *114*, 11348–11381.

[44] M. Brecher, J. Zhang, H. Li, *Virol. Sin.* **2013**, *28*, 326–336.

[45] A. E. Aleshin, S. A. Shiryaev, A. Y. Strongin, R. C. Liddington, *Protein Sci.* **2007**, *16*, 795–806.

[46] M. Brecher, Z. Li, B. Liu, J. Zhang, C. A. Koetzner, A. Alifarag, S. A. Jones, Q. Lin, L. D. Kramer, H. Li, *PLOS Pathog.* **2017**, *13*, e1006411.

[47] C. G. Noble, C. C. Seh, A. T. Chao, P. Y. Shi, *J. Virol.* **2012**, *86*, 438–446.

[48] Q. Li, C. Kang, *Int. J. Mol. Sci.* **2020**, *21*, 2527.

[49] W.-N. Chen, K. V. Loscha, C. Nitsche, B. Graham, G. Otting, *FEBS Lett.* **2014**, *588*, 2206–2211.

[50] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, U. Hommel, *Nat. Struct. Mol. Biol.* **2006**, *13*, 372–373.

[51] D. Ekonomiuk, A. Caflisch, *Protein Sci.* **2009**, *18*, 1003–1011.

[52] X. Chen, K. Yang, C. Wu, C. Chen, C. Hu, O. Buzovetsky, Z. Wang, X. Ji, Y. Xiong, H. Yang, *Cell Res.* **2016**, *26*, 1260–1263.

[53] J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang, R. Hilgenfeld, *Science* **2016**, *353*, 503–505.

[54] A. Roy, L. Lim, S. Srivastava, Y. Lu, J. Song, *PLoS One* **2017**, *12*, e0180632.

[55] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.

[56] M. A. M. Behnam, C. Nitsche, S. M. Vechi, C. D. Klein, *ACS Med. Chem. Lett.* **2014**, *5*, 1037–1042.

[57] N. J. Braun, J. P. Quek, S. Huber, J. Kouretova, D. Rogge, H. Lang-Henkel, E. Z. K. Cheong, B. L. A. Chew, A. Heine, D. Luo, T. Steinmetzer, *ChemMedChem* **2020**, *15*, 1439–1452.

[58] L. Zhu, J. Yang, H. Li, H. Sun, J. Liu, J. Wang, *Biochem. Biophys. Res. Commun.* **2015**, *461*, 677–680.

[59] H. Maus, S. J. Hammerschmidt, G. Hinze, F. Barthels, V. H. Pérez Carrillo, U. A. Hellmich, T. Basché, T. Schirmeister, *Eur. J. Med. Chem.* **2023**, *258*, 115573.

[60] H. Qianzhu, E. H. Abdelkader, I. D. Herath, G. Otting, T. Huber, *ACS Sensors* **2022**, *7*, 44–49.

[61] O. J. Brady, S. I. Hay, *Lancet* **2019**, *394*, 1991–1992.

[62] D. M. Berube, in *Pandemic Resil. Lessons We Should Have Learn. from Zika*, **2023**, pp. 1–25.

[63] S. M. Amberg, A. Nestorowicz, D. W. McCourt, C. M. Rice, *J. Virol.* **1994**, *68*, 3794–3802.

[64] M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman, J. A. Hardy, *ACS Chem. Biol.* **2013**, *8*, 2744–2752.

[65] H. A. Rothan, H. C. Han, T. S. Ramasamy, S. Othman, N. A. Rahman, R. Yusof, *BMC Infect. Dis.* **2012**, *12*, 314.

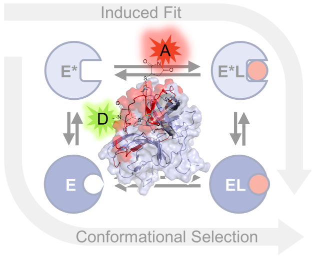
[66] C. Baronti, G. Piorkowski, R. N. Charrel, L. Boubis, I. Leparc-Goffart, X. de Lamballerie, *Genome Announc.* **2014**, *2*, e00500-14.

[67] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *J. Med. Chem.* **2019**, *62*, 11359–11382.

[68] F. von Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister, U. A. Hellmich, *FEBS Lett.* **2019**, *593*, 2204–2213.

[69] S. J. Hammerschmidt, S. Huber, N. J. Braun, M. Lander, T. Steinmetzer, C. Kersten, *Arch. Pharm. (Weinheim).* **2023**, *356*, e2200518.

Entry for the Table of Contents



Unlocking the mystery of protein-ligand interactions is crucial for drug discovery, and this article delves into the mechanisms behind them. Using cutting-edge smFRET technology, the study investigates the binding of ZIKV and DENV2 NS2B-NS3 proteases to small molecule inhibitors and distinguishes between two models – Induced Fit and Conformational Selection. The results shed new light on the kinetics of conformational changes and may advance progress in drug design.

Supporting Information

Conformational Selection and Induced Fit: The Behavior of Two Homologous Proteases

Hannah Maus[a], Gerald Hinze[b], Stefan J. Hammerschmidt[a], Tanja Schirmeister*[a], Thomas Basché*[b]

[a] Prof. Dr. Tanja Schirmeister
Institute of Pharmaceutical and Biomedical Sciences (IPBS)
Johannes Gutenberg-University
Staudingerweg 5, 55128 Mainz, Germany
E-mail: schirmei@uni-mainz.de

[b] Prof. Dr. Thomas Basché
Department of Chemistry
Johannes Gutenberg-University
Johann-Joachim-Becher Weg-18-20, 55128 Mainz, Germany
E-mail: basche@uni-mainz.de

Supporting information for this article is given via a link at the end of the document.

Table of Contents

Experimental Procedures

Protein Constructs and Site-directed Mutagenesis

ZIKV. To generate the ZIKV protease constructs, we used a pET-11a vector containing the sequence of the French Polynesia ZIKV strain (GenBank ID: KJ776791.2)[1]. This vector had been previously used by our team.[2] The construct consisted of two domains, NS2Bcf (amino acids 1423-1467) and NS3pro (amino acids 1503-1688), which were linked together by a glycine-rich linker (G4SG4). We inserted a R95A mutation in NS2B to prevent autocatalytic cis-cleavage. [3] The construct included an *N*-terminal hexahistidine (His6) tag followed by a tobacco etch virus protease (TEV) cleavage site. To obtain the ZIKV NS2B/NS3 ^{5}ZiPro construct, we performed site-directed mutagenesis using the Kapa HiFi PCR kit (Kapa Biosystems Inc., Woburn, MA, USA) and the gZiPro template. We used ten primers with the following sequences (amino acids in NS2B are indicated with an *):

(i) C80S: 5'-CTACTCCGGTCCGTGGAAACTGG-3' and 5'-GGACCGGAGTAGCTAACCAGGTC-3'

(ii) C143S: 5'-GGACAAATCCGGCCGTGTTATTGG-3' and 5'-CGGCCGGATTTGTCCAGAATCGG-3'

(iii) C178S: 5'-GTGGAGTCCTTTGAACCGAGCATG-3' and 5'-GTTCAAAGGACTCCACCGGGGTTTC-3'

(iv) S84*C: 5'-GTGACTTCTGCCTGGTTGAAGACG-3' and 5'-CCAGGCAGAAGTCACCGCTTTCG-3'

(v) S160C: 5'-GAACGGTTGCTACGTTAGCGCGATC-3' and 5'-CTAACGTAGCAACCGTTCTTGATCACAAC-3'

The PCR was performed with an annealing temperature gradient ranging from 65 to 75 °C, increasing by 0.33 °C per cycle. Out of approximately 15 clones, we obtained the correct sequence after screening for variants that did not have all the desired mutations.

DENV.[2,4] To introduce the S79C (NS2B) and S158C (NS3pro) mutations into the DENV NS2B-NS3 protease, we utilized a pET15b vector containing the DENV2 NS2B^{cf}-NS3p^{ro} gene (GenBank ID: AY037116.1). This vector already had two-point mutations (I30A and L31A in NS3) incorporated to enhance protein solubility. Additionally, it featured an *N*-terminal hexahistidine tag linked to NS2B^{cf} via a thrombin cleavage site.

For site-directed mutagenesis, we employed the Kapa HiFi PCR kit (Kapa Biosystems Inc., Woburn, MA, USA) and the following primers:

(i) S79C: 5'-GATCCTGTCAATTACGATCTGCGAAGATG-3' and 5'-GACATGCTACCATCTTCGCAGATCGTAATTG-3'

(ii) S158C: 5'-GGATCGGGCAGCTACCGCTGATTTCTGCTTG-3' and 5'-CGTAGGCACCACAACGGGTCACGACACCATTACC-3'.

Protein Expression and Purification

The quintuple mutant of glycine-linked ZIKV NS2B^{cf}-NS3^{pro} and the double mutant of glycine-linked DENV NS2B^{cf}-NS3^{pro} were expressed and purified following previously established methods.[5,6] In summary, the respective expression vectors (pET11a and pET15b) were transformed into Escherichia coli (E. coli) BL21 Gold (DE3) cells made competent using calcium chloride (Agilent Technologies, Santa Clara, CA, USA). The transformed cells were cultured in LB medium supplemented with 100 mg·L^{-1} ampicillin at 37 °C until reaching an optical density (OD$_{600}$) of approximately 0.8. Overexpression of the target proteins was induced by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG) and incubating the cultures at 18 °C for approximately 20 hours. The cells were then harvested by centrifugation at 9000 rpm and 4 °C for 15 minutes.

For purification, the cell pellets were resuspended in a lysis buffer containing 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.1 vol% TritonX‑100, RNase, DNase, lysozyme, and 1 mM DTT. The cells were lysed by sonication (10 cycles of 45 seconds each at 60% power) while maintaining constant cooling with ice (Sonoplus HD 2200; Bandelin, Berlin, Germany). The crude lysate was clarified by centrifugation (20,000 rpm at 4 °C for 1 hour) and subjected to purification using immobilized metal affinity chromatography (IMAC) with a HisTrap HP 5ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany) on an ÄKTA start FPLC device (Cytiva). After washing with 5 column volumes (CV) of IMAC buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole), the His$_6$-tagged protein was eluted using a linear gradient of IMAC buffer B containing 250 mM imidazole. The eluted NS2B-NS3 protein was concentrated using Vivaspin 10 MWCO spin concentrators (Sartorius AG, Göttingen, Germany) and further purified using size exclusion chromatography (HiLoad 16/600 Superdex 75 column, GE Healthcare) in SEC buffer A (20 mM Tris-HCl, pH 8.0, 150 mM NaCl) or SEC buffer B for the ZIKV NS2B-NS3 (with an additional 1 mM DTT). Fractions containing NS2B/NS3 were concentrated, rapidly frozen in liquid nitrogen, and stored at ‑80 °C. The purity and identity of NS2B/NS3 were confirmed using SDS-PAGE with Coomassie blue staining, followed by fluorometric enzyme activity assays.

The quintuple mutant of glycine-linked ZIKV NS2B^{cf}-NS3^{pro} and the double mutant of glycine linked DENV NS2B^{cf}-NS3^{pro}, were expressed and purified as described previously. In brief, the respective expression vectors (pET11a and pET15b) were transformed into calcium chloride competent Escherichia coli (E. coli) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). After reaching an optical density (OD600) of ~0.8 in LB medium containing 100 mg· L‑1 ampicillin at 37 °C, overexpression was induced by the addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG) at 18 °C for ~20 h. Cell pellets were harvested by centrifugation at 9000 rpm at 4 °C for 15 min. For purification, cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.1 vol% TritonX‑100, RNase, DNase, lysozyme, 1 mM DTT) and lysed by 10 cycles à 45 s of sonication at 60% power under constant cooling with ice (Sonoplus HD 2200; Bandelin, Berlin, Germany). The crude lysate was cleared by centrifugation (20,000 rpm at 4 °C for 1 h) and purified by immobilized metal affinity chromatography (IMAC) using a HisTrap HP 5ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany) on an ÄKTA start FPLC device (Cytiva). After a washing step with 5 column volumes (CV) of IMAC buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole), elution of the His$_6$ tagged protein was performed with a linear gradient in IMAC buffer B with 250 mM imidazole. Eluted NS2B-NS3 was concentrated using Vivaspin 10 MWCO spin concentrators (Sartorius AG, Göttingen, Germany) and subjected to a size exclusion chromatography step (HiLoad 16/600 Superdex 75 column, GE Healthcare) in SEC buffer A (20 mM Tris-HCl, pH 8.0, 150 mM NaCl) or SEC buffer B for the ZIKV NS2B-NS3 (with additional 1 mM DTT). Fractions containing NS2B/NS3 were concentrated, flash‑frozen in liquid nitrogen, and stored at ‑80 °C. The purity and identity of NS2B/NS3 were confirmed by SDS‑PAGE with Coomassie blue staining and subsequent fluorometric enzyme activity assays.

Fluorescent Dye Labeling of the NS2B-NS3 Proteases

Dye labeling was performed as described previously.[4,7,8] To prepare the purified protein for labeling, Vivaspin protein concentrators with a molecular weight cutoff (MWCO) of 10.0 kDa (GE Healthcare, Chicago, IL, USA) were utilized. The protein was rebuffered by exchanging the solution with labeling buffer containing 20 mM Tris-HCl pH 6.8, 150 mM NaCl, and 0.5 mM TCEP. The rebuffering process took place at 4 °C and lasted for 30 minutes. Subsequently, the TCEP was removed from the protein solution using a Vivaspin protein concentrator in the same buffer without TCEP. For labeling, an excess of ATTO 488 (≈2-fold) and ATTO 643 (≈2.5-fold) dyes were added to the protein solution. The samples were incubated in the dark at room temperature with gentle shaking for 2 hours. During this incubation, a Michael addition reaction occurred between the thiol group of cysteine in the protein and the β-position of the α,β-unsaturated carbonyl moiety (maleimide anchor) of the dyes, resulting in the formation of stable thioether bonds. Following labeling, the labeled protease was dialyzed overnight at 4 °C in a dialysis buffer containing 20 mM Tris-HCl pH 6.8, 150 mM NaCl, and 1 mM CHAPS. This step allowed for the removal of excess dyes and unreacted components. Finally, the labeled protease was purified using gel filtration on a HiLoad 16/600 Superdex 75 pg column (GE Healthcare, Chicago, IL, USA) with a buffer consisting of 20 mM Tris-HCl pH 8.0 and 150 mM NaCl. The gel filtration step helped separate the labeled protease from any remaining impurities and ensured its purity for subsequent experiments.

smFRET Experiments

SmFRET Experiments were performed as described before.[4,7,8] The FRET-pair labeled NS2B-NS3 proteases were studied in custom-made sample cells. These cells consisted of a glass coverslip coated with Ni-NTA modified poly(ethylene glycol) (Ni_01, MicroSurfaces Inc., Minneapolis, MN, USA), and a cylinder of protein repellent plastic (Protein LoBind Tube, Eppendorf SE, Hamburg, Germany) securely attached to it. In each sample cell 150 µL of the respective solution containing FRET-labeled NS2B-NS3 protease were filled. For DENV, the concentration was approximately 100 pM, while for ZIKV, it was around 1000 pM. The protease solutions were prepared in a buffer consisting of 50 mM Tris-HCl pH 9.0, 1 mM CHAPS, and 1 mM TCEP. To dissolve the ligands, DMSO was used, which resulted in a 10 vol% DMSO content in the sample cell. For measurements without ligands, 10 vol% DMSO were also included in the sample cell. The concentration of competitive inhibitors was varied between 2 µM and 10 µM. Additionally, an allosteric inhibitor was examined at a concentration of 100 µM. The protease solutions, with or without inhibitors, were added to the sample cells, and incubation was carried out for 3 minutes, to immobilize the proteins on the Ni-NTA modified glass slide.

To excite the fluorophores, we employed a custom-built confocal microscope equipped with a pulsed white light fiber laser (10 MHz, SC OEM, YSL Photonics, China). The excitation pulses were centered around 502 nm (FWHM ~ 1.5 nm) via an acousto-optical tunable filter (AOTF-VIS-DR, Fianium, United Kingdom). The laser light passed through a single-mode fiber, was recollimated, and focused onto the sample surface by a microscope objective (plan apochromat, 100×, NA = 1.4, oil immersion, Zeiss, Germany). The excitation power was set to 50 nW in front of the objective. The emitted fluorescence from the FRET-labeled NS2B-NS3 proteases was collected through the same objective. Excitation and emission were separated using a dichroic mirror (ZT491 rdcxrxt-UF1, CHROMA, USA). The fluorescence light passed through another dichroic mirror (ZT640rdc-UF1, CHROMA, USA) in the detection beam path, which transmitted red light while reflecting shorter wavelength light. These two spectrally separated beams were then focused onto two avalanche photodiodes (APDs): the acceptor channel (A: SPCM-AQRH-15, PerkinElmer, United States), and the donor channel (D: PDM 50ct, MPD, Italy). Appropriate spectral filtering (acceptor channel A: ET655lp, CHROMA, United States; donor channel D: FF01-550/88-25, Semrock, USA ensured separate detection of donor and acceptor fluorescence. The two detector APDs were connected to a HydraHarp 400 module (PicoQuant, Germany), which recorded the individual photons' absolute and relative arrival time (relative to the excitation pulse). After alignment of the microscope fluorescence images of 10 x 10 µm size were captured from the coated glass slide. The localized red spots corresponded to the donor-acceptor labeled proteases (see Figure 3 B). Subsequently, time traces were recorded for each single protease and analysed as described below.

Data Analysis and Modeling

Analysis of smFRET Data

Data analysis was performed in several steps (1-4). Time-tagged photons of donor and acceptor emission were recorded to identify the open/closed states of the proteases by their differing FRET efficiencies. For this purpose, both the intensities of the two detection channels as well as the TCSCP times (fluorescence lifetimes of the two fluorophores) are suitable,. An automated evaluation of the large number of time traces was prohibited by the high variability of intensities and fluctuations, which ultimately made an almost complete manual inspection and evaluation necessary.

1) With the help of self-developed software, the photon stream from single proteases could be visualized as intensity time traces of donor and acceptor fluorescence. Different bin times could be set, and markers were placed to indicate the transitions between open/closed states. Afterward, an assignment to open or closed states was performed based on intensity and fluorescence lifetime changes in both channels, resulting in a sequence of dwell times of open and closed states for each individual time trace.

2) Some of the data traces were evaluated employing a hidden Markov model algorithm [M. Pirchi, R. Tsukanov, R. Khamis, T.E. Tomov, Y. Berger, D.C. Khara, H. Volkov, G. Haran, E. Nir, J. Phys. Chem. B **2016**, 120, 13065-13075] with subsequent calculation of states by means of the Viterbi algorithm implemented in MATLAB (MathWorks, R2017b). However, this approach neglected the available information from fluorescence lifetimes and was very vulnerable to signal fluctuations that were not caused by FRET efficiency fluctuations.

3) With the aim of Matlab routines, dwell times in open or closed states from individual time traces were grouped depending on protease and ligand concentration. Notably, each protease/ligand combination was measured 2–4 times on different days.

4) For each protease/ligand combination, empirical cumulative distribution functions (eCDF) were calculated from the dwell times, neglecting the very last state of each individual time trace sinceit was artificially terminated by dye bleaching.

Data Interpretation and Modeling

The mean rates $\langle k_{12} \rangle$ and $\langle k_{21} \rangle$ were determined from each of the eCDFs.

$\mathrm{E}_1^{\mathrm{open}} \xrightarrow{k_{12}} \mathrm{E}_2^{\mathrm{closed}}$ and $\mathrm{E}_2^{\mathrm{closed}} \xrightarrow{k_{21}} \mathrm{E}_1^{\mathrm{open}}$

If presence of only a single rate constant k_x, the eCDF functions could be described with $\mathrm{eCDF}(\tau) = 1 - \exp(-k_{\mathrm{x}} \cdot \tau)$. However, non-exponential distributions were observed. For their modeling, it is assumed that the activation energies which determine the rate constants *via* the Arrhenius equation are normally distributed:

$\mathrm{V}_{\mathrm{Ea}}(\mathrm{E_a}) \propto \exp\left(-\frac{(\mathrm{E_a}-\mathrm{E_{a,0}})^2}{2\sigma^2}\right)$ and $k(\mathrm{E_a}) = k_0 \exp\left(-\frac{\mathrm{E_a}}{\mathrm{R\,T}}\right)$

This approach generates a log-normal distribution of rates. In the fits performed, the parameters $\mathrm{E_{a,0}}$ and k_0 were combined because otherwise, the fits would be overdetermined: $\mathrm{E}'_{\mathrm{a}} = \mathrm{E_a} - \mathrm{E_{a,0}}$ and $k(\mathrm{E_a}) = k_0 \exp\left(-\frac{\mathrm{E}'_{\mathrm{a}}+\mathrm{E_{a,0}}}{\mathrm{R\,T}}\right) = k(\mathrm{E}'_{\mathrm{a}}) = k'_0 \exp\left(-\frac{\mathrm{E}'_{\mathrm{a}}}{\mathrm{R\,T}}\right)$. A separate parameterization of $\mathrm{E_{a,0}}$ and k_0 would only be possible *via* a temperature dependence. Mean rates were determined *via* arithmetic means:

$\langle k_{\mathrm{x}} \rangle = \int_{-\infty}^{+\infty} k\, \mathrm{V}_k(k) \mathrm{d}k$

Experimental Results

Lists of Time Traces of Immobilized ZIKV and DENV2 NS2B-NS3 Proteases

Table S1. List of time traces recorded for the immobilized ZIKV NS2B-NS3 protease.

sample	date	total number of time traces	evaluable number of time traces	single state	single state		multiple states	distribution	
					open	closed		single	multiple
ZIKV NS2B-NS3 apo	19.08.	96	63	53	50	3	10	84%	16%
	23.08.	103	68	56	54	2	12	82%	18%
	25.08.	86	28	25	22	3	3	89%	11%
	13.09.	166	107	87	84	3	20	81%	19%
		451	266	221	210	11	45	83%	17%
ZIKV NS2B-NS3 +2 µM $\mathbf{I_c^{ZIKV}}$	19.08.	89	68	21	19	2	47	31%	69%
	23.08.	118	43	14	10	4	29	33%	67%
	25.08.	117	48	25	21	4	23	52%	48%
	13.09.	87	54	14	14	0	40	26%	74%
		358	213	74	64	10	139	35%	65%
ZIKV NS2B-NS3 +5 µM $\mathbf{I_c^{ZIKV}}$	19.08.	89	62	13	12	1	49	21%	79%
	23.08.	118	42	9	3	6	33	21%	79%
	25.08.	114	51	12	11	1	39	24%	76%
	13.09.	79	55	12	11	1	43	22%	78%
		400	210	46	37	9	164	22%	78%
ZIKV NS2B-NS3 +10 µM $\mathbf{I_c^{ZIKV}}$	14.09.	55	39	10	7	3	29	26%	74%
	15.09.	84	47	7	5	2	40	15%	85%
	16.09.	94	54	7	5	2	47	13%	87%
		233	140	24	17	7	116	17%	83%
ZIKV NS2B-NS3 apo +1 mM TCEP	14.09.	123	79	54	52	2	25	68%	32%
	15.09.	128	70	44	40	4	26	63%	37%
	16.09.	81	50	29	27	2	21	58%	42%
		332	199	127	119	8	72	64%	36%
ZIKV NS2B-NS3 +1 mM TCEP +100 µM $\mathbf{I_a}$	14.09.	60	41	37	36	1	4	90%	10%
	15.09.	77	49	46	46	0	3	94%	6%
	16.09.	68	50	40	40	0	10	80%	20%
		205	140	123	122	1	17	88%	12%
Σ		1979	1168						

Table S2. List of time traces recorded for the immobilized DENV NS2B-NS3 protease.

sample	date	total number of time traces	evaluable number of time traces	single state	single state		multiple states	distribution	
					open	closed		single	multiple
DENV NS2B-NS3 apo	25.07	188	97	52	46	6	45	54%	46%
	12.08.	101	62	29	28	1	33	47%	53%
	16.08.	176	110	52	47	5	58	47%	53%
	12.09.	130	82	46	36	10	36	56%	44%
		595	351	179	157	22	172	51%	49%
DENV NS2B-NS3 +2 µM $\mathbf{I_c^{DENV}}$	10.08.	175	65	24	20	4	41	37%	63%
	11.08.	137	79	22	19	3	57	28%	72%
	16.08.	120	69	12	5	7	57	17%	83%
	12.09.	104	56	23	13	10	33	41%	59%
		536	269	81	57	24	188	30%	70%
DENV NS2B-NS3 +5 µM $\mathbf{I_c^{DENV}}$	17.08.	307	75	28	24	4	47	37%	63%
	12.09.	134	54	15	9	6	39	28%	72%
		441	129	43	33	10	86	33%	67%
Σ		1619	752						

Distributions of Energy Transfer Efficiencies

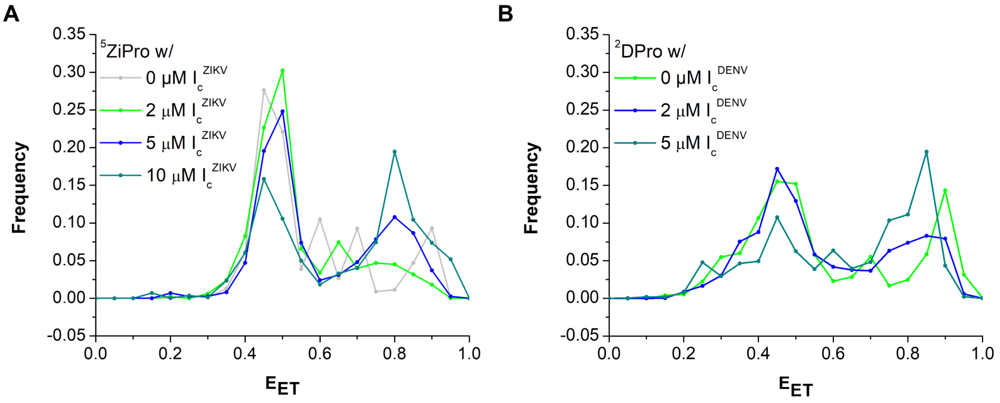


Figure S1. Distributions of energy transfer efficiencies of immobilized ^{5}ZiPro (**A**) and ^{2}DPro (**B**).

References

[1] C. Baronti, G. Piorkowski, R. N. Charrel, L. Boubis, I. Leparc-Goffart, X. de Lamballerie, *Genome Announc.* **2014**, *2*, e00500-14.

[2] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *J. Med. Chem.* **2019**, *62*, 11359–11382.

[3] F. von Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister, U. A. Hellmich, *FEBS Lett.* **2019**, *593*, 2204–2213.

[4] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2023**, *32*, e4526.

[5] S. J. Hammerschmidt, S. Huber, N. J. Braun, M. Lander, T. Steinmetzer, C. Kersten, *Arch. Pharm. (Weinheim).* **2023**, *356*, e2200518.

[6] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.

[7] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *J. Phys. Chem. B* **2021**, *125*, 6837–6846.

[8] H. Maus, S. J. Hammerschmidt, G. Hinze, F. Barthels, H. P. Victor, U. A. Hellmich, T. Basch, T. Schirmeister, V. H. Pérez Carrillo, U. A. Hellmich, T. Basché, T. Schirmeister, *Eur. J. Med. Chem.* **2023**, *258*, 115573.

Author Contributions

Hannah Maus: data curation – lead, investigation – lead, writing of original draft – lead, Gerald Hinze: formal analysis- lead, proofreading and revision of manuscript – lead, Stefan J. Hammerschmidt: proofreading and revision of manuscript – equal, Tanja Schirmeister: project administration – equal, funding – equal, proofreading and revision of manuscript – equal, Thomas Basché: project administration – equal, funding – equal, proofreading and revision of manuscript – lead.

5.5 Designing Photo Affinity Tool Compounds for the Investigation of the DENV NS2B-NS3 Protease Allosteric Binding Pocket

5.5.1 Context, Project Summary, and Own Contributions

Allosteric inhibitors of the DENV protease have been developed to target the protease by binding to a site other than the active site.[105,108,121,135,249] This leads to a conformational change and inhibition of protease activity.[106,284] However, these inhibitors' precise mechanism of action and binding site on the protease still need to be better understood. While crystal structures of competitive inhibitors show the protease in the closed conformation[84,116], methods in solution have demonstrated that allosteric inhibitors prefer the open conformation.[106,284] Noble *et al.* and Yildiz *et al.* identified a potential allosteric binding pocket around the amino acid Ala125.[84,103] Further investigations using site-directed mutagenesis and maleimide labeling confirmed the binding of benzo[*d*]thiazole-based allosteric inhibitors in this pocket.[108] To better understand the binding site of allosteric inhibitors, photoaffinity labeling (PAL) offers a valuable technique. PAL involves attaching a photoreactive group to a ligand, which UV light can activate to form a covalent bond with the target protein. This enables the identification of the ligand's binding site.[234,236]

In this study, we designed and synthesized tool compounds for PAL targeting the allosteric pocket of the DENV protease. These compounds consisted of a benzo[*d*]thiazole, representing the affinity unit, 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid as the photoreactive group, and a ligation handle (thiol or norbornene) which allowed click chemistry with a reporter (maleimide or tetrazine functionalized dye or biotin) after photoaffinity labeling of the protease (Figure 26). Dye labeling and biotin-mediated pull-down were used to detect the binding of inhibitors on an SDS-PAGE gel. Additionally, LC-MS/MS analysis of the labeled protein after digestion demonstrated the interaction of the allosteric inhibitor with Gly121, which confirmed the inhibitor's binding to the proposed allosteric pocket. Molecular docking validated that the interaction of the allosteric inhibitor with Gly121 was likely. The results were further supported by maleimide blockage of the proposed allosteric binding pocket after site-directed mutagenesis of Thr122 and the resulting decreased affinity of the inhibitor.

Our study demonstrated different PAL tool compound design and synthesis strategies and confirmed the binding pocket location through mass spectrometry and maleimide blockage assays. These findings provide a foundation for the development of potent allosteric inhibitors.

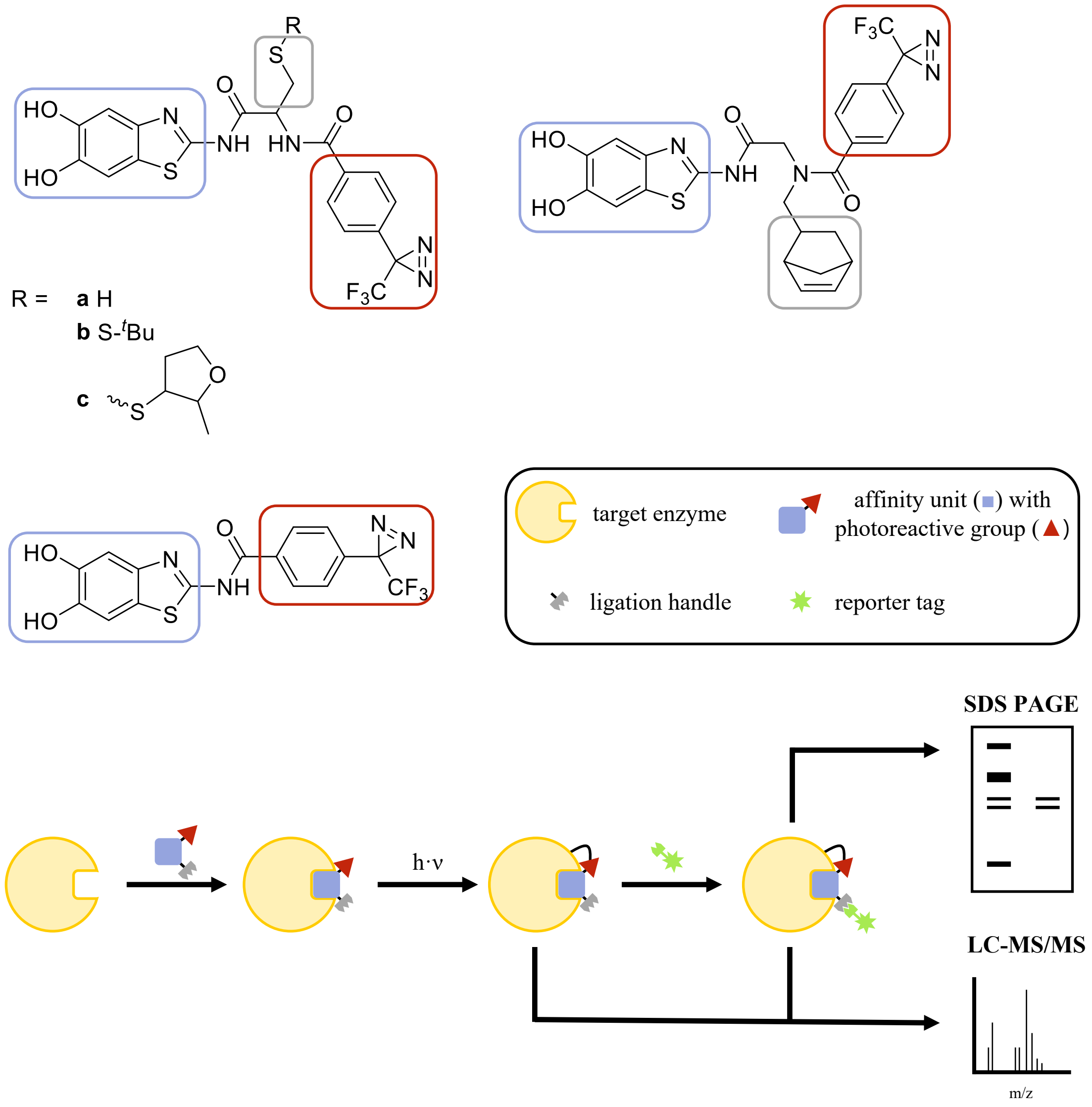


Figure 26. Bi- and trifunctional PAL tool compounds carrying an affinity unit (blue), a photoreactive group (red), and in the case of trifunctional compounds, a ligation handle (yellow). The affinity unit allows non-covalent interaction with the target enzyme. After irradiation, a covalent bond is formed. The covalent protein-ligand complex can be detected after reporter tag attachment in an SDS-PAGE gel or by LC-MS/MS.

Own contributions: tool compound synthesis, fluorometric enzyme assay, PAL experiments, SDS-PAGE gel analysis, molecular docking experiments, and writing of the original draft & editing of the manuscript.

Contributions from other authors: protein expression & purification, mass spectrometry experiments, and manuscript editing.

Manuscript submitted to RSC Medicinal Chemistry (impact factor: 3.47).

5.5.2 Manuscript

The following manuscript from page 220 to page 243 contains unpublished data.

Designing Photoaffinity Tool Compounds for the Investigation of the DENV NS2B-NS3 Protease Allosteric Binding Pocket

Hannah Maus[a], Andrea Gellert[a], Olivia R. Englert[a], Jia-Xuan Chen[b], Tanja Schirmeister[a], Fabian Barthels*[a]

[a] Johannes Gutenberg-University Mainz, IPBS, Staudingerweg 5, 55128 Mainz, Germany

[b]. Johannes Gutenberg-University Mainz, IMB, Ackermannweg 4, 55128 Mainz, Germany

* Corresponding author, barthels@uni-mainz.de

Abstract. Dengue virus (DENV) infection still lacks specific antiviral therapy, making the NS2B-NS3 protease an attractive target for drug development. However, allosteric inhibitors that bind to a site other than the active site still need to be better understood. In this study, we designed and synthesised tool compounds for photoaffinity labelling (PAL) to investigate the binding site of allosteric inhibitors on the DENV protease. These tool compounds contained an affinity moiety, a photoreactive group, and a reporter tag for detection. Upon irradiation, the photoreactive group formed a covalent bond with the protease, allowing for binding site identification. SDS-PAGE-based assays confirmed the qualitative binding of the designed inhibitors to the allosteric pocket, and pull-down experiments validated the interaction. Tryptic protein digestion following liquid chromatography/mass spectrometry analysis further supported the binding of the inhibitor to the proposed pocket revealing photo-attachment to an NS3 loop close to the C-terminus. These results enhance our understanding of allosteric inhibitors and their mechanism of action against the DENV protease. The developed tool compounds and PAL are potent tools for future drug discovery efforts and investigations targeting the DENV protease.

Introduction

Dengue virus (DENV) is a globally prevalent arbovirus that infects millions yearly, resulting in significant morbidity and mortality.[1,2] DENV belongs to the Flaviviridae family of viruses, including other important human pathogens such as Zika and yellow fever virus. Currently, no specific antiviral therapy is available to treat DENV infection, and the management of the disease is primarily symptomatic.[3,4]

Flaviviruses have a small single-stranded RNA genome that is translated into a polyprotein.[5,6] The cleavage of the polyprotein into seven structural and three non-structural proteins is essential for viral replication. Host proteases and the viral NS2B-NS3 protease carry it out.[7–9] The DENV NS2B-NS3 protease has emerged as a promising target for antiviral drug development. It is highly conserved across different viral serotypes and shows high homology between all flaviviral species.[10–12]

The viral protease exists in solution in an equilibrium between at least two different conformations, termed open and closed, which differ mainly in positioning the NS2B cofactor concerning the NS3 protease.[13–16]

Over the past decade, significant efforts have been made to identify and develop inhibitors of the DENV protease.[17–23] Competitive inhibitors, which compete with the natural substrate for binding to the protease's active site, have been extensively studied as potential drugs against the protease. The NS2B-NS3 protease is a chymotrypsin-like serine protease that prefers basic amino acids (Arg, Lys) in the substrate binding pocket.[24] This active site binding pocket is relatively shallow and difficult to address by small molecules. Therefore, mainly substrate-like inhibitors with dibasic or polybasic character are reported as competitive inhibitors.[8] In addition to the charged character and the associated poor bioavailability of competitive inhibitors, one major limitation of competitive inhibitors is their lack of inhibitory selectivity, as they can also inhibit host proteases that share a similar active site.[25]

Allosteric inhibitors of the DENV protease have been developed to overcome this issue. They bind to a site on the protein other than the active site, causing a conformational change that inhibits the protease activity.[26,27] In recent years, allosteric NS2B-NS3 inhibitors with different structures have been described in the literature.[5,19,20,23,28,29]

Despite the promise of allosteric inhibitors, their mechanism of action and binding site on the protease remains poorly understood. While several crystal structures for competitive inhibitors show the protease in the closed conformation[14,18], for allosteric inhibitors, it has been shown exclusively by methods in solution that they favour the protease's open conformation.[26,30] The co-crystal structures of DENV NS2B-NS3 protease in complex with a set of allosteric inhibitors were published[10] but are controversial discussed in the scientific community, which suggests that probably not an allosteric inhibitor but rather the C-terminus of NS3 was placed in the allosteric pocket.[31–33] A potential allosteric binding pocket of the NS2B-NS3 protease was described by Noble and Yildiz and is localised around the amino acid Ala-125.[11,14] Site-directed mutagenesis and maleimide labelling in this binding pocket reduced the binding affinity of benzo[*d*]thiazole-based allosteric inhibitors, suggesting the binding of these inhibitors in this pocket.[19]

To better understand the binding site of allosteric inhibitors, photoaffinity labelling (PAL) might provide a valuable technique that can identify the binding site of a ligand on a protein.[34,35] PAL involves the covalent attachment of a photoreactive group to a ligand, which can then be activated by UV light to form a covalent bond with the target protein.[35,36] This allows for the identification of the binding site of the ligand on the protein.

A suitable photoaffinity probe should ideally have a similar structure to known inhibitors and have comparable affinity and activity levels.[35,37] The photoreactive group should be able to be activated at a wavelength that does not destroy biological molecules such as proteins but only forms highly reactive intermediates (300–400 nm). Known classical photo crosslinkers include, for example, diazirines or aryl azides.[35,37] A photoaffinity probe requires two mandatory functionalities in a molecule, and one optional functionality can be introduced to report binding. Mandatory is an affinity moiety responsible for reversible binding to the target molecule. This can be based on a known inhibitor of the target protein. In addition, a photoreactive group is required, which, upon irradiation, forms a permanent covalent bond with the target molecule (protein). Finally, a reporter tag can be introduced, which is used to detect the formed covalent bond with a fluorescent dye binding or to enable the enrichment of labelled proteins using a pull-down enrichment. Reporter tags can be, for example, groups that are suitable for bioorthogonal click reactions.[35]

Here, we describe the design and synthesis of tool compounds for PAL that address the allosteric pocket of DENV protease. Using various reporter tags and click chemistry, we could detect the specific binding of the inhibitors using an SDS-PAGE gel. The analysis of labelled protein by LC/MS confirmed the binding of an inhibitor to the proposed binding pocket around Ala-125.

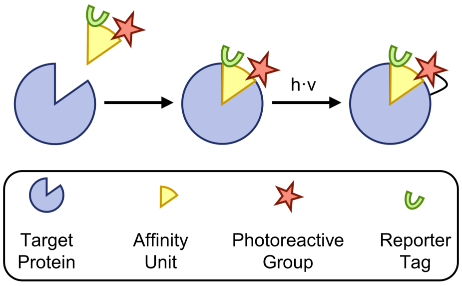


Figure 1. Schematic representation of the reaction of a PAL tool compound with a target protein upon irradiation to form a covalent protein-inhibitor complex.

Results and Discussion

Design and Synthesis of PAL Tool Compounds

Design. Based on literature-known allosteric inhibitors **1a–c**[20], different tool compounds suitable for the PAL of DENV protease were designed. As a common structural feature, all tool compounds have a benzo[*d*]thiazole building block responsible for the compound´s affinity towards the protease. As the photoreactive building block, 4-(3-(Trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid was used. Compound **2** represents the most straightforward combination of these two building blocks. However, this compound lacks a reporter tag to detect binding to the protein in an SDS-PAGE gel by a dye or biotin labelling. Therefore, additional tool compounds (compounds **3a–c** and **4**) were designed, each containing a reporter group that enables the detection of binding of these compounds to the target protein in an SDS-PAGE gel. In compounds **3a–c**, the free thiol of a cysteine moiety is used as a reporter to label it after irradiation with maleimide-bearing dyes. In addition to the tool compound with free cysteine **3a**, the cysteine was protected with two different reductively cleavable protecting groups, namely SS^{t}Bu **3b** and disulfanyl-2-methyl tetrahydrofuran (SS-Mot, **3c**). The latter can be cleaved faster and with higher conversion rates than the commonly used SS^{t}Bu protecting group.[38] The introduction of the thiol-protecting group guarantees the stability of the tool compounds during PAL irradiation and can reductively be removed before maleimide labelling of the inhibitor-protein conjugate. The norbornene reporter group in compound **4** can be labelled with tetrazine-bearing fluorescent dyes by an inverse electron demand Diels-Alder reaction (IEDDA).[39,40] In addition, this group was used to pull down protein-inhibitor conjugate by labelling with tetrazine-biotin and enrichment on streptavidin-bearing magnetic beads.

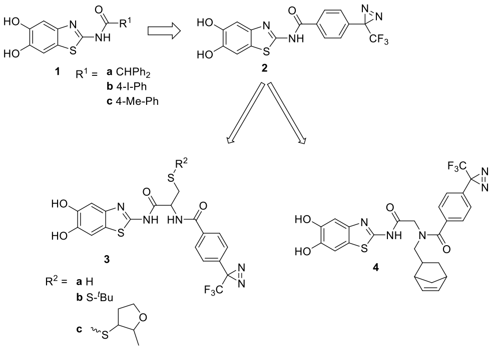


Figure 2. Design of PAL tool compounds.

Synthesis. The building block 5,6-dimethoxybenzo[*d*]thiazol-2-amine **6** was synthesised starting from 3,4-dimethoxyaniline **5** as described previously.[19,41] A base-catalysed TBTU-coupling of 5,6-dimethoxybenzo[*d*]thiazol-2-amine **6** with 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid and subsequent cleavage of the methyl ethers with BBr_3 gave compound **2**.

Compounds **3a–c** were obtained in 5-step synthesis starting from 5,6-dimethoxybenzo[*d*]thiazole **6**. Base-catalysed TBTU-coupling with Boc-cysteine(SS^{t}Bu) provided intermediate **7**. The Boc group was removed under acidic conditions. The intermediate hydrochloride obtained was reacted in a base-catalysed TBTU-coupling with 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid to give the dimethoxy-protected intermediate **8**. The dimethoxy- and the SS^{t}Bu protecting groups were cleaved in dichloromethane using ethanethiol and aluminium chloride, yielding compound **3a** with a free cysteine moiety. Different protecting groups were introduced using *S-tert*-butyl methanesulfonothioate or *S*-(2-methyltetrahydrofuran-3-yl)methanesulfonothioate, and final tool compounds **3b** and **3c** were obtained.

Scheme 1. Synthesis of PAL tool compounds. **A** Synthesis of building block 5,6-dimethoxybenzo[*d*]thiazol-2-amine **6**. Reagents and conditions a: KSCN, Br_2, acetic acid, 16 h, r.t. **B** Synthesis of compound **2**. Reagents and conditions a: 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid, TBTU, DIPEA, ethyl acetate, 2 d, 0 °C – r.t.; b: BBr_3 in DCM, 16 h, -78 °C – r.t. **C** Synthesis of PAL tool compounds with thiol reporter group **3a–c**. Reagents and conditions a: Boc-Cys(SS^{t}Bu)-OH, TBTU, HOBt, DIPEA, ethyl acetate, 2 d, 0 °C – r.t.; b: HCl in dioxane, 2 h, r.t.; c: 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid, TBTU, HOBt, DIPEA, ethyl acetate, 2 d, 0 °C – r.t.; d: ethanethiol, $AlCl_3$ in DCM, 2 h, 0 °C – r.t.; e: *S-tert*-butyl methanesulfonothionate or *S*-(2-methyltetrahydrofuran-3-yl)methanesulfonothionate, methanol, triethylamine, 2 h, r.t. **D** Synthesis of compound **4**. Reagents and conditions a: (Fmoc)-Gly-OH, triethylamine, DCM, 16 h, r.t.; b: 2x 40% piperidine in DMF, 10 min, r.t.; c: norbornenecarbaldehyde, acetic acid cat., DMF, 2 h, r.t.; d: sodium borohydride, THF/MeOH (1:1), 2 h, r.t.; e: TFA/DCM (1:1), 3 h, r.t.; f: di-*tert*-butyldicarbonat, triethylamine, water, 3 h, r.t.; g: 5,6-dimethoxybenzo[*d*]thiazol-2-amine, TBTU, DIPEA, ethyl acetate, 2 d, 0 °C – r.t; h: HCl in dioxane, 2 h, r.t.; i: 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid, TBTU, DIPEA, ethyl acetate, 2 d, 0 °C – r.t.; j: BBr_3 in DCM, 2 h, -78 °C – r.t

Compound **4** was obtained in a solid phase reaction using 2-chlorotrityl chloride resin. Fmoc-protected glycine was covalently bound base-catalysed to the resin. A 40% piperidine solution in DMF removed the Fmoc protecting group. Norbornene carboxaldehyde was then used to form the imine **12** *in situ*. Sodium borohydride reduced the imine to the secondary amine **13**. Subsequently, the intermediate product was cleaved from the solid phase by TFA and DCM (1:1). Before a base-catalysed TBTU-coupling with 5,6-dimethoxybenzo[*d*]thiazole was carried out, the secondary amine **14** was protected using Boc anhydride. The Boc group was cleaved after the coupling under acidic conditions. Compound **4** was then obtained by coupling with 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid and cleaving the aryl methyl ethers using BBr_3.

Photochemical Reactivity of the PAL Tool Compounds

To determine the optimal irradiation conditions, the PAL tool compounds were first irradiated without protein in a buffered solution at 365 nm, and the products and intermediates formed were analysed by LC/MS. In addition to the intensity, variable parameters during irradiation are the distance between the radiation source and the sample solution. The latter was fixed at 2 cm for all experiments. The duration of irradiation was limited to a maximum of 30 minutes to prevent possible damage to the protease and the associated loss of function.[34,42,43] To determine the optimal conditions, exemplarily, a solution of a PAL tool compound **3b** in buffer (50 mM Tris pH 9.0, 1 mM CHAPS) with 10 vol% DMSO was irradiated in a quartz cuvette under argon atmosphere. A sample was analysed by LC/MS before the start of irradiation and after 30 s, 1 min, 5 min, 10 min, 15 min, 20 min, 25 min, resp. 30 min of irradiation time. When the diazirine is irradiated, nitrogen is released, and a carbene is formed. This subsequently reacts with the protein or the solvent (water) to form a covalent bond. In addition to this stable photo adduct, the more long-living linear diazo compound is formed as an unwanted side product upon irradiation (Figure 3A).[44,45] The percentage of the reactant, the photo adduct with water and the linear by-product was determined from the relative peak areas of the UV spectra at 254 nm. Figure 3B shows the mass fractions of the described adducts after different irradiation times for tool compound **3b**. It is shown that the proportion of the reactant decreases with increasing irradiation time, while the ratio of the photo adduct with water and the linear form increases. However, a complete conversion of the reactant could not be achieved after 30 min. Yet, a sub-stochiometric conversion is sufficient to form protein-inhibitor adducts since the tool compound is used in excess in protein labelling. In addition to the pulsed irradiation, continuous irradiations (without interruption) were carried out for the tool compounds in the buffer. Slightly different results were obtained. The optimised irradiation conditions defined for protein labelling are listed in Table 1.

No irradiation conditions could be established for compound **3a** since complete decomposition of the compound was observed after an irradiation period of 30 s, demonstrating the necessity of a protecting group to ensure sufficient stability during irradiation.

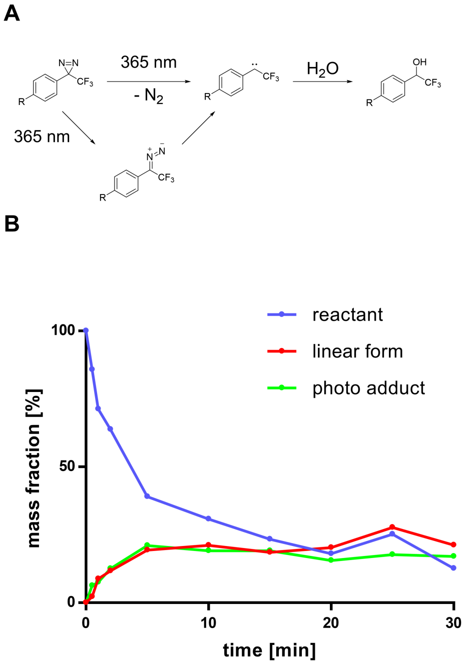
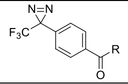


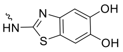
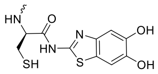
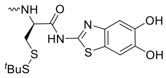
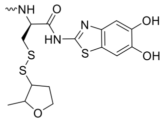
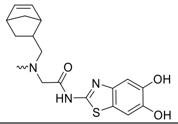
Figure 3. Photochemical reactivity of diazirines. **A** Decay of the diazirine during irradiation. **B** Mass fraction of the reactant, the linear form and the photo adduct upon irradiation of compound **3b** in buffer (50 mM Tris HCl pH 9.0, 1 mM CHAPS).

Reductive Deprotection of PAL Tool Compounds with Thiol-Protecting Group

Compounds **3b** and **3c** are derivatives of compound **3a** with a thiol-protecting group, which ensures the stability of the compounds during irradiation, but subsequently has to be cleaved off for labelling with a reporter dye. Both groups are reductively cleavable. TCEP was used as the reducing agent, offering the advantage over DTT that residues of the reducing agent do not interfere with the subsequent maleimide click reaction with fluorescent dyes. Compounds **3b** and **3c** were incubated with 1 mM TCEP in Tris buffer at pH 6.8 and pH 9.0, respectively, and the cleavage of the protecting group was monitored by LC/MS. The percentage of free thiol as the peak area of the UV spectrum at 254 nm was determined after 30 min, 1 h, and 2 h.

Table 1. Optimised irradiation conditions of PAL tool compounds.



#	R	Intensity	Duration
2		10% (1 W·cm^{-2})	30 min
3a		-	-
3b		6% (0.6 W·cm^{-2})	15 min
3c		6% (0.6 W·cm^{-2})	5 min
4		6% (0.6 W·cm^{-2})	15 min

For compound **3b**, which carries an SS^{t}Bu protecting group, a mass fraction of about 10% free thiol was detected at pH 9.0 after 30 min. The proportion could not be significantly increased by more prolonged incubation. At pH 6.8, no cleavage of the protecting group could be detected even after 2 h. The reductive cleavage of the SS-Mot protecting group is reported in the literature to be faster and more efficient than the cleavage of the SS^{t}Bu protecting group.[38] We confirmed this with compound **3c,** which privileged the SS-Mot protecting strategy for our application. At pH 9.0, complete cleavage was detected after 30 min. Reducing the pH to 6.8 resulted in 75% cleavage, indicating that a higher pH of 9.0 is more advantageous.

Table 2. Reductive cleavage of protecting group of compounds **3b** and **3c** by TCEP in buffer (50 mM Tris pH 9.0, 1 mM CHAPS, 1 mM TCEP).

		fraction of free thiol after				
#	**pH**	**15 min**	**30 min**	**1 h**	**2 h**	**4 h**
3b	6.8	n.d.	0%	0%	0%	4%
	9.0	n.d	10%	14%	9%	7%
3c	6.8	76%	85%	n.d.	n.d.	n.d.
	9.0	94%	85%	n.d.	n.d.	n.d.

Inhibitory Activity of PAL Tool Compounds

The inhibitory activity of the tool compounds was investigated by a fluorometric assay. Cleavage of the substrate Boc-Gly-Arg-Arg-AMC by the NS2B-NS3 protease results in the formation of free AMC and, consequently, the increase in fluorescence (excitation 380 nm, detection 460 nm). The increase in fluorescence per unit time (slope) can be considered a measure of the activity of the protease. Inhibition of the protease resulted in a decrease in the fluorescence increase. From a dilution series of the tool compounds, IC$_{50}$ values were determined from the relative inhibition compared to a control measurement with DMSO. For example, the fluorescence increases within 10 min in the presence of seven different concentrations of compound **3c** and the resulting IC$_{50}$ curves are shown in Figure 4. The IC$_{50}$ values of the different tool compounds were between 10 µM and 20 µM matching the parent compounds **1a–c** potency (Table 3). Consequently, the affinity was only slightly affected by the structural variations of the different tool compounds.

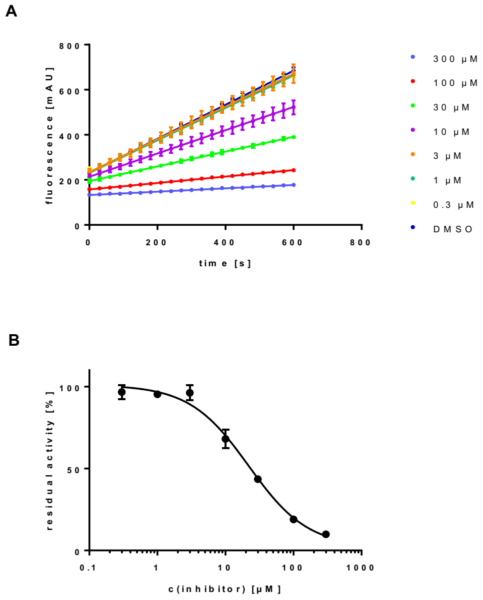


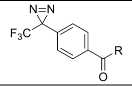
Figure 4. Inhibitory activity of compound **3c**. **A** Fluorescence increase over time for different concentrations of **3c**. **B** IC$_{50}$ curve resulting from A. A four parament Hill fit determined the IC$_{50}$ to be 20.0 ± 2.0 µM.

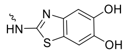
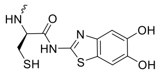
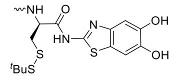
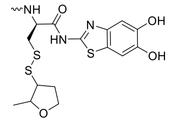
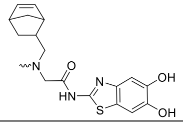
Detection of Photoaffinity Labelling by Click Reaction

For the following experiments, compounds **3c** and **4** were considered. Compound **2** does not carry a reporter group, so click reaction cannot detect its covalent binding to the protease. Compound **3c** was advantageous compared to compound **3b** due to the more efficient protecting group cleavage.

For photoaffinity labelling of DENV NS2B-NS3 protease, a 10-fold excess of the respective inhibitor was used compared to the protein concentration. The reporters integrated into the respective tool compound can demonstrate the successful formation of a covalent adduct between the PAL tool compound and protein. Fluorescent dyes with appropriate functionality were used for this purpose. Compounds **3b** and **3c** have a thiol as a reporter group, which maleimides can address.

Table 3. IC_{50} values for PAL tool compounds.



#	R	IC_{50} (µM)
2		18.8 ± 2.0*
3a		13.4 ± 2.1
3b		10.2 ± 1.5
3c		20.0 ± 2.0
4		17.0 ± 1.3

* The IC_{50} was obtained from five different inhibitor concentrations (0, 0.3, 1, 3, 10, 30 µM).

Since the DENV protease is free of native cysteine residues, labelling only occurs if the inhibitor previously introduced a free thiol into the inhibitor-protein complex. To avoid non-specific maleimide labelling of the protease, the sample solution was buffered to pH 6.8 using centrifuge filters after irradiation and cleavage of the protective group. On the one hand, this suppresses the labelling of lysin and histidine residues in the protein by maleimide labels. On the other hand, free inhibitor, which is not covalently bound to the protein, is removed from the solution. In addition, the excess reducing agents could be removed, which would interfere with the dye labelling of the thiol with maleimide. Maleimide functionalised Cy5 was added to the sample solution (2–6 eq.), and the reaction mixture was incubated for 2 h in the dark. This allowed the click reaction between the free thiol and the maleimide to form a covalent bond. Dye labelling of the protein and, thus, successful photoaffinity labelling of the protein was demonstrated using an SDS-PAGE gel (Figure 5). For this purpose, the gel was scanned with an excitation and detection wavelength specific to the Cy5 label (633 nm/670 nm). In addition, the protein was stained with Coomassie brilliant blue to monitor the total protein content in the gel. The intensity of the bands of the unirradiated sample without inhibitor (blind$^-$), the unirradiated sample with inhibitor (blind$^+$), and the irradiated sample (irr$^+$) were compared (Figures 5A and B). The Coomassie staining confirmed that the same amount of protein was loaded in each case, and the protein was stable under the irradiation conditions. The fluorescent Typhoon scan (Cy5) showed that a weak fluorescent band is visible without irradiation. The sample of the protein-inhibitor solution after irradiation, on the other hand, shows a clear band at the level of the NS2B-NS3 protease. The labelling by compound **4** showed similar results (see SI), confirming specific PAL labelling by the designed tool compounds.

Even though non-specific Cy5-maleimide labelling could not be suppressed entirely, this experiment clearly shows that the formation of a covalent protein-inhibitor adduct can be formed by irradiation and subsequently detected by dye labelling.

In addition to Cy5 labelling, we demonstrated the successful formation of the covalent protein-inhibitor complex by enrichment using biotin-streptavidin pull-down in an SDS-PAGE gel. For this purpose, the irradiated sample of compound **4** and the DENV protease was mixed with tetrazine-functionalised biotin and incubated overnight for norbornene IEDDA ligation. Subsequently, the protein-inhibitor complexes covalently bound to biotin were immobilised on magnetic beads carrying streptavidin. Several washing steps then removed unlabelled protein and free inhibitor. To detach the complexes from the magnetic spheres, these were taken up in Lämmli buffer with biotin and SDS in excess for elution. After heating to 95 °C for 10 min, the sample was applied to an SDS-PAGE gel. A single protein band indicated the successful formation of the covalent inhibitor-protein complex and showed that only intact NS2B-NS3 were enriched from the mixture. Several closely spaced bands are visible in the gel for the negative sample (Figure 5C, line ii), which contained only non-irradiated DENV NS2B-NS3 protein previously attributed to auto-cleavage of the protease.[46]

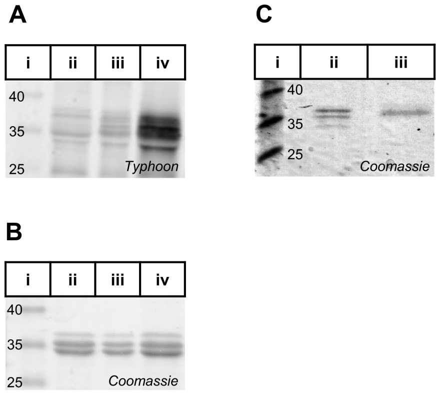


Figure 5. Scan of SDS-PAGE gels of PAL experiments with DENV NS2B-NS3 protease. **A** Typhoon Scan of the gel. i) marker; ii) control blind$^-$: non-irradiated sample of DENV NS2B-NS3 after labelling with Cy5; iii) control blind$^+$: non-irradiated sample of DENV NS2B-NS3 with compound **3c** after labelling with Cy5; iv) Sample irr$^+$: irradiated sample of DENV NS2B-NS3 with compound **3c** after labelling with Cy5. **B** Scan of the Coomassie-stained gel. i) marker; ii) blind$^-$: DENV NS2B-NS3; iii) blind$^+$: DENV NS2B-NS3 + **3c**; iv) irr$^+$: DENV NS2B-NS3 + **3c**. **C** Scan of the Coomassie-stained gel after Biotin-Streptavidin pull-down of the DENV NS2B-NS3 with compound **4**. i) marker; ii) DENV NS2B-NS3; iii) DENV NS2B-NS3 + **4** after Biotin-Streptavidin pull-down.

Mass spectrometry of Covalent Protein-Ligand Complexes

The evidence on forming the covalent protein-ligand complex in the PAL experiment using SDS-PAGE gel is essential for further investigations. However, detecting covalent binding alone does not reveal where the binding pocket addressed by the ligand is located in the protein. To obtain this information, samples of the covalent protein-ligand complexes from the PAL experiments were first proteolytically digested, and the tryptic peptide fragments obtained were then analysed by liquid chromatography-mass spectroscopy. Sequence coverage of more than 99% was obtained for the DENV protease. For prototypic inhibitor **2**, this experimental setup could detect covalent modification of the protein at position G121 (Figure 6A). G121 is located in the loop$_{117–121}$, part of the allosteric binding pocket proposed by Noble et al.[14] (Figure 6B). Docking experiments of compound **2**, performed analogously to Wu et al.[23],proposed the location of the trifluormethyldiazirine near G121 (Figure 6C), supporting PAL and LC/MS results.

To further support these results, maleimide blockage assays which have already been described for the class of benzo[*d*]thiazole inhibitors[19,20], were performed with compound **2**. For this purpose, a wildtype (WT) cysteine mutant T122C was used (Figure 6B), which allowed blockage of the proposed allosteric binding pocket with *N*-benzyl maleimide (BMI). A comparison of the affinities of compound **2** with and without blockage of T122C (Figure 6D) confirmed the binding pocket, which was also found in the PAL experiments. Blocking the proposed allosteric pocket with BMI increased the IC_{50} of compound **2** (21.81 ± 1.82 µM) approximately by a factor of 2, compared to the unblocked binding pocket of T122C (11.64 ±2.41 µM). The unaffected IC_{50} of the WT in the presence of BMI (19.42 ± 4.06 µM) compared to the IC_{50} of the WT without BMI (18.82 ± 1.96 µM) shows that the increase in IC_{50} is not due to the presence of BMI. The reduction of the affinity of compound **2** for the protein by blocking the binding pocket located next to T122C confirms the interaction of the inhibitor with the proposed allosteric binding pocket.

Conclusions

Here we describe the synthesis of different PAL tool compounds for the allosteric binding pocket of the DENV NS2B-NS3 protease. PAL allows the location of the interaction between a ligand and a target to be identified. For this purpose, a covalent bond is formed between the ligand carrying a photoreactive group and the target protein by irradiation at 365 nm. In addition to the prototypic compound **2,** consisting of an affinity unit and a photoreactive group, further compounds were synthesised, allowing the formation of the covalent protein-ligand complex to be detected by fluorescent labelling or pull-down with a reporter tag. The demonstration of the successful formation of the complex upon irradiation is an essential prerequisite for further experiments and encouraged us to search for the binding site of the ligand in the protein. Using protein digestion and mass spectrometry, we were thus able to show that the class of allosteric inhibitors studied here, based on a benzo[*d*]thiazole, bind into the binding pocket around Ala-125 proposed by Noble et al.[14] In this way, our study not only demonstrated different strategies for the design and synthesis of PAL tool compounds but also confirmed the location of the allosteric binding pocket by mass spectrometry of the labelled protein and maleimide blockage assays. Our results can form the basis for developing new potent allosteric inhibitors. The demonstrated interaction of benzo[*d*]thiazole-based inhibitors with the region around G121/T122C offers new starting points for the introduction of functional groups into this class of inhibitors that enable, for example, a covalent interaction with T122 or T120. In this way, it might be possible to design covalent allosteric inhibitors.

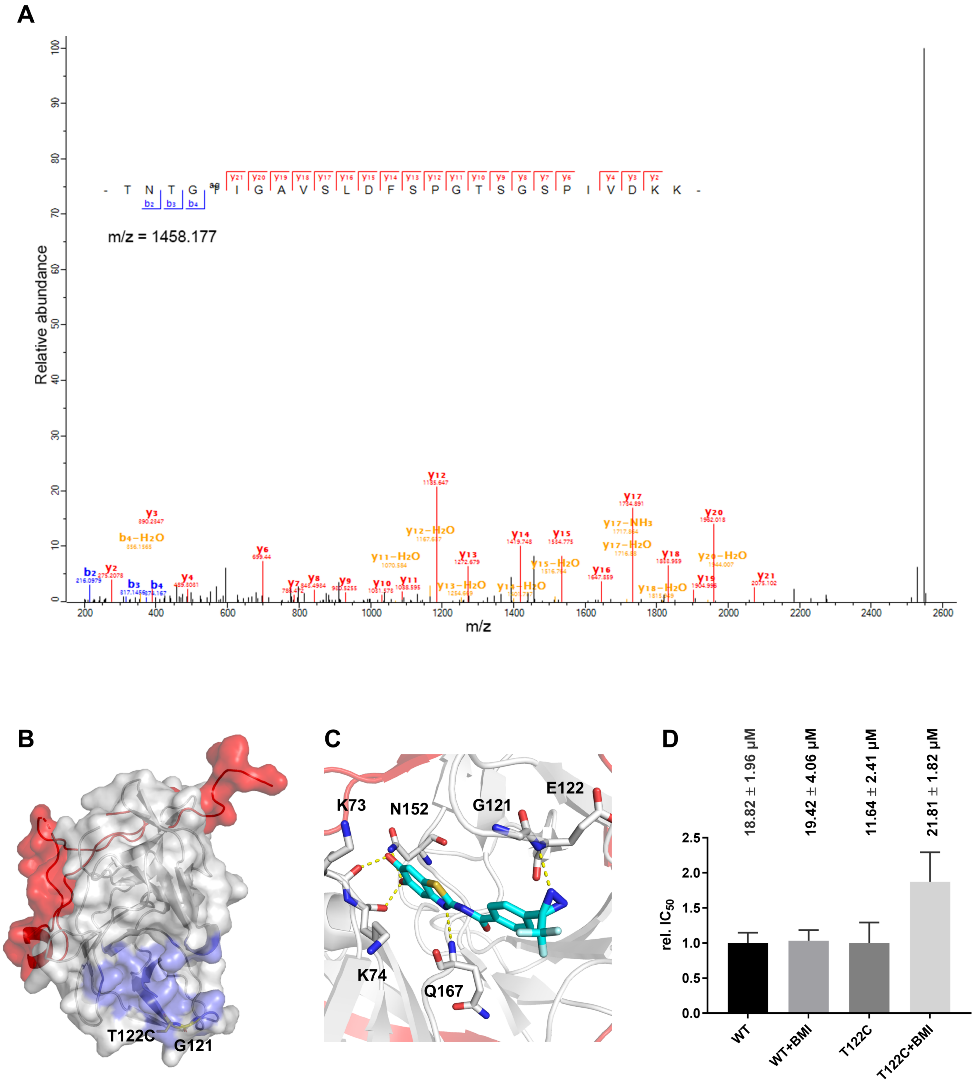


Figure 6. Identification of the binding pocket of compound **2**. **A** Results of mass spectroscopic studies of the digested protein after covalent modification by compound **2** by PAL experiments. **B** Structure of the dengue virus protease (PDB: 2FOM).[13] NS2B is highlighted in red, NS3 in grey and the allosteric pocket in blue. The modified amino acid G121 and the mutated amino acid T122C are shown as yellow sticks. **C** The by docking proposed binding pose of compound **2** in the allosteric pocket of the closed conformation of the DENV3 protease (PDB: 3U1I)[14]. Compound **2** is shown in cyan. NS2B is highlighted in red, NS3 grey. Amino acids forming polar interactions are labelled and highlighted with sticks. Note that at position 122, there is an E in DENV3 instead of T. **D** Results of the maleimide blockade experiments with compound **2**. Relative IC_{50} values for the wildtype (WT) used for the PAL experiments and T122C mutant of the wildtype (T122C), which was used for maleimide blockade experiments with (+) and without *N*-benzyl maleimide (BMI). The absolute IC_{50} values are given as numerical values. The quotient of the respective absolute IC_{50} for the corresponding protein with and without BMI was formed to provide relative IC_{50} values

Material and Methods

Photochemical Studies

Laser and experimental setup. The photoactivation studies were performed according to procedures as reported by Brunner et al.[45] and Seifert et al.[47] using a high-power UV LED (λ_{max} = 365 nm; 5 W; UV LED smart; Opsytec Dr. Gröbel GmbH) with the "standard" optic for a small spot diameter and a focused beam profile. Please refer to the homepage of the UV light source supplier (https://www.opsytec.com/products/uv-led-light-sources) for further details. Irradiation experiments were conducted within a ventilated fume hood in quartz cuvettes (d = 10 mm), including a tiny magnetic stirring bar, placed 2 cm from the light source for effective photoactivation. Quartz cuvettes were sealed with a septum so the experiments could be carried out under an argon atmosphere.

PAL Experiments for Gel-Assay. Experiments were carried out in Tris buffer (50 mM Tris pH 9.0, 1 mM CHAPS) containing 10 vol% of DMSO to ensure the solubility of PAL tool compounds. PAL tool compounds were used at a concentration of 100 µM. The protease was used at a final concentration of 10 µM. The buffer was degassed with Argon and filtered before use. The samples with protein were irradiated under the settings described in Table 1 after an incubation time of 5 min. After irradiation, incubation was continued for another 10 min to allow a complete reaction of the carbene. TCEP dissolved in buffer was then added to the sample at a final concentration of 1 mM. Using Amicon Ultra centrifuge filters (Millipore, cutoff 10 kDa, volume 0.5 mL), the buffer was exchanged to remove excess TCEP and adjust the pH to 6.8 (50 mM Tris pH 6.8, 1 mM CHAPS). Subsequently, an excess of 2−6 eq. Cy5 maleimide or 10 eq. Cy5 tetrazine dissolved in DMF was added. Dye labelling was done in the dark at room temperature within one hour for Cy5 maleimide and overnight for Cy5 tetrazine. The labelled protein was then rewashed using the centrifuge filters to remove excess dye. Lämmli buffer was added to 10 µL of the resulting solution, which was heated to 95 °C for 10 min and applied to an SDS-PAGE gel. Control samples (blind$^-$ and blind$^+$) were treated identically without irradiation.

PAL Experiments for Pull-Down. After irradiation, a 100-fold excess of tetrazine-biotin was added to the samples and incubated overnight at room temperature. Subsequently, streptavidin immobilised on magnetic beads (100 µL, Pierce, ThermoScientific) was added. After 5 min incubation, the buffer solution containing excess biotin-tetrazine was removed, and the spheres were washed three times with 200 µL buffer. Before applying to the gel, 10 µL of Lämmli buffer supplemented with biotin (final concentration 10 µM) was added to the dry beads and heated to 95 °C for 10 min.

SDS-PAGE gels. For gel electrophoresis, separation gels with 15% acrylamide and stacking gels with 4% acrylamide were used. The gels were produced with a thickness of 1 mm. The running buffer was 25 mM Tris pH 6.8, 192 mM glycine and 0.1% SDS. An 80-100 V voltage was applied for 10 min, then increased to 120-160 V. The gels were run at this voltage for 10 min. The marker PageRuler Prestained Protein Ladder from Thermo Fisher Scientific was applied as a size reference.

The SDS-PAGE gels were scanned on the AMERSHAM BIOSCIENCE TYPHOON TRIO+ scanner. The programme TYPHOON SCANNER CONTROL version 5.0 was used. For Cy5, an emission filter at 670 nm (670 BP) with a width of 30 nm was used. A red laser with a wavelength of 633 nm and a gain of 400 to 600 V was used for excitation. The gels were stained with Coomassie brilliant blue staining solution for one to two hours. After subsequent destaining of the gels with acetic acid, a colourimetric scan of the gels was taken.

Enzyme Assay

Fluorometric enzyme assays were performed as described previously.[19,20] Briefly, 250 nM of the purified protease in buffer (50 mM Tris pH 9.0, 1 mM CHAPS) with 5 µL substrate (100 µM Boc-Gly-Arg-Arg-AMC) in DMSO and 10 µL of the corresponding inhibitor in DMSO resulting in a total volume of 200 µL was added to a 96 well plate. The fluorescence of the released AMC was measured at 380 nm excitation and 460 nm emission. The percentage activity of the protease with the addition of the inhibitors was determined as the proportion of the slope with respect to the slope of the DMSO control.

AMC correction. To determine the AMC correction factors, the fluorescence of 1 µM AMC $F(AMC_{DMSO})$ was compared with the fluorescence of 1 µM AMC in the presence of the corresponding inhibitor $F(AMC_{inhibitor})$ concentration. For this purpose, 185 µL buffer (50 mM Tris pH 9.0, 1 mM CHAPS) with 5 µL AMC (final concentration 1 µM) and 10 µL of the corresponding inhibitor in DMSO were mixed, and the fluorescence was determined (λ_{ex} =380 nm, λ_{em} = 460 nm). The residual activities determined from the slopes in the fluorometric assay were multiplied by the correction factor of $F(AMC_{DMSO})/F(AMC_{inhibitor})$ to obtain the corrected residual activities.

Maleimide Blockage Assay. A stock solution of the corresponding protein (50 µM) in buffer (50 mM Tris pH 9.0, 1 mM CHAPS) was mixed with a 1.5-fold excess of *N*-benzyl maleimide (75 µM) in DMSO (1 vol%) and incubated on ice for 15 min. Subsequently, the fluorometric assay was performed as described above.

Mass spectrometry of Covalent Protein-Ligand Complexes

Protein in-gel digestion. The Coomassie-stained gel band was excised and cut into small cubes, followed by destaining in 50% ethanol/25 mM ammonium bicarbonate. The proteins were then reduced in 10 mM DTT at 56 °C and alkylated by 50 mM iodoacetamide in the dark at room temperature. Afterwards, proteins were digested by trypsin in 50 mM ammonium bicarbonate overnight at 37 °C. Following peptide extraction sequentially using 30% and 100% acetonitrile, the sample volume was reduced in a centrifugal evaporator to remove residual acetonitrile. The resultant peptide solution was purified by solid phase extraction in C18 StageTips.[48]

Liquid chromatography-tandem mass spectrometry. Peptides were separated in the EASY-nLC 1200 UHPLC system (Thermo Fisher Scientific) using an in-house packed 55 cm analytical column (inner diameter: 75 µm; ReproSil-Pur 120 C18-AQ 1.9 µm silica particles, Dr. Maisch GmbH) by online reversed-phase chromatography through a 90 min gradient of 2.4–32% acetonitrile with 0.1% formic acid at a nanoflow rate of 250 nl/min. Electrospray ionisation sprayed The eluted peptides directly into an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). Mass spectrometry measurement was conducted in data-dependent acquisition mode using a top10 method with one full scan (resolution: 60,000, scan range: 300-1650 m/z, target value: 3 × 106, maximum injection time: 60 ms) followed by ten fragmentation scans via higher energy collision dissociation (HCD; normalised collision energy: 30%; resolution: 30,000, target value: 1 × 105, maximum injection time: 60 ms, isolation window: 1.4 m/z). Only precursor ions of +2 to +7 charge state were selected for fragmentation scans. Additionally, precursor ions already isolated for fragmentation were dynamically excluded for 25 s.

Mass spectrometry data analysis. The raw data file was processed by the MaxQuant software package (version 2.1.3.0)[49] using its built-in Andromeda search engine[50]. Spectral data were searched against a target-decoy database consisting of the forward and reverse sequences of DENV NS2B-NS3 and a list of 246 common contaminants. Trypsin/P specificity was assigned. Carbamidomethylation of cysteine was set as a fixed modification. Oxidation of methionine, acetylation of the protein *N*-terminus and the TPD-labelled ligand of compound **2** ($C_{16}H_9F_3N_2O_3S$) on all 20 possible natural amino acids were assigned as variable modifications. A maximum number of five modifications per peptide was allowed. A maximum of two missed cleavages were tolerated. The minimum peptide length was set to seven amino acids. The false discovery rate (FDR) was set to 1% for both peptide and protein identifications. The processed MS/MS spectra were visualised and annotated in the Viewer module of MaxQuant.[51]

Molecular Docking

The crystal structure of DENV-3 protease in complex with the aldehyde inhibitor Bz-nKKR-H (PDB: 3U1I) was used for molecular docking.[14] Possible docking modes between the ligand and the protease were studied using the FlexX docking approach of the LeadIT 2.1.6 suite (BioSolveIt, Germany). The MOE software was used for the energy minimisation of compound structures (Molecular Operating Environment, 2012.10). All water and ligand molecules were neglected during the docking procedure. The binding site was defined on a proper protein pocket, which was shown to be a specific allosteric binding site for other noncompetitive inhibitors near the catalytic site.[52]

Chemistry

All solvents and reagents were obtained from commercial suppliers (Sigma-Aldrich, Acros Organics, Alfa Aesar, TCI, BLDpharm or Apollo Scientific) and used without prior purification, if not otherwise stated. Reactions at −78 °C were conducted in a dry ice/acetone cooling bath. Purifications by flash chromatography were performed on silica gel (0.015−0.040 mm, Machery-Nagel). Reaction progress was monitored by thin-layer chromatography (TLC) using Machery-Nagel Alugram Xtra Sil G/UV254 silica 60 plates. Compound visualisation on these plates was attained by irradiation at 254 nm. ^1H and ^{13}C NMR spectra were recorded on a Bruker Fourier 300, a Bruker Avance-III HD (^1H NMR, 300 MHz; ^{13}C NMR, 75.5 MHz), or a Bruker Avance-II (^1H NMR 600 MHz; ^{13}C NMR 151 MHz). All chemical shifts are referenced to the signal of the residual solvent ($CDCl_3$, 7.26 and 77.16 ppm; DMSO-d_6, 2.50 and 39.52 ppm; methanol-d_4, 3.31 and 49.00 ppm for ^1H NMR and ^{13}C NMR, respectively) and reported in parts per million (ppm) relative to tetramethylsilane (TMS). Multiplicities of the corresponding NMR signals are given using the following abbreviations: br = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, and combinations thereof.

Analytical HPLC analysis was performed on an HP Agilent 1100 series HPLC system using an Agilent Poroshell 120 EC-C18 (150 × 2.10 mm, 4 µm) column at 40 °C oven temperature and a detection wavelength of 254 nm. The mobile phase consisted of mixtures of acetonitrile, Milli-Q-grade water, and 10% of a 0.1% formic acid solution in Milli-Q-grade water. The purity of the final compounds was confirmed by HPLC analysis and was higher than 95% in all cases.

Method A. *O*-(benzotriazole-1-yl)-*N*,*N*,*N'*,*N'*-tetramethyluronium tetrafluoroborate (TBTU; 1.1 eq.), *N*-hydroxy-benzotriazole (HOBt; 1.1 eq.) and the respective carboxylic acid (1 eq.) were suspended in ethyl acetate. Under ice cooling at 0 °C, *N*,*N*-diisopropylethylamine (DIPEA; 2.5 eq.) was added. After 30 minutes, the corresponding amine (1 eq.) was added at RT and stirred for at least 24 hours at RT.
To terminate the reaction, water was added, the two phases were separated using a separating funnel, and the aqueous phase was washed three times with ethyl acetate. The combined organic extracts were extracted three times with hydrochloric acid (2 M), once with deionised water, three times with saturated sodium hydrogen carbonate solution and once with saturated saline solution and dried over sodium sulfate and filtered. The solvent was removed under reduced pressure, and the residue was lyophilised for further drying. The crude products were purified by column chromatography.

Method B. Approximately 4 mL HCl in dioxane (4 M) was added to the respective Boc-protected amine and stirred for two hours. Petroleum ether (40 °C – 60 °C) was then added, and the precipitated product was filtered off through a frit and washed with petroleum ether (40 °C – 60 °C).

Method C. Ethanethiol (1 mL) and $AlCl_3$ (10 eq.) were dissolved in DCM under an argon atmosphere. Under ice cooling at 0 °C, the corresponding aryl methyl ether (1 eq.) was added, and the mixture was slowly brought to RT and stirred for 24 hours.

The reaction was then terminated by adding hydrochloric acid (2 M), and the two resulting phases were separated using a separating funnel. The aqueous phase was washed three times with DCM, and the combined organic extracts were extracted with saturated saline, dried over sodium sulfate, and filtered. The solvent was removed under reduced pressure, and the residue was lyophilised for further drying. The crude products were purified by column chromatography.

Method D. The aryl methyl ether was dissolved in dry dichloromethane under an argon atmosphere. It was cooled to –78°C, and an excess (2–10 eq.) of BBr_3 in DCM was added. After a complete reaction, termination was done by adding methanol and excess solvent was removed. The crude products were purified by column chromatography.

Method E. A solution of a thiol (1 eq.) and the corresponding methane sulfonothionate (1.5 eq.) with triethylamine (1 eq) in ethanol was stirred overnight at room temperature. After removing the solvent, the crude product was purified by column chromatography.

5,6-Dimethoxy benzo[d]*thiazole-2-amine (**6**)* was synthesised following Trapani et al.[41] 3,4-Dimethoxyaniline (5 g, 32.7 mmol, 1 eq.) and potassium thiocyanate (6.98 g, 71.9 mmol, 2.2 eq.) were suspended in acetic acid, and bromine (1.64 mL, 32.7 mmol, 1 eq.) was slowly added under cooling at a temperature below 20 °C. The reaction mixture was stirred at RT for one day and then filtered. The residue was heated under reflux for 20 minutes in water and 1 mL hydrochloric acid conc. and then filtered. The boiling in the hydrochloric acid solution was repeated with the residue, and the filtrates obtained were combined. When the filtrates cooled, **6** crystallised out of the solution as hydrochloride. Colourless solid; yield: 2.58 g (1.1 mmol, 38%); m_p = 272 – 274 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 9.66 (s, 2H), 7.55 (s, 1H), 7.13 (s, 1H), 3.78 (d, J = 9,4 Hz, 6H); ^{13}C (75 MHz, DMSO-d_6) δ [ppm] = 169.2, 149.7, 147.1, 132.7, 114.9, 106.4, 98.9, 56.6, 56.4.

N-*(5,6-dihydroxybenzo*[d]*thiazol-2-yl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (**2**)* was synthesised following method A and D. A: 4-(3-(Trifluoromethyl)-3*H*-diazirin-3-yl)benzoic acid (40 mg, 0.22 mmol, 1.1 eq.), TBTU (71 mg, 0.22 mmol, 1.1 eq.), DIPEA (0.14 mL, 0.82 mmol, 4 eq.), **6** (40 mg, 0.20 mmol, 1 eq), DMF (5 mL). D: crude product from A (30 mg, 0.07 mmol, 1 eq.), BBr_3 in DCM 1 M (0.43 mL, 0.43 mmol, 6 eq.). The crude product was purified by column chromatography (DCM/MeOH 49:1 + 0.1% TFA). Yellow solid; yield: 27.0 mg (0.06 mmol 58%); m_p = 238 – 240 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 8.21 (d, J = 8.4 Hz, 2H), 7.44 (d, J = 8.4 Hz, 2H), 7.26 (s, 1H), 7.13 (s, 1H); ^{13}C (75 MHz, DMSO-d_6) δ [ppm] = 164.8, 158.7, 158.2, 145.8, 144.2, 134.1, 131.5, 129.2, 126.6, 123.6, 121.7, 120.0, 106.5, 105.9, 27.9; m/z for $[C_{16}H_{10}F_3N_4O_3S]^+$ calc. 395.0, found 394.9; purity: 97%.

tert-*Butyl-(*S*)-(3-(tert-butyldisulfanyl)-1-((5,6-dimethoxybenzol*[d]*thiazole-2-yl)amino-1-oxopropan-2-yl)carbamate (**7**)* was synthesised following method A. *N*-(*tert*-butoxycarbonyl)-*S*-(*tert*-butylthio)-L-cysteine (0.254 g, 0.82 mmol, 1 eq.), TBTU (0.29 g, 0.902 mmol, 1.1 eq.), HOBt (0.122 g, 0.902 mmol, 1.1 eq.), DIPEA (0.35 mL, 2.05 mmol, 2.5 eq.), **24** (0.202 g, 0.820 mmol, 1 eq.). No further purification of this intermediate was carried out. Colourless solid; yield: 0.383 g (0.76 mmol, 92%); m_p = 107 – 109 °C; $[\alpha]_{589}^{25}$ = - 20 (c = 0.005 $\frac{g}{mL}$; MeOH); ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 12.44 (s, 1H), 7.55 (s, 1H), 7.33 (d, J = 7.9 Hz, 1H), 7.30 (s, 1H), 4.50 (d, J = 7.3 Hz, 1H), 3.82 (d, J = 6.4 Hz, 6H), 3.12 – 3.03 (m, 2H), 1.40 – 1.35 (m, 9H), 1.29 – 1.28 (m, 9H); ^{13}C NMR (75 MHz, DMSO-d_6) δ [ppm] = 156.7, 155.6, 149.4, 147.5, 143.0, 123.4, 104.2, 104.0, 79.1, 56.4, 56.2, 54.5, 48.2, 42.4, 38.7, 30.0, 28.6.

*(*S*)-*N*-(3-(*tert*-Butyl-disulfanyl)-1-((5,6-dimethoxybenzo*[d]*thiazole-2-yl)amino)-1-oxopropan-2-yl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (**8**)* was synthesised following method B and A. B: **7** (0.3 g, 1 eq), HCl in Dioxane 4 M (2 mL) A: 4-(3-(Trifluormethyl)-3*H*-diazirin-3-yl)benzoic acid (0.10 g, 0.43 mmol, 1 eq.), TBTU (0.152 g, 0.473 mmol, 1.1 eq.), HOBt (0.064 g, 0.473 mmol, 1.1 eq.), DIPEA (0.18 mL, 1.075 mmol, 2.5 eq.), crude product from B (0.17 g, 0.43 mmol, 1 eq.). No further purification of this intermediate was carried out. Colourless solid; yield: 0.215 g (0.35 mmol, 79%); decomp. > 62 °C; $[\alpha]_{589}^{25}$ = - 28 (c = 0.005 $\frac{g}{mL}$; MeOH); ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 12.60 (s, 1H), 9.07 (d, J = 7.5 Hz, 1H), 8.09 – 7.98 (m, 2H), 7.55 (s, 1H), 7.46 – 7.38 (m, 2H), 7.29 (s, 1H), 5.05 – 4.93 (m, 1H), 3.82 (d, J = 5.8 Hz, 6H), 3.32 (s, 2H), 1.30 (s, 9H); ^{13}C (75 MHz, DMSO-d_6) δ [ppm]= 169.8, 166.1, 149.4, 147.5, 142.9, 135.7, 131.1, 129.0, 126.9, 123.4, 104.2, 104.0, 56.4, 56.2, 54.0, 48.3, 41.9, 30.0.

*(*S*)-*N*-(1-((5,6-Dihydroxy-benzo*[d]*thiazole-2-yl)-amino)-3-mercapto-1-oxopropan-2-yl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (**3a**)* was synthesised following method C. **8** (0.1 g, 0.15 mmol, 1 eq.), $AlCl_3$ (0.2 g, 1.5 mmol, 10 eq.) und 1 mL ethanethiol. The crude product was purified by HPLC. Colourless solid; yield: 0.045 g (0.09 mmol, 56%); m_p = 152 – 154 °C; $[\alpha]_{589}^{25}$ = + 22 (c = 0.005 $\frac{g}{mL}$; MeOH); ^{1}H NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.37 (s, 1H), 9.19 (s, 2H), 8.97 (d, J = 7.4 Hz, 1H), 8.08 – 8.03 (m, 2H), 7.43 (d, J = 8.1 Hz, 2H), 7.21 (s, 1H), 7.09 (s, 1H), 4.80 – 4.75 (m, 1H), 3.05 – 3.00 (m, 1H), 2.97 – 2.92 (m, 1H), 2.61 – 2.57 (m, 1H); ^{13}C NMR (151 MHz, DMSO-d_6) δ [ppm] = 169.2, 165.8, 158.4, 158.1, 155.2, 145.6, 144.0, 141.9, 135.2, 130.7, 128.6, 128.4, 126.5, 124.5, 122.7, 121.9, 120.9, 106.5, 106.4, 56.7, 40.1, 28.3, 28.0, 25.4; m/z for $[C_{19}H_{15}F_3N_5O_4S_2]^+$ calc. 498.1, found 498.2; purity: 99%.

*(*S*)-*N*-(3-(*tert*-Butyl disulfanyl)-1-((5,6-dihydroxybenzo*[d]*thiazol-2-yl)amino)-1-oxopropan-2-yl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide (**3b**)* was synthesised following method E. **3a** (0.045 g, 0.088 mmol, 1 eq.) (*S*)-*tert*-Butylmethan-sulfonothionate (0.022 g, 0.133 mmol, 1.5 eq.), triethylamine (1 eq.), ethanol. The crude product was purified by HPLC. Colourless solid; yield: 10.5 mg (0.02 mmol, 20%); m_p = 135 – 137 °C; $[\alpha]_{589}^{25}$ = + 2,5 (c = 0.005 $\frac{g}{mL}$; MeOH); ^{1}H NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.48 (s, 1H), 9.08 (d, J = 7.5 Hz, 1H), 8.03 (d, J = 8.5 Hz, 2H), 7.43 (d, J = 8.2 Hz, 2H), 7.22 (s, 1H), 7.09 (s, 1H), 4.97 – 4.90 (m, 1H), 3.30 – 3.19 (m, 3H), 3.12 – 3.07 (m, 1H), 1.30 (s, 9H); ^{13}C NMR (151 MHz, DMSO-d_6) δ [ppm] = 169.2, 165.6, 158.3, 158.0, 155.2, 145.6, 144.0, 141.9, 135.2, 130.7, 128.6, 126.5, 122.7, 122.0, 120.9,

106.5, 106.4, 53.5, 47.9, 45.7, 41.4, 40.1, 29.6, 28.3, 28.0, 27.7, 8.7; m/z for $[C_{23}H_{23}F_3N_5O_4S_3^+]$ calc. 586.1, found 585.9; purity: 97%.

N-*((2S)-1-((5,6-Dihydroxy benzo[*d*]thiazole-2-yl) amino)-3-((2-methyl tetrahydrofuran-3-yl) disulfanyl)-1-oxopropan-2-yl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide* (***3c***) was synthesised following method E. **3a** (0.026 g, 0.052 mmol, 1 eq.), *S*-(2-methyltetra-hydrofuran-3-yl) methansulfonothionate (0.022 g, 0.133 mmol, 1.5 eq.), triethylamine (1 eq.), ethanol. The crude product was purified by HPLC. Colourless solid; yield: 13 mg (0.02 mmol, 41%); m_p = 99 – 101 °C; ^{1}H NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.57 – 12.50 (m, 1H), 9.14 – 9.09 (m, 1H), 8.03 (d, J = 8.5 Hz, 2H), 7.44 (d, J = 8.1 Hz, 2H), 7.22 (s, 1H), 7.09 (s, 1H), 5.02 – 4.94 (m, 1H), 4.03 – 3.95 (m, 1H), 3.82 – 3.78 (m, 1H), 3.74 – 3.71 (m, 1H), 3.36 – 3.31 (m, 1H), 3.24 – 3.17 (m, 1H), 2.34 – 2.24 (m, 1H), 2.04 – 1.97 (m, 1H), 1.92 – 1.85 (m, 1H), 1.26 – 1.20 (m, 1H), 1.20 – 1.12 (m, 3H); ^{13}C NMR (151 MHz, DMSO-d_6) δ [ppm] = 165.6, 155.2, 145.6, 144.0, 141.9, 135.2, 130.7, 128.5, 126.5, 122.7, 120.9, 120.1, 106.4, 80.0, 79.8, 76.5, 76.5, 65.8, 65.8, 65.2, 53.2, 53.1, 53.1, 53.1, 40.1, 32.6, 32.6, 32.2, 32.1, 28.3, 28.0, 19.8, 19.7, 16.7, 16.5; m/z for $[C_{24}H_{23}F_3N_5O_5S_3^+]$ calc. 614.1, found 613.9; purity: 96%.

(Bicyclo[2.2.1]hept-5-en-2-ylmethyl)glycine (***14***) 2Cl-Trt-Cl resin (1 g, 1 mmol) was placed in a syringe. Then Fmoc-Gly-OH (1 g, 3 mmol) was dissolved in DCM and drawn up with the syringe together with TEA (1 mL). The mixture was shaken at RT for 16 hours. Subsequently, piperidine (40%) in DMF was added twice to remove the Fmoc group and incubated for 10 minutes. Norbornenecarboxyaldehyde (1 mL) in DMF (4 mL) with a drop of acetic acid was then added to the syringe and incubated for one hour. The 2ClTrt-Cl resin was transferred to a round bottom flask and suspended in a mixture of THF/MeOH (1:1). Sodium borohydride (200 mg) was added in portions. The resin was then transferred to the syringe and washed with MeOH and THF. The mixture was shaken with TFA/DCM (1:1) for 3 hours at RT. DCM was then removed under reduced pressure, and the product was lyophilised with an acetonitrile/water mixture. Colourless solid; yield: 0.3 g (1.66 mmol, 49%); decomp. > 110 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 8.95 (s, 1H), 8.54 – 7.59 (m, 1H), 6.60 – 5.70 (m, 2H), 3.88 (d, J = 9.4 Hz, 1H), 3.81 – 3.41 (m, 1H), 3.14 – 2.91 (m, 1H), 2.86 – 2.73 (m, 1H), 2.64 (d, J = 7.2 Hz, 1H), 2.37 (dtd, J = 11.5, 7.6, 3.9 Hz, 1H), 1.95 – 1.60 (m, 1H), 1.35 (dq, J = 8.3, 2.1 Hz, 1H), 1.28 – 1.13 (m, 2H), 0.56 (ddd, J = 11.7, 4.4, 2.5 Hz, 1H); ^{13}C NMR (75 MHz, DMSO-d_6) δ [ppm] = 168.6, 168.6, 138.4, 137.4, 136.4, 132.4, 51.5, 49.5, 47.8, 47.6, 45.2, 44.3, 44.2, 42.4, 41.7, 35.9, 35.7, 31.2, 30.7, 9.0.

N-*(Bicyclo[2.2.1]hept-5-en-2-ylmethyl)*-N-*(*tert-*butoxycarbonyl)glycine* (***15***)*.* **14** (270 mg, 1.5 mmol) was dissolved in DCM (20 mL), and then Boc_2O (261 mg, 1.5 mmol) and TEA (675 µL) were added and stirred at RT for 3 hours. The product was extracted with HCl (1 M) three times and filtered through a mini-silica column. The solvent was removed under reduced pressure, and the crude product was used without further purification. Yellowish oil; yield: 0.302 g (1.07 mmol, 72%); ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 7.23 – 6.99 (m, 2H), 4.90 – 4.79 (m, 2H), 4.40 – 4.13 (m, 1H), 3.86 (d, J = 40.9 Hz, 3H), 3.60 (d, J = 33.3 Hz, 2H), 2.71 (d, J = 49.2 Hz, 1H), 2.42 (s, 4H), 2.38 (s, 4H), 2.25 (d, J = 7.8 Hz, 2H), 2.14 (s, 2H); ^{13}C NMR (75 MHz, DMSO-d_6) δ [ppm] = 171.8, 146.6, 137.3, 136.9, 85.9, 79.3, 79.1, 52.8, 51.8, 51.7, 49.5, 44.3, 44.2, 42.2, 41.8, 40.4, 40.1, 37.9, 30.4, 30.1, 28.5, 28.5, 28.3, 27.3.

tert-*Butyl(bicyclo[2.2.1] hept-5-en-2-ylmethyl)(2-((5,6-dimethoxybenzo[*d*]thiazole-2-yl)amino)-2-oxoethyl)carbamate* (***16***) was synthesised following method A. 5,6-dimethoxybenzo[*d*]thiazole-2-amine **6** (0.130 g, 1 eq.), TBTU (0.166 g, 1.1 eq.), DIPEA (0.300 mL, 2.5 eq.), **15** (0.150 g, 0.53 mmol, 1 eq.). The crude product was purified by column chromatography (CH/EtOAc 2:1). Yellowish oil; yield: 0.22 g (0.46 mmol, 88%); ^{1}H NMR (300 MHz, DMSO-d_6) δ [ppm] = 11.26 (s, 2H), 6.94 (d, J = 13.4 Hz, 2H), 5.74 (dt, J = 16.0, 2.5 Hz, 1H), 3.90 (dd, J = 10.7, 6.4 Hz, 2H), 3.87 – 3.74 (m, 1H), 3.65 (d, J = 3.2 Hz, 6H), 3.30 – 3.05 (m, 1H), 2.92 – 2.73 (m, 1H), 2.56 – 2.47 (m, 1H), 1.76 (d, J = 8.6 Hz, 1H), 1.61 – 1.46 (m, 1H), 1.19 (s, 6H), 1.13 (s, 3H), 1.09 (s, 1H), 1.00 – 0.91 (m, 2H); ^{13}C NMR (75 MHz, DMSO-d_6) δ [ppm] = 158.1, 149.8, 148.0, 136.9, 136.2, 122.1, 102.8, 102.0, 56.4, 56.2, 53.6, 49.7, 44.4, 42.3, 41.8, 38.1, 38.0, 30.4, 30.0, 28.3, 26.9, 14.2.

N-*(Bicyclo[2.2.1]hept-5-en-2-ylmethyl)*-N-*(2-((5,6-dimethoxybenzo[*d*]thiazole-2-yl)amino)-2-oxoethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide* (***17***) was synthesised following method B and A. The crude product was purified by column chromatography (CH/EtOAc 6:1). Yellow solid; yield: 0.26 g (0.44 mmol, 79%); m_p = 137 – 139 °C; ^{1}H NMR (300 MHz, CD_2Cl_2) δ [ppm] = 7.50 – 7.26 (m, 1H), 6.87 (q, J = 6.5, 5.0 Hz, 2H), 6.58 (d, J = 8.0 Hz, 2H), 6.54 – 6.49 (m, 2H), 5.47 – 5.14 (m, 2H), 4.73 – 4.58 (m, 1H), 3.74 (d, J = 15.2 Hz, 1H), 3.40 (q, J = 7.1 Hz, 1H), 3.19 (d, J = 3.4 Hz, 6H), 2.84 – 2.71 (m, 1H), 2.53 – 2.41 (m, 1H), 2.02 (d, J = 10.0 Hz, 1H), 0.81 (s, 1H), 0.56 (dd, J = 13.3, 6.2 Hz, 3H), 0.19 (q, J = 7.7, 7.1 Hz, 1H); ^{13}C NMR (75 MHz, CD_2Cl_2) δ [ppm] = 176.6, 167.2, 157.9, 149.8, 148.0, 140.7, 137.1, 131.0, 127.8, 127.1, 126.6, 122.6, 103.0, 102.5, 60.3, 56.2, 56.0, 54.1, 53.8, 53.4, 53.1, 52.7, 44.3, 42.3, 41.8, 30.6, 30.3, 27.1, 20.9, 14.0.

N-*(Bicyclo[2.2.1]hept-5-en-2-ylmethyl)*-N-*(2-((5,6-bihydroxybenzo[*d*]thiazol-2-yl)amino)-2-oxoethyl)-4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzamide* (***4***) was synthesised following method D. **17** (0.1 g, 0.17 mmol, 1 eq.), DCM (10 mL), BBr_3 in DCM 1 M (1 mL, 1 mmol, 6 eq.). The crude product was purified by HPLC. Colourless solid; yield: 8.3 mg (0.01 mmol, 8%); decomp. > 134 °C; ^{1}H NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.26 (d, J = 11.5 Hz, 1H), 12.05 (d, J = 20.5 Hz, 1H), 9.18 (s, 2H), 7.58 – 7.54 (m, 1H), 7.49 – 7.45 (m, 1H), 7.42 – 7.37 (m, 1H), 7.33 – 7.29 (m, 1H), 7.21 (d, J = 3.0 Hz, 1H), 7.10 – 7.04 (m, 1H), 6.12 – 6.06 (m, 1H), 6.07 – 6.00 (m, 1H), 4.39 (s, 1H), 4.24 – 4.12 (m, 1H), 3.45 – 3.38 (m, 1H), 3.36 – 3.26 (m, 1H), 3.21 – 3.08 (m, 1H), 3.05 – 2.95 (m, 1H), 2.89 – 2.75 (m, 1H), 2.74 – 2.55 (m, 2H), 2.31 (s, 1H), 1.87 – 1.80 (m, 1H), 1.78 – 1.69 (m, 1H), 1.64 – 1.58 (m, 1H), 1.29 – 1.21 (m, 1H), 1.20 – 1.15 (m, 1H), 1.09 – 1.03 (m, 1H), 0.58 – 0.51 (m, 1H)*; ^{13}C NMR (151 MHz, DMSO-d_6) δ [ppm] = 170.1, 169.7, 155.3, 145.6, 143.9, 137.9, 137.3, 136.9, 136.7, 136.5, 136.2, 128.6, 128.4, 128.0, 127.8, 127.4, 126.7, 121.8, 106.5, 106.4, 54.7, 53.9, 50.5, 49.6, 49.3, 44.2, 44.0, 43.9, 43.7, 43.2, 41.7, 41.4, 41.2, 40.1, 37.8, 37.3, 36.7, 36.5, 30.2, 30.1, 29.9, 29.8*; m/z for [$C_{22}H_{20}F_3N_6O_5S_2^+$] calc. 558.1, found 558.0; purity: 97%.

*Endo/exo norbornene complicates the evaluation of NMR spectra.

Author Contributions

HM: Data Curation, Investigation, Writing – Original Draft, Writing – Review & Editing; AG: Data Curation, Writing – Review & Editing; ORE: Data Curation, Writing – Review & Editing; JXC: Data Curation, Investigation, Writing – Review & editing; TS: Resources, Supervision, Writing -Review & Editing; FB: Project administration, Conceptualization, Writing – Review & Editing.

Conflicts of Interest

There are no conflicts to declare.

Acknowledgements

We thank Stefan J. Hammerschmidt for providing the proteins.

Notes and References

[1] H. Harapan, A. Michie, M. Mudatsir, R. T. Sasmono, A. Imrie, *BMC Res. Notes* **2019**, *12*, 4–9.

[2] WHO, "Dengue and severe dengue," can be found under https://apps.who.int/iris/bitstream/handle/10665/204161/Fact_Sheet_WHD_2014_EN_1629.pdf, **2020**.

[3] J. G. Rigau-Pérez, G. G. Clark, D. J. Gubler, P. Reiter, E. J. Sanders, A. Vance Vorndam, *Lancet* **1998**, *352*, 971–977.

[4] B. Troost, J. M. Smit, *Curr. Opin. Virol.* **2020**, *43*, 9–21.

[5] A. Mukhametov, E. I. Newhouse, N. Ab, J. A. Saito, M. Alam, *J. Mol. Graph. Model.* **2014**, *52*, 103–113.

[6] S. I. Yun, Y. M. Lee, *J. Microbiol.* **2017**, *55*, 204–219.

[7] A. E. Shannon, K. J. Chappell, M. J. Stoermer, S. Y. Chow, W. M. Kok, D. P. Fairlie, P. R. Young, *Protein Expr. Purif.* **2016**, *119*, 124–129.

[8] V. Boldescu, M. A. M. Behnam, N. Vasilakis, C. D. Klein, *Nat. Rev. Drug Discov.* **2017**, *16*, 565–586.

[9] M. A. M. Behnam, C. Nitsche, V. Boldescu, C. D. Klein, *J. Med. Chem.* **2016**, *59*, 5622–5649.

[10] Y. Yao, T. Huo, Y.-L. Lin, S. Nie, F. Wu, Y. Hua, J. Wu, A. R. Kneubehl, M. B. Vogt, R. Rico-Hesse, Y. Song, *J. Am. Chem. Soc.* **2019**, *141*, 6832–6836.

[11] M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman, J. A. Hardy, *ACS Chem. Biol.* **2013**, *8*, 2744–2752.

[12] H. A. Rothan, H. C. Han, T. S. Ramasamy, S. Othman, N. A. Rahman, R. Yusof, *BMC Infect. Dis.* **2012**, *12*, 314.

[13] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, U. Hommel, *Nat. Struct. Mol. Biol.* **2006**, *13*, 372–373.

[14] C. G. Noble, C. C. Seh, A. T. Chao, P. Y. Shi, *J. Virol.* **2012**, *86*, 438–446.

[15] C. Nitsche, *Biophys. Rev.* **2019**, *11*, 157–165.

[16] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *J. Phys. Chem. B* **2021**, *125*, 6837–6846.

[17] Z. Yin, S. J. Patel, W. L. Wang, W. L. Chan, K. R. Ranga Rao, G. Wang, X. Ngew, V. Patel, D. Beer, J. E. Knox, N. L. Ma, C. Ehrhardt, S. P. Lim, S. G. Vasudevan, T. H. Keller, *Bioorganic Med. Chem. Lett.* **2006**, *16*, 40–43.

[18] N. J. Braun, J. P. Quek, S. Huber, J. Kouretova, D. Rogge, H. Lang-Henkel, E. Z. K. Cheong, B. L. A. Chew, A. Heine, D. Luo, T. Steinmetzer, *ChemMedChem* **2020**, *15*, 1439–1452.

[19] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *J. Med. Chem.* **2019**, *62*, 11359–11382.

[20] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.

[21] M. A. M. Behnam, C. Nitsche, S. M. Vechi, C. D. Klein, *ACS Med. Chem. Lett.* **2014**, *5*, 1037–1042.

[22] M. A. M. Behnam, D. Graf, R. Bartenschlager, D. P. Zlotos, C. D. Klein, *J. Med. Chem.* **2015**, *58*, 9354–9370.

[23] H. Wu, S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W. E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weißbrich, R. Spannaus, E. G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister, J. Bodem, *Antimicrob. Agents Chemother.* **2015**, *59*, 1100–1109.

[24] S. M. Amberg, A. Nestorowicz, D. W. McCourt, C. M. Rice, *J. Virol.* **1994**, *68*, 3794–3802.

[25] M. Brecher, J. Zhang, H. Li, *Virol. Sin.* **2013**, *28*, 326–336.

[26] M. Brecher, Z. Li, B. Liu, J. Zhang, C. A. Koetzner, A. Alifarag, S. A. Jones, Q. Lin, L. D. Kramer, H. Li, *PLOS Pathog.* **2017**, *13*, e1006411.

[27] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2023**, *32*, e4526.

[28] S. Pambudi, N. Kawashita, S. Phanthanawiboon, M. D. Omokoko, P. Masrinoul, A. Yamashita, K. Limkittikul, T. Yasunaga, T. Takagi, K. Ikuta, T. Kurosu, *Biochem. Biophys. Res. Commun.* **2013**, *440*, 393–398.

[29] M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman, J. A. Hardy, *ACS Chem. Biol.* **2013**, *8*, 2744–2752.

[30] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Sci.* **2022**, *32*, e4526.

[31] R. Grinter, "[ccp4bb] Questionable Ligand Density: 6MO0, 6MO1, 6MO2," can be found under https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47072.html, **n.d.**

[32] S. Voss, C. Nitsche, *Bioorganic Med. Chem. Lett.* **2020**, *30*, 126965.

[33] M. A. M. Behnam, C. D. P. Klein, *Biochimie* **2020**, *174*, 117–125.

[34] M. Hashimoto, Y. Hatanaka, *European J. Org. Chem.* **2008**, *2008*, 2513–2523.

[35] E. Smith, I. Collins, *Future Med. Chem.* **2015**, *7*, 159–183.

[36] Y. Sadakane, Y. Hatanaka, *Anal. Sci.* **2006**, *22*, 209–218.

[37] D. P. Murale, S. C. Hong, M. M. Haque, J. S. Lee, *Proteome Sci.* **2017**, *15*, 1–34.

[38] S. Chakraborti, T. Chakraborti, N. S. Dhalla, *Proteases in Human Diseases*, Springer Singapore, Singapore, **2017**.

[39] M. Vrabel, P. Kölle, K. M. Brunner, M. J. Gattner, V. López-Carrillo, R. de Vivie-Riedle, T. Carell, *Chem. - A Eur. J.* **2013**, *19*, 13309–13312.

[40] B. L. Oliveira, Z. Guo, G. J. L. Bernardes, *Chem. Soc. Rev.* **2017**, *46*, 4895–4950.

[41] G. Trapani, M. Franco, A. Latrofa, A. Reho, G. Liso, *Eur. J. Pharm. Sci.* **2001**, *14*, 209–216.

[42] Z. Song, Q. Zhang, *Org. Lett.* **2009**, *11*, 4882–4885.

[43] N. Pozdnyakov, H. E. Murrey, C. J. Crump, M. Pettersson, T. E. Ballard, C. W. Am Ende, K. Ahn, Y. M. Li, K. R. Bales, D. S. Johnson, *J. Biol. Chem.* **2013**, *288*, 9710–9720.

[44] F. Kotzyba-Hibert, I. Kapfer, M. Goeldner, *Angew. Chemie Int. Ed. English* **1995**, *34*, 1296–1312.

[45] J. Brunner, H. Senn, F. M. Richards, *J. Biol. Chem.* **1980**, *255*, 3313–3318.

[46] F. Hammerstein, L. M. Lauth, S. Hammerschmidt, A. Wagner, T. Schirmeister, U. A. Hellmich, *FEBS Lett.* **2019**, *593*, 2204–2213.

[47] T. Seifert, M. Malo, T. Kokkola, K. Engen, M. Fridén-Saxin, E. A. A. Wallén, M. Lahtela-Kakkonen, E. M. Jarho, K. Luthman, *J. Med. Chem.* **2014**, *57*, 9870–9888.

[48] J. Rappsilber, Y. Ishihama, M. Mann, *Anal. Chem.* **2003**, *75*, 663–670.

[49] J. Cox, M. Mann, *Nat. Biotechnol.* **2008**, *26*, 1367–1372.

[50] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, *J. Proteome Res.* **2011**, *10*, 1794–1805.

[51] S. Tyanova, T. Temu, A. Carlson, P. Sinitcyn, M. Mann, J. Cox, *Proteomics* **2015**, *15*, 1453–1456.

[52] R. Othman, T. S. Kiat, N. Khalid, R. Yusof, E. I. Newhouse, J. S. Newhouse, M. Alam, N. A. Rahman, *J. Chem. Inf. Model.* **2008**, *48*, 1582–1591.

Supporting Information

Designing Tool Compounds for Photoaffinity Labelling: Investigating the Binding Pocket of Allosteric Inhibitors for DENV Protease

Hannah Maus[a], Andrea Gellert[a], Olivia Englert[a], Jia-Xuan Chen[b], Tanja Schirmeister[a], Fabian Barthels*[,a]

[a] Johannes Gutenberg-University Mainz, IPBS, Staudingerweg 5, 55128 Mainz, Germany

[b] Johannes Gutenberg-University Mainz, IMB, Ackermannweg 4, 55128 Mainz, Germany

* Corresponding author, barthels@uni-mainz.de

Table of Contents

Scans of SDS-PAGE gels

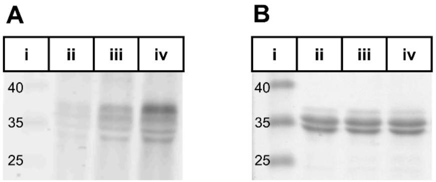


Figure S1. Scans of SDS-PAGE gels of PAL experiments with DENV NS2B-NS3 protease. **A** Typhoon scan of the gel. i) marker; ii) blind$^-$: DENV NS2B-NS3; iii) blind$^+$: DENV NS2B-NS3 + **4**; iv) irr$^+$: DENV NS2B-NS3 + **4**. **B** Scan of the Coomassie stained gel. i) marker; ii) blind$^-$: DENV NS2B-NS3; iii) blind$^+$: DENV NS2B-NS3 + **4**; iv) irr$^+$: DENV NS2B-NS3 + **4**.

NMR Spectra

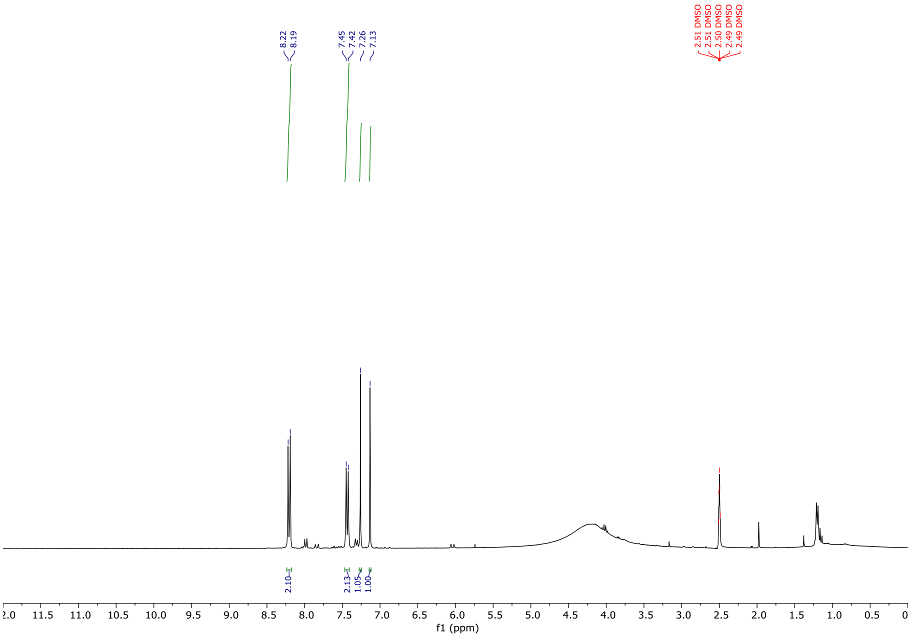


Figure S2. ^{1}H NMR spectrum of compound **2**.

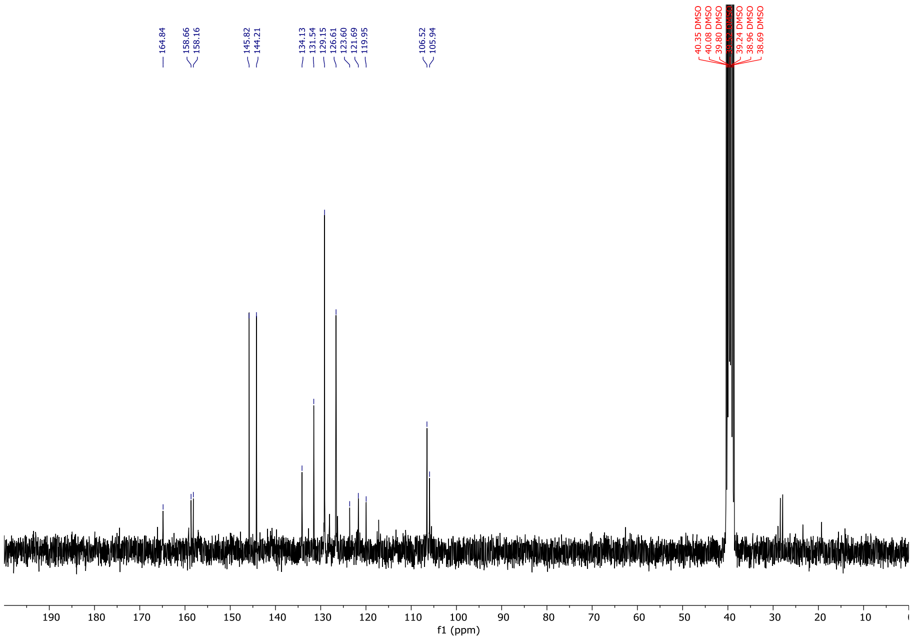


Figure S3. ^{13}C NMR spectrum of compound **2**.

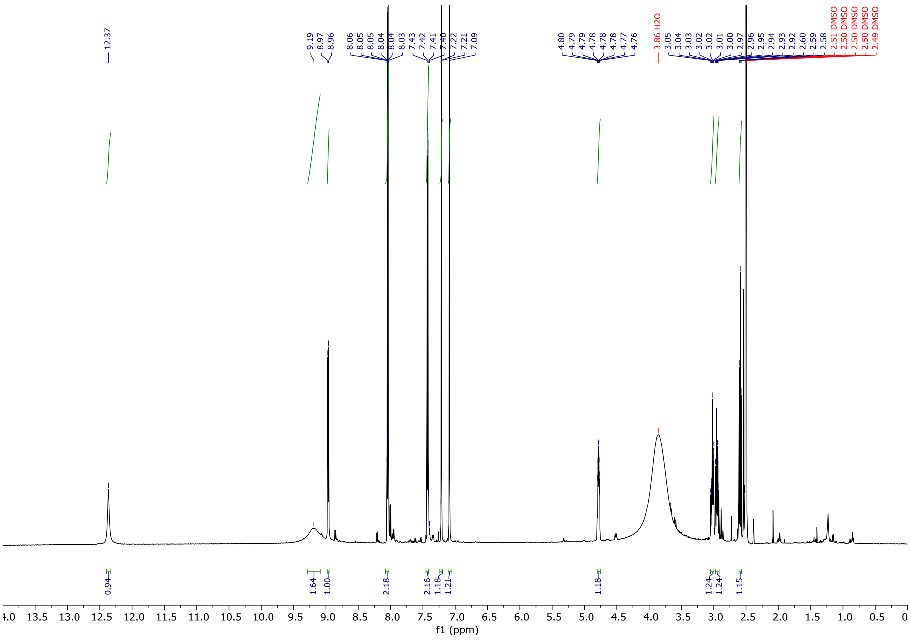


Figure S4. ^{1}H NMR spectrum of compound **3a**.

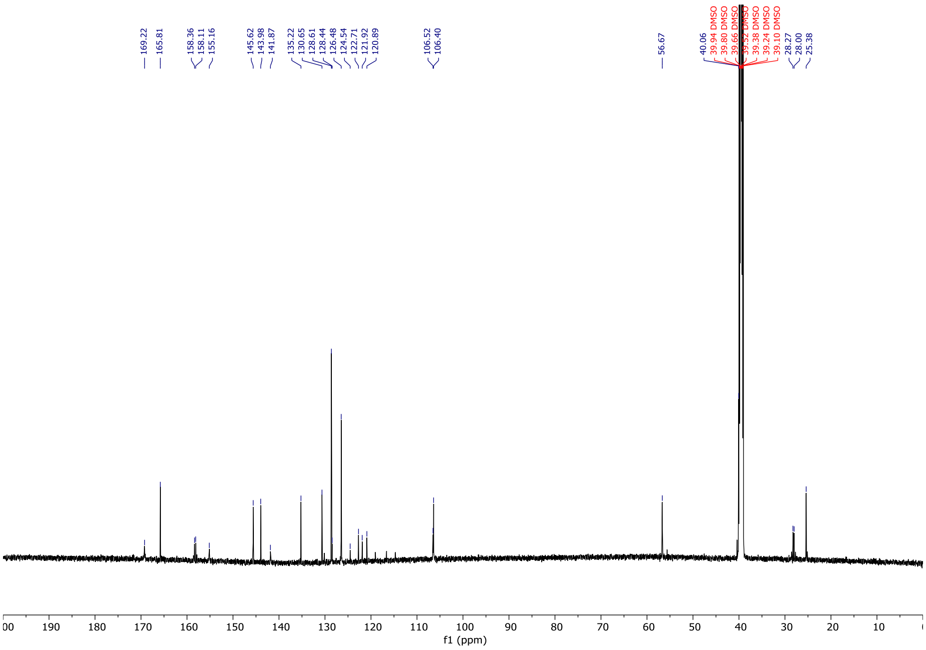


Figure S5. ^{13}C NMR spectrum of compound **3a**.

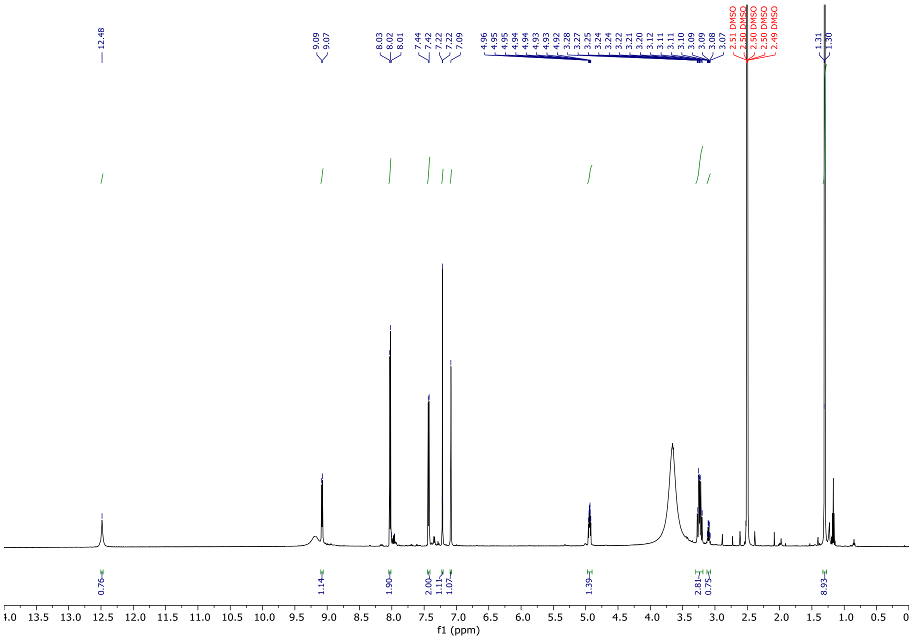


Figure S6. ^{1}H NMR spectrum of compound **3b**.

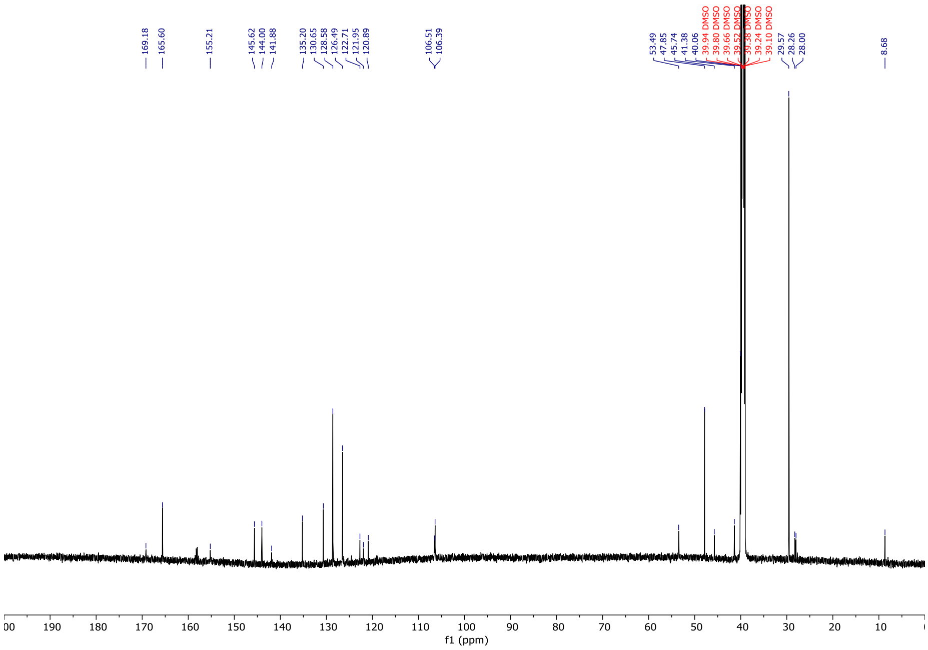


Figure S7. ^{13}C NMR spectrum of compound **3b**.

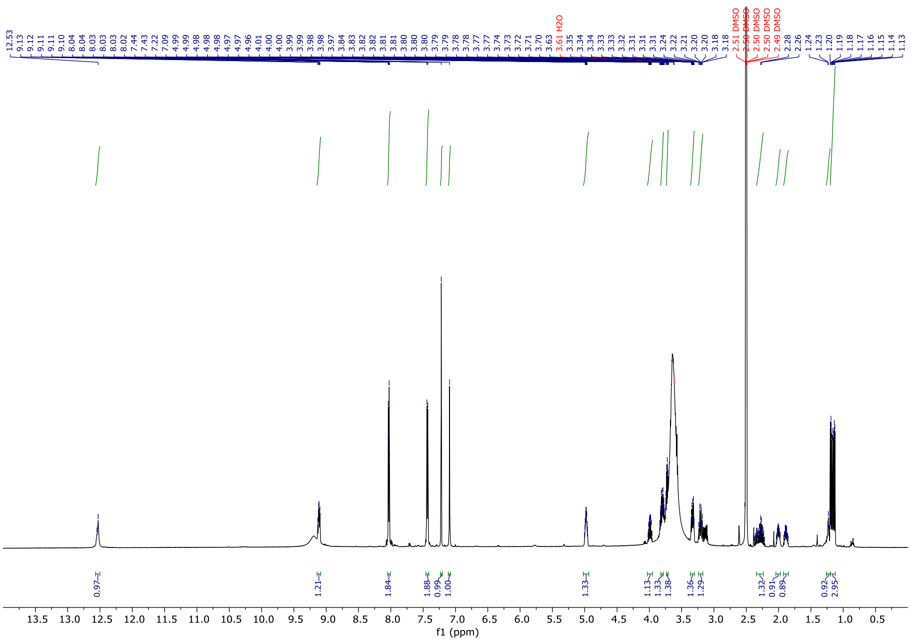


Figure S8. ^1H spectrum of compound **3c**.

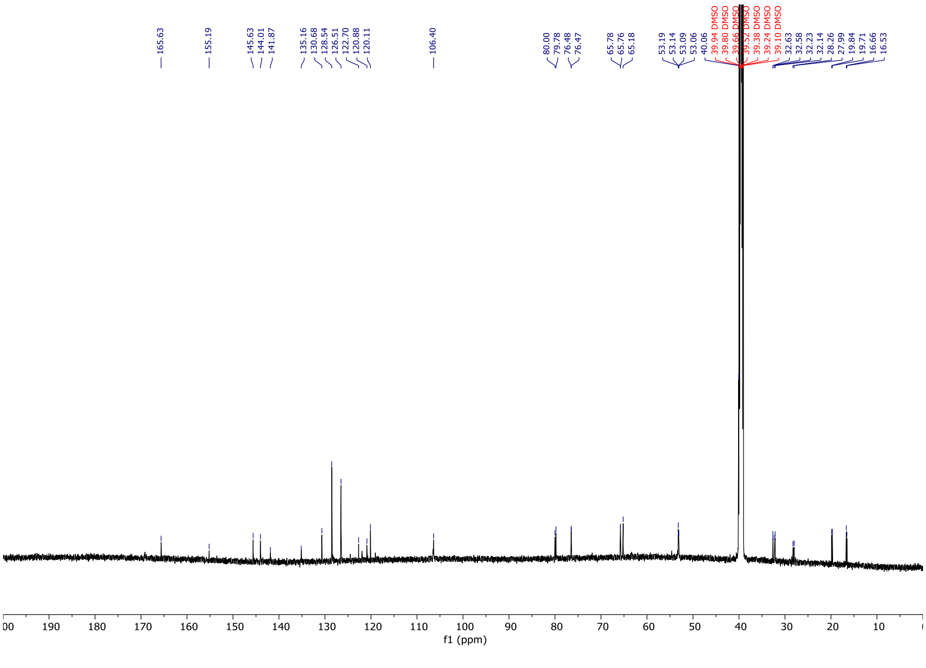


Figure S9. ^{13}C spectrum of compound **3c**.

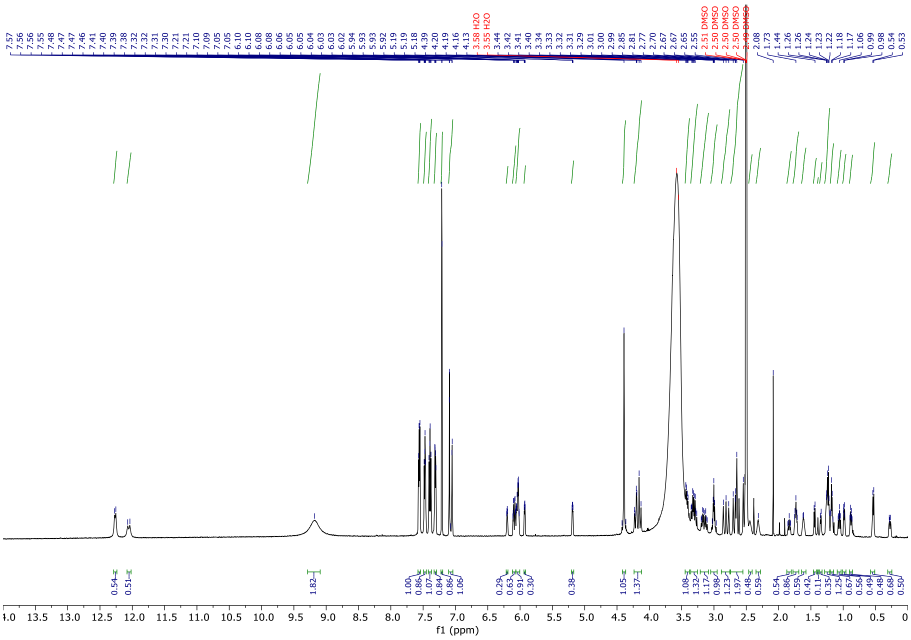


Figure S10. ^{1}H spectrum of compound **4**.

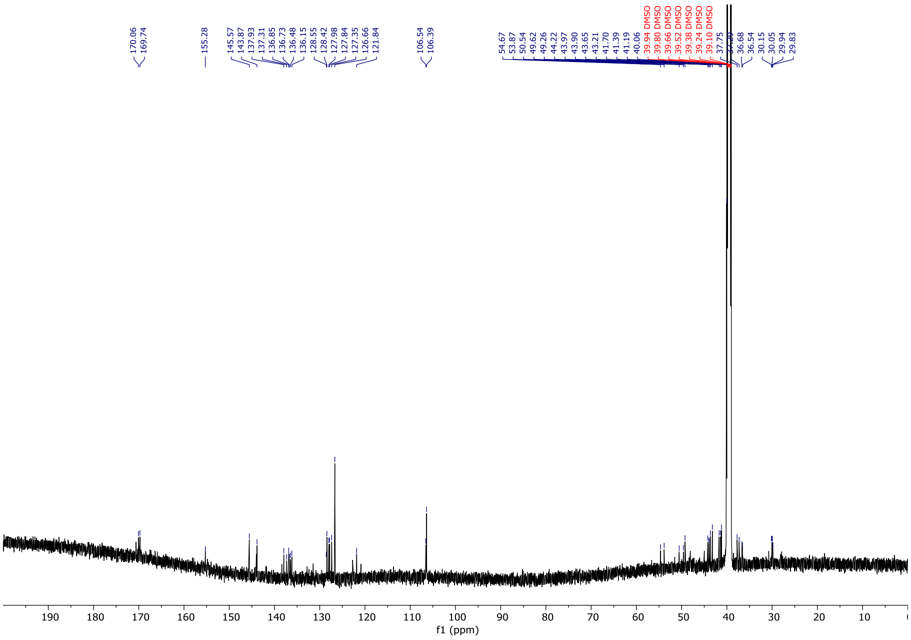


Figure S11. ^{13}C NMR spectrum of compound **4**.

HPLC Chromatograms

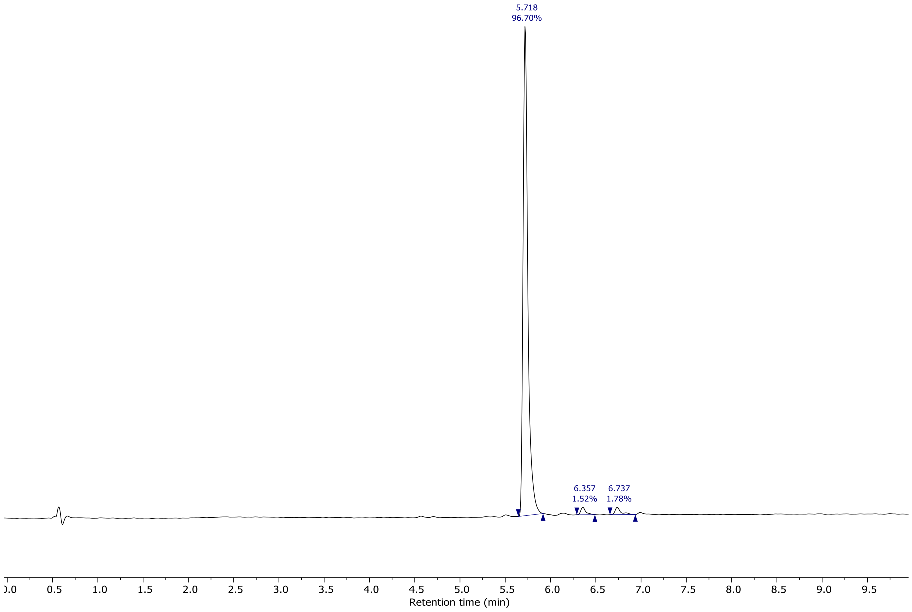


Figure S12. HPLC chromatogram of compound **2** at a wavelength of 254 nm.

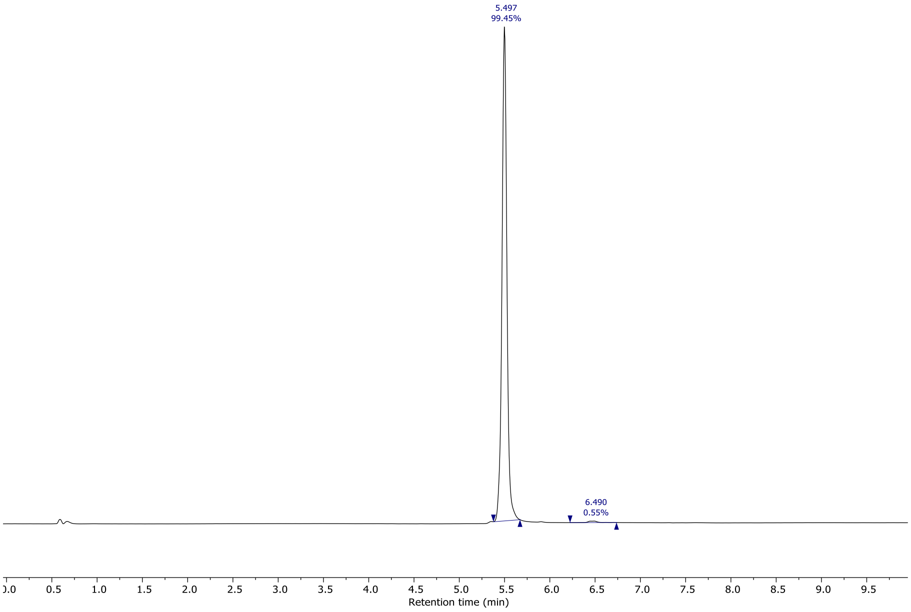


Figure S13. HPLC chromatogram of compound **3a** at a wavelength of 254 nm.

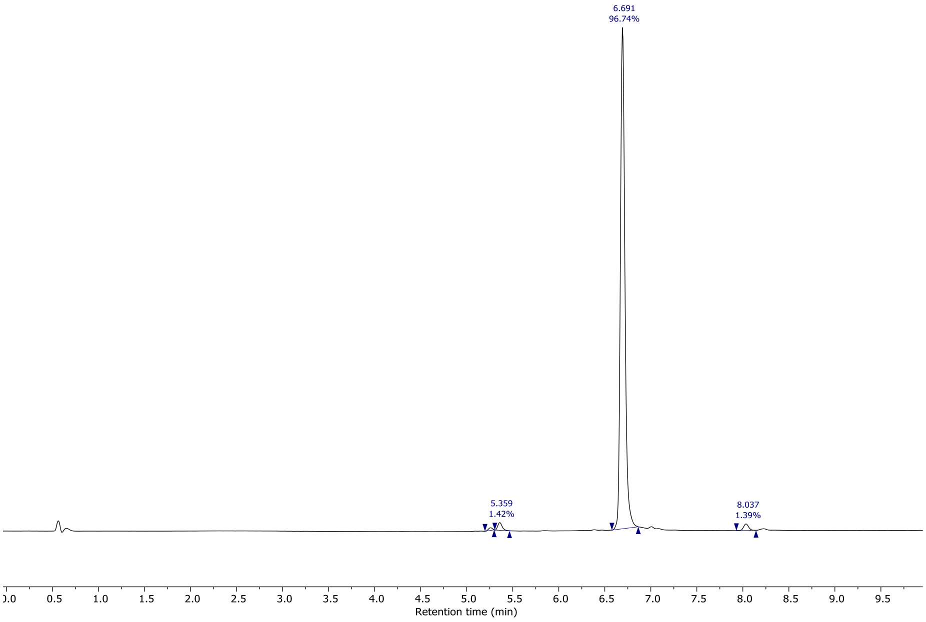


Figure S14. HPLC chromatogram of compound **3b** at a wavelength of 254 nm.

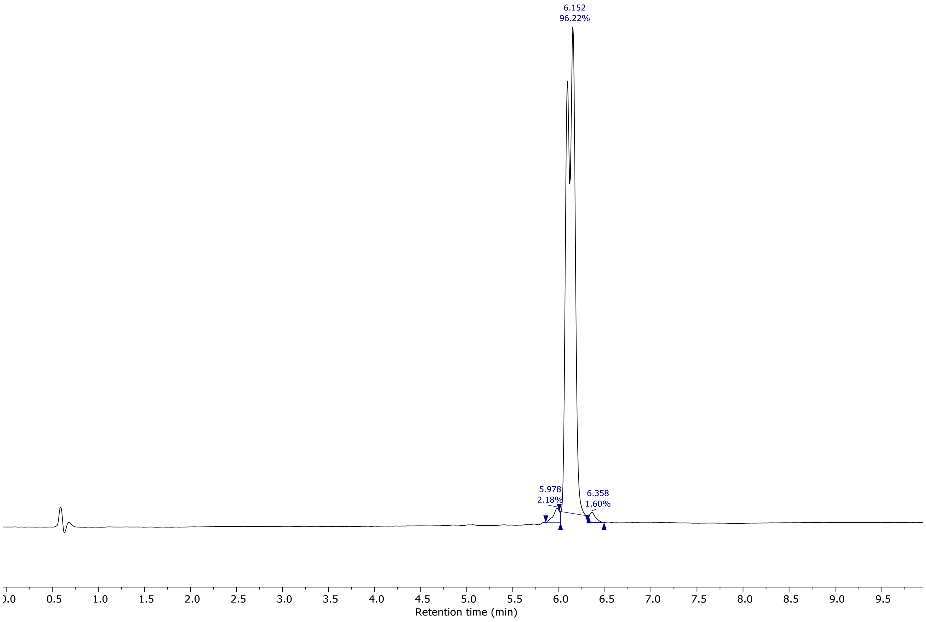


Figure S15. HPLC chromatogram of compound **3c** at a wavelength of 254 nm. Double peak due to the formation of diastereomers since SSMot-protecting group was used as a racemate.

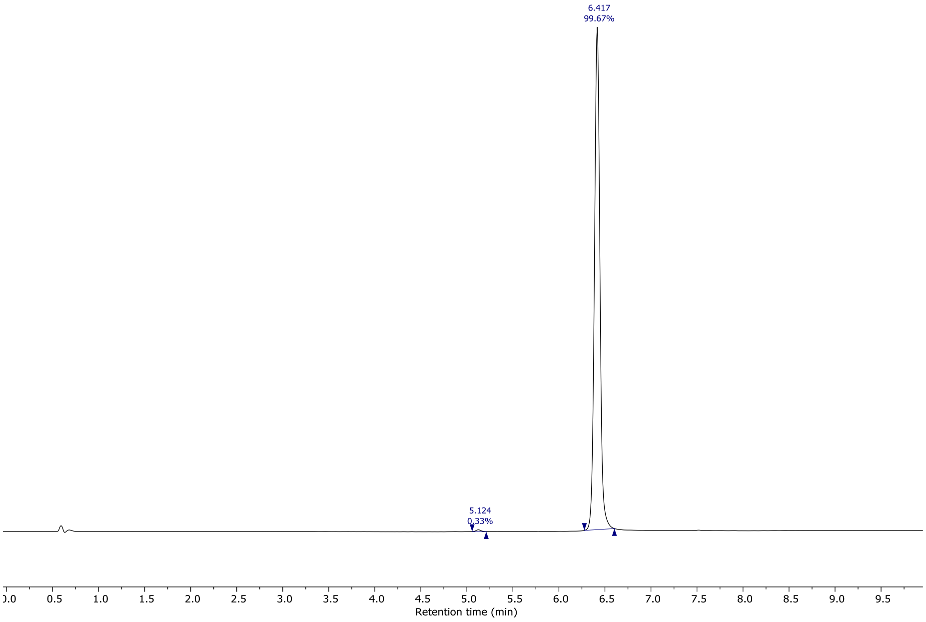


Figure S16. HPLC chromatogram of compound **4** at a wavelength of 254 nm.

6 Conclusions and Outlook

This work, divided into two projects, addressed different aspects of the inhibitor design and protein-ligand interactions of ZIKV and DENV NS2B-NS3 proteases.

Project 1. In the first project, competitive and allosteric inhibitors were developed that address and inhibit the NS2B-NS3 protease.

Starting from previously published allosteric inhibitors based on benzo[*d*]thiazole, three different approaches were pursued to optimize these inhibitors. First, truncated inhibitors containing amino benzo[*d*]thiazole linked to various benzoic acids were developed, showing improved LE compared to the lead structures. Second, based on molecular docking results, which showed that the (*R*)- and (*S*)-enantiomers of the lead structure address different binding sites in the protein, Y-shaped inhibitors were designed to address both binding pockets simultaneously. Last, scaffold hopping was used to exchange the benzo[*d*]thiazole scaffold and invert the amide bond. This resulted in improved stability of the inhibitors. The most promising inhibitors from these three series had IC_{50} values in the low micromolar range.

A fragment-based drug design approach obtained the competitive inhibitors. From a computer-assisted screening of 340 small chemical fragments, a series of α-cyanoacrylates and α-cyanoacrylamides were designed, of which the best inhibitors had K_i values between 7.8 and 20 μM and selectivity toward other proteases. The most promising compound reduced ZIKV infection with an EC_{50} of 35 μM.

The evaluation of the inhibitory capacity of the synthesized compounds was based on a fluorometric assay. This can be biased by interference with the compounds. The underlying assay was optimized to make the development of potential inhibitors more reliable in the future. Instead of the previously used AMC- or FRET-based substrates, substrates based on NBD were developed, synthesized, and their suitability for the fluorometric assay was verified. Compared to the previous fluorophores, NBD's excitation and detection wavelengths are shifted to higher wavelengths, which reduced typical assay interferences.

Project 2. In this project, the influence of ligands on the conformation of the ZIKV and DENV proteases was investigated using smFRET. Dye-labeled proteases were observed in solution or immobilized on a modified glass slide. Conclusions about the conformational dynamics of the protease were drawn by calculating the resulting FRET efficiencies and cross-correlation functions.

First, the DENV protease was shown to be in an equilibrium of different conformations without ligands. The addition of a competitive inhibitor resulted in the stabilization of the closed conformation of the protease. This was independent of the dyes used and the label positions in the enzyme. Since the protease without ligands is predominantly in the open conformation, additional stabilization of this conformation by the allosteric inhibitor was not directly visible. Unambiguous detection of the conformation stabilized by allosteric inhibitors was only possible using a competition assay developed as part of this work.

The study of the ZIKV protease using smFRET necessitated a fivefold mutant because, unlike the DENV protease, the ZIKV protease contains three native cysteines that prevent site-specific labeling. Different methods showed that the fivefold mutant could be used representatively for the wild type of the protein. This demonstrated that not only cysteine-free enzymes are suitable for investigation with smFRET after maleimide dye labeling. Similar to the DENV protease, it was shown for the ZIKV protease that allosteric inhibitors favor the open conformation, and competitive inhibitors favor the closed conformation of the protease. In addition to smFRET, other methods revealed different influences of competitive and allosteric inhibitors on the protease. In thermal shift assays, the different inhibitors induced an opposite shift in melting temperatures, attributed to the stabilization of different conformations. In ^{19}F-NMR experiments with 7F-tryptophan labeled protein, W50 was found to be the reporter for the protease conformation. While the other tryptophan residues were only slightly affected by the presence of the competitive inhibitor, W50 showed a shift. However, no effect on W50 was observed for the allosteric inhibitor, which, as already described for the DENV protease, is probably due to further stabilization of the already dominant open conformation.

The immobilization of the proteases before smFRET experiments allowed a more extended observation period of individual proteases, thus allowing the determination of kinetic parameters such as residence times and transition rates. These parameters were used to conclude the underlying interaction mechanism between ligands and proteins. For this purpose, we considered how the parameters changed by adding different concentrations of competitive inhibitors. Our results showed that the ZIKV and DENV proteases have opposite mechanisms in ligand interaction despite their sequence similarity. While the ZIKV protease follows the induced fit model, a conformational selection model was demonstrated for the DENV protease.

In a final project, the binding site of the allosteric inhibitors was investigated in more detail. For this purpose, different PAL tool compounds were designed and synthesized, which allowed the detection of successful photoaffinity labeling of the protein with the inhibitor after irradiation by an SDS-PAGE gel. LC-MS/MS analysis of labeled protein showed the interaction of the allosteric inhibitor with Gly121, which fits the allosteric binding pocket proposed by different groups[106,107,125], and the molecular docking results.

Outlook. Although further development of the allosteric inhibitors has improved their properties and an improvement in inhibitory activity has also been achieved, the latter is relatively moderate compared with the previously published benzo[*d*]thiazole inhibitors[108]. Consequently, it is assumed that further replacement of individual functional groups will not substantially improve inhibitory activity. Instead, other strategies should be pursued. The results of the PAL experiments showed an interaction of the allosteric inhibitors with Gly121. The Gly121 residue is surrounded by two threonine residues, namely Thr120 and Thr122. An interesting approach would be explicitly addressing them by functional groups that allow covalent binding. For example, boronic acids, already used as warheads for serine and threonine proteases, would be conceivable.[128,129,287–290] Thus, a new class of covalent allosteric inhibitors could be generated. Additionally, the adjacency of the allosteric and active binding sites in the NS2B-NS3 protease offers the possibility for dual inhibitors that combine the structural features of competitive and allosteric inhibitors.

The smFRET experiments were performed with an artificial construct of the proteases in which a truncated NS2B^{cf} is covalently bound to the protease subunit of the NS3 protein by a Gly$_4$-Ser-Gly$_4$ linker (linked construct). This construct is used by various groups working on flaviviral NS2B-NS3 proteases.[84,94,97,117,127,281,291–296] Nevertheless, there are studies on the influence of the linker on the proteases. Among others, Phoo *et al.* and Hill *et al.* showed that the non-covalently linked proteases mainly adopt the closed conformation without a ligand.[99,113,114,276,297] In contrast, the linked construct prefers the open conformation, as we observed it.[99,112,113,283,284,297,298] Accordingly, repeating the smFRET experiments for additional constructs would be interesting. Constructs already described include bivalent (completely unlinked), enzymatically cleavable (linked by amino acids that are autocatalytically cleaved), and the construct covalently linked by a glycine-rich linker, which was used within this work.[99,297] But not only the linker could influence the conformation of the proteases. As described above, all constructs consist of a truncated NS2B^{cf} and the NS3 protease subunit of the NS3 protein. NS2B, in its full-length, is a transmembrane protein and allows the localization of the NS2B-NS3 protease in the membrane of the ER. Thus, in addition to investigating the influence of the linker, the effect of anchoring on the membrane should also be investigated. For this purpose, the membrane proteins can be anchored in the lipid bilayer of liposomes, which in turn are immobilized on a glass slide.[196] However, the expression of full-length NS2B has not yet been described in the literature.

For competitive inhibitors, we were able to show the mechanism by which they interact with the ZIKV and DENV proteases, respectively. However, two different competitive inhibitors were used in this work. Consequently, repeating the experiments with the same inhibitor that addresses both proteases would be helpful to conclude with certainty that the proteases themselves and not the different inhibitors are responsible for the different mechanisms. However, it should be noted that the proteases already showed differences in their behavior without an inhibitor. The interaction mechanism of allosteric inhibitors could not be investigated so far because the allosteric inhibitor led to a decrease in signal density due to the stabilization of the open conformation. A strategy with three dyes would be appropriate to study the binding mechanism of allosteric inhibitors.[196] In addition to the donor- and acceptor-labeled protease, the inhibitor is labeled with a third dye, which allows the observation of the addressed conformation. A similar experiment was described by Kim *et al*. for the MBP. They used Cy3 and Cy5 to label the protein and labeled the ligand with Cy7.[232] A similar experimental setup could be envisioned for the flaviviral proteases.

Not only in the smFRET experiments but also in the ^{19}F NMR experiments, we were able to show the influence of competitive inhibitors on the ZIKV protease. However, the effects of allosteric inhibitors remained hidden in these experiments as well. A similar approach as we described for the competition smFRET assay could be followed to visualize the influence of allosteric inhibitors on the protease by ^{19}F-NMR.

Lastly, although PAL confirmed the potential allosteric binding pocket and demonstrated an interaction of the allosteric inhibitors with Gly121, a description of the exact location of the inhibitors in the binding pocket cannot be derived from this information and can still only be predicted by molecular docking. A detailed description of the allosteric inhibitors in the binding pocket would be possible by crystallization.

All in all, within this work, a wide variety of aspects of the protein-ligand interaction of the NS2B-NS3 protease of ZIKV and DENV could be elucidated by different methods. These findings contribute to the general understanding of the proteins and can form the basis of further experiments.

7 Bibliography

[1] N. Kitadai, S. Maruyama, *Geoscience Frontiers* **2018**, *9*, 1117–1153.

[2] N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, *Cell* **2009**, *138*, 774–786.

[3] B. Willenberg, K. Vollmar Hughes, *Home Economics Guide* **1983**, 1916.

[4] P. E. Wright, H. J. Dyson, *Current Opinion in Structural Biology* **2009**, *19*, 31–38.

[5] T. Pawson, P. Nash, *Genes & Development* **2000**, *14*, 1027–1047.

[6] X. Deupi, B. Kobilka, **2007**, pp. 137–166.

[7] R. Weatherman, R. Fletterick, T. Scanlan, *Annual Review Biochemistry* **1999**, *68*, 559–581.

[8] R. L. Baldwin, *Protein Science* **2008**, *9*, 207–207.

[9] D. Wu, G. LaRosa, M. Simon, *Science* **1993**, *261*, 101–103.

[10] G. G. Hammes, *Enzyme Catalysis and Regulation*, Elsevier, **1982**.

[11] C. H. Henkels, J. C. Kurz, C. A. Fierke, T. G. Oas, *Biochemistry* **2001**, *40*, 2777–2789.

[12] L. Birnbaumer, M. Birnbaumer, in *Biomembranes*, Wiley-VCH Verlag GmbH, Weinheim, Germany, **2008**, pp. 153–252.

[13] C. Londos, C. Sztalryd, J. T. Tansey, A. R. Kimmel, *Biochimie* **2005**, *87*, 45–49.

[14] M. Ptashne, *Nature* **1986**, *322*, 697–701.

[15] P. H. von Hippel, *Annual Review of Biophysics and Biomolecular Structure* **2007**, *36*, 79–105.

[16] R. K. Montange, R. T. Batey, *Annual Review of Biophysics* **2008**, *37*, 117–133.

[17] A. J. Barrett, *Current Protocols in Protein Science* **2000**, *21*, DOI 10.1002/0471140864.ps2101s21.

[18] R. Fotedar, L. Diederich, A. Fotedar, *Progress in cell cycle research* **1996**, *2*, 147–163.

[19] S. J. Martin, *Pathobiology of Human Disease: A Dynamic Encyclopedia of Disease Mechanisms* **2014**, *16*, 145–152.

[20] A. Di Bacco, G. Gill, *Cell Cycle* **2006**, *5*, 2310–2313.

[21] U. Jenal, T. Fuchs, *EMBO Journal* **1998**, *17*, 5658–5669.

[22] D. E. Goll, W. C. Kleese, A. Szpacenko, *Animal Growth Regulation* **1989**, 141–183.

[23] M. R. Maurizi, *Experientia* **1992**, *48*, 178–201.

[24] A. J. S. Hawkins, *Functional Ecology* **1991**, *5*, 222–233.

[25] S.-J. Li, M. Hochstrasser, *Nature* **1999**, *398*.

[26] E. Di Cera, *IUBMB Life* **2009**, *61*, 510–515.

[27] J. D. Colbert, S. P. Matthews, G. Miller, C. Watts, *European Journal of Immunology* **2009**, *39*, 2955–2965.

[28] M. J. Gorman, S. M. Paskewitz, *Insect Biochemistry and Molecular Biology* **2001**, *31*, 257–262.

[29] V. Quesada, G. R. Ordóñez, L. M. Sánchez, X. S. Puente, C. López-Otín, *Nucleic Acids Research* **2009**, *37*, 239–243.

[30] S. Chakraborti, T. Chakraborti, N. S. Dhalla, *Proteases in Human Diseases*, Springer Singapore, Singapore, **2017**.

[31] A. A. Agbowuro, W. M. Huston, A. B. Gamble, J. D. A. Tyndall, *Medicinal Research Reviews* **2018**, *38*, 1295–1331.

[32] C. Nitsche, S. Holloway, T. Schirmeister, C. D. Klein, *Chemical Reviews* **2014**, *114*, 11348–11381.

[33] M. Brecher, J. Zhang, H. Li, *Virologica Sinica* **2013**, *28*, 326–336.

[34] Y. Lin, K. Fang, Y. Zheng, H. L. Wang, J. Wu, *Journal of Travel Medicine* **2022**, *29*, 1–11.

[35] M. H. Collins, S. W. Metz, *Clinical Therapeutics* **2017**, *39*, 1519–1536.

[36] O. J. Brady, P. W. Gething, S. Bhatt, J. P. Messina, J. S. Brownstein, A. G. Hoen, C. L. Moyes, A. W. Farlow, T. W. Scott, S. I. Hay, *PLoS Neglected Tropical Diseases* **2012**, *6*, e1760.

[37] WHO, “Dengue and severe dengue,” can be found under https://apps.who.int/iris/bitstream/handle/10665/204161/Fact_Sheet_WHD_2014_EN_1629.pdf, **2020**.

[38] ECDC (European Centre for Disease Prevention and Control), “Dengue virus disease cases, June 2022-May 2023,” can be found under https://www.ecdc.europa.eu/en/publications-data/dengue-virus-disease-cases-june-2022-may-2023, **2023**.

[39] S. Bhatt, P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, J. M. Drake, J. S. Brownstein, A. G. Hoen, O. Sankoh, M. F. Myers, D. B. George, T. Jaenisch, G. R. William Wint, C. P. Simmons, T. W. Scott, J. J. Farrar, S. I. Hay, *Nature* **2013**, *496*, 504–507.

[40] G. Le Goff, J. Revollo, M. Guerra, M. Cruz, Z. Barja Simon, Y. Roca, J. Vargas Florès, J. P. Hervé, *Parasite* **2011**, *18*, 277–280.

[41] T. W. Scott, A. C. Morrison, **2010**, pp. 115–128.

[42] L. H. Chen, M. E. Wilson, *Tropical Diseases, Travel Medicine and Vaccines* **2016**, *2*, 15.

[43] V. Wiwanitkit, *The Journal of Infection in Developing Countries* **2009**, *4*, 051–054.

[44] T. Boussemart, P. Babe, G. Sibille, C. Neyret, C. Berchel, *Journal of Perinatology* **2001**, *21*, 255–257.

[45] S. I. Yun, Y. M. Lee, *Journal of Microbiology* **2017**, *55*, 204–219.

[46] M. R. Duffy, T.-H. Chen, W. T. T. Hancock, A. M. Powers, J. L. Kool, R. S. S. Lanciotti, M. Pretrick, M. Marfel, S. Holzbauer, C. Dubray, L. Guillaumot, A. Griggs, M. Bel, A. J. Lambert, J. Laven, O. Kosoy, A. Panella, B. J. Biggerstaff, M. Fischer, E. B. Hayes, *New England Journal of Medicine* **2009**, *360*, 2536–2543.

[47] M. Aubry, A. Teissier, M. Huart, S. Merceron, J. Vanhomwegen, C. Roche, A.-L. Vial, S. Teururai, S. Sicard, S. Paulous, P. Desprès, J.-C. Manuguerra, H.-P. Mallet, D. Musso, X. Deparis, V.-M. Cao-Lormeau, *Emerging Infectious Diseases* **2017**, *23*, 2014–2015.

[48] L. R. Petersen, D. J. Jamieson, A. M. Powers, Ma. A. Honein, *The New England journal of medicine* **2016**, *374*, 1552–1563.

[49] World Health Organization (WHO), "WHO guidelines for the prevention of sexual transmission of Zika virus," **2020**.

[50] G. Gutiérrez-Bugallo, L. A. Piedra, M. Rodriguez, J. A. Bisset, R. Lourenço-de-Oliveira, S. C. Weaver, N. Vasilakis, A. Vega-Rúa, *Nature Ecology & Evolution* **2019**, *3*, 561–569.

[51] S. Leta, T. J. Beyene, E. M. De Clercq, K. Amenu, M. U. G. Kraemer, C. W. Revie, *International Journal of Infectious Diseases* **2018**, *67*, 25–35.

[52] D. Musso, C. Roche, E. Robin, T. Nhan, A. Teissier, V.-M. Cao-Lormeau, *Emerging Infectious Diseases* **2015**, *21*, 359–361.

[53] C. J. Gregory, T. Oduyebo, A. C. Brault, J. T. Brooks, K.-W. Chung, S. Hills, M. J. Kuehnert, P. Mead, D. Meaney-Delman, I. Rabe, E. Staples, L. R. Petersen, *The Journal of Infectious Diseases* **2017**, *216*, S875–S883.

[54] H. Harapan, A. Michie, R. T. Sasmono, A. Imrie, *Viruses* **2020**, *12*, 829.

[55] A. Balmaseda, S. N. Hammond, L. Pérez, Y. Tellez, S. I. Saborío, J. C. Mercado, R. Cuadra, J. Rocha, M. A. Pérez, S. Silva, C. Rocha, E. Harris, *American Journal of Tropical Medicine and Hygiene* **2006**, *74*, 449–456.

[56] L. C. Katzelnick, L. Gresh, M. E. Halloran, J. C. Mercado, G. Kuan, A. Gordon, A. Balmaseda, E. Harris, *Science* **2017**, *358*, 929–932.

[57] M. Atif, M. Azeem, M. R. Sarwar, A. Bashir, *Infection* **2016**, *44*, 695–705.

[58] M. S. Majumder, E. Cohn, D. Fish, J. S. Brownstein, *Bulletin of the World Health Organization* **2016**, DOI 10.2471/BLT.16.171009.

[59] J. Mlakar, M. Korva, N. Tul, M. Popović, M. Poljšak-Prijatelj, J. Mraz, M. Kolenc, K. Resman Rus, T. Vesnaver Vipotnik, V. Fabjan Vodušek, A. Vizjak, J. Pižem, M. Petrovec, T. Avšič Županc, *New England Journal of Medicine* **2016**, *374*, 951–958.

[60] CDC, "ZIKA and Pregnancy," DOI 10.1016/j.reprotox.2018.07.020can be found under https://www.cdc.gov/pregnancy/zika/testing-follow-up/zika-syndrome-birth-defects.html, **2019**.

[61] WHO, "Zika virus disease," can be found under https://www.who.int/health-topics/zika-virus-disease#tab=tab_1, **2023**.

[62] T. M. Burns, *Seminars in Neurology* **2008**, DOI 10.1055/s-2008-1062261.

[63] V. Cao-Lormeau, A. Blake, S. Mons, S. Lastere, C. Roche, J. Vanhomwegen, T. Dub, L. Baudouin, A. Teissier, P. Larre, A. Vial, C. Decam, V. Choumet, S. Halstead, H. Willison, L. Musset, J. Manuguerra, P. Despres, E. Fournier, H. Mallet, D. Musso, A. Fontanet, J. Neil, F. Ghawché, *Lancet* **2016**, *387*, 1531–1539.

[64] CDC, "Zika and Guillain-Barré Syndrome," can be found under https://www.cdc.gov/zika/healtheffects/gbs-qa.html, **2016**.

[65] A. Uncini, N. Shahrizaila, S. Kuwabara, *Journal of Neurology, Neurosurgery and Psychiatry* **2017**, *88*, 266–271.

[66] N. M. Silva, N. C. Santos, I. C. Martins, *Tropical Medicine and Infectious Disease* **2020**, *5*, DOI 10.3390/tropicalmed5040150.

[67] S. J. Thomas, I. K. Yoon, *Human Vaccines and Immunotherapeutics* **2019**, *15*, 2295–2314.

[68] S. B. Halstead, *Vaccine* **2017**, *35*, 6355–6358.

[69] L. Villar, G. H. Dayan, J. L. Arredondo-García, D. M. Rivera, R. Cunha, C. Deseda, H. Reynales, M. S. Costa, J. O. Morales-Ramírez, G. Carrasquilla, L. C. Rey, R. Dietze, K. Luz, E. Rivas, M. C. Miranda Montoya, M. Cortés Supelano, B. Zambrano, E. Langevin, M. Boaz, N. Tornieporth, M. Saville, F. Noriega, *New England Journal of Medicine* **2015**, *372*, 113–123.

[70] M. S. Mustafa, V. Rasotgi, S. Jain, V. Gupta, *Medical Journal Armed Forces India* **2015**, *71*, 67–70.

[71] J. E. Hazlewood, B. Tang, K. Yan, D. J. Rawle, J. J. Harrison, R. A. Hall, J. Hobson-Peters, A. Suhrbier, *Vaccines* **2022**, *10*, 85.

[72] E. Fernandez, M. S. Diamond, *Current Opinion in Virology* **2017**, *23*, 59–67.

[73] J. Torresi, G. Ebert, M. Pellegrini, *Human Vaccines & Immunotherapeutics* **2017**, *13*, 1059–1072.

[74] D. F. Barreto-Vieira, D. Couto-Lima, F. C. Jácome, G. C. Caldas, O. M. Barth, *Memorias do Instituto Oswaldo Cruz* **2020**, *115*, 1–7.

[75] A. Mukhametov, E. I. Newhouse, N. Ab, J. A. Saito, M. Alam, *Journal of Molecular Graphics and Modelling* **2014**, *52*, 103–113.

[76] B. D. Lindenbach, C. M. Rice, *Fields Virology, 5th Edition, Chapter 33*, **2007**.

[77] V. Boldescu, M. A. M. Behnam, N. Vasilakis, C. D. Klein, *Nature Reviews Drug Discovery* **2017**, *16*, 565–586.

[78] M. A. M. Behnam, C. Nitsche, V. Boldescu, C. D. Klein, *Journal of Medicinal Chemistry* **2016**, *59*, 5622–5649.

[79] A. E. Shannon, K. J. Chappell, M. J. Stoermer, S. Y. Chow, W. M. Kok, D. P. Fairlie, P. R. Young, *Protein Expression and Purification* **2016**, *119*, 124–129.

[80] M. G. Guzman, D. J. Gubler, A. Izquierdo, E. Martinez, S. B. Halstead, *Nature Reviews Disease Primers* **2016**, *2*, 1–26.

[81] M. D. Fernandez-Garcia, M. Mazzon, M. Jacobs, A. Amara, *Cell Host and Microbe* **2009**, *5*, 318–328.

[82] S. Mukhopadhyay, R. J. Kuhn, M. G. Rossmann, *Nature Reviews Microbiology* **2005**, *3*, 13–22.

[83] C. Klein, *Pharmazie in Unserer Zeit* **2010**, *39*, 50–53.

[84] C. G. Noble, C. C. Seh, A. T. Chao, P. Y. Shi, *Journal of Virology* **2012**, *86*, 438–446.

[85] S. M. Amberg, A. Nestorowicz, D. W. McCourt, C. M. Rice, *Journal of Virology* **1994**, *68*, 3794–3802.

[86] R. Yusof, S. Clum, M. Wetzel, H. M. K. Murthy, R. Padmanabhan, *Journal of Biological Chemistry* **2000**, *275*, 9963–9969.

[87] R. Perera, R. J. Kuhn, *Current Opinion in Microbiology* **2008**, *11*, 369–377.

[88] H. A. Rothan, H. C. Han, T. S. Ramasamy, S. Othman, N. A. Rahman, R. Yusof, *BMC Infectious Diseases* **2012**, *12*, 314.

[89] N. D. Rawlings, A. J. Barrett, A. Bateman, *Nucleic Acids Research* **2009**, *38*, 227–233.

[90] N. D. Rawlings, D. P. Tolle, A. J. Barrett, *Nucleic Acids Research* **2004**, *32*, 160–164.

[91] J. F. Bazan, R. J. Fletterick, *Virology* **1989**, *171*, 637–639.

[92] T. J. Chambers, C. S. Hahn, R. Galler, C. M. Rice, *Annual Rev. Microbiology* **1990**, *44*, 649–688.

[93] I. E. Gouvea, M. A. Izidoro, W. A. S. Judice, M. H. S. Cezari, G. Caliendo, V. Santagada, C. N. D. dos Santos, M. H. Queiroz, M. A. Juliano, P. R. Young, D. P. Fairlie, L. Juliano, *Archives of Biochemistry and Biophysics* **2007**, *457*, 187–196.

[94] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, U. Hommel, *Nature Structural & Molecular Biology* **2006**, *13*, 372–373.

[95] A. Poulsen, C. Kang, T. Keller, *Current Pharmaceutical Design* **2014**, *20*, 3422–3427.

[96] C. B. Kang, S. Gayen, W. Wang, R. Severin, A. S. Chen, H. A. Lim, C. S. B. Chia, A. Schüller, D. N. P. Doan, A. Poulsen, J. Hill, S. G. Vasudevan, T. H. Keller, *Antiviral Research* **2013**, *97*, 137–144.

[97] J. Li, S. P. Lim, D. Beer, V. Patel, D. Wen, C. Tumanut, D. C. Tully, J. A. Williams, J. Jiricek, J. P. Priestle, J. L. Harris, S. G. Vasudevan, *Journal of Biological Chemistry* **2005**, *280*, 28766–28774.

[98] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic acids research* **2000**, *28*, 235–242.

[99] W. W. Phoo, Y. Li, Z. Zhang, M. Y. Lee, Y. R. Loh, Y. B. Tan, E. Y. Ng, J. Lescar, C. Kang, D. Luo, *Nature Communications* **2016**, *7*, 1–8.

[100] I. Schechter, A. Berger, *Biochemical and Biophysical Research Communications* **1967**, *27*, 157–162.

[101] L. Hedstrom, *Current Protocols in Protein Science* **2001**, *26*, DOI 10.1002/0471140864.ps2110s26.

[102] A. K. Timiri, B. N. Sinha, V. Jayaprakash, *European Journal of Medicinal Chemistry* **2016**, *117*, 125–143.

[103] M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman, J. A. Hardy, *ACS Chemical Biology* **2013**, *8*, 2744–2752.

[104] R. Othman, T. S. Kiat, N. Khalid, R. Yusof, E. I. Newhouse, J. S. Newhouse, M. Alam, N. A. Rahman, *Journal of Chemical Information and Modeling* **2008**, *48*, 1582–1591.

[105] S. Pambudi, N. Kawashita, S. Phanthanawiboon, M. D. Omokoko, P. Masrinoul, A. Yamashita, K. Limkittikul, T. Yasunaga, T. Takagi, K. Ikuta, T. Kurosu, *Biochemical and Biophysical Research Communications* **2013**, *440*, 393–398.

[106] M. Brecher, Z. Li, B. Liu, J. Zhang, C. A. Koetzner, A. Alifarag, S. A. Jones, Q. Lin, L. D. Kramer, H. Li, *PLOS Pathogens* **2017**, *13*, e1006411.

[107] M. Yildiz, S. Ghosh, J. A. Bell, W. Sherman, J. A. Hardy, *ACS Chemical Biology* **2013**, *8*, 2744–2752.

[108] B. Millies, F. Von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Göppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U. A. Hellmich, T. Schirmeister, *Journal of Medicinal Chemistry* **2019**, *62*, 11359–11382.

[109] W. W. Phoo, Z. Zhang, M. Wirawan, E. J. C. Chew, A. B. L. Chew, J. Kouretova, T. Steinmetzer, D. Luo, *Antiviral Research* **2018**, *160*, 17–24.

[110] E. Harper, A. Berger, *Biochemical and Biophysical Research Communications* **1972**, *46*, 1956–1960.

[111] A. Sandner, T. Hüfner-Wulsdorf, A. Heine, T. Steinmetzer, G. Klebe, *Journal of Medicinal Chemistry* **2019**, *62*, 9753–9771.

[112] L. Zhu, J. Yang, H. Li, H. Sun, J. Liu, J. Wang, *Biochemical and Biophysical Research Communications* **2015**, *461*, 677–680.

[113] M. E. Hill, M. Yildiz, J. A. Hardy, *Biochemistry* **2019**, *58*, 776–787.

[114] Y. M. Kim, S. Gayen, C. B. Kang, J. Joy, Q. Huang, A. S. Chen, J. L. K. Wee, M. J. Y. Ang, H. A. Lim, A. W. Hung, R. Li, C. G. Noble, L. T. Lee, A. Yip, Q. Y. Wang, C. S. B. Chia, J. Hill, P. Y. Shi, T. H. Keller, J. Liang, K. Wee, M. Jin, Y. Ang, H. A. Lim, A. W. Hung, R. Li, C. G. Noble, L. T. Lee, A. Yip, Q. Y. Wang, C. San, B. Chia, J. Hill, P. Y. Shi, T. H. Keller, *Journal of Biological Chemistry* **2013**, *288*, 12891–12900.

[115] J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang, R. Hilgenfeld, *Science* **2016**, *353*, 503–505.

[116] N. J. Braun, J. P. Quek, S. Huber, J. Kouretova, D. Rogge, H. Lang-Henkel, E. Z. K. Cheong, B. L. A. Chew, A. Heine, D. Luo, T. Steinmetzer, *ChemMedChem* **2020**, *15*, 1439–1452.

[117] Z. Zhang, Y. Li, Y. R. Loh, W. W. Phoo, A. W. Hung, C. Kang, D. Luo, *Science* **2016**, *354*, 1597–1600.

[118] K. Yang, H., Chen, X., Ji, X., Xiong, Y., Yang, **2017**, DOI 10.2210/pdb5GXJ/pdb.

[119] Y. Yao, T. Huo, Y.-L. Lin, S. Nie, F. Wu, Y. Hua, J. Wu, A. R. Kneubehl, M. B. Vogt, R. Rico-Hesse, Y. Song, *Journal of the American Chemical Society* **2019**, *141*, 6832–6836.

[120] R. Grinter, "[ccp4bb] Questionable Ligand Density: 6MO0, 6MO1, 6MO2," can be found under https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47072.html, **n.d.**

[121] S. Voss, C. Nitsche, *Bioorganic and Medicinal Chemistry Letters* **2020**, *30*, 126965.

[122] M. A. M. Behnam, C. D. P. Klein, *Biochimie* **2020**, *174*, 117–125.

[123] X. Chen, K. Yang, C. Wu, C. Chen, C. Hu, O. Buzovetsky, Z. Wang, X. Ji, Y. Xiong, H. Yang, *Cell Research* **2016**, *26*, 1260–1263.

[124] A. Roy, L. Lim, S. Srivastava, Y. Lu, J. Song, *PLOS ONE* **2017**, *12*, e0180632.

[125] C. G. Noble, Y. L. Chen, H. Dong, F. Gu, S. P. Lim, W. Schul, Q. Y. Wang, P. Y. Shi, *Antiviral Research* **2010**, *85*, 450–462.

[126] K. Chappell, M. Stoermer, D. Fairlie, P. Young, *Current Medicinal Chemistry* **2008**, *15*, 2771–2784.

[127] D. Leung, K. Schroder, H. White, N. Fang, M. J. Stoermer, G. Abbenante, J. L. Martin, P. R. Young, D. P. Fairlie, *The Journal of Biological Chemistry* **2001**, *276*, 45762–45771.

[128] Z. Yin, S. J. Patel, W. L. Wang, W. L. Chan, K. R. Ranga Rao, G. Wang, X. Ngew, V. Patel, D. Beer, J. E. Knox, N. L. Ma, C. Ehrhardt, S. P. Lim, S. G. Vasudevan, T. H. Keller, *Bioorganic and Medicinal Chemistry Letters* **2006**, *16*, 40–43.

[129] Z. Yin, S. J. Patel, W. Wang, G. Wang, W. Chan, K. R. R. Rao, J. Alam, D. A. Jeyaraj, X. Ngew, V. Patel, D. Beer, S. P. Lim, S. G. Vasudevan, T. H. Keller, **2006**, *16*, 36–39.

[130] M. A. M. Behnam, C. Nitsche, S. M. Vechi, C. D. Klein, *ACS Medicinal Chemistry Letters* **2014**, *5*, 1037–1042.

[131] C. Bechtler, C. Lamers, *RSC Medicinal Chemistry* **2021**, *12*, 1325–1351.

[132] P. G. Dougherty, A. Sahni, D. Pei, *Chemical Reviews* **2019**, *119*, 10241–10287.

[133] Y. Takagi, K. Matsui, H. Nobori, H. Maeda, A. Sato, T. Kurosu, Y. Orba, H. Sawa, K. Hattori, K. Higashino, Y. Numata, Y. Yoshida, *Bioorganic and Medicinal Chemistry Letters* **2017**, *27*, 3586–3590.

[134] N. J. Braun, J. P. Quek, S. Huber, J. Kouretova, D. Rogge, H. Lang-Henkel, E. Z. K. Cheong, B. L. A. Chew, A. Heine, D. Luo, T. Steinmetzer, *ChemMedChem* **2020**, *15*, 1439–1452.

[135] H. Wu, S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W. E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weißbrich, R. Spannaus, E. G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister, J. Bodem, *Antimicrobial Agents and Chemotherapy* **2015**, *59*, 1100–1109.

[136] S. A. Shiryaev, A. V. Cheltsov, K. Gawlik, B. I. Ratnikov, A. Y. Strongin, *ASSAY and Drug Development Technologies* **2011**, *9*, 69–78.

[137] K. J. Pacholarz, R. A. Garlish, R. J. Taylor, P. E. Barran, *Chemical Society Reviews* **2012**, *41*, 4335–4355.

[138] S. Mandal, M. Moudgil, S. K. Mandal, *European Journal of Pharmacology* **2009**, *625*, 90–100.

[139] M. A. Williams, T. Daviter, *Protein-Ligand Interactions*, Humana Press, Totowa, NJ, **2013**.

[140] X. Du, Y. Li, Y. L. Xia, S. M. Ai, J. Liang, P. Sang, X. L. Ji, S. Q. Liu, *International Journal of Molecular Sciences* **2016**, *17*, 1–34.

[141] C. Bissantz, B. Kuhn, M. Stahl, *Journal of Medicinal Chemistry* **2010**, *53*, 5061–5084.

[142] D. W. Banner, **2003**.

[143] J. P. Goddard, J. L. Reymond, *Current Opinion in Biotechnology* **2004**, *15*, 314–322.

[144] J. Sumranjit, S. J. Chung, *Molecules* **2013**, *18*, 10425–10451.

[145] D. Robinette, N. Neamati, K. B. Tomer, C. H. Borchers, *Expert Review of Proteomics* **2006**, *3*, 399–408.

[146] V. Canbay, U. auf dem Keller, *Current Opinion in Chemical Biology* **2021**, *60*, 89–96.

[147] I. L. H. Ong, K.-L. Yang, *The Analyst* **2017**, *142*, 1867–1881.

[148] A. Carmel, M. Zur, A. Yaron, E. Katchalski, *FEBS Letters* **1973**, *30*, 11–14.

[149] H. Tuppy, U. Wiesbauer, E. Wintersberger, *Hoppe-Seyler´s Zeitschrift für physiologische Chemie* **1962**, *329*, 278–288.

[150] M. Zimmerman, E. Yurewicz, G. Patel, *Analytical Biochemistry* **1976**, *70*, 258–262.

[151] S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. Chung, N. P. Coussens, J. L. Dahlin, X. Xu., *Assay Guidance Manual*, Eli Lilly & Company And The National Center For Advancing Translational Sciences, **2004**.

[152] M. Bernetti, M. Masetti, W. Rocchia, A. Cavalli, *Annual Review of Physical Chemistry* **2019**, *70*, 143–171.

[153] M. G. Acker, D. S. Auld, *Perspectives in Science* **2014**, *1*, 56–73.

[154] J. P. Goddard, J. L. Reymond, *Trends in Biotechnology* **2004**, *22*, 363–370.

[155] H. Bisswanger, *Perspectives in Science* **2014**, *1*, 41–55.

[156] B. T. Burlingham, T. S. Widlanski, *Journal of Chemical Education* **2003**, *80*, 214–218.

[157] C. Yung-Chi, W. H. Prusoff, *Biochemical Pharmacology* **1973**, *22*, 3099–3108.

[158] M. DIXON, *The Biochemical journal* **1953**, *55*, 170–171.

[159] P. J. Butterworth, *BBA - Enzymology* **1972**, *289*, 251–253.

[160] F. H. Niesen, H. Berglund, M. Vedadi, *Nature Protocols* **2007**, *2*, 2212–2221.

[161] K. R. Rupesh, A. Smith, P. E. Boehmer, *Biochemical and Biophysical Research Communications* **2014**, *454*, 604–608.

[162] K. Huynh, C. L. Partch, *Current Protocols in Protein Science* **2015**, *79*, DOI 10.1002/0471140864.ps2809s79.

[163] S. N. Krishna, C.-H. Luan, R. K. Mishra, L. Xu, K. A. Scheidt, W. F. Anderson, R. C. Bergan, *PLoS ONE* **2013**, *8*, e81504.

[164] M. W. Pantoliano, E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane, F. R. Salemme, *Journal of Biomolecular Screening* **2001**, *6*, 429–440.

[165] K. Gao, R. Oerlemans, M. R. Groves, *Biophysical Reviews* **2020**, *12*, 85–104.

[166] J. J. Bowling, W. R. Shadrick, E. C. Griffith, R. E. Lee, in *Special Topics in Drug Discovery*, InTech, **2016**.

[167] M.-C. Lo, A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, *Analytical Biochemistry* **2004**, *332*, 153–159.

[168] M. Vivoli, H. R. Novak, J. A. Littlechild, N. J. Harmer, *Journal of Visualized Experiments* **2014**, 1–13.

[169] D. E. Scott, C. Spry, C. Abell, *Differential Scanning Fluorimetry as Part of a Biophysical Screening Cascade*, **2016**.

[170] A. O. Magnusson, A. Szekrenyi, H. J. Joosten, J. Finnigan, S. Charnock, W. D. Fessner, *FEBS Journal* **2019**, *286*, 184–204.

[171] C. G. Alexander, R. Wanner, C. M. Johnson, D. Breitsprecher, G. Winter, S. Duhr, P. Baaske, N.

Ferguson, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2014**, *1844*, 2241–2250.

[172] M. Maschberger, “Exploring protein stability by nanoDSF,” can be found under www. nanotemper-technologies. com., **n.d.**

[173] M. U. D. Ahmad, A. Fish, J. Molenaar, S. Sreeramulu, C. Richter, N. Altincekic, H. Schwalbe, H. Wienk, A. Perrakis, *Journal of Visualized Experiments* **2021**, *2021*, 1–18.

[174] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer US, Boston, MA, **2006**.

[175] D. Bruce, E. Cardew, S. Freitag-Pohl, E. Pohl, *Journal of Visualized Experiments* **2019**, *2019*, 1–11.

[176] H. Mazal, G. Haran, *Current Opinion in Biomedical Engineering* **2019**, *12*, 8–17.

[177] V. Kanelis, J. D. Forman-Kay, L. E. Kay, *IUBMB Life* **2001**, *52*, 291–302.

[178] Y. Shi, *Cell* **2014**, *159*, 995–1014.

[179] Y. Cheng, N. Grigorieff, P. A. Penczek, T. Walz, *Cell* **2015**, *161*, 438–449.

[180] H. Frauenfelder, S. G. Sligar, P. G. Wolynes, *Science* **1991**, *254*, 1598–1603.

[181] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972.

[182] D. E. Koshland, *Proceedings of the National Academy of Sciences* **1958**, *44*, 98–104.

[183] F. J. Meyer-Almes, *European Biophysics Journal* **2016**, *45*, 245–257.

[184] J.-P. Changeux, S. Edelstein, *F1000 Biology Reports* **2011**, *3*, 1–15.

[185] J. Guo, H. X. Zhou, *Chemical Reviews* **2016**, *116*, 6503–6515.

[186] D. D. Boehr, R. Nussinov, P. E. Wright, *Nature Chemical Biology* **2009**, *5*, 789–796.

[187] G. G. Hammes, Y.-C. Chang, T. G. Oas, *Proceedings of the National Academy of Sciences* **2009**, *106*, 13737–13741.

[188] D. Michel, *Biochimie* **2016**, *128–129*, 48–54.

[189] T. Förster, *Annalen der Physik* **1948**, *437*, 55–75.

[190] E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Y. Chung, X. Michalet, S. Weiss, *Science* **2018**, *359*, DOI 10.1126/science.aan1133.

[191] X. Michalet, S. Weiss, M. Jäger, *Chemical Reviews* **2006**, *106*, 1785–1813.

[192] R. M. Clegg, *Current Biology* **1995**, *6*, 103–110.

[193] C. A. Royer, *Chemical Reviews* **2006**, *106*, 1769–1784.

[194] H. Sahoo, P. Schwille, *ChemPhysChem* **2011**, *12*, 532–541.

[195] L. Stryer, *Annual Review Biochemistry* **1978**, *47*, 819–846.

[196] R. Roy, S. Hohng, T. Ha, *Nature Methods* **2008**, *5*, 507–516.

[197] A. A. Deniz, M. Dahan, J. R. Grunwell, T. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss, P. G. Schultz, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 3670–3675.

[198] R. B. Best, K. A. Merchant, I. V. Gopich, B. Schuler, A. Bax, W. A. Eaton, *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 18964–18969.

[199] K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, W. A. Eaton, *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 1528–1533.

[200] B. Schuler, W. A. Eaton, *Current Opinion in Structural Biology* **2008**, *18*, 16–26.

[201] J. Widengren, V. Kudryavtsev, M. Antonik, S. Berger, M. Gerken, C. A. M. Seidel, *Analytical Chemistry* **2006**, *78*, 2039–2050.

[202] W. E. Moerner, D. P. Fromm, *Review of Scientific Instruments* **2003**, *74*, 3597–3619.

[203] D. Axelrod, *Traffic* **2001**, *2*, 764–774.

[204] E. L. Elson, D. Magde, *Biopolymers* **1974**, *13*, 1–27.

[205] D. Magde, E. Elson, W. W. Webb, *Physical Review Letters* **1972**, *29*, 705–708.

[206] D. Magde, E. L. Elson, W. W. Webb, *Biopolymers* **1974**, *13*, 29–61.

[207] P. Schwille, **2001**, *34*.

[208] S. T. Hess, W. W. Webb, *Biophysical Journal* **2002**, *83*, 2300–2317.

[209] O. Krichevsky, G. Bonnet, *Reports on Progress in Physics* **2002**, *65*, 251–297.

[210] K. Bacia, S. A. Kim, P. Schwille, *Nature Methods* **2006**, *3*, 83–89.

[211] K. Bacia, P. Schwille, *Nature Protocols* **2007**, *2*, 2842–2856.

[212] E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, C. A. M. Seidel, *Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection*, Elsevier Inc., **2010**.

[213] J. Widengren, E. Schweinberger, S. Berger, C. A. M. Seidel, *Journal of Physical Chemistry A* **2001**, *105*, 6851–6866.

[214] R. Kühnemuth, C. A. M. Seidel, *Single Molecules* **2001**, *2*, 251–254.

[215] T. Torres, M. Levitus, *Journal of Physical Chemistry B* **2007**, *111*, 7392–7400.

[216] M. Margittai, J. Widengren, E. Schweinberger, G. F. Schröder, S. Felekyan, E. Haustein, M. König, D. Fasshauer, H. Grubmüller, R. Jahn, C. A. M. Seidel, *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 15516–15521.

[217] M. Dimura, T. O. Peulen, C. A. Hanke, A. Prakash, H. Gohlke, C. A. Seidel, *Current Opinion in Structural Biology* **2016**, *40*, 163–185.

[218] L. M. Hirvonen, K. Suhling, *Measurement Science and Technology* **2017**, *28*, DOI 10.1088/1361-6501/28/1/012003.

[219] K. Suhling, L. M. Hirvonen, W. Becker, S. Smietana, H. Netz, J. Milnes, T. Conneely, A. Le Marois, O. Jagutzki, F. Festy, Z. Petrášek, A. Beeby, *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2019**, *942*, 162365.

[220] G. Gouridis, G. K. Schuurman-Wolters, E. Ploetz, F. Husada, R. Vietrov, M. De Boer, T. Cordes, B. Poolman, *Nature Structural and Molecular Biology* **2015**, *22*, 57–64.

[221] J. M. H. Goudsmits, D. J. Slotboom, A. M. Van Oijen, *Nature Communications* **2017**, *8*, 1–10.

[222] M. Yang, N. Livnat Levanon, B. Acar, B. Aykac Fas, G. Masrati, J. Rose, N. Ben-Tal, T. Haliloglu, Y. Zhao, O. Lewinson, *Nature Chemical Biology* **2018**, *14*, 715–722.

[223] M. De Boer, G. Gouridis, R. Vietrov, S. L. Begg, G. K. Schuurman-Wolters, F. Husada, N. Eleftheriadis, B. Poolman, C. A. McDevitt, T. Cordes, *eLife* **2019**, *8*, 1–28.

[224] C. Ratzke, F. Berkemeier, T. Hugel, *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 161–166.

[225] C. Ratzke, B. Hellenkamp, T. Hugel, *Nature Communications* **2014**, *5*, DOI 10.1038/ncomms5192.

[226] F. H. Schopf, M. M. Biebl, J. Buchner, *Nature Reviews Molecular Cell Biology* **2017**, *18*, 345–360.

[227] B. Hellenkamp, P. Wortmann, F. Kandzia, M. Zacharias, T. Hugel, *Nature Methods* **2017**, *14*, 176–182.

[228] C. Ratzke, M. Mickler, B. Hellenkamp, J. Buchner, T. Hugel, *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 16101–16106.

[229] N. S. Hatzakis, *Biophysical Chemistry* **2014**, *186*, 46–54.

[230] D. Bucher, B. J. Grant, J. A. McCammon, *Biochemistry* **2011**, *50*, 10530–10539.

[231] C. Tang, C. D. Schwieters, G. M. Clore, *Nature* **2007**, *449*, 1078–1082.

[232] E. Kim, S. Lee, A. Jeon, J. M. Choi, H. S. Lee, S. Hohng, H. S. Kim, *Nature Chemical Biology* **2013**, *9*, 313–318.

[233] A. Singh, E. R. Thornton, F. H. Westheimer, *Journal of Biological Chemistry* **1962**, *237*, PC3006–PC3008.

[234] E. Smith, I. Collins, *Future Medicinal Chemistry* **2015**, *7*, 159–183.

[235] L. Dubinsky, B. P. Krom, M. M. Meijler, *Bioorganic and Medicinal Chemistry* **2012**, *20*, 554–570.

[236] D. P. Murale, S. C. Hong, M. M. Haque, J. S. Lee, *Proteome Science* **2017**, *15*, 1–34.

[237] J. Das, *Chemical Reviews* **2011**, *111*, 4405–4417.

[238] S. Zhao, L. Zhao, X. Zhang, C. Liu, C. Hao, *European Journal of Medicinal Chemistry* **2016**, *123*, 514–522.

[239] A. M. Sadaghiani, S. H. Verhelst, M. Bogyo, *Current Opinion in Chemical Biology* **2007**, *11*, 20–28.

[240] P. P. Geurink, L. M. Prely, G. A. van der Marel, R. Bischoff, H. S. Overkleeft, in *Peptide-Based Materials*, **2011**, pp. 85–113.

[241] H. Ovaa, P. F. Van Swieten, B. M. Kessler, M. A. Leeuwenburgh, E. Fiebiger, A. M. C. H. Van den Nieuwendijk, P. J. Galardy, G. A. Van der Marel, H. L. Ploegh, H. S. Overkleeft, *Angewandte Chemie - International Edition* **2003**, *42*, 3626–3629.

[242] M. Vrabel, P. Kölle, K. M. Brunner, M. J. Gattner, V. López-Carrillo, R. de Vivie-Riedle, T. Carell, *Chemistry - A European Journal* **2013**, *19*, 13309–13312.

[243] J. E. Moore, W. H. Ward, *Journal of the American Chemical Society* **1956**, *78*, 2414–2418.

[244] S. B. Gunnoo, A. Madder, *ChemBioChem* **2016**, *17*, 529–553.

[245] K. Renault, J. W. Fredy, P.-Y. Renard, C. Sabot, *Bioconjugate Chemistry* **2018**, *29*, DOI 10.1021/acs.bioconjchem.8b00252.

[246] Z. Song, Q. Zhang, *Organic Letters* **2009**, *11*, 4882–4885.

[247] K. A. Woll, W. P. Dailey, R. G. Eckenhoff, *Identification of General Anesthetic Target Protein-*

Binding Sites by Photoaffinity Labeling and Mass Spectrometry, Elsevier Inc., **2018**.

[248] O. Jahn, K. Eckart, H. Tezval, J. Spiess, *Analytical and Bioanalytical Chemistry* **2004**, *378*, 1031–1036.

[249] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorganic & Medicinal Chemistry* **2021**, *47*, 116392.

[250] T. R. Weikl, F. Paul, *Protein Science* **2014**, *23*, 1508–1518.

[251] C. Nitsche, L. Zhang, L. F. Weigel, J. Schilz, D. Graf, R. Bartenschlager, R. Hilgenfeld, C. D. Klein, *Journal of Medicinal Chemistry* **2017**, *60*, 511–516.

[252] H. Lee, J. Ren, S. Nocadello, A. J. Rice, I. Ojeda, S. Light, G. Minasov, J. Vargas, D. Nagarathnam, W. F. Anderson, M. E. Johnson, *Antiviral Research* **2017**, *139*, 49–58.

[253] M. A. M. Behnam, D. Graf, R. Bartenschlager, D. P. Zlotos, C. D. Klein, *Journal of Medicinal Chemistry* **2015**, *58*, 9354–9370.

[254] N. Gruba, J. I. Rodriguez Martinez, R. Grzywa, M. Wysocka, M. Skoreński, M. Burmistrz, M. Łęcka, A. Lesner, M. Sieńczyk, K. Pyrć, *FEBS Letters* **2016**, *590*, 3459–3468.

[255] N. Kühl, D. Graf, J. Bock, M. A. M. Behnam, M. M. Leuthold, C. D. Klein, *Journal of Medicinal Chemistry* **2020**, *63*, 8179–8197.

[256] N. Kühl, M. M. Leuthold, M. A. M. Behnam, C. D. Klein, *Journal of Medicinal Chemistry* **2021**, *64*, 4567–4587.

[257] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Advanced Drug Delivery Reviews* **1997**, *23*, 3–25.

[258] C. A. Lipinski, *Drug Discovery Today: Technologies* **2004**, *1*, 337–341.

[259] F. Benmansour, I. Trist, B. Coutard, E. Decroly, G. Querat, A. Brancale, K. Barral, *European Journal of Medicinal Chemistry* **2017**, *125*, 865–880.

[260] N. Frimayanti, S. M. Zain, V. S. Lee, H. A. Wahab, R. Yusof, N. A. Rahman, *In Silico Biology* **2011**, *11*, 29–37.

[261] J. P. Quek, S. Liu, Z. Zhang, Y. Li, E. Y. Ng, Y. R. Loh, A. W. Hung, D. Luo, C. B. Kang, E. Yihui, Y. Ru, A. W. Hung, D. Luo, C. B. Kang, *Antiviral Research* **2020**, *175*, 104707.

[262] C. Nitsche, C. Steuer, C. D. Klein, *Bioorganic and Medicinal Chemistry* **2011**, *19*, 7318–7337.

[263] C. Steuer, C. Gege, W. Fischl, K. H. Heinonen, R. Bartenschlager, C. D. Klein, *Bioorganic and Medicinal Chemistry* **2011**, *19*, 4067–4074.

[264] B. C. Chenna, L. Li, D. M. Mellott, X. Zhai, J. L. Siqueira-Neto, C. Calvet Alvarez, J. A. Bernatchez, E. Desormeaux, E. Alvarez Hernandez, J. Gomez, J. H. McKerrow, J. Cruz-Reyes, T. D. Meek, *Journal of Medicinal Chemistry* **2020**, *63*, 3298–3316.

[265] T. T. Baird, C. S. Craik, *Handbook of Proteolytic Enzymes* **2013**, *3*, 2594–2600.

[266] A. Jadhav, R. S. Ferreira, C. Klumpp, B. T. Mott, C. P. Austin, J. Inglese, C. J. Thomas, D. J. Maloney, B. K. Shoichet, A. Simeonov, *Journal of Medicinal Chemistry* **2010**, *53*, 37–51.

[267] J. Comley, *Drug Discovery World Summer 2003* **2003**, 91–98.

[268] Y. Liu, W. Kati, C.-M. Chen, R. Tripathi, A. Molla, W. Kohlbrenner, *Analytical Biochemistry* **1999**, *267*, 331–335.

[269] N. Thorne, D. S. Auld, J. Inglese, *Current Opinion in Chemical Biology* **2010**, *14*, 315–324.

[270] F. E. Jernigan, D. S. Lawrence, *Chemical Communications* **2013**, *49*, 6728.

[271] S. G. Deeks, M. Smith, M. Holodniy, J. O. Kahn, *Journal of the American Medical Association* **1997**, *277*, 145–153.

[272] S. Jo, H. Kim, S. Kim, D. H. Shin, M. Kim, *Chemical Biology & Drug Design* **2019**, *94*, 2023–2030.

[273] J.-M. Delaissé, Y. Eeckhout, G. Vaes, *Biochemical and Biophysical Research Communications* **1984**, *125*, 441–447.

[274] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, S. Weiss, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 6264–6268.

[275] C. Bräuchle, D. C. Lamb, J. Michaelis, *Single Particle Tracking and Single Molecule Energy Transfer*, Wiley, Weinheim, **2009**.

[276] L. De La Cruz, T. H. D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber, G. Otting, *Journal of the American Chemical Society* **2011**, *133*, 19205–19215.

[277] T. C. Chou, P. Talalay, *Advances in Enzyme Regulation* **1984**, *22*, 27–55.

[278] Q. Li, C. Kang, *International Journal of Molecular Sciences* **2020**, *21*, 2527.

[279] Y. Li, Z. Zhang, W. W. Phoo, Y. R. Loh, W. Wang, S. Liu, M. W. Chen, A. W. Hung, T. H. Keller, D. Luo, C. B. Kang, *Structure* **2017**, *25*, 1242-1250.e3.

[280] Y. Li, Z. Zhang, W. W. Phoo, Y. R. Loh, R. Li, H. Y. Yang, A. E. Jansson, J. Hill, T. H. Keller, K. Nacro, D. Luo, C. B. Kang, *Structure* **2018**, *26*, 555-564.e3.

[281] M. C. Mahawaththa, B. J. G. Pearce, M. Szabo, B. Graham, C. D. Klein, C. Nitsche, G. Otting, *Antiviral Research* **2017**, *142*, 141–147.

[282] Y. Li, W. W. Phoo, Y. R. Loh, Z. Zhang, E. Y. Ng, W. Wang, T. H. Keller, D. Luo, C. Kang, *FEBS Letters* **2017**, *591*, 2338–2347.

[283] C. Götz, G. Hinze, A. Gellert, H. Maus, F. von Hammerstein, S. J. Hammerschmidt, L. M. Lauth, U. A. Hellmich, T. Schirmeister, T. Basché, *The Journal of Physical Chemistry B* **2021**, *125*, 6837–6846.

[284] H. Maus, G. Hinze, S. J. Hammerschmidt, T. Basché, T. Schirmeister, *Protein Science* **2022**, *32*, e4526.

[285] N. Pozzi, A. D. Vogt, D. W. Gohara, E. Di Cera, *Current Opinion in Structural Biology* **2012**, *22*, 421–431.

[286] H. Maus, S. J. Hammerschmidt, G. Hinze, F. Barthels, V. H. Pérez Carrillo, U. A. Hellmich, T. Basché, T. Schirmeister, *European Journal of Medicinal Chemistry* **2023**, *258*, 115573.

[287] P. Müller, M. Meta, J. L. Meidner, M. Schwickert, J. Meyr, K. Schwickert, C. Kersten, C. Zimmer, S. J. Hammerschmidt, A. Frey, A. Lahu, S. de la Hoz-Rodríguez, L. Agost-Beltrán, S. Rodríguez, K. Diemer, W. Neumann, F. V. Gonzàlez, B. Engels, T. Schirmeister, *International Journal of Molecular Sciences* **2023**, *24*, 7226.

[288] W. Yang, X. Gao, B. Wang, *Medicinal Research Reviews* **2003**, *23*, 346–368.

[289] Y. Tan, J. Wu, L. Song, M. Zhang, C. J. Hipolito, C. Wu, S. Wang, Y. Zhang, Y. Yin, *International Journal of Molecular Sciences* **2021**, *22*, DOI 10.3390/ijms222312958.

[290] R. Smoum, A. Rubinstein, V. M. Dembitsky, M. Srebnik, *Chemical Reviews* **2012**, *112*, 4156–4220.

[291] A. K. Bera, R. J. Kuhn, J. L. Smith, *Journal of Biological Chemistry* **2007**, *282*, 12883–12892.

[292] G. Robin, K. Chappell, M. J. Stoermer, S. H. Hu, P. R. Young, D. P. Fairlie, J. L. Martin, *Journal of Molecular Biology* **2009**, *385*, 1568–1577.

[293] K. J. Chappell, T. A. Nall, M. J. Stoermer, N. X. Fang, J. D. A. Tyndall, D. P. Fairlie, P. R. Young, *Journal of Biological Chemistry* **2005**, *280*, 2896–2903.

[294] K. J. Chappell, M. J. Stoermer, D. P. Fairlie, P. R. Young, *Journal of Biological Chemistry* **2006**, *281*, 38448–38458.

[295] S. A. Shiryaev, B. I. Ratnikov, A. V. Chekanov, S. Sikora, D. V. Rozanov, A. Godzik, J. Wang, J. W. Smith, Z. Huang, I. Lindberg, M. A. Samuel, M. S. Diamond, A. Y. Strongin, *Biochemical*

Journal **2006**, *393*, 503–511.

[296] T. A. Nall, K. J. Chappell, M. J. Stoermer, N. X. Fang, J. D. A. Tyndall, P. R. Young, D. P. Fairlie, *Journal of Biological Chemistry* **2004**, *279*, 48535–48542.

[297] W. W. Phoo, A. El Sahili, Z. Z. Zhang, M. W. Chen, C. W. Liew, J. Lescar, S. G. Vasudevan, D. Luo, *Antiviral Research* **2020**, *182*, 104900.

[298] R. Hilgenfeld, J. Lei, L. Zhang, **2018**, pp. 131–145.

Curriculum Vitae