Beyond Binding Affinity – Elucidating Protease Inhibition Mechanisms from a Biophysical Perspective

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

Apotheker Stefan Josef Hammerschmidt

geb. in Zweibrücken

Mainz, 2023

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

Dean: [redacted]

Name of the 1st reviewer: [redacted]

Name of the 2nd reviewer: [redacted]

Doctoral research: From 01. January 2018 to 17. April 2023

Date of the doctoral examination: _____

D77 (Dissertation Universität Mainz)

Declaration of Authorship

I, Stefan Josef Hammerschmidt, declare that this thesis entitled, “Beyond Binding Affinity – Elucidating Protease Inhibition Mechanisms from a Biophysical Perspective” and the work presented in it are my own. I hereby declare that:

- I did this work completely 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.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted parts from the work of others, the source is given. Except for such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Place, Date: Mainz, _____ Signature: _____

“Life itself is your teacher, and you are in a state of constant learning.”

— Bruce Lee

Acknowledgments

[redacted]

[redacted]

Table of Contents

Abbreviations

^{5}ZiPro	quintuple-mutant of ZIKV NS2B/NS3
ADE	antibody-dependent enhancement
ADME	absorption, distribution, metabolism, and elimination
AEX	exchange chromatography
BMI	*N*-benzylmaleimide
bZiPro	bivalent expressed ZIKV NS2B/NS3
C	capsid
CADD	computer-aided drug design
Can	collagen adhesin
CD	circular dichroism
Clf	clumping factors
cpd	compound
CUB	Cls/Clr, urchin embryonic growth factor, bone morphogenetic protein-1
CWSS	cell wall sorting signal
DEER	pulsed double electron-electron resonance
DENV	Dengue virus
DESC	differentially expressed in squamous cell carcinoma
ds	double-stranded
DSC	differential scanning calorimetry
DSF	differential scanning fluorimetry
E	envelope
eDNA	extracellular DNA
E_{ET}	energy transfer efficiency
EMI	*N*-ethylmaleimide
EPR	electron paramagnetic resonance
ER	endoplasmic reticulum
eDENV4	enzymatic cleavable linked DENV4 full-length $NS2B_{cf}$/NS3
eZiPro	enzymatic cleavable linked ZIKV $NS2B_{cf}/NS3_{pro}$
FDA	Food and Drug Administration
Fnbp	fibronectin-binding proteins
FRET	Förster resonance energy transfer
GPI	glycosyl-phosphatidylinositol
gZiPro	glycine-linked ZIKV $NS2B_{cf}/NS3_{pro}$
HAT	human airway trypsin-like
HCV	hepatitis C virus
HGF	hepatocyte growth factor

His_6	hexahistidine
IDR	intrinsically disordered region
IgG	immune globulin G
IOMC	Institute of Organic Chemistry & Macromolecular Chemistry
IPSL	3-(2-iodoacetamido)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy
IR	infrared
Isd	iron-regulated surface determinant
ITC	isothermal titration calorimetry
JEV	Japanese encephalitis virus
LDLA	low-density lipoprotein receptor class A
LE	ligand efficiency
LipE	lipophilic ligand efficiency
LLE	lipophilic ligand efficiency
MAM	meprin, A5 antigen, receptor protein phosphatase
MASP	membrane-anchored serine proteases
MSCRAMM	microbial surface components recognizing adhesive matrix molecule
MST	microscale thermophoresis
MT-SP1	membrane-type serine proteinase 1
MW	molecular weight
N	binding stoichiometry
nanoSPR	nano surface plasmon resonance
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
n_{proton}	number of transferred protons
NS	nonstructural
$NS2B_{cf}$	cofactor region of NS2B
$NS3_{pro}$	protease domain of NS3
NTA	Ni-nitrilotriacetic acid
NTD	neglected tropical diseases
PDB	RCSB protein databank
pre-MSCRAMM	precursor surface protein
prM	precursor membrane
RMSD	root mean square deviation
RNA	ribonucleic acid
ROC	receiver operator characteristics
rt.	room temperature
S	spike protein
S. aureus	*Staphylococcus aureus*

SAR	structure-activity relationship
SARS-CoV-2	severe acute respiratory syndrome corona virus 2
SATM	signal anchor transmembrane domain
SEA	sperm protein, enteropeptidase, and agrin
Sec	secretory pathway
smFRET	single-molecule Förster resonance energy transfer
SML	SrtA-mediated ligation
SP	serine protease domain
SpA	protein A
SPR	surface plasmon resonance
SRCR	group A scavenger receptor cysteine-rich
SrtA	sortase A
ss	single-stranded
ST14	suppressor of tumorgenicity 14
T_m	melting temperature
TMPRSS14	transmembrane protease serine 14
TTSP	type II transmembrane serine proteases
$-T\Delta S$	temperature-dependent binding entropy
uPA	and urokinase-type plasminogen activator
WNV	West Nile virus
WT	wild type
YFV	Yellow fever virus
ZIKV	Zika virus
ΔC_P	the isobaric binding heat capacity
ΔG	Gibbs energy
ΔH	binding enthalpy

Abstract

Proteases play essential roles in nature and are involved in protein maturation, digestion, forming extracellular matrices and their degradation, and immune escape. Inhibition of such proteases often represents a promising strategy to counter related diseases. Detailed knowledge beyond binding affinity is a prerequisite to developing more potent inhibitors. Characterizing the modes of action relies on biophysical techniques that can provide structural information on binding modes and determine binding kinetics or thermodynamics. Furthermore, they can trace conformational changes. In the course of this work, inhibitors for several disease-related proteases were characterized using a toolset of biophysical techniques. The target proteases were selected for investigation based on their distinct functions and catalysis mechanisms. In this regard, the flaviviral NS3 proteases feature an NS2B cofactor, can adopt at least two conformations, and are essential for polyprotein processing and maturation. *Staphylococcus aureus* sortase A is a transpeptidase that relies on a Ca^{2+} cofactor and is vital to forming extracellular matrices. The human matriptase displays a roughly symmetric substrate binding site and degrades extracellular matrices by activating matrix metalloproteases. Thereof, the following projects were derived:

Project 1. Two strategies were examined to elucidate the mechanistic basis of entropic inhibitor optimization: The macrocyclization of Zika virus NS2B/NS3 protease inhibitors and the introduction of higher ligand symmetry of human matriptase inhibitors to enable multiple equivalent binding modes. The various opportunities to derive data from well-designed ITC experiments prompted us to develop a set of teaching experiments, demonstrating the versatility of ITC experiment execution and analysis. To guide medicinal scientists through ITC-assisted ligand development, we implemented the webserver *ITCcalc*.

Project 2. The Dengue virus and Zika Virus NS2B/NS3 proteases display an allosteric binding site. Several biophysical methods, such as X-ray co-crystallization, MST, ITC, DSF, CD, EPR, fluorometric assays, smFRET, and ^{19}F-NMR, were utilized to give in-depth information on conformational changes that are induced upon ligand binding.

Project 3. During virus maturation, the NS2B/NS3 protease is autocatalytically cleaved at the NS2B/NS3 junction. To date, the biological significance of this highly conserved process is unknown. In this project, we investigated the mechanisms of this process to understand the regulatory functions of autocatalytic cleavage. To this end, we used fluorometric assays, CD, extensive site-directed mutagenesis, analytical SDS-PAGE, and protein MS.

Project 4. Due to their antibacterial properties, metal ions are widely applied as additives in surface disinfectants, alloys of medicinal instruments, textiles, and wound dressings. SrtA constitutes an exciting anti-virulence target to counter antibiotic-resistant *Staphylococcus aureus* lineages. In this work, we found several metal cations to activate, inhibit, or modulate SrtA activity in FRET-based transpeptidase assays. Furthermore, their mode of action behind the activity modulation of SrtA was elucidated. Tb^{3+} FRET assays, ITC, DSF, MST, and X-ray crystallography were used for this purpose.

Kurzdarstellung

Proteasen spielen in der Natur viele wesentliche Rollen und sind an Proteinreifung, Verdauung, der Ausbildung extrazellulärer Matrices und deren Abbau sowie an der Immunflucht maßgeblich beteiligt. Die Hemmung solcher Proteasen stellt häufig eine vielversprechende Strategie zur Bekämpfung der damit verbundenen Krankheiten dar. Detaillierte Kenntnisse über die Bindungsaffinität hinaus sind eine Voraussetzung für die Entwicklung wirksamerer Inhibitoren. Die Charakterisierung der Wirkungsweisen beruht auf biophysikalischen Techniken, die strukturelle Informationen über die Bindungsmodi liefern und die Bindungskinetik oder -thermodynamik bestimmen können. Außerdem können sie Konformationsänderungen verfolgen. Im Rahmen dieser Arbeit wurden Inhibitoren für mehrere krankheitsrelevante Proteasen mit Hilfe eines Instrumentariums biophysikalischer Techniken charakterisiert. Die Zielproteasen wurden auf der Grundlage ihrer unterschiedlichen Funktionen und Katalysemechanismen für die Untersuchung ausgewählt. Die flaviviralen NS3-Proteasen weisen einen NS2B-Kofaktor auf, können mindestens zwei Konformationen annehmen und sind für die Verarbeitung und Reifung von Polyproteinen unerlässlich. Die Sortase A von *Staphylococcus aureus* ist eine Transpeptidase, die auf einen Ca^{2+}-Kofaktor angewiesen und für die Bildung extrazellulärer Matrices unerlässlich ist. Die menschliche Matriptase weist eine annähernd symmetrische Substratbindungsstelle auf und baut extrazelluläre Matrices durch Aktivierung von Matrixmetalloproteasen ab. Daraus wurden die folgenden Projekte abgeleitet:

Projekt 1. Um die mechanistischen Grundlagen der entropischen Inhibitoroptimierung zu erhellen, wurden zwei Strategien untersucht: Die Makrozyklisierung von Zika-Virus NS2B/NS3-Proteaseinhibitoren und die Einführung höherer Ligandensymmetrie von humanen Matriptaseinhibitoren, um mehr äquivalente Bindemodi zu ermöglichen. Die vielfältigen Möglichkeiten, Daten aus gut konzipierten ITC-Experimenten zu gewinnen, haben uns veranlasst, eine Reihe von Lehrexperimenten zu entwickeln, die die Vielseitigkeit der Durchführung und Analyse von ITC-Experimenten demonstrieren. Um medizinische Wissenschaftler durch die ITC-gestützte Ligandenentwicklung zu führen, haben wir den Webserver *ITCcalc* implementiert.

Projekt 2. Die NS2B/NS3-Proteasen des Dengue- und Zika-Virus weisen eine allosterische Bindungsstelle auf. Verschiedene biophysikalische Methoden wie Röntgen-Co-Kristallisation, MST, ITC, DSF, CD, EPR, fluorometrische Assays, smFRET und ^{19}F-NMR wurden eingesetzt, um detaillierte Informationen über die Konformationsänderungen zu erhalten, die durch die Ligandenbindung induziert werden.

Projekt 3. Während der Virusreifung wird die NS2B/NS3-Protease autokatalytisch an der NS2B/NS3-Verbindung gespalten. Bis heute ist die biologische Bedeutung dieses hochkonservierten Prozesses unbekannt. In diesem Projekt haben wir die Mechanismen dieses Prozesses untersucht, um die regulatorischen Funktionen der autokatalytischen Spaltung zu verstehen. Zu diesem Zweck verwendeten wir fluorometrische Assays, CD, umfangreiche ortsgerichtete Mutagenese, analytische SDS-PAGE und Protein-MS.

Projekt 4. Aufgrund ihrer antibakteriellen Eigenschaften werden Metallionen in großem Umfang als Zusatzstoffe in Oberflächendesinfektionsmitteln, Legierungen von medizinischen Instrumenten, Textilien und Wundverbänden eingesetzt. SrtA ist ein spannendes Anti-Virulenz-Target, um antibiotikaresistente *Staphylococcus aureus*-Stämme zu bekämpfen. In dieser Arbeit haben wir mehrere Metallkationen gefunden, die die Aktivität von SrtA in FRET-basierten Transpeptidase-Assays aktivieren, hemmen oder modulieren. Darüber hinaus wurde die Wirkungsweise hinter der Aktivitätsmodulation von SrtA aufgeklärt. Zu diesem Zweck wurden Tb^{3+}-FRET-Tests, ITC, DSF, MST und Röntgenkristallographie eingesetzt.

1. Introduction

Since all projects are connected by the target molecules and applied biophysical methods, the first part of the introduction gives a short overview of the herein-used proteases. Their physiological functions, pathophysiological relevance, architecture, the mechanism of the catalyzed enzymatic reaction, and further aspects of interest are illuminated. This part is followed by the introduction of the four main projects that were addressed in this work. Since isothermal titration calorimetry (ITC) represents the biophysical technique predominantly used during this thesis, a short chapter (*3. Project 1: Entropy Optimization in Medicinal Chemistry*) is dedicated to the historical development of this versatile methodology and its applications.

1.1. Proteases

1.1.1. Dengue and Zika NS2B/NS3 Proteases

Global Burden and Symptoms of Dengue and Zika Infections. The global burden of neglected tropical diseases (NTD) has dramatically risen in recent decades.[1] Mosquitoes are to blame for many of them and are known to transmit a concerning number of diseases. Prominent representatives are Zika and Dengue fever, both caused by the eponymous Zika (ZIKV) and Dengue virus (DENV), transmitted by mosquitoes of the *Aedes aegypti* and *Aedes albopictus* genus as their main vectors.[2,3] Through global warming, increasing globalization, and international traveling, these vectors are distributed from the tropical and subtropical regions to more tempered countries (Figure 1).[3]

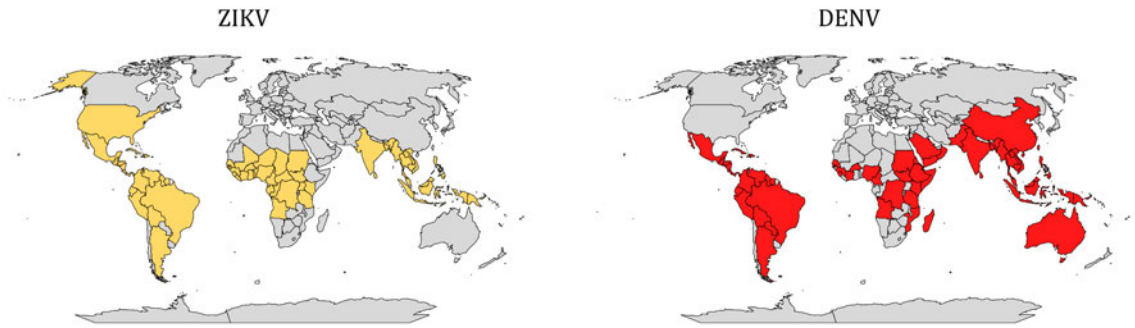


Figure 1. *Global distribution of ZIKV (yellow) and DENV (red) by country.* Affected countries are colored according to data from Collins & Metz.[4]

ZIKV causes rather epidemic outbreaks like the most recent in India 2021[5] and the Americas with over 460,000 suspected cases in 2015/2016.[6] In 2021, only 22,800 cases of ZIKV cumulated. In contrast, several countries reported increasing DENV cases in 2019, with 5.2 million registered cases. Since most infections of ZIKV and DENV take an asymptomatic course or show mild flu-like symptoms, the actual numbers of infections are expected to be extremely high, with half of the

world population at risk of DENV and an estimated 100–400 million annual infections.[7,8] The widespread ZIKV outbreak in 2016 and the correlation of ZIKV to neurological disorders such as the Guillain-Barré syndrome and microcephaly in neonates, the WHO declared ZIKV a public health emergency of international concern.[9–12] The ~20% of Dengue patients that develop symptoms suffer headaches, pain behind the eyes, muscle and joint pains, nausea, vomiting, and rash. Dengue fever can reach a critical phase 3–7 days after illness onset, called severe Dengue. With a potentially fatal outcome, this phase is characterized by severe pain, persistent vomiting, and bleedings from mucous membranes in the nose, mouth, and intestines.[13] Based on different surface proteins, DENVs are divided into four serotypes DENV (DENV1–DENV4).[14] Recovery from infection is expected to result in lifelong immunity against that serotype. However, antibody-dependent enhancement (ADE) increases the risk of developing severe Dengue after subsequent infection by different serotypes.[15] To date, no specific treatment for DENV and ZIKV infections is available, limiting the possibilities to merely symptomatic therapies and vector control measures. Although a tetravalent vaccine against DENV (Dengvaxia®, Sanofi-Pasteur) is approved by the Food and Drug Administration (FDA), it is limited to patients who have already recovered from DENV infection. Dengvaxira® shows different efficacies against the specific serotypes and pronounced side effects in seronegative patients.[16,17] The discovery of a putative fifth serotype in 2007 in Malaysia also aggravated vaccine development.[18] However, several vaccines against DENV and ZIKV currently undergo clinical trials.[19–21]

Roles of NS2B/NS3 in the Flaviviral Replication Cycle. Together with other flaviviruses like the West Nile virus (WNV), the Yellow fever virus (YFV), and the Japanese encephalitis virus (JEV), DENV and ZIKV are single-stranded (+) sense ribonucleic acid (RNA) viruses of the Flaviviridae family.[22,23] In the viral replication cycle, the ~11,000 base pairs comprising RNA is translated into a single ~3000 amino acids containing precursor polyprotein. Proper processing is critical to release the three structural (capsid (C), precursor membrane (prM), and envelope (E)) and the seven nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). Polyprotein processing is performed by the host proteases furin and signalase and the virus-specific protease domain of NS3 ($NS3_{pro}$), whose activity relies on the short C-terminal hydrophilic cofactor region of the adjacent NS2B ($NS2B_{cf}$).[24,25] Besides that, NS2B enables correct folding, enhanced solubility of $NS3_{pro}$, and anchors NS3 to the membrane of the endoplasmic reticulum.[26–30] When this essential function of NS2B/NS3 is abolished, the virus replication is interrupted (Figure 2). Thus, the NS2B/NS3 proteases are promising drug targets to counter flaviviral infections.[25,26,31]

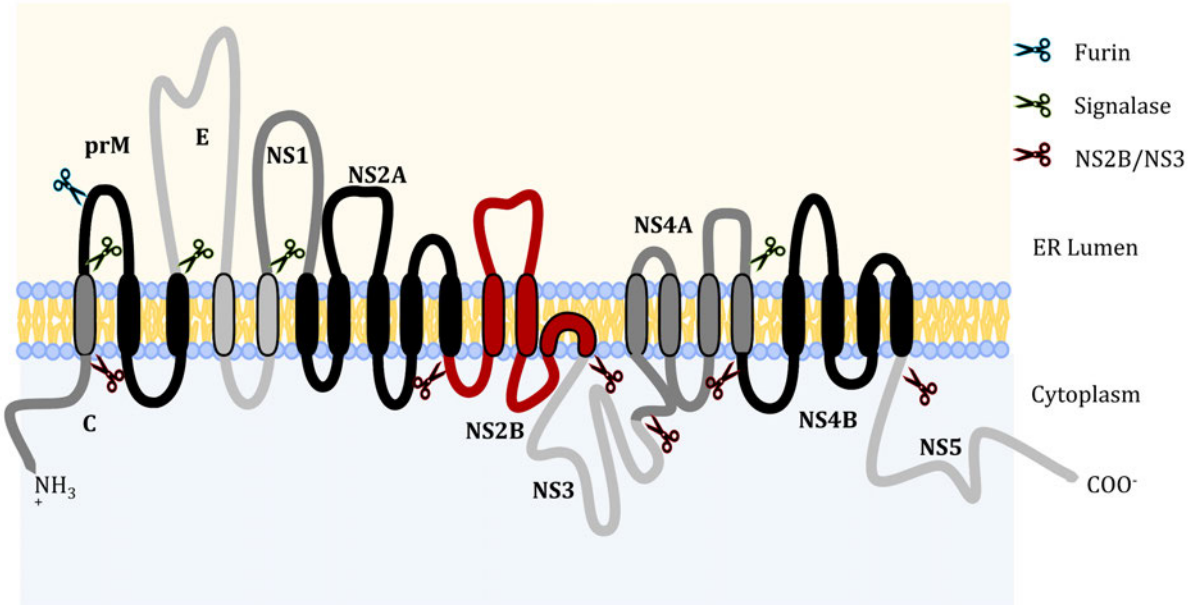


Figure 2. *Schematic depiction of flaviviral precursor polypeptide processing.* The polypeptide is located at the endoplasmatic reticulum's (ER) membrane. Scissors indicate cleavage sites of furin (blue), signalase (green), and $NS2B_{cf}/NS3_{pro}$ (red). The Graphic was designed following Nitsche *et al.*[26]

Binding site and Catalytic Reaction. Based on the evolutionary relationships to trypsin and chymotrypsin and the catalytic triad, the MEROPS database allocates the NS2B/NS3 proteases to the clan PA, the subclan PA(S), and the family S7.[32] The DENV NS2B/NS3 protease (or flavivirin) has the identifier S07.001. The ZIKV NS2B/NS3 can be found with the identifier S07.003. The clan PA comprises chymotrypsin A-like endopeptidases with a catalytic triad in the order His, Asp, and Ser or Cys. The subclan PA(S) specifies serine proteases. Accordingly, the catalytic triad of NS2B/NS3 is composed of His51, Asp75, and Ser135.[33] The binding site for substrates consists of shallow grooves on the protein surface forming the S3–S1' pockets, but with only flat S1 pockets, contrasting what can be seen for serine proteases of other families of, e.g., the human matriptase (Figure 3).[34–36] In the P1 position of the substrate, both DENV and ZIKV prefer basic residues, with a preference for Arg over Lys.[37] Whereas DENV prefers a second Arg in the P2 position, ZIKV has a slightly higher cleavage activity for Lys.[38] The S3 and S4 pockets are less selective. Gly, Ala, and Lys can be found on natural substrates. In the P1' position, small amino acids like Gly and Ser are preferred.[39]

A crystal structure with the auto-processed C-terminus of NS2B in the binding pocket revealed the binding mode of the substrate Gly-Lys-Arg (Figure 3, RCSB protein databank[40] (PDB)-ID: 5GJ4).[41] The positively charged P1 Arg is stabilized by the negatively charged Asp129 and the backbone oxygen of Tyr130 in the S1 pocket. The side chains of Ser81 and Asp83 recognize the negative charge of the P2 Lys. The P3 Gly carbonylic oxygen interacts rather unspecifically with the oxygen of Gly151 and the nitrogen of Gly153 in the S3 site. Noteworthy, the substrate in this crystal

structure is still connected to NS2B. This shortened construct probably leads to distorted P3 Gly and P4 Thr residues that do not interact with the $NS3_{pro}$ domain.[41]

The proteolytic activity of the protease is promoted by the nucleophilic attack of the hydroxyl oxygen of Ser135 to the P1 carbonylic carbon of the substrate's scissile amide bond. Asp71 enhances the nucleophilicity of Ser135, which increases the basicity of His51 adjacent to Ser135. As a serine protease of the clan PA(S) (subclan PA(S)), the catalytic Ser residue of NS2B/NS3 performs the nucleophilic attack in its protonated state. A tetrahedral intermediate is formed that harbors the negative charge at the substrate's carbonyl oxygen. This negative charge is stabilized by the oxyanion hole of the protease formed by Ala132, Gly133, and Thr134 (Figure 3A).[42] The polarized His51 can now accept the proton of Ser135 to form an imidazolium ion. The negatively charged Asp71 stabilizes the resulting positive charge of the imidazolium. The tetrahedral intermediate releases the newly generated N-terminus of the substrate and leaves the catalytic Ser acylated. Hydrolysis of this ester releases the new C-terminus, and Ser135 is regenerated to start a new catalysis cycle (Figure 4).

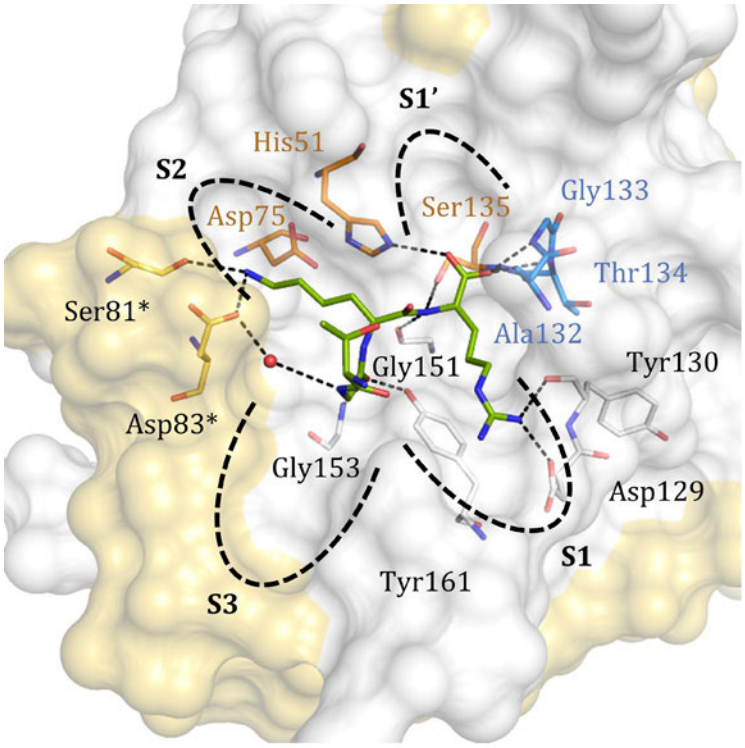


Figure 3. *The binding site of the ZIKV NS2B/NS3 protease.* The binding pocket of NS2B/NS3 in the closed state after autocatalysis of the linker (PDB-ID: 5GJ4).[41] NS3 is shown as a white surface, NS2B is colored yellow. Amino acids of the catalytic triad are shown as orange sticks, and the amino acids forming the oxyanion hole are blue. Binding sub-pockets are indicated according to Schechter&Berger.[43] The binding mode of the natural substrate TGKR (green sticks) was revealed. Polar interactions are depicted as black dashed lines. Only those amino acids forming polar interactions are displayed as thin lines for a clear view. Residues of NS2B are indicated with an asterisk. The figure was created using PyMOL.[44]

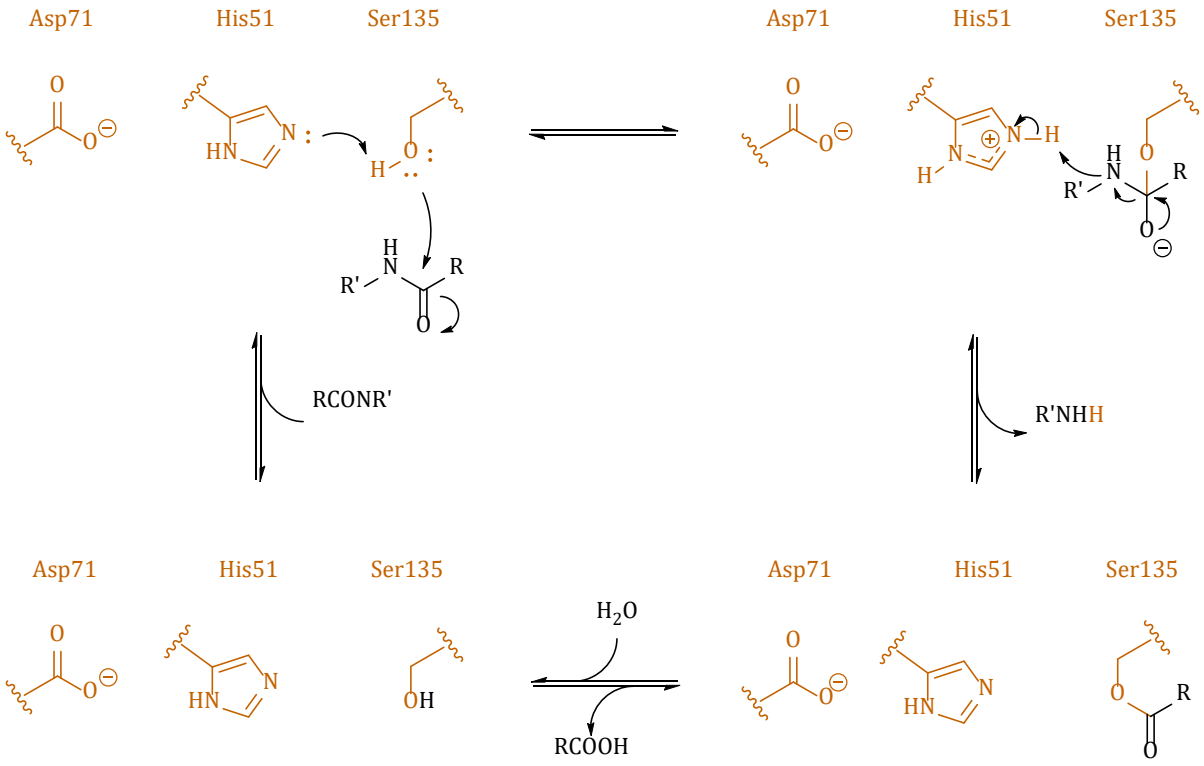


Figure 4. *Proteolytic catalysis mechanism of the ZIKV NS2B/NS3 protease.* The amino acids of the ZIKV NS2B/NS3 are orange-colored, and the substrate is black. Mechanism exemplarily for proteases of the subclan PA(S).[32]

Conformational Flexibility of NS2B/NS3. $NS3_{pro}$ adopts a chymotrypsin-like fold and harbors the His51-Asp75-Ser135 catalytic triad,[45] while the NS2B cofactor is highly flexible and can adopt at least two distinct conformations. In the catalytically active *closed* conformation, NS2B is wrapped around NS3, forming a *β*-hairpin that contributes to substrate recognition by partly forming the S2 and S3 binding sub-pockets (Figure 3).[41,43,46,47] In the inactive *open* conformation, NS2B is largely disordered but stays bound to NS3 although there is an autocleavage site between NS2B and NS3 (Figure 2).[34,48–50] Both conformations can be found in equilibrium in solution with the relative abundance highly influenced by the type of expression construct, buffer pH, and ionic strength.[51–53] Complexed with inhibitors or substrates, exclusively co-crystals of the *closed* conformation were obtained to date (Figure 5A). The open conformation was obtained in the apo-state (Figure 5B).

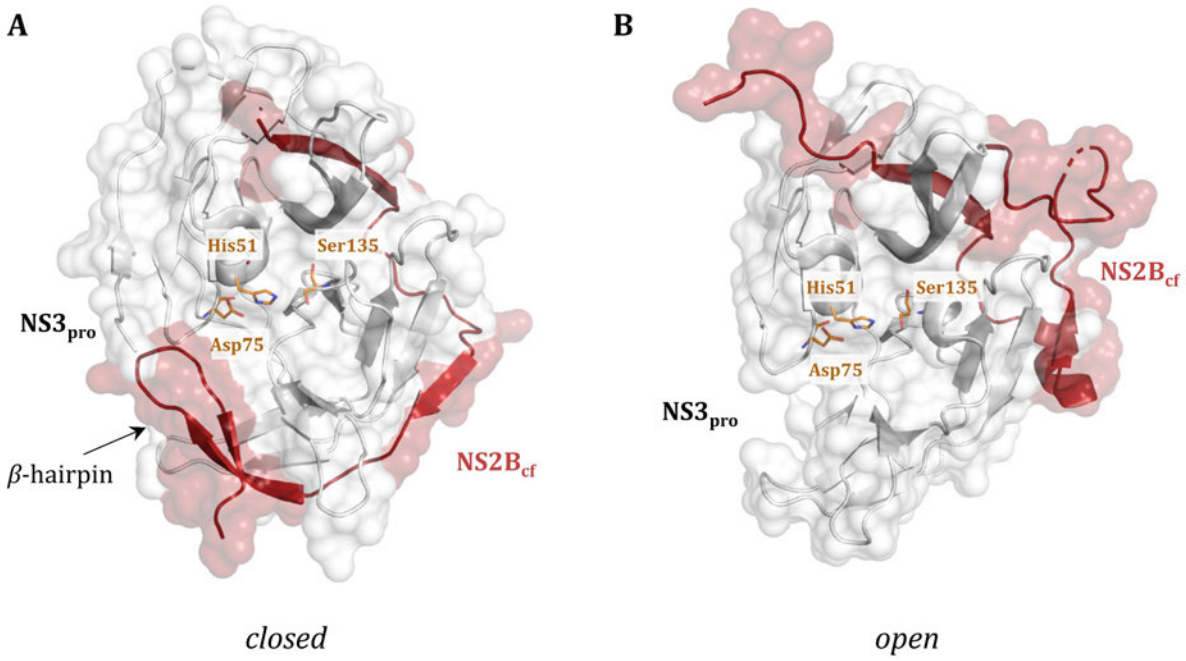


Figure 5. *The crystallographically determined conformations of NS2B/NS3.* (**A**) Structure of the *closed* conformation of DENV3 NS2B/NS3 (PDB-ID: 3U1I).[54] (**B**) Structure of the *open* conformation of DENV2 NS2B/NS3 (PDB-ID: 2FOM).[34] The cartoon and the transparent surface of the NS2B cofactor are colored red, and NS3 is white. Amino acids of the catalytic triad are orange. The figure was created using PyMOL.[44]

Obstacles in Inhibitor development and optimization. Although several active site-directed inhibitors with considerably high potencies were reported, developing inhibitors of NS2B/NS3 as drug candidates is considered challenging.[31,55] This has several reasons: (i) The NS2B/NS3 proteases share a shallow water exposed catalytic site.[34–36] Hence, bound inhibitors retain large solvent-exposed areas and cannot tightly fit into deep grooves. (ii) Since the proteases prefer basic P1 and P2 residues, potent inhibitors often feature multibasic scaffolds.[31,36,56–58] Therefore, membrane permeabilities are limited, and most inhibitors lack potent *in vivo* activity.[59–62] (iii) The conformational flexibility with two crystallized conformations and the high flexibility of the open conformation complicate rational drug design. All active site-directed inhibitors were proven to bind into the closed conformation. However, whether the closed state is yielded by conformational selection or induced-fit mechanisms remains elusive.[34,54,63,64] In this regard, a high entropic penalty could impede the transition of the flexible open conformation to the more rigid closed conformation.[54] (iv) The development of allosteric inhibitors is hampered by the lack of structural information on the inhibitor-bound state. However, studies with disulfide traps and a split-luciferase assay suggest that they stabilize a conformation other than the closed one.[52,53,65]

1.1.2. The Human Matriptase MT-SP1

Classification of Membrane-Anchored Serine Proteases. The human matriptase, also known as transmembrane protease serine 14 (TMPRSS14), suppressor of tumorgenicity 14 (ST14), prostamin and membrane-type serine proteinase 1 (MT-SP1), is one of currently 20 known human membrane-anchored serine proteases (MASP) and belongs to the type II transmembrane serine proteases (TTSP).[66–68] The TTSPs share a characteristically conserved N-terminal signal anchor that acts as a transmembrane domain. Other membrane-anchored serine proteases harbor the transmembrane region at their C-terminus (type I transmembrane seine proteases) or are bound to the membrane by a C-terminal glycosyl-phosphatidylinositol (GPI) linkage.[69] A variable stem region connects the transmembrane domain to the C-terminal trypsin-like serine protease region (Figure 6).[67]

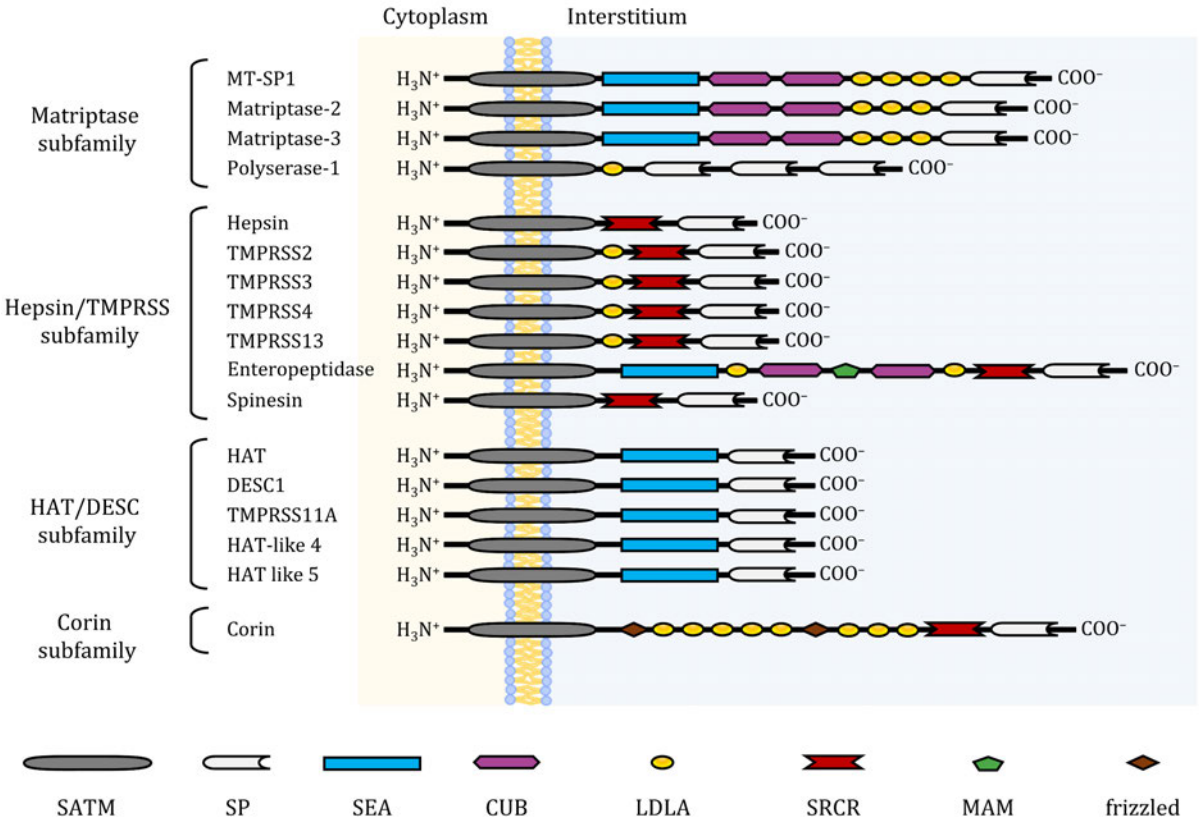


Figure 6. *Classification of TTSPs and the arrangement of the contained domains.* All proteases of the TTSP family share an N-terminal signal anchor transmembrane domain (SATM, gray) and a C-terminal serine protease domain (SP, white) that are connected by variable stem regions formed by several domains: sea urchin sperm protein, enteropeptidase and agrin (SEA, blue), Cls/Clr, urchin embryonic growth factor, bone morphogenetic protein-1 (CUB, purple); low-density lipoprotein receptor class A (LDLA, yellow), group A scavenger receptor cysteine-rich (SRCR, red), meprin, A5 antigen, receptor protein phosphatase (MAM, green) and frizzled receptors (frizzled, brown). The figure is modified following Martin & List.[69]

Their location at the cell surface enables MASPs to interact with extracellular soluble proteins, matrix components, and surface proteins of other cells. TTSPs often feature receptor domains like LDLA, scavenger receptor cysteine-rich (SRCR), or meprin, A5 antigen, receptor protein phosphatase (MAM). Therefore, TTSPs are expected to fulfill functions in cell signaling pathways *via* their intracellular N-terminal signal anchor. Because of these mechanisms, several TTSPs are associated with cancer.[69] Based on the arrangement of the domains, the phylogeny of the serine protease domain, and the chromosomal localization, the TTSPs can be divided into four subfamilies: (1) the matriptase subfamily, (2) the hepsin/TMPRSS subfamily, (3) the human airway trypsin-like (HAT)/differentially expressed in squamous cell carcinoma (DESC) subfamily and (4) the Corin subfamily. Noteworthy, although not belonging to the hepsin/TMPRSS subfamily, the human matriptase is still referred to as TMPRSS14 (in this work denoted as MT-SP1). The MEROPS database assigns the MT-SP1 protease domain to the clan PA, the subclan PA(S), the family S1, subfamily A, and the identifier S01.302.[32]

Clinical Relevance of MT-SP1. As a protease highly expressed in breast cancer and involved in matrix degradation, MT-SP1 is associated with tumor progression and metastasis.[70,71] Also, several substrates of MT-SP1 were found to promote tumorigeneses, such as hepatocyte growth factor (HGF) and urokinase-type plasminogen activator (uPA).[72,73] Both proteins influence the degradation of extracellular matrices, thereby promoting tissue remodeling, cancer progression, and metastasis.[48] Thus, inhibitors of MT-SP1 may represent promising anticancer drugs.[72–75] In addition to these roles, MT-SP1 is essential for epithelial barrier formation, hair follicle development, and thymic functions.[76] Due to its role as an activator of metalloproteinases, MT-SP1 could also be targeted to modify osteoarthritis.[77] MT-SP1 is expressed in several tissues like the lungs, esophagus, stomach, small intestines, colon and intestines, kidneys, vagina, and prostate.[78] Concerning the recent pandemic of severe acute respiratory syndrome corona virus 2 (SARS-CoV-2), MT-SP1 could be involved in priming the SARS-CoV-2 spike (S) protein to promote cell entry. Since SARS-CoV-2 infections are promoted by entering epithelial cells, his hypothesis is supported by the crucial roles of MT-SP1 for epithelial barrier function and its high expression in the lungs and the gastrointestinal tract.[79,80] Through the MT-SP1 mediated activation of the tissue factor-triggered clotting cascade, MT-SP1 could also play roles in thromboembolic complications that occasionally occur during severe SARS-CoV-2 infections.[81–83]

Structural Biology and Catalytic Reaction of MT-SP1. MT-SP1 is expressed as a zymogen with low intrinsic activity. In a *trans*-catalyzed mechanism, it is activated by proteolytic cleavage between Arg15 and Val16 to generate the enzymatic active protease domain that remains bound to the stem region *via* a disulfide bridge between Cys5 and Cys122.[84] The protease domain forms two β-barrels with the catalytic triad embedded in between them. (Figure 7A). As a member of the PA clan, the trypsin-like MT-SP1 catalyzes proteolysis by the catalytic triad His57-Asp102-Ser195 in a mechanism similar to NS2B/NS3 (Figure 4). The substrate specificity for basic amino acids in the P1 position is determined by the negatively charged Asp189 at the bottom of the S1 subsite (Figure 7).[85] The preference for smaller amino acids, such as Gly and Ser, in P2 can be explained by Phe99, sterically hindering larger substrate residues.[86] Interestingly, Phe and Thr could also be accommodated by the S2 pocket. A crystal structure revealed the rotational flexibility of the benzyl ring of Phe99, thereby adjusting the pocket size to the P2 residue.[85,87] The fused S3/S4 subsite is shallow and forms a binding area for larger hydrophobic residues interacting with Trp215. Asp96 and the backbone oxygen of Phe97 create a negatively charged surface suitable for coordinating a basic residue in P3 or P4, but not for both positions simultaneously.[85,88–90]

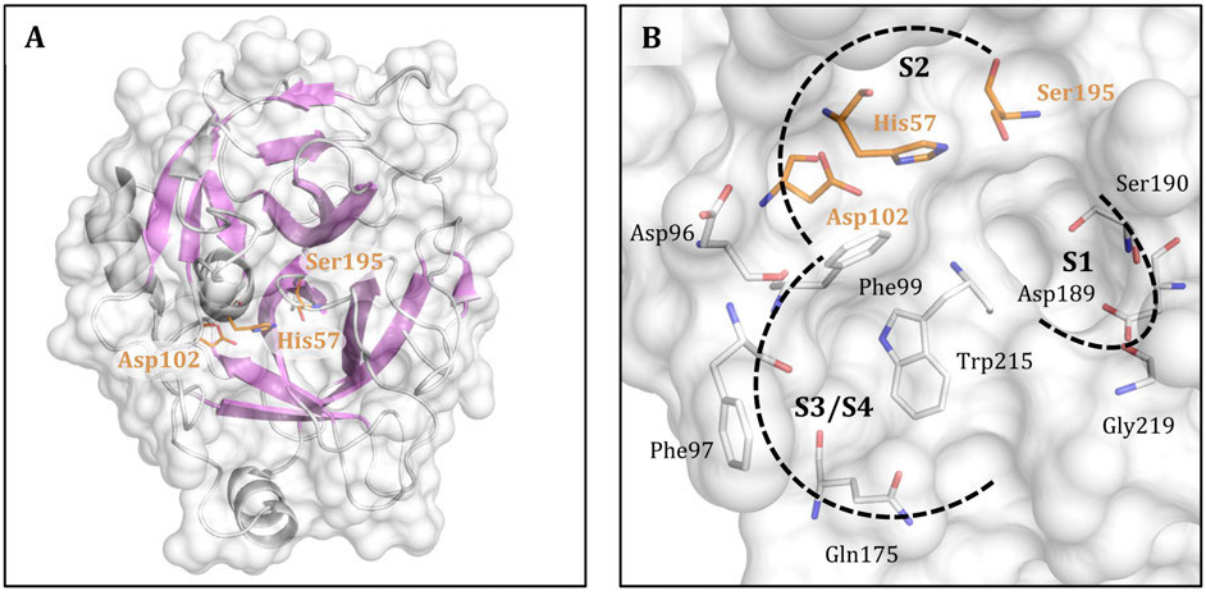


Figure 7. *Structure and Binding site of the MT-SP1.* (**A**) X-ray structure of MT-SP1 (PDB-ID: 4O9V)[91] with an orange-colored catalytic triad (His57, Asp102, Ser195) in between two β-barrels (β-sheets are colored violet). (**B**) The binding site of MT-SP1. For a clear view, only those amino acids forming the binding pocket (sticks with white carbon atoms) and the catalytic triad (orange sticks, red labels) are shown as sticks. The binding sub-pockets are indicated according to Schechter&Berger[43] with black dashed lines (PDB-ID: 4O9V)[91]. The figure was created using PyMOL.[44]

1.1.3. *Staphylococcus aureus* Sortase A

Classification and Biological Function of Sortase A. Based on their amino acid sequence, sortases can be divided into the six classes A–F.[92,93] Hereof, the *Staphylococcus aureus* (*S. aureus)* sortase A (SrtA) is the role model for the class A sortases and is the best-characterized specimen that has been studied for over 20 years.[94] *S. aureus* SrtA anchors several microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) to the outer surface of the bacterial cell wall.[94] These MSCRAMMs are associated with host cell adherence and immune evasion functions.[95,96] Given these vital roles, it is comprehensive that genes, encoding sortases, have been identified for ~1000 Gram-positive bacteria.[97] The MEROPS database classifies the *S. aureus* SrtA (hereafter denoted simply as SrtA) as a member of the clan CL, family C60, subfamily A, and the identifier C60.001.[32] The clan CL consists of cysteine proteases of the sortase family (C60) and other bacterial transpeptidases (family C82).[98]

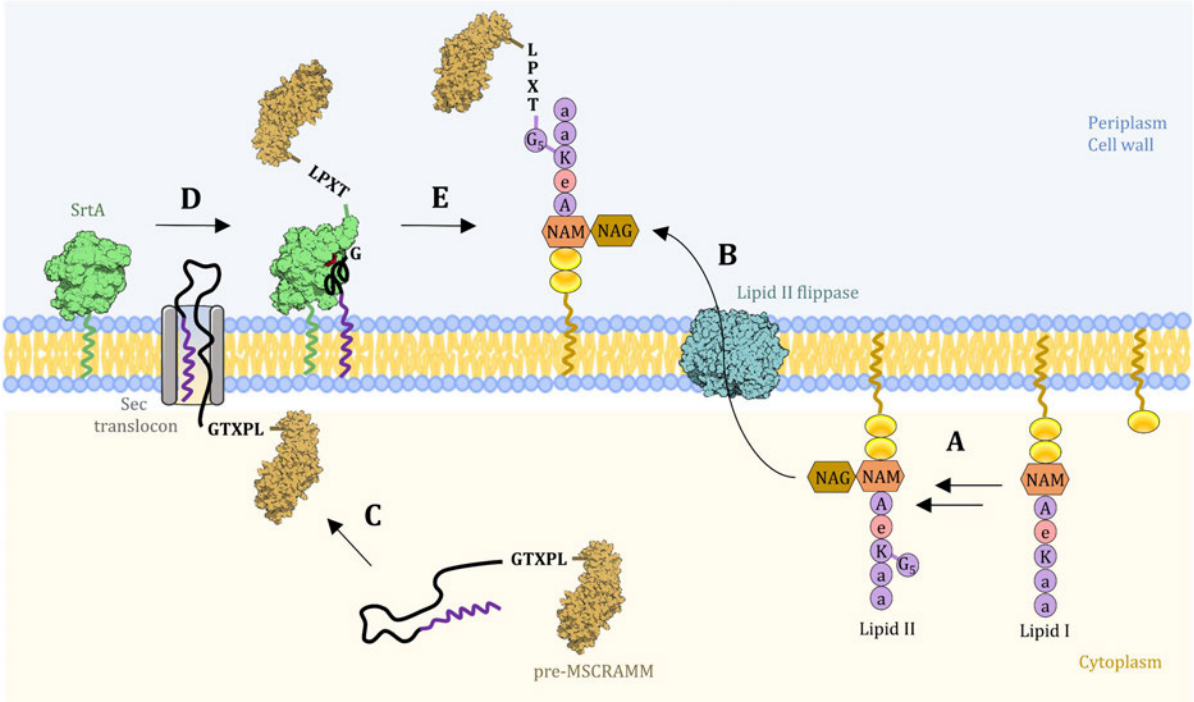


Figure 8. *Schematic cell wall sorting reaction of an MSCRAMM by SrtA transpeptidase activity*. (**A**) The ubiquitous building block lipid II is synthesized from lipid I in a multi-step process. (**B**) The transmembrane-anchored lipid II is transported by lipid II flippase (PDB-ID: 6NC6)[99] to the cell surface. (**C**) Surface protein precursor of microbial surface components recognizing adhesive matrix molecules (pre-MSCRAMM, beige; exemplarily clumping factor A, PDB-ID: 1N67)[100] are synthesized in the bacterial cell and harbor a C-terminal secretory pathway (Sec) signal peptide (bright purple) and the sorting signal LPXTG as part of the cell wall sorting signal (CWSS). (**D**) The sec translocon retains the pre-MSCRAMM mainly in the cytosol until it is recognized by SrtA (green, PDB-ID apo-protein: 1IJA,[101] PDB-ID substrate bound: 2KID)[102] as a substrate. The covalent intermediate of the transpeptidase reaction is formed as a thioester of SrtA Cys184 to the carbonyl functionality of the LPXT threonine. (**E**) The pentaglycine of lipid II attacks the thioester to yield the membrane-anchored lipid II-MSCRAMM conjugate, which is further incorporated into the peptidoglycan cell wall. This figure was created using QuteMol to visualize the protein surfaces.[103]

As members of the clan CL, all sortases harbor an Arg-His-Cys catalytic triad.[104,105] SrtA is a type II membrane transpeptidase that recognizes precursor surface proteins (pre-MSCRAMMs) by their sorting signal LPXTG. Subsequently, a multistep transpeptidase reaction is catalyzed to link these surface proteins to peptidoglycans (Figure 8).[94] The ubiquitous peptidoglycan building block lipid II is the acceptor of SrtA-mediated transpeptidation and is assembled from the bactoprenol-conjugate lipid I in a multistep intracellular process (Figure 8A).[106,107] Lipid II is membrane-bound and transported to the periplasm by lipid II flippase but remains membrane-anchored (Figure 8B).[99] As SrtA substrates, pre-MSCRAMMs feature the C-terminal CWSS consisting of (i) the sorting signal LPXTG, (ii) the hydrophobic stretch, and (iii) a positively charged tail.[95,105,108,109] Furthermore, they harbor an N-terminal signal peptide for Sec translocon-mediated secretion and export to the periplasm (Figure 8C). They remain primarily intracellular until they are recognized as substrates by SrtA. SrtA cleaves the scissile bond of the sorting signal LPXT↓G between Thr and Gly. The thioacyl intermediate (Figure 8D) persists until the newly generated C-terminus is attached to the pentaglycine cross-bridge of the peptidoglycan precursor lipid II (Figure 8E).[94] Such surface proteins linked to peptidoglycans are subsequently incorporated into the cell wall to assemble the extracellular matrix and to fulfill a variety of biological functions, essential *S. aureus* virulence.[110]

Relevance of Biofilm Formation in *S. aureus* Infections. Many Gram-positive bacteria can form three-dimensional communities with an extracellular matrix of linked polymers called biofilm. The extracellular matrix consists mainly of polysaccharides, proteins, and extracellular DNA (eDNA) released by programmed cell death.[111,112] Biofilms can adhere to soft tissues such as epithelia and endothelia in the lung and intestines. Furthermore, even adhesion to plain, hard surfaces, such as teeth, or inanimate surfaces, such as medical implants, is feasible.[113] Biofilms can cause severe medical complications since biofilm-forming bacteria are more resistant to antibiotics and immune reactions than more susceptible planktonic cells of the same species.[112,114] Biofilms of *S. aureus* contain several proteinogenic virulence factors that are MSCRAMMs and substrates of SrtA. They play crucial roles in the biofilm life cycle, such as adhesion to the host matrix, damaging host tissues, providing nutrients, and immune system evasion.[110] Biofilms also pose physical barriers to biocidal drugs and immune cells and allow intracellular bacterial communication.[115–117] Colonialization is enabled through the adhesion of the bacterial cells to host matrices, mediated by adhesins like collagen adhesin (Can) and fibronectin-binding proteins A and B (FnbpA, FnbpB).[110,118,119] Dispersion of single cells is also reduced by clumping factors A and B (ClfA and ClfB).[120] They mediate the attachment to plastic surfaces and platelet cells, which can also be incorporated into biofilms. By fibrinogen binding, these factors also activate blood clotting.[121] Among the exotoxins released by *S. aureus* is α-hemolysin, leading to apoptosis *via* pore formation in host cell membranes.[122] The lysed host cells release hemoglobin, which is part of the

nutrition of *S. aureus* and can be acquired by the iron-regulated surface determinant (Isd) system. The SrtA substrates IsdA and IsdB are relevant for hemoglobin binding, separation of heme from the protein backbone, and heme transport.[123,124] One protein facilitating evasion of the immune system is protein A (SpA), which binds immune globulin G (IgG) antibodies, thus allowing to evade phagocytosis.[125] Besides these biological functions, biofilm formation is attributed to higher antibiotic resistance, as antibiotics cannot penetrate the physical barriers posed by biofilms, leading to subtherapeutic local concentrations that drive the development of resistance mechanisms.[126] Furthermore, biofilm maturation involves bacterial cells in dormant states as persister cells with lowered metabolic rates, thereby increasing their tolerance to antibiotic chemotherapy.[126,127] Given all these virulence factors, it is comprehensive that silencing of SrtA in infected immune-deficient mice prevented lethal outcomes.[126] Inhibition or knockout of SrtA leads to deficiency in displaying surface proteins[128] but does not interfere with bacterial growth.[94,129] Thus, inhibition of SrtA is a promising strategy to lower the virulence of *S. aureus*. In combination with traditional antibiotics, anti-virulence drugs like SrtA inhibitors could pave the way to enhance antibiotic efficiency by simultaneously preventing the development of new resistances due to the low selective pressure they impose.[130,131] Additionally, SrtA inhibitors can harm pathogens, thereby sparing the microbiota and overcoming physical resistance by enhancing bacteria exposure to antibiotics and immune cells.[130] The reduction of adhesion also enables the mechanical cleansing of bacteria from tissues.[132] Since SrtA is placed on the outside of the bacterial cell membrane, it is well accessible for anti-virulence agents.[133]

Mechanism of the Transpeptidase Reaction of SrtA. Three amino acids of SrtA are involved in the catalytic transpeptidase reaction cycle, His120, Cys184, and Arg197. The exact catalysis mechanism of the transpeptidase is yet not fully elucidated. To date, the binding mode of lipid II remains elusive. A main controversy in the catalytic mechanism concerns the protonation states of His120 and Cys184.[134,135] Measurements of the pk_a values of the two amino acids revealed the population of the zwitterionic state to be less than 0.1% under native conditions. This finding could further be supported by a geometric analysis of the residues that found the distance of 7 Å between both residues to be too far for ionic interactions.[101,135,136] In papain-like proteases (clan CA), an Arg or Asn residue activates His, the closest ionizable residue to the catalytic Cys. In contrast, the catalytic Cys184 of SrtA resides between His120 and Arg197, the nearest ionizable residue of Cys184.[136] Other Cys proteases that feature such an uncharged catalytic dyad are those of the clan CD (e.g., caspase and legumain). To compensate for the low nucleophilicity of the Cys thiol, they require a prominent oxyanion hole that stabilizes the anionic charge of the transient tetrahedral. Thereby, the energy level of the intermediate is lowered so that the nucleophilicity of the Cys thiol is sufficient to attack the P1 carbonyl carbon.[137–139] The absence of such a pronounced oxyanion hole in SrtA leads to the conclusion that the catalytic residues must lead the nucleophilic

attack in their imidazolium-thiolate form.[140] Thus, the active state of the enzyme constitutes a lowly populated form, leading to a low catalysis rate of such reversely protonated enzymes.[141,142] The currently accepted mechanism proceeds through a ping-pong catalysis mechanism (Figure 9).[134,143,144]

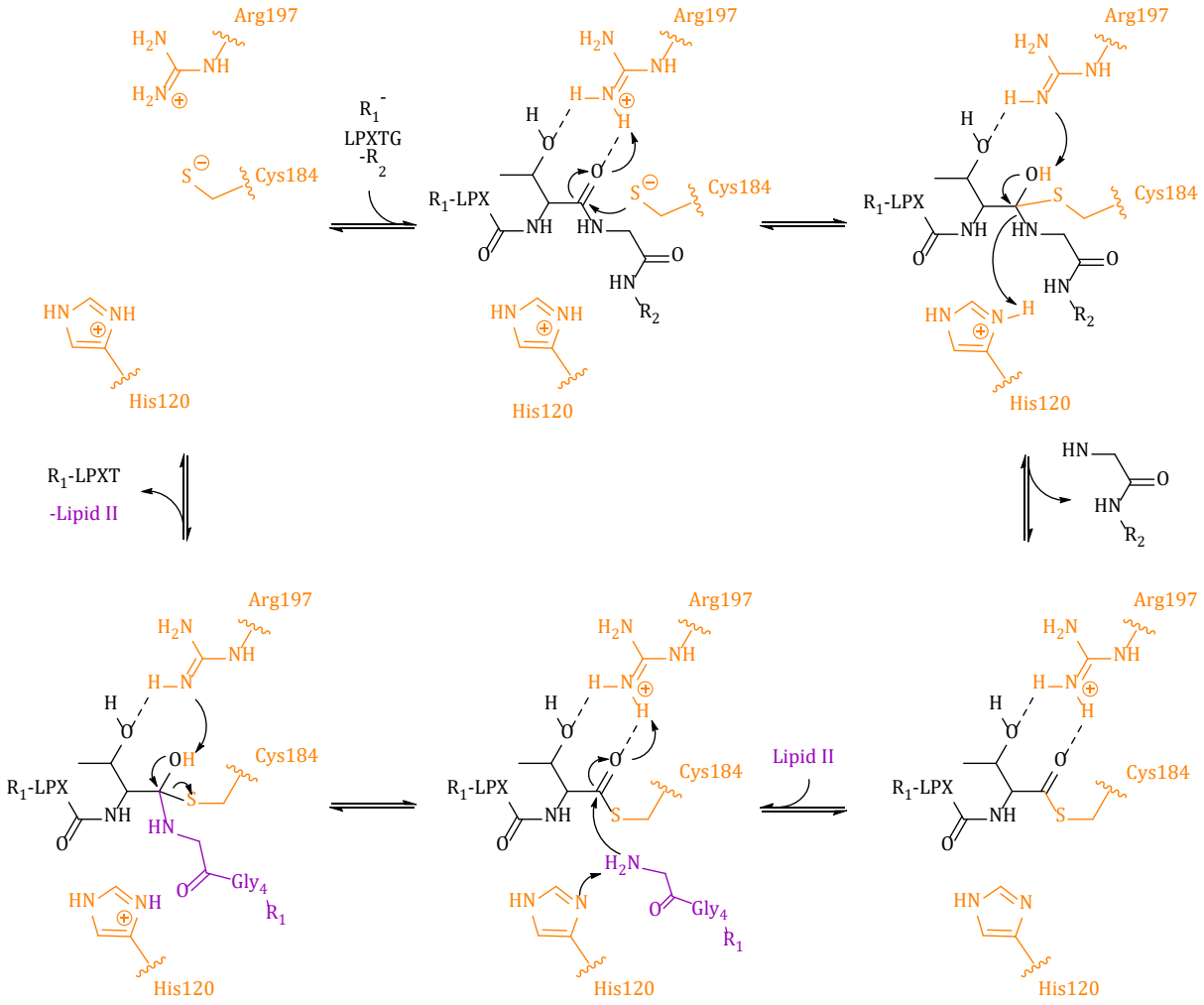


Figure 9. *Mechanism of the transpeptidase mechanism of SrtA*. In the catalytic transpeptidase reaction cycle, the substrate's LPXTG motive (black) is attacked by the nucleophile Cys184 thiolate (orange). The thiohemiketal intermediate decays to the thioacylated Cys184 and the shed Gly residue. After the acceptor for LPXT, lipid II (purple), attacks the thioester, the LPXT-lipid II conjugate is released. SrtA is regenerated to perform a new catalytic cycle. The figure was created and modified following Jacobitz *et al.*[109]

Arg197 acts as a donor and acceptor of protons during the catalysis, stabilizing the oxyanion intermediate.[105] The orientation of Arg197 is further stabilized by interacting with the P1 Thr side chain.[145,146] The first step in the catalytic transpeptidase reaction is recognizing the sorting signal by SrtA and binding the LPXTG motif into the active site. Subsequently, the statistically low-populated Cys184 thiolate performs a nucleophilic attack at the P1 Thr carbonyl carbon. A thiohemiketal is formed as a tetrahedral transition state.[109,144] The negative charge is stabilized by

the oxyanion hole consisting of the Thr183 side chain and Arg197, which can provide a proton.[145] The scissile amide bond between Thr and Gly is cleaved when a proton from the His120 imidazolium is transferred, resulting in an uncharged thioacyl intermediate.[94,147] The resulting thioester persists until a nucleophilic attack of the lipid II pentaglycine's terminal amine facilitates diacylation. A second thiohemiketal is formed, subsequently releasing the LPXT-lipid II conjugate and the regenerated active site of SrtA.[144,146]

Structural Biology of SrtA. Besides the 18 kDa catalytic domain, the primary sequence of SrtA exhibits an N-terminal signal peptide to facilitate the Sec translocon mediated externalization and a hydrophobic transmembrane domain, anchoring SrtA to the cell surface.[147] The catalytic domain (amino acids 60–206) forms an eight-stranded *β*-barrel. The rigid *β*-sheet formed by *β*4, *β*7, and *β*8 harbors the His120-Cys184-Arg197 catalytic triad and is surrounded by the *β*3/*β*4 loop, the *β*6/*β*7 loop, and the *β*7/*β*8 loop (Figure 10A).[136] For catalytic activity, SrtA requires the binding of a calcium ion cofactor, whose binding site is formed by the acidic amino acids Glu105, Glu108, Asp112, and Glu171 of the *β*3/*β*4 and the *β*6/*β*7 loops. The *disordered loop* *β*6/*β*7 gets fixed *via* Glu171 coordination to Ca^{2+} and adopts a partly closed conformation (Figure 10B).[101] The sorting signal displaces the *β*7/*β*8 loop, which adopts a partly open conformation in the complexed state and is therefore referred to as the *dynamic loop* (Figure 10C).[148,149] In the complex structure with the Ca^{2+} cofactor and the sorting signal, the disordered *β*6/*β*7 loop adopts a fully closed conformation and forms a short 3_{10}-helix.[150] This helix interacts with the sorting signal's P3 Pro and the P4 Leu, residing on a platform formed by the C-termini of the *β*4 and *β*7 strands. The binding pocket is formed by the *β*4/*β*5 loop and the N-terminal region of the long *β*6/*β*7 loop. The dynamic *β*7/*β*8 loop opens fully (Figure 10D).[102]

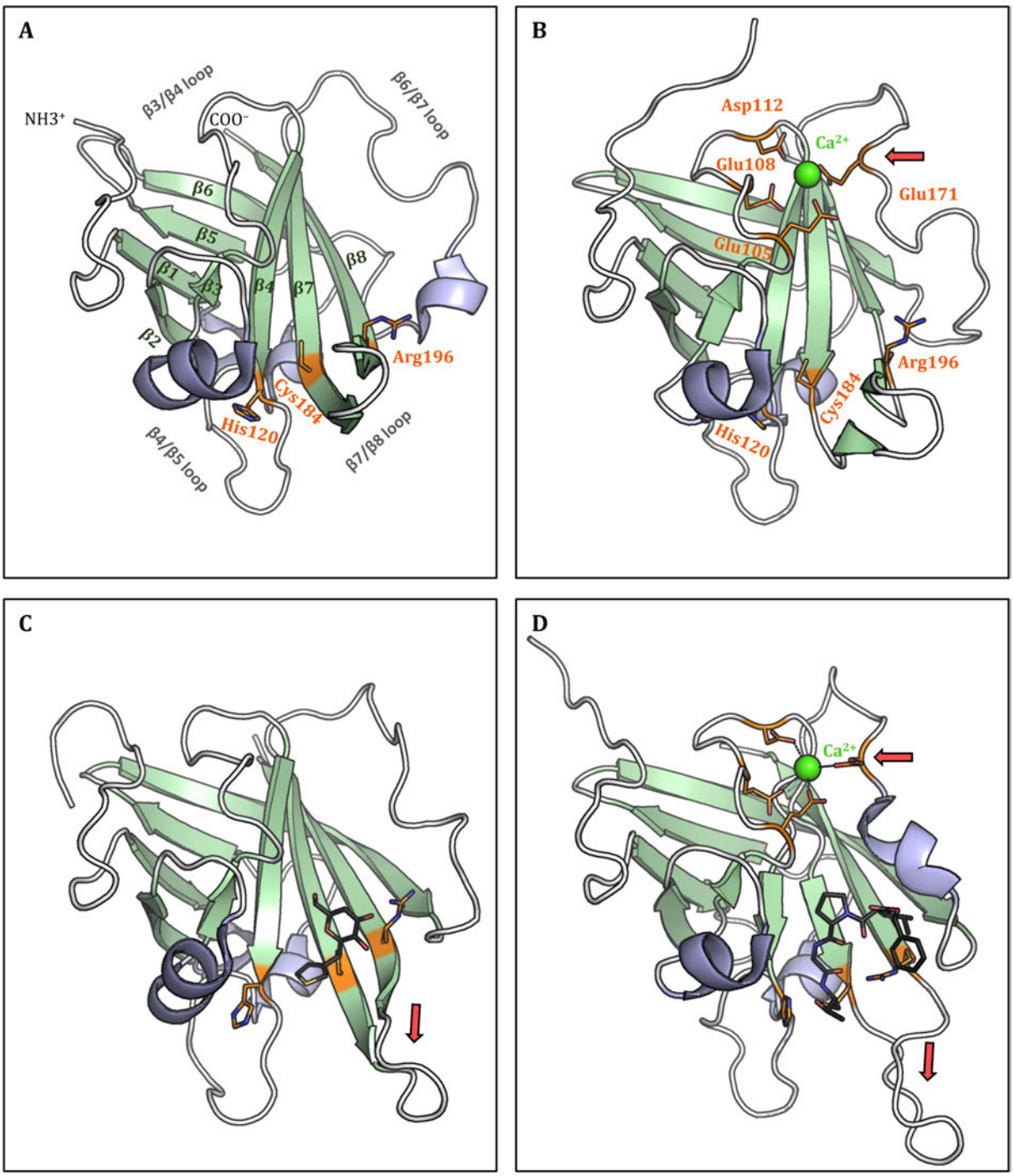


Figure 10. *Structural rearrangement of SrtA upon cofactor and substrate binding*. (**A**) The β-strands β1–β8 (green) form the bottom and top of the eight-stranded β-barrel surrounded by flexible loops (white) connecting the β-strands. The helices are colored blue. The catalytic triad (orange sticks) sits at the top of the β-sheet formed by β4, β7, and β8. Hence, the active site is dominated by the flexible loop regions of the β3/β4, the β6/β7, and the β7/β8 loop (PDB-ID: 1T2P).[136] (**B**) In the Ca^{2+} bound state (PDB-ID: 1IJA; with manually added calcium ion and rotated Glu105 and Glu171),[101] a calcium ion (green) is coordinated by four acidic residues of the β3/β4 and the β6/β7 loops (Glu105, Glu108, Asp112, and Glu171, depicted as orange sticks). The disordered β6/β7 loop is fixed by Ca^{2+} coordination and adopts a partly closed conformation (red arrow). (**C**) In the sorting signal bound state (PDB-ID: 6R1V; here, a small covalent inhibitor shown as black carbon atoms)[148], the dynamic β7/β8 loop gets displaced and adopts a more open conformation. (**D**) In the Ca^{2+} and sorting signal bound state (PDB-ID: 2KID),[102] the sorting signal-mimetic ligand (black carbon atoms) sits on an interface, formed by the β4/β5 loop, the β6/β7 loop and the N-terminal region of the long β7/β8 loops shape the active site. Figure created using PyMOL.[44]

Allosteric Activation of SrtA through Calcium Ion Binding. In crystal structures of the apo form of SrtA, the $\beta7/\beta8$ loop was only poorly resolved and showed substantially different conformations between asymmetric subunits, suggesting a high degree of flexibility (Figure 11A).[136] Hence, it can be concluded that this loop is an intrinsically disordered region (IDR).[151] IDR were found to adopt well-defined structures upon ligand binding, are susceptible to allosteric mechanisms, and can hence promote activation *via* dock and coalesce mechanisms.[152–154] Differential scanning calorimetry (DSC) with and without Ca^{2+} found the melting temperature (T_m) to be increased by 3.4 K with Ca^{2+}, indicating a higher degree of order in the SrtA structure.[149] Nuclear magnetic resonance (NMR) studies confirmed this by revealing that substrate and Ca^{2+} binding lead to rigidified flexible loops.[149,151,155] Upon binding, Ca^{2+} allosterically activates SrtA by fixating the C-terminal part of the disordered loop. In the respective NMR solution structure, the disordered loop's N-terminus was the only part that displayed lower order parameters.[101] Coordination of Glu171 leaves the $\beta3/\beta4$ loop mainly rigid. However, a fundamental conformational change can be observed for the disordered $\beta6/\beta7$ loop that partly closes to preorganize for binding of the sorting signal (Figure 11B).[101,109,151,155] By interactions of the newly formed 3_{10}-helix and the P3 Pro and P4 Leu residues, the affinity of the substrate is increased.[149,156] Conversely, when the sorting signal binds to the apo structure, the dynamic $\beta7/\beta8$ loop gets displaced by the ligand and stretches to open the binding pocket in an induced-fit mechanism (Figure 11C).[102,155] The complexion of the Ca^{2+} cofactor and the sorting signal promotes a disorder-to-order-transition of the disordered loop. The initial IDR adopts a highly rigidized conformation *via* a conformational selection mechanism (Figure 11D).[102,149,151,155] The lipid II polyglycine tail recognition is poorly understood, as neither X-ray nor NMR structures are available for SrtA. However, the interaction of the dynamic loop in LPAT bound SrtA with triglycine could be confirmed with NMR shift mapping.[102] Hence, in analogy to *S. aureus* SrtB, the opening of the dynamic loop is believed to form a pocket for subsequent binding of the pentaglycine residue of lipid II, although the $\beta7/\beta8$ loops of SrtA and SrtB are highly different (Figure 11E).[157]

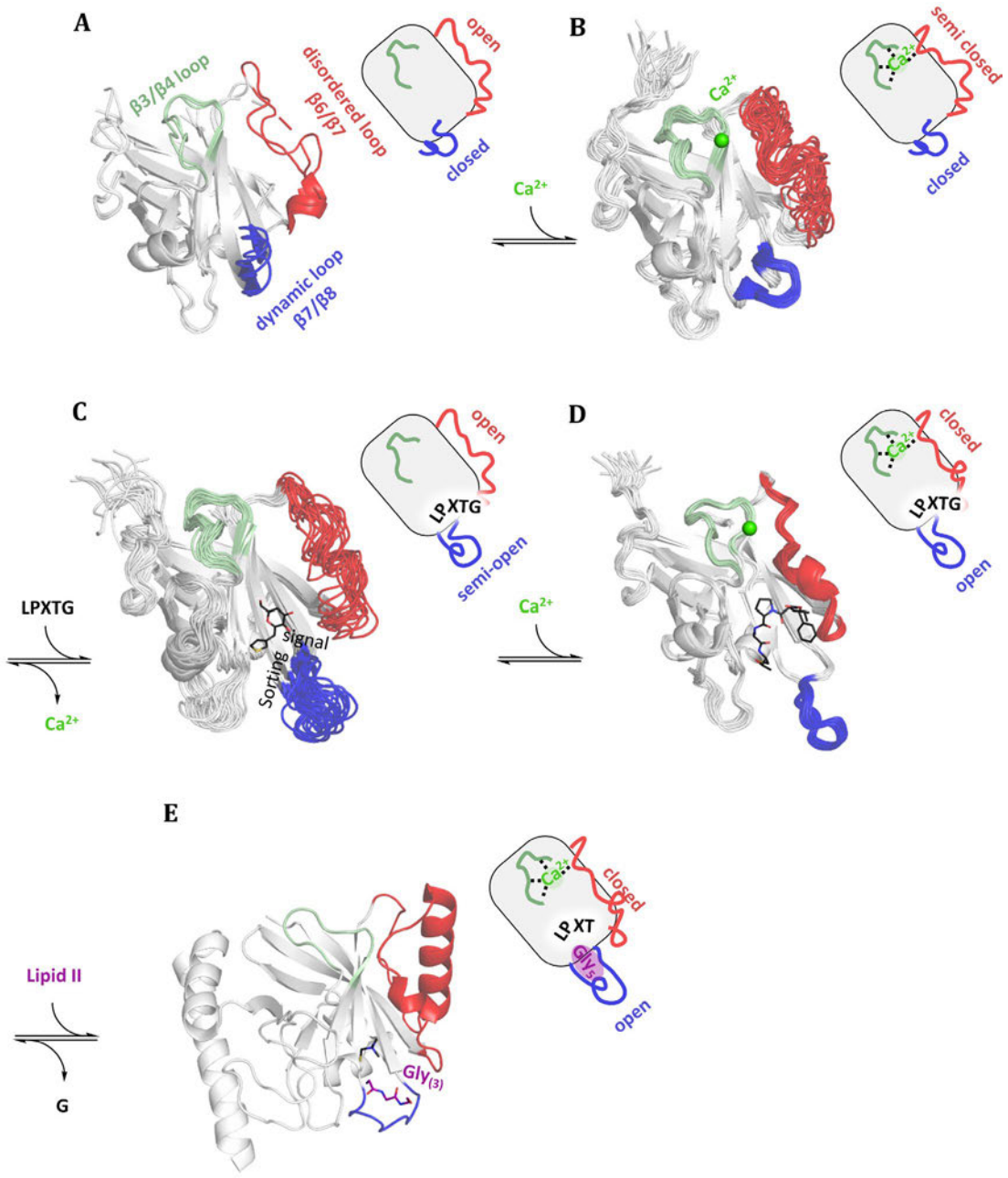


Figure 11. *Cooperativity of Ca^{2+}, sorting signal, and lipid II in ligand binding of SrtA.* Left: X-ray (**A**, **E**) or NMR ensembles (**B**–**D**) of SrtA (white, **A**–**D**) or SrtB (white, **E**) in different complexed states. Right: The respective schematic presentation of the disordered (red) and the dynamic loop's (blue) conformations. (**A**) The overlay of the three subunits of the asymmetric unit of the apo SrtA reveals an open conformation of the disordered and a closed conformation of the dynamic loop in the absence of ligands (PDB-ID: 1T2P).[136] (**B**) Calcium ions (green) coordinate the $\beta 3/\beta 4$ loop (pale green) and the C-terminal part of the disordered loop, leading to a partly closed conformation, pre-organizing it for binding the sorting signal. The N-terminal region stays highly flexible (PDB-ID: 1IJA).[101] (**C**) Binding of the sorting signal or substrate mimetic inhibitor (black) to SrtA partly opens the dynamic loop. In contrast, the disordered loop stays in the open conformation and is still highly flexible as the positive charges of the Ca^{2+}-coordinating amino acids are repulsive (PDB-ID: 6R1V).[148] (**D**) When Ca^{2+} and the sorting signal are bound to SrtA, a disorder-to-order transition of the disordered loop takes place to adopt a rigid closed conformation. The dynamic loop stretches fully to form a binding interface for a second substrate (PDB-ID: 2KID).[102] (**E**) Although this structure has not yet been solved for SrtA, one can assume that the closely related *Staphylococcus aureus* SrtB displays a similar behavior of polyglycine tail binding. Herein, Gly_3 was co-crystallized in the pocket, generated by the opening of the dynamic loop (PDB-ID: 1QXA).[157] The figure was created using PyMOL.[44]

1.2. Projects and Objectives

1.2.1. Project 1: Entropy Optimization in Medicinal Chemistry

History and Theory of Isothermal Titration Calorimetry. Direct measurements of reaction enthalpies were first conducted before the 19th century. In 1783, experiments performed on an ice calorimeter were reported by Lavoisier and Laplace. They wrapped an ice jacket around the reaction cell and collected the water from thawed ice to calculate quantities of released heat. A second ice layer insulated the device. However, it could only be used during cold winters with temperatures around the freezing point.[158] In the first diathermal electric calorimeters, where temperature changes of the devices were monitored, the heat capacity of heated materials and solutions needed to be known for calculating the enthalpies. This was particularly interesting for reactions accompanied by significant changes in heat capacity. Corrections for the heat exchange between the calorimeter and its environment were also mandatory.[159] Endeavors to overcome these issues led to the first isothermal calorimeters, where a Peltier element and a Joule heater were used to control the calorimeter's temperature. A reference cell was introduced to compensate for heat exchange with the environment.[160] However, isothermal calorimeters were limited to reactions in which the equilibrium was reached in the minute timescale since maintaining temperature in the presence of fluctuating heat generation rates proved difficult.[159,160] The first ITC was described in 1968.[159,161] The Peltier element was used to constantly cool the reaction vessel so that the power supply of a variable Joule heater could be monitored to obtain apparent enthalpies (Figure 12).

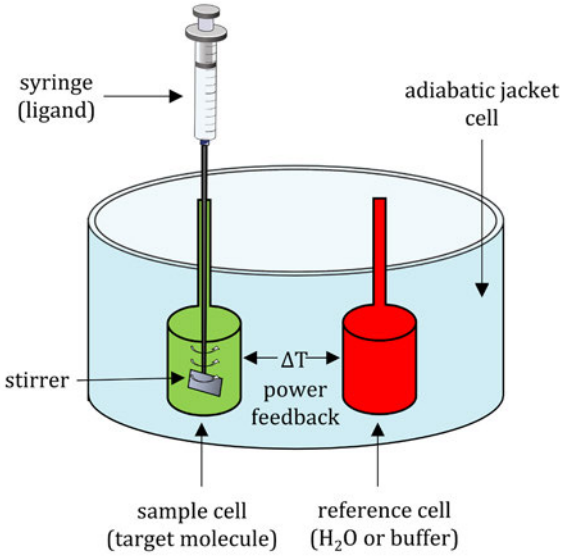


Figure 12. *Schematic representation of an ITC device.* The instrument records the heat energy per time required to maintain the same temperature in the sample to the reference cell. Both cells are in a thermostatically controlled adiabatic jacket cell.

Widespread research of binding interactions using ITC emerged after Ernesto Freire *et al.* introduced the device in their publication entitled 'Isothermal Titration Calorimetry' in 1990.[162] Encouraged by commercially available ITC instruments, this technique was quickly accessible to a broad spectrum of researchers with annually increasing numbers of publications.[163] Ongoing optimization of ITC devices led to several improvements. With higher sensitivity in data acquisition, reaction volumes could be reduced in so-called microcalorimeters down to 0.1–1.4 mL to limit sample consumption.[164,165] A second benefit of reduced sample volumes was the faster homogenization of the solutions, enabling the determination of binding kinetics by analyzing equilibration times.[166] Liquid handling automation allows scheduling several ITC experiments and enables measurements in higher throughput.[167] With user-independent workflows, data collected with automated systems are less susceptible to handling errors and provide results with higher reproducibility.[168,169] A set of well-designed ITC experiments can determine the complete thermodynamic binding profile. The K_d can be derived from the slope at the inflection point of the binding isotherm, which is correlated to the Gibbs energy (ΔG) *via* Equation (1), with R as the universal gas constant and T as the absolute temperature in Kelvin.

$$\Delta G = R \cdot T \cdot \ln (K_d) \tag{1}$$

The binding enthalpy (ΔH) can be obtained from the released heat of the first injections when all added ligand is bound. The temperature-dependent binding entropy (–TΔS) is then calculated with the Gibbs Equation (2).[170]

$$\Delta G = \Delta H - T\Delta S \tag{2}$$

A great benefit of ITC is also the direct determination of the binding stoichiometry (N) from the molar ratio at the inflection point of the isotherm. ITC experiments performed at varying temperatures enable calculating the isobaric binding heat capacity (ΔC_P) *via* Equation (3).

$$\Delta C_P = \left(\frac{\Delta\Delta H}{\Delta T}\right)_P \tag{3}$$

The value of ΔC_P depends on changes in solvation, flexibility, and conformational changes, occurring during the binding event.[171–173] Thus, ΔC_P can be used to assess potential conformational changes. In their absence, a more negative ΔC_P can indicate hydrophobic residues being buried.[174,175] In medicinal chemistry, hit-to-lead optimization campaigns mainly aim at improving binding affinity towards the target and increasing selectivity over off-targets by maintaining the drug-like properties of the ligands. Once the scaffold is determined, conducting structure-activity relationship (SAR) studies is a straightforward way to achieve this. By modifying terminal groups of the ligands comparatively, trends in binding affinity can be combined to yield more promising lead structures. To not rely entirely on serendipity, structural information of the target and the

ligand's binding mode can assist rational drug design in addressing specific residues or binding pockets. Experimentally determined structures can be complemented with computer-aided drug design (CADD) methods.[176–178] During drug development, gains in affinity tend to be achieved by entropic benefits rather than enthalpic interactions. Additional lipophilic interactions are regularly accompanied by increasing bulkiness, higher molar masses, and higher logP values of the compounds. Hence, the ligand metrics like ligand efficiency (LE) and lipophilic ligand efficiency (LLE or LipE) are decreased.[179–181] It is thus more promising to start with relatively small ligands exhibiting a high LLE instead of more affine, larger ligands with a lower LLE. During optimization campaigns, they provide greater latitude for modification while adhering to Lipinski's rule of 5 (molecular weight (MW) < 500 Da; logP < 5),[182] essential to achieve adequate absorption, distribution, metabolism, and elimination (ADME) properties.[183]

In this regard, the first project of this thesis elucidates two strategies for entropic ligand optimization without increasing ligand size. Optimizing ligand preorganization of peptidomimetic ZIKV NS2B/NS3 protease inhibitors *via* macrocyclization was investigated using ITC, CADD, and X-ray. The second strategy is increasing the number of equal binding modes of trivalent inhibitors of MT-SP1 by increasing the ligand symmetry. The thermodynamic impact of these modifications is quantified by ITC and discussed using molecular docking and affinity calculations of the incremental affinities per single binding mode. A versatile set of low-budget teaching experiments was designed to educate medicinal chemists in ligand optimization using advanced ITC experimentation controlled by ligand metrics. The open-accessible web server *ITCcalc* was launched to help analyze the ITC data. Thus, *ITCcalc* maximizes the obtained information by correcting the thermodynamic binding profiles regarding buffer ionization enthalpies, calculating the number of transferred protons (n_{proton}), and ΔC_P. Moreover, *ITCcalc* can calculate commonly used ligand efficiency metrics to support decision-making in hit prioritization and hit-to-lead optimization.

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

(1) **Hammerschmidt, S. J.**; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B / NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm.* **2023**; 356(4):e2200518. https://doi.org/10.1002/ardp.202200518

Own contribution: Enzyme expression and purification, NS2B/NS3 inhibition assays, ITC experiments, writing of the original draft & editing of the manuscript.

Contributions from other authors: Inhibitor synthesis, X-ray crystallography, quantum chemical investigations writing of experimental parts & editing of the manuscript.

(2) **Hammerschmidt, S. J.**[#]; [redacted] Improving binding entropic by ligand symmetry – a case study with human matriptase. *Manuscript submitted to RSC Medicinal Chemistry.*

Own contribution: Establishment of enzyme expression and purification, establishment of matriptase inhibition assays, establishment of ITC experiments, writing of the original draft & editing of the manuscript.

Contributions from other authors: Enzyme expression and purification, matriptase inhibition assays, ITC experiments, synthesis of inhibitors, molecular docking studies, writing parts of the chapter "Synthesis of trivalent Inhibitors", writing of the experimental section "Chemistry" and "Molecular Docking Studies" & editing of the manuscript.

(3) **Hammerschmidt, S. J.**; [redacted] Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc. *Manuscript submitted to J. Chem. Educ.*

Own contribution: ITC experiments and evaluation, writing of the original draft & editing of the manuscript.

Contributions from other authors: Programming of the *ITCcalc* webserver, writing & editing of the manuscript.

1.2.2. Project 2: Chasing the Binding Conformation of NS2B/NS3 Protease Inhibitors

The proteases of DENV and ZIKV feature at least two conformations shown by crystallization experiments (Figure 5, DENV open/close: PDB-ID: 2FOM[34], 3U1I[54]; ZIKV open/close PDB-ID: 5GXJ[184], 5LC0[185]). Whereas the $NS3_{pro}$ domain stays in the chymotrypsinogen fold and the N-terminal region of the $NS2B_{cf}$ stays rigid, the relative position of the C-terminal area of $NS2B_{cf}$ shifts dramatically between both conformations. Thereby, the C-terminal part of $NS2B_{cf}$ forms either a β-hairpin in the closed conformation, contributing to the substrate recognition by partly forming the S2 sub-pocket, or is highly flexible and only poorly resolved in X-ray structures. For ZIKV NS2B/NS3, a "pre-open" (PDB-ID: 5T1V[186]) and a "super-open" conformation (PDB-ID: 5TFN[187]) were reported, indicating the possibility of even more conformations or high flexibility of the open state, instead of being a well-defined conformation (Figure 13).

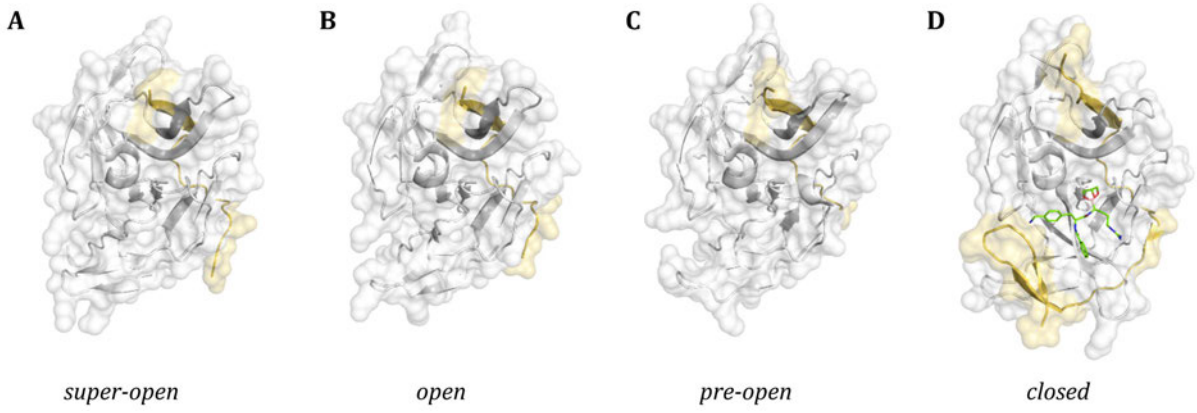


Figure 13. *Conformational flexibility of the open conformation of ZIKV NS2B/NS3.* (**A**–**C**) The *super-open* (**A**, PDB-ID: 5TFN[187]), the *open* (**B**, PDB-ID: 5GXJ[184]), and the *pre-open* conformation (**C**, PDB-ID: 5T1V[186]) differ mainly in the positioning of the NS3 C-terminus. In both the super-open and the open conformation, the NS3 C-terminus points to the substrate binding site. In the *pre-open* conformation, it is faced away and interacts with the strand that connects the two β-barrels of $NS3_{pro}$. (**D**) The closed conformation (PDB-ID: 5LC0[185]) is characterized by the well-defined β-hairpin formed by the C-terminus of $NS2B_{cf}$.

Whereas co-crystals of NS2B/NS3 with active site-directed inhibitors always yielded the closed conformation, crystal structures of the open conformation were exclusively achieved as apo-structures. Despite the widely accepted assumption that allosteric inhibitors of NS2B/NS3 bind and stabilize the open conformation, no direct experimental proof for that claim has been provided to date. Using disulfide traps and a split-luciferase assay, the absence of a signal indicating the closed conformation was interpreted as stabilization of the open conformation.[52,53,65] Recently, the co-crystal structures of DENV NS2B/NS3 in the open conformation in complex with a set of allosteric inhibitors **1**–**3** (PDB-ID: 6MO0, 6MO1, and 6MO2) were published.[188] However, after some controversy, it turned out that probably not an allosteric inhibitor but rather the C-terminus of NS3 was placed in the allosteric pocket (Figure 14).[31,64,189]

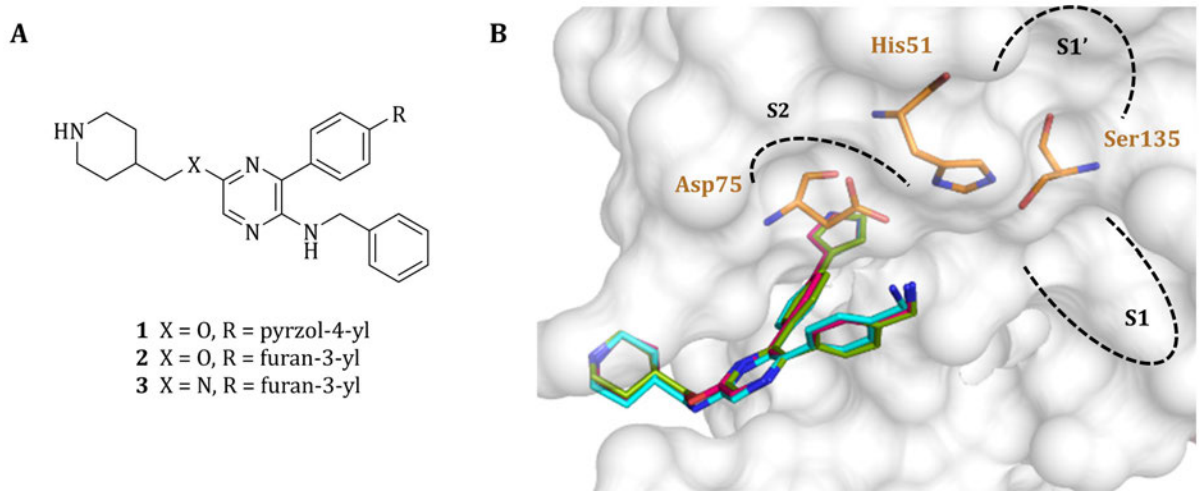


Figure 14. *Non-competitive inhibitors **1**–**3** of DENV and ZIKV NS2B/NS3 proteases.* (**A**) 2D-Structures of the non-competitive inhibitors **1**–**3**. (**B**) Postulated binding modes of inhibitors **1**–**3** (cpd **1**: magenta sticks, cpd **2**: green sticks, cpd **3**: cyan sticks; PDB-IDs: 6MO2, 6MO0, and 6MO1, respectively)[188] in the allosteric pocket of DENV2 NS2B/NS3 (red surface/white surface). For a clear view, only the surface of 6MO0 is shown. Figure created using PyMOL.[44]

In NMR studies, only competitive inhibitors have been investigated so far, demonstrating the closed conformation to be stabilized.[190–194] Hence, neither is there proof for the conformation stabilized by allosteric inhibitors, nor is their dynamic effect on the ZIKV and DENV protease's conformational equilibrium investigated in detail.[190] Therefore, this project is dedicated to determining the allosteric binding pocket and the conformation that is stabilized by the binding of the allosteric inhibitors developed in the group of [redacted].[195,196] To this means, several biophysical methods such as Electron paramagnetic resonance (EPR), microscale thermophoresis (MST), nano surface plasmon resonance (nanoSPR), ITC, SwitchSENSE and X-ray crystallography were used that failed to or did not yet lead to successfully determining the binding mode of allosteric inhibitors.

Altogether, several challenging and time-consuming techniques proved unsuitable for detecting the binding behavior of allosteric NS2B/NS3 inhibitors. Nevertheless, binding-site verification of the allosteric inhibitors was successful. Additionally, different biophysical methods could be employed to observe conformational transitions of NS2B/NS3.

The following unpublished data, the publication, and the submitted manuscript present the results obtained in this project:

Elucidating the mode of action of allosteric Dengue- and Zika NS2B/NS3 Protease Inhibitors.

(Results and discussion of unpublished data obtained in *Project 2: Chasing the Binding Conformation of NS2B/NS3 Protease Inhibitors* not included in the publication and manuscript.)

Contributions of others: The contributions of all collaborateurs to this unpublished data is stated in at the respective section.

(4) Millies, B.; Von Hammerstein, F.; Gellert, A.; **Hammerschmidt, S.**; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; Kersten, C.; Kiefer, W.; Bodem, J.; Hildt, E.; Windbergs, M.; Hellmich, U. A.; Schirmeister, T. 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

Own contribution: Enzyme expression and purification, site-directed mutagenesis of DENV2 NS2B/NS3 T122C, fluorometric enzyme activity assays, maleimide coupling, enzyme inhibition assays, writing of experimental parts & editing of the manuscript.

Contributions from other authors: Enzyme expression and purification, site-directed mutagenesis, fluorometric enzyme activity assays, enzyme inhibition assays, Inhibitor synthesis, docking studies, molecular dynamics studies, circular dichroism spectroscopy, cell permeability studies, metabolism studies, antiviral activity assays, writing of the original draft & editing of the manuscript.

(5) [redacted] The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR. *Manuscript submitted to Eur. J. Med. Chem.*

Own contribution: Enzyme expression and purification, site-directed mutagenesis, ^{19}F-Labeling of the proteases, ^{19}F-NMR experiments, writing parts of the original draft & editing of the manuscript.

Contributions from other authors: smFRET labeling and experiments, DSF, nanoDSF, CD-spectroscopy, fluorometric enzyme activity assays, ^{19}F-NMR experiments, writing of the original draft & editing of the manuscript.

1.2.3. Project 3: Zika Virus NS2B/NS3 Autocleavage Behavior

Besides the proteolytic processing of other viral proteins in the flaviviral precursor polyprotein, there is also an autocleavage site located between flaviviral NS2B and NS3 (Figure 15A).[26] This autocleavage site is exciting since it is highly conserved among flaviviruses.[197] The fact that even after autocleavage, NS3 stays tightly connected with NS2B raises questions about the biological relevance of this site. For early biochemical research of this protease, a shortened $NS2B_{cf}/NS3_{pro}$ construct linked with a flexible glycine-rich linker ($Gly_4SerGly_4$) was shown to retain catalytic activity (Figure 15B).[198]

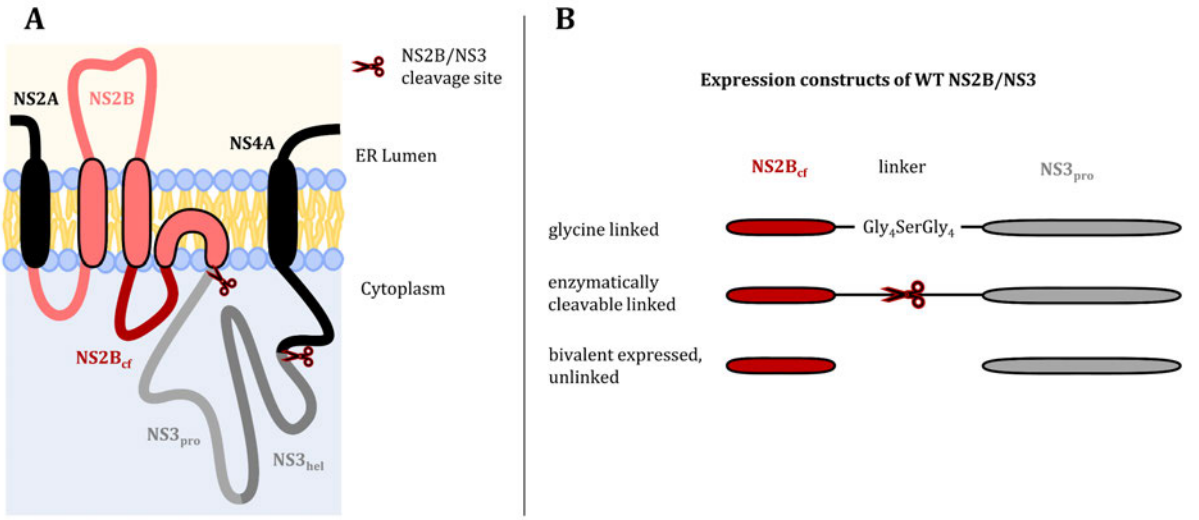


Figure 15. *Expression constructs of flaviviral NS2B/NS3.* (**A**) Schematic depiction of NS2B (red) and NS3 (gray) in the flaviviral precursor polypeptide. The autocleavage site is indicated with a red scissor. The Graphic was designed following Nitsche *et al.*[26] (**B**) Architecture of the three expression constructs of NS2B/NS3 with either a glycine-rich linker, an enzymatically cleavable linker resembling the native cleavage site, or the bivalent expressed, unlinked construct. as designed earlier.[41,198,199]

In contrast to bivalently expressed unlinked NS2B/NS3 constructs, however, evidence was provided that the conformational equilibrium of the glycine-linked protease is shifted toward the open conformation, reducing the catalytic activity of the protease.[53,199] This was achieved by analysis of the behavior of conformational disulfide traps.[53] Therefore, six different constructs of DENV2 NS2B/NS3 were designed by introducing Cys pairs in the glycine-linked and a bivalent expressed construct. The positions of these Cys residues were chosen so that they can form disulfide bridges and trap either the open (I73*C, P106C), the closed (S75*C/K117C), or in both (A125C/V162C) conformations. After allowing disulfides to form, the content of free cysteine residues was determined to assess the prevalence of the respective conformation of the protease in solution. The glycine-linked protease yielded fewer disulfides for the closed traps than the unlinked construct. For the open trap, no significant difference could be observed. Regarding the WT proteases, the enzymatic activity of the linked construct was lower than that of the unlinked protease, further supporting the hypothesis of a higher population of the inactive open

conformation in the presence of the glycine-rich linker.[53,200] In crystallization experiments with linked and unlinked DENV4 NS2B/NS3 constructs, the open conformation was only yielded for the glycine-linked protease underlining this influence of the glycine-rich linker.[199] To better understand the role of the autocleavage site and its biological relevance, this project is dedicated to investigating the mechanism behind this cleavage process on the example of ZIKV NS2B/NS3.

The following publication and the unpublished and not yet submitted manuscript are part of the results obtained in this project:

(6) von Hammerstein, F.; Lauth L. M.; **Hammerschmidt, S.**; Wagner, A.; Schirmeister, T.; Hellmich, U. A. *Cis* autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs. *FEBS Lett*. **2019**; 593(16):2204–13. https://doi.org/10.1002/1873-3468.13507

Own contribution: Protein expression and purification, CD-spectroscopy, tryptophan-fluorescence melting curves, SDS-Page analysis, western blotting, enzyme activity assay, writing parts & editing of the manuscript.

Contributions from other authors: Protein expression and purification, cloning of constructs, CD-spectroscopy, SDS-Page analysis, analytical gel filtration, enzyme activity assay, densitometric analysis, CD-spectroscopy, tryptophan-fluorescence melting curves, writing of the original draft of the manuscript.

(7) **Hammerschmidt, S. J.**; [redacted] Promiscuity in the autocleavage behavior of the Zika virus NS2B/NS3 protease. *Unpublished and not yet submitted manuscript.*

Own contribution: Protein expression and purification, site-directed mutagenesis, SDS-PAGE, densitometric analysis, enzyme activity assay, writing of the original draft & editing of the manuscript.

Contributions from other authors: Protein mass spectroscopy, editing of the manuscript.

1.2.4. Project 4: Cation Interactions With SrtA

Obstacles in anti-SrtA drug development. The fast emergence of bacterial strands, resistant to multiple antibiotics, poses a major healthcare challenge. Beyond countering this development by developing conventional antibiotics, the potential value of anti-virulence drugs has been recognized in recent years.[130,201–203] SrtA is a well-studied transpeptidase relevant for the virulence of *S. aureus via* biofilm production (see *1.1.3 Staphylococcus aureus Sortase A*).[94,126,128] The various virulence factors assembled into the extracellular matrix by SrtA and its location on the bacterial cell membrane's surface make it a readily accessible and promising anti-virulence drug target. However, there are some hindrances to drug development, hampering the progression of innovative SrtA-targeting anti-virulence drugs.[133] (i) The catalytic Cys184 is reversely protonated, leading to a population below 0.1% in the nucleophilic zwitterionic state. Due to the lack of a pronounced oxyanion hole, the Cys184 is less reactive than in related cysteine proteases[101,135,140] and challenging to target with protease-specific electrophilic warheads.[204,205] (ii) The high flexibility of the β6/7 and β7/8 loops, forming the binding site of the sorting signal, is dramatically reduced upon ligand binding. A high entropic penalty accompanies this rigidification. Hence, even the natural LPETG substrate has a high K_M value of 5.5 mM.[101,102,155,206] (iii) For sufficient transpeptidase activity, local concentrations of both LPETG substrates and Gly_5 acceptors are high in the environment of SrtA, leading to weak reversible inhibitors since they need to compete for binding sites.[147]

SrtA in Bioengineering. Beyond its exciting role as an anti-virulence target to suppress biofilm production, SrtA is also prominent in biochemical engineering due to its suitability to connect proteins (known as *sortagging*) or to perform backbone cyclization reactions.[93,207,208] Since SrtA transpeptidase activity does not rely on membrane anchoring, an N-terminal truncated Δ59 SrtA construct is commonly used for biochemical applications and *in vitro* experiments.[101,209] In recent years, SrtA-mediated ligation (SML) has been widely used to facilitate protein-to-protein fusion,[210] peptide and protein cyclizations,[93,211] immobilization of biocatalysts onto solid surfaces,[212] preparation of complex glycoconjugates,[213] antibody–drug conjugation,[214,215] and *in vivo* protein modification.[216] Backbone cyclization of proteins with spatially close flexible termini endows several advantages over their linear analogs.[217] Cyclic proteins are less susceptible to proteolytic degradation by exopeptidases and can exhibit higher enzymatic activity due to lower conformational variability.[218,219] Also, cyclization can improve the oral bioavailability of therapeutic peptides.[220] Chemical cyclization is also possible but can come in low yields.[93,221] For suitable peptides of at least 19 amino acids, including the sorting signal and the oligoglycine, biotechnological cyclization with SrtA is possible with almost quantitative yields.[211] With shorter peptidic sequences and higher substrate concentrations, the main products are dimers and

oligomers.[211] Since SrtA activity relies on Ca^{2+} as a cofactor, the reaction buffers are limited to phosphate, carbonate, EDTA-, and EGTA-free solutions. This issue led to the design of calcium-independent mutants of SrtA by replacing some of the chelating Glu with Lys, also fixing the $\beta 6/\beta 7$ loop.[222,223] The best variant, however, reached only ~80% of the native enzyme's activity.[223]

Cation interactions with SrtA. The catalytic activity of SrtA depends on its Ca^{2+}-cofactor, enhancing catalytic activity 8-fold.[101] Some other ions were tested in activity assays to find Mg^{2+} and Mn^{2+} to reach ~50% of the Ca^{2+} activity.[101,209] Interestingly, the catalytic activity of SrtA without Ca^{2+} as the control measurement differs highly between those studies. Therefore, it can be assumed that residual Ca^{2+} from the enzyme purification was present in the reaction, causing this activity. Whereas the study with lower control activities found no activation of SrtA in the presence of Co^{2+} and Zn^{2+},[101] careful inspection of this data indicates that these cations may inhibit SrtA activity. This discrepancy led us to conclude that some ions can bind to the allosteric binding site, thereby replacing Ca^{2+} but not inducing a proper folding of SrtA.

Given the hindrances in inhibitor development, more knowledge about interactions of cations with the allosteric calcium binding site could reveal cations that bind to SrtA but act as allosteric inhibitors or partial agonists. This knowledge could benefit the development of new alloys and metal-containing materials for medical instruments or implants with antibacterial properties. In addition, innovative drugs such as chelators with an LPXTG backbone could be used, which are degraded by SrtA and thus release the inhibitory cations. Furthermore, mechanistic and structural information on the allosteric modulation of SrtA activity could provide a valuable basis for designing allosteric inhibitors for SrtA. From the bioengineering perspective, cofactors alternative to calcium ions could find broad applications in media unsuitable for calcium ions.

The following unpublished and not yet submitted manuscript presents the results obtained in this project:

(8) **Hammerschmidt, S. J.**; [redacted] Interactions of Staphylococcus Aureus Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors. *Unpublished and not yet submitted manuscript.*

Own contribution: Establishment of enzyme expression and purification, mutagenesis of SrtA active site mutants, Gibson Assembly of SrtA dimers, enzyme activity assays, ITC experiments, MST experiments, DSF experiments, X-ray crystallography, writing of the original draft & editing of the manuscript.

Contributions from other authors: Solving the crystal structure of SrtA in complex with cpd **3** and Mn^{2+}, Tb^{3+}-FRET assay, Synthesis of cpds **2** and **3**, Synthesis of the Lanthanide-Chlorides, Writing of the Experimental Section “Chemistry”.

2. List of Publications and Manuscripts

2.1. Publications and Manuscripts as Part of this Doctoral Thesis

2.1.1. Publications and Manuscripts of Project 1: Entropy Optimization in Medicinal Chemistry

(1) **Hammerschmidt, S. J.**; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B / NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm.* **2023**; 356(4):e2200518. https://doi.org/10.1002/ardp.202200518

(2) **Hammerschmidt, S. J.**[#]; [redacted] Improving binding entropic by ligand symmetry – a case study with human matriptase. *Manuscript submitted to RSC Medicinal Chemistry.*

(3) **Hammerschmidt, S. J.**; [redacted] Advanced Isothermal Titration Calorimetry for Medicinal Chemists with *ITCcalc*. *Manuscript submitted to J. Chem. Educ.*

2.1.2. Publications and Manuscripts of Project 2: Chasing the Binding Conformation of NS2B/NS3 Protease Inhibitors

(4) Millies, B.; Von Hammerstein, F.; Gellert, A.; **Hammerschmidt, S.**; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; Kersten, C.; Kiefer, W.; Bodem, J.; Hildt, E.; Windbergs, M.; Hellmich, U. A.; Schirmeister, T. 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

(5) [redacted] **Hammerschmidt, S. J.**[#]; [redacted] The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR. *Manuscript submitted to Eur. J. Med. Chem.*

2.1.3. Publications and Manuscripts of Project 3: Zika Virus NS2B/NS3 Autocleavage Behavior

(6) von Hammerstein, F.; Lauth L. M.; **Hammerschmidt, S.**; Wagner, A.; Schirmeister, T.; Hellmich, U. A. *Cis* autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs. *FEBS Lett*. **2019**; 593(16):2204–13. https://doi.org/10.1002/1873-3468.13507

(7) **Hammerschmidt, S. J.**; [redacted] Promiscuity in the autocleavage behavior of the Zika virus NS2B/NS3 protease. *Unpublished and not yet submitted manuscript.*

2.1.4. Publications of Project 4: Cation Interactions With SrtA

(8) **Hammerschmidt, S. J.**; [redacted] Interactions of Staphylococcus Aureus Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors. *Unpublished and not yet submitted manuscript.*

2.2. Research Articles Beyond this Doctoral Thesis

(9) Barthels, F.; Marincola, G.; Marciniak, T.; Konhäuser, M.; **Hammerschmidt, S.**; Bierlmeier, J.; Distler, U.; Wich, P. R.; Tenzer, S.; Schwarzer, D.; Ziebuhr, W.; Schirmeister, T. 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

(10) Maus, H.; Barthels, F.; **Hammerschmidt, S. J.**; Kopp, K.; Millies, B.; Gellert, A.; Ruggieri, A.; Schirmeister, T. SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B/NS3 Proteases. *Bioorganic Med. Chem.* **2021**; 47(August):116392. https://doi.org/10.1016/j.bmc.2021.116392

(11) Amendola, G.; Ettari, R.; Previti, S.; Di Chio, C.; Messere, A.; Di Maro, S.; **Hammerschmidt, S. J.**; Zimmer, C.; Zimmermann, R. A.; Schirmeister, T.; Zappalà, M.; Cosconati, S. 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.

(12) Götz, C.; Hinze, G.; Gellert, A.; Maus, H.; von Hammerstein, F.; **Hammerschmidt, S. J.**; Lauth, L. M.; Hellmich, U. A.; Schirmeister, T.; Basché, T. 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

(13) Welker, A.; Kersten, C.; Müller, C.; Madhugiri, R.; Zimmer, C.; Müller, P.; Zimmermann, R. A.; **Hammerschmidt, S.**; Maus, H.; Ziebuhr, J.; Sotriffer, C.; Schirmeister, T. Structure-Activity Relationships of Benzamides and Isoindolines Designed as SARS-CoV Protease Inhibitors Effective against SARS-CoV-2. *ChemMedChem* **2021**; 16(2):340–354. https://doi.org/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. https://doi.org/10.1002/cmdc.202000987

(14) Vilela, G. G.; Silva, W. F. dos S.; Batista, V. de M.; Silva, L. R.; Maus, H.; **Hammerschmidt, S. J.**; Costa, C. A. C. B.; Moura, O. F. da S.; de Freitas, J. D.; Coelho, G. L.; Brandão, J. de A.; Anderson, L.; Bassi, Ê. J.; Xavier de Araújo-Júnior, J.; Schirmeister, T.; da Silva-Júnior, E. F. da. Fragment-Based Design of α-Cyanoacrylates and α-Cyanoacrylamides Targeting Dengue and Zika NS2B/NS3 Proteases. *New J. Chem.* **2022**; 46(42):20322–20346. https://doi.org/10.1039/d2nj01983c

(15) Barthels, F.; Meyr, J.; **Hammerschmidt, S. J.**; Marciniak, T.; Räder, H. J.; Ziebuhr, W.; Engels, B.; Schirmeister, T. 2-Sulfonylpyrimidines as Privileged Warheads for the Development of *S. aureus* Sortase A Inhibitors. *Front. Mol. Biosci.* **2022**; 8(January):1–17. https://doi.org/10.3389/fmolb.2021.804970

(16) Previti, S.; Ettari, R.; Calcaterra, E.; Di Chio, C.; Ravichandran, R.; Zimmer, C.; **Hammerschmidt, S.**; Wagner, A.; Bogacz, M.; Cosconati, S.; Schirmeister, T.; Zappalà, M. Development of Urea-Bond-Containing Michael Acceptors as Antitrypanosomal Agents Targeting Rhodesain. *ACS Med. Chem. Lett.* **2022**; 13(7):1083–1090. https://doi.org/10.1021/acsmedchemlett.2c00084

(17) Volynets, G. P.; Barthels, F.; **Hammerschmidt, S. J.**; Moshynets, O. V.; Lukashov, S. S.; Starosyla, S. A.; Vyshniakova, H. V.; Iungin, O. S.; Bdzhola, V. G.; Prykhod'ko, A. O.; Syniugin, A. R.; Sapelkin, V. M.; Yarmoluk, S. M.; Schirmeister, T. Identification of Novel Small-Molecular Inhibitors of *Staphylococcus aureus* Sortase A Using Hybrid Virtual Screening. *J. Antibiot.* **2022**;75(6):321–332. https://doi.org/10.1038/s41429-022-00524-8

(18) Barthels, F.; **Hammerschmidt, S. J.**; Fischer, T. R.; Zimmer, C.; Kallert, E.; Helm, M.; Kersten, C.; Schirmeister, T. A Low-Cost 3D-Printable Differential Scanning Fluorometer for Protein and RNA Melting Experiments. *HardwareX* **2022**; 11:e00256. https://doi.org/10.1016/j.ohx.2022.e00256

(19) Brinkmann, S.; Semmler, S.; Kersten, C.; Patras, M. A.; Kurz, M.; Fuchs, N.; **Hammerschmidt, S. J.**; Legac, J.; Hammann, P. E.; Vilcinskas, A.; Rosenthal, P. J.; Schirmeister, T.; Bauer, A.; Schäberle, T. F. Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues. *ACS Chem. Biol.* **2022**; 17(3):576–589. https://doi.org/10.1021/acschembio.1c00861

(20) Previti, S.; Ettari, R.; Calcaterra, E.; Di Maro, S.; **Hammerschmidt, S. J.**; Müller, C.; Ziebuhr, J.; Schirmeister, T.; Cosconati, S.; Zappalà, M. Structure-Based Lead Optimization of Peptide-Based Vinyl Methyl Ketones as SARS-CoV-2 Main Protease Inhibitors. *Eur. J. Med. Chem.* **2022**; 247:115021. https://doi.org/10.1016/j.ejmech.2022.115021.

(21) Maus, H.; Hinze, G.; **Hammerschmidt, S. J.**; Basché, T.; Schirmeister, T. A Competition smFRET Assay to Study Ligand-induced Conformational Changes of the Dengue Virus Protease. *Protein Sci.* **2023**; 32(1):e4526. https://doi.org/10.1002/pro.4526.

(22) Müller, P.; Meta, M.; Meidner, J.L.; Schwickert, M.; Meyr, J.; Schwickert, K.; Kersten, C.; Zimmer, C.; **Hammerschmidt, S. J.**; Frey, A.; Lahu, A.; de la Hoz-Rodríguez, S.; Agost-Beltrán, L.; Santiago Rodríguez, S.; Diemer, K.; Neumann, W.; González, F. V.; Engels, B.; Schirmeister, T. Investigation of the Compatibility between Warheads and Peptidomimetic Sequences of Protease Inhibitors—A Comprehensive Reactivity and Selectivity Study. *Int J Mol Sci.* **2023**; 24(8):7226. https://doi.org/10.3390/ijms24087226.

2.3. Unpublished Research Article Manuscripts Beyond this Doctoral Thesis

(23) [redacted] **Hammerschmidt, S.**: [redacted] Peptidyl Nitroalkene Inhibitors of Main Protease (M^{pro}) rationalized by Computational/Crystallographic Investigations as Antivirals against SARS-CoV-2. *Manuscript submitted to Nat. Commun.*

(24) [redacted] **Hammerschmidt, S. J.**; [redacted] Enzymatic *in vitro* screening of a series of 190 cpds identifies thiosemicarbazones as promising inhibitors of SARS-CoV-2 M^{pro} protein: Molecular docking interactions. *Manuscript submitted to Arch. Pahrm.*

(25) [redacted] **Hammerschmidt, S. J.**; [redacted] Next Generation of Fluorometric Protease Assays: 7-Nitrobenz-2-oxa-1,3-diazol-4-yl-amides (NBD-Amides) as Class-Spanning Protease Substrates. *Unpublished and not yet submitted manuscript.*

2.4. Review Articles Beyond this Doctoral Thesis

(26) Müller, P.[#]; Maus, H. [#]; **Hammerschmidt, S. J.**[#]; Knaff, P. M.; Mailänder, V.; Schirmeister, T.; Kersten, C. Interfering with Host Proteases in SARS-CoV-2 Entry as a Promising Therapeutic Strategy. *Curr. Med. Chem.* **2021**; 29(4):635–665. https://doi.org/10.2174/0929867328666210526111318

(27) 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**; 233–295. https://doi.org/10.1007/978-3-030-71165-8_13

2.5. Other Publications Beyond this Doctoral Thesis

(28) **Hammerschmidt, S. J.**; Arzneibuch-Kommentar, Ph.Eur. 9.8 Lactulose, **2019**. ISBN 978-3-8047-4082-2

(29) **Hammerschmidt, S. J.**; Arzneibuch-Kommentar, Ph.Eur. 10.1 Sulfamethizol, **2020**. ISBN 978-3-8047-4082-2

(30) **Hammerschmidt, S. J.**; Müller, P.; Schirmeister, T. SARS-CoV-PL^{pro}-Inhibitoren als mögliche Breitspektrum-Virostatika. *BIOspektrum* **2021**; 27(3):254–256. https://doi.org/10.1007/s12268-021-1576-6

(31) Rodrigues, É.; Maus, H.; **Hammerschmidt, S.**; Araújo-Júnior, J.; Schirmeister, T.; Silva-Júnior, E. A new furin-based peptidomimetic as an inhibitor of NSB2-NS3 protease from Zika virus (ZIKV). In *Proceedings of 7th International Electronic Conference on Medicinal Chemistry.* **2021**; 11431. https://doi.org/10.3390/ECMC2021-11431

3. Project 1: Entropy Optimization in Medicinal Chemistry

3.1. Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs

3.1.1. Context, Project Summary, and Own Contribution

In preceding works, the group of [redacted] synthesized and evaluated peptidomimetic active site-directed inhibitors of flaviviral NS2B/NS3 proteases.[47,224] Since these potent inhibitors harbor di- or tribasic scaffolds, their membrane permeability and *in vivo* activity were low.[59–62] In a crystal structure of the closely related WNV NS2B/NS3 protease in complex with such a multibasic inhibitor (PDB-ID: 2YOL), the ligand adopts a horseshoe-like distorted conformation with both termini in a proximal orientation (Figure 16A).[224] This prompted researchers in the group of [redacted] to perform macrocyclization of their inhibitors to enhance binding affinity. Thereby, they strived to optimize binding enthalpy through reduced ligand preorganization and yielding inhibitors with improved membrane permeability and potentially higher selectivity.[56,57,225] This was done in analogy to the macrocyclic inhibitors of the hepatitis C virus (HCV) NS3/4A protease paritaprevir, danoprevir, and ciluprevir.[226] Since the HCV NS3/4A protease also displays an open S1 pocket, similar to NS2B/NS3 (see *1.1.1 Dengue and Zika NS2B/NS3 Proteases*), it is possible to perform a cyclization from the P1 to the P3 residue. In contrast to the HCV inhibitors, where the side chains of P1 and P3 could be used for macrocyclization, the conformation of NS2B/NS3 inhibitor's P3 Lys points to the Phe84 backbone oxygen and away from the P1 side chain (Figure 16A). Therefore, macrocyclization was performed from the P1 side chain to the P3 backbone (Figure 16B).

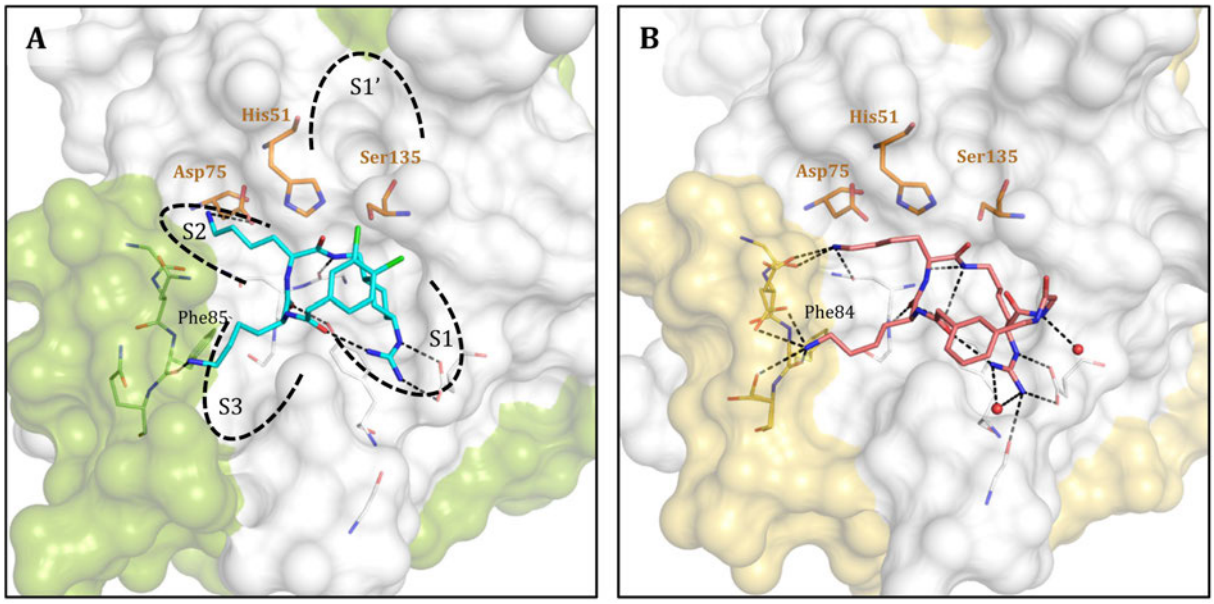


Figure 16. *Binding conformations of a linear reference cpd and a macrocycle.* NS3 is shown as a white surface, NS2B is green (WNV) or yellow (ZIKV). Amino acids of the catalytic triad are shown as orange sticks, and the amino acids forming the oxyanion hole are blue. Binding sub-pockets are indicated according to Schechter&Berger.[43] (**A**) A linear reference cpd (cyan) is bound to WNV NS2B/NS3 (PDB-ID: 2YOL).[224] (**B**) Macrocyclic cpd **4** in the ZIKV binding site (PDB-ID: 6Y3B)[57] displaying the binding mode of the natural substrate Thr-Gly-Lys-Arg (black sticks). Polar interactions are depicted as black dashed lines. For a clear view, only those amino acids forming polar interactions are displayed as thin lines. This figure was created using PyMOL.[44]

Indeed, macrocyclization led to a series of highly potent inhibitors, and the co-crystal structure of the lead cpd was solved, demonstrating the retained orientation of the inhibitor (Figure 16B). However, the thermodynamic basis of these improvements has not been fully explored. Herein, we determined the thermodynamic binding profiles of the macrocyclic cpd **4** and the linear reference cpds **5**–**7**, synthesized by NIKLAS J. BRAUN of the group of [redacted] Interestingly, the results indicate that cpds **4** and **7** were entropically less favored than inhibitors **5** and **6** (Figure 17). Determination of ΔC_P revealed highly negative values that could represent the buried hydrophobic residues during the shift of the conformational equilibrium of NS2B/NS3 towards the closed conformation upon binding the inhibitors.

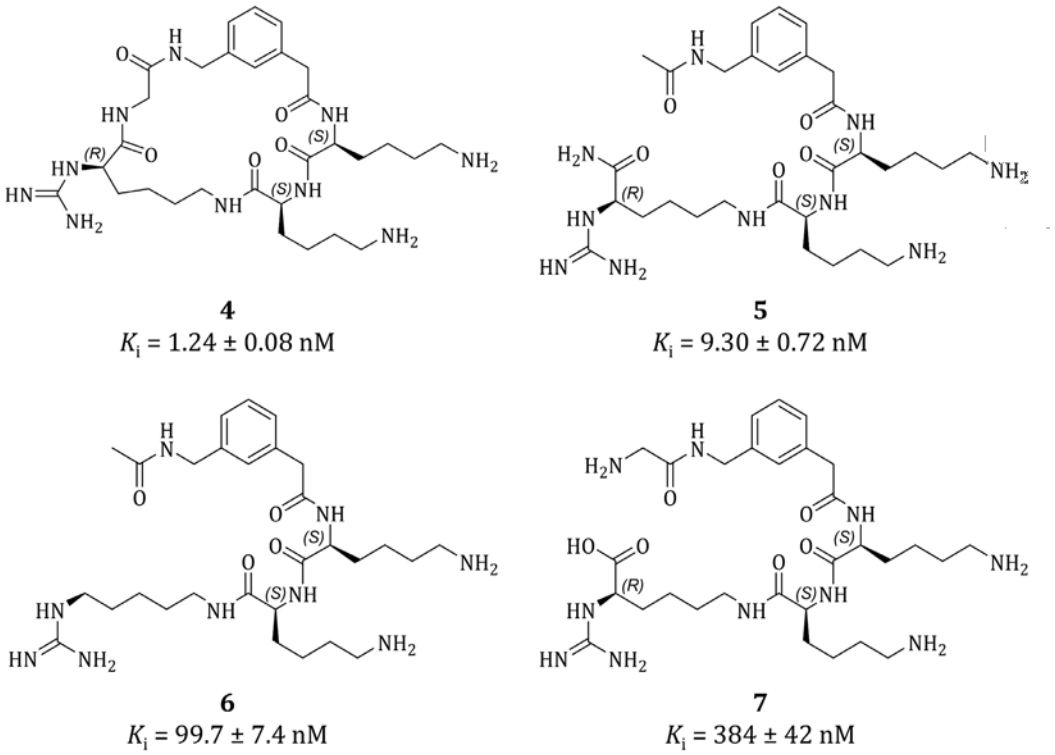


Figure 17. *The 2D-Structure of cpds **4**–**7**, subjected to investigations in this Publication.* The macrocyclic cpd **4** is the most potent inhibitor. MD-Simulations showed that cpd **7** is also very rigid, probably due to a stable salt bridge that forms between the carboxy and the amine terminus of the Gly residue.

To shed some light on these results, we performed MD simulations. Cpds **5** and **6** exhibited higher root mean square deviation (RMSD) values throughout the simulations, whereas **4** and **7** were more rigid. This finding demonstrated that cpds **5** and **6** retain flexibility in the complexed state. A ligand torsion distribution analysis revealed higher torsion energies in the bound states of inhibitors **4** and **7** relative to **5** and **6**, indicating a less favorable binding conformation fixed in macrocyclization of cpd **4** and by an intramolecular salt-bridge of **7**. Finally, we solved co-crystal structures of cpds **5**–**7**, displaying lower electron densities in the flexible parts of the inhibitors, supporting the hypothesis of higher ligand flexibility to be accountable for the observed entropic penalty. Based on the crystal structures, we normalized B'-Factors using BANΔIT[48] and compared them to the RMSD values of the MD simulations of the protease but did not find significant differences between the analyzed inhibitors.

(1) **Hammerschmidt, S. J.**; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B / NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm.* **2023**; 356(4):e2200518. https://doi.org/10.1002/ardp.202200518

Own contribution: Enzyme expression and purification, NS2B/NS3 inhibition assays, ITC experiments, writing of the original draft & editing of the manuscript.

Contributions from other authors: Inhibitor synthesis, X-ray crystallography, quantum chemical investigations writing of experimental parts & editing of the manuscript.

This work has been published in *Archiv der Pharmazie* (impact factor: 4.61).

The appended 'Supporting Information' represents an abridged version. The full version can be accessed online at doi: 10.1002/ardp.202200518.

3.1.2. Publication

The following publication quoted (within " and ") from page 41 to page 84 is exactly the same as the manuscript cited on page 40. "

Received: 2 October 2022 | Revised: 15 November 2022 | Accepted: 17 November 2022

DOI: 10.1002/ardp.202200518

FULL PAPER

ARCH PHARM DPhG
Archiv der Pharmazie

Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs

Stefan J. Hammerschmidt[1] | Simon Huber[2] | Niklas J. Braun[2] | Marc Lander[1] | Torsten Steinmetzer[2] | Christian Kersten[1]

[1]Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Mainz, Germany

[2]Institute of Pharmaceutical Chemistry, Philipps-University, Marburg, Germany

Correspondence
Christian Kersten, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany.
Email: kerstec@uni-mainz.de

Funding information
LOEWE-Zentrum für Translationale Medizin und Pharmakologie, Grant/Award Number: DRUID (Novel Drug Targets against Poverty-Related; Deutsche Forschungsgemeinschaft, Grant/Award Number: INST 247/921-1 FUGG

Abstract

Cyclization of small molecules is a widely applied strategy in drug design for ligand optimization to improve affinity, as it eliminates the putative need for structural preorganization of the ligand before binding, or to improve pharmacokinetic properties. In this work, we provide a deeper insight into the binding thermodynamics of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its linear analogs. Characterization of the thermodynamic binding profiles by isothermal titration calorimetry experiments revealed an unfavorable entropy of the macrocycle compared to the open linear reference ligands. Molecular dynamic simulations and X-ray crystal structure analysis indicated only minor benefits from macrocyclization to fixate a favorable conformation, while linear ligands retained some flexibility even in the protein-bound complex structure, possibly explaining the initially surprising effect of a higher entropic penalty for the macrocyclic ligand.

KEYWORDS
crystallization, macrocycles, molecular dynamics, thermodynamics, Zika NS2B/NS3 protease

1 | INTRODUCTION

1.1 | Zika virus (ZIKV) infection

ZIKV infections gained worldwide interest since they became endemic in Brazil in 2015. In this outbreak, approximately 400,000 to 1.6 million people were infected. Besides the transmission via its main vectors *Aedes aegypti* and *Aedes albopictus*, whose occurrences are typically associated with flaviviral distribution, it was also discovered to be spread by sexual intercourse.[1–5] A typical ZIKV infection is accompanied by mild flu-like symptoms and a high prevalence of asymptomatic cases. As a result of widespread ZIKV infection cases all over South America and its correlation to neurological disorders such as the Guillain–Barré syndrome and microcephaly in neonates, the WHO declared ZIKV a public health emergency in 2016.[6–10]

1.2 | NS2B/NS3 protease

In the viral replication cycle, the viral genome is translated into a precursor-polyprotein processed by the viral NS2B/NS3 serine protease and host proteases into the three structural (Capsid, Membrane, and Envelope) and the seven nonstructural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5).[11] When this crucial function of NS2B/NS3 is inhibited, viral replication is abolished. Therefore, the NS2B/NS3 protease resembles an attractive drug target to counter ZIKV infections.[12–15] For the proteolytic activity of flaviviral proteases, the NS3 protease domain relies on its NS2B cofactor, which can adopt at least two conformations. In the inactive and so-called *open* conformation, NS2B is loosely bound to the NS3 protease domain and mainly disordered.[16,17] In the catalytically active *closed* conformation, NS2B is wrapped around NS3 and contributes to substrate recognition by forming

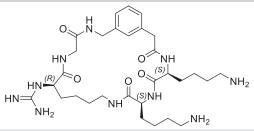
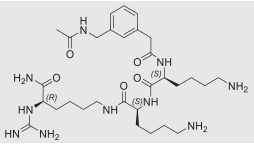
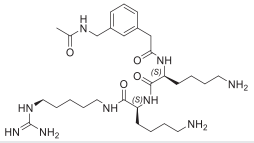
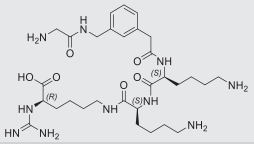
parts of the S2 and S3 binding pockets.[18–20] In solution, both conformations coexist in equilibrium.[21–23] Upon ligand or substrate binding to the active site, exclusively cocrystals of the closed conformation were obtained so far.

1.3 | Inhibitor design and optimization

Due to the shallow shape of the NS2B/NS3 active site[24] and its preference for basic P1 and P2 residues, the development of drug candidates targeting this site is considered to be very challenging.[25] Consequently, most active site-directed ZIKV NS2B/NS3 protease inhibitors are based on peptidic or peptidomimetic scaffolds with often limited bioavailability featuring multibasic scaffolds.[12,26–29] Although they reached considerably high affinities, these compounds exhibited only poor antiviral potency most likely caused by limited membrane permeabilities.[30–33] The crystal structure of a linear substrate analog inhibitor in complex with the closely related West Nile virus NS2B/NS3 protease revealed a horseshoe-like backbone conformation with both termini in close proximity to each other (Protein Data Bank[34] [PDB]-ID: 2YOL; Supporting Information: Figure S1).[35] This gave rise to a macrocyclization approach not only intended to optimize the inhibitory potency but also increase membrane permeability resulting in compound **1** (Table 1).[27] In the cocrystal structure of inhibitor **1** in complex with ZIKV NS2B/NS3 (PDB-ID: 6Y3B, Figure 1), the P1 guanidine residue is located at the bottom of the S1 pocket and forms an ionic interaction with Asp129. The carbonyl oxygen of Gly159 forms a hydrogen bond directly to the guanidine as well as a water-mediated hydrogen bond. Other H_2O-mediated interactions are formed by Asp129 to the backbone nitrogen of the glycine linker. An intramolecular interaction between the P4 carbonyl oxygen and the terminal guanidine nitrogen stabilizes the inhibitor's conformation. This interaction can also be found in the linear reference inhibitor (Supporting Information: Figure S1). The hydroxy group of Tyr161 stabilizes the position of the backbone of inhibitor **1** by forming hydrogen bonds between its hydroxy group to the nitrogen of the P1–P2 amide and the backbone carbonyl oxygen of P3 Lys that further interacts with the Gly153 nitrogen. The side chain of the P2 Lys forms polar contacts to the NS2B Asp83 and NS3 Asn152 side chains and to the Gly82 backbone in the S2 pocket. The side chain of P3 Lys is stabilized by polar interactions with the carbonyl oxygen of Phe84 and the S3 forming side chains of Asp83 and Ser85. It is noteworthy that the NS2B residues Asp83 and Ser85 are found in two conformations based on their occupancies in the crystal structure. Therefore, Asp83 can contribute to the stabilization of both the P2 Lys and the P3 Lys side chain.

The complex formation of an oligo-peptidic ligand with a protein is often accompanied by the fixation of rotatable bonds upon binding.[18] This reduction of conformational degrees of freedom is usually accompanied by an entropic penalty and, thus, lowers binding affinity. In ligand optimization attempts, a widely accepted strategy to increase affinity is the rigidification of rotatable bonds to preorganize the ligand in the binding conformation via (macro-) cyclization.[18,36] Although it is tempting to assume that mainly entropic effects lead to affinity enhancement by cyclization, also enthalpic benefits have been reported, which has led to some controversy about the underlying physical principles.[37-40] In the case of peptidomimetics, macrocyclization,

TABLE 1 Affinity and buffer-corrected thermodynamic binding profiles obtained by the fluorometric enzyme activity assay and ITC

Compound	Structure	K_i (nM)	K_d (nM)	$\Delta G°$ (kJ·mol^{-1})	$\Delta H°$ (kJ·mol^{-1})	$-T\Delta S°$ (kJ·mol^{-1})	n_{proton}	ΔC_P (kJ·mol^{-1})
1		1.24 ± 0.08	110 ± 7	−40.1 ± 1.7	−74.8	34.7	0.17	−3.77
2		9.30 ± 0.72	2120 ± 440	−32.7 ± 0.3	−53.1	20.4	0.30	−2.16
3		99.7 ± 7.4	2700 ± 90	−31.9 ± 0.2	−58.5	26.6	−0.10	−2.90
4		384 ± 42	8220 ± 250	−29.5 ± 0.1	−77.3	47.8	0.24	N.d.

Abbreviations: ITC, isothermal titration calorimetry; N.d., not determined.

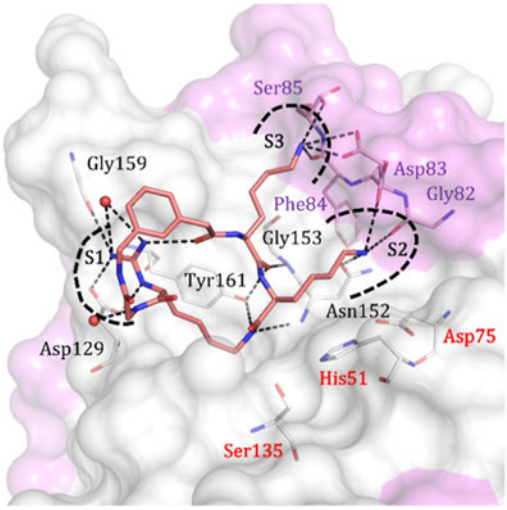


FIGURE 1 Binding mode of compound **1** in complex with the ZIKV NS2B/NS3 protease (PDB-ID: 6Y3B). Inhibitor **1** is depicted as sticks with salmon-colored carbon atoms. For a clear view, only amino acids forming polar interactions (black dashed lines) are shown as lines with violet (NS2B) and white carbon atoms (NS3). Amino acids forming the catalytic triad are labeled in red.

bridging, and cross-linking represent promising strategies to achieve more potent ligands with several benefits.[27,41,42] Besides higher proteolytic stability,[28] properly restrained molecules are less likely to adopt conformations to fit the binding pockets of off-target proteins; thus, macrocyclization offers a route to achieve more selective compounds.[43] Additionally, by derivatization of polar terminal groups and reduction of the number of rotatable bonds, macrocyclization offers a convincing benefit to membrane permeability.[29,30,44,45] Although the preorganization of ligands is a regularly utilized technique, the driving forces behind the improved affinities remain largely elusive. In rational drug design campaigns, isothermal titration calorimetry (ITC) is an indispensable tool to characterize binding thermodynamics. It allows to quantify the enthalpic (ΔH) and entropic ($-T\Delta S$) contributions to the overall binding free energy (ΔG) and to evaluate the impact of structural modifications on these parameters.

In this work, the thermodynamic effects of macrocyclization in inhibitor **1** compared to its noncyclic analogs **2**–**4** (Table 1) were investigated using a fluorometric enzyme activity assay, ITC, molecular dynamic (MD) simulations, and X-ray crystallography.

2 | RESULTS AND DISCUSSION

2.1 | Inhibitory activities

A fluorometric enzyme inhibition assay was employed to redetermine the inhibitory potencies of the compounds under elucidation. The obtained K_i values are in good agreement with previous results[26,27]: **1**: 1.24 ± 0.08 nM, **2**: 9.30 ± 0.72 nM; **3**: 99.7 ± 7.4 nM; and **4**: 384 ± 42 nM (Table 1).

2.2 | Thermodynamic binding profiles

In ITC experiments, the same affinity trend was observed but with significantly higher K_d values compared to K_i (Table 1), as occasionally observed when comparing different techniques.[46,47] Interestingly, there are large differences in the factors between K_i and K_d (ca. 87 for inhibitor **1**, 228 for inhibitor **2**, 27 for inhibitor **3**, and 21 for inhibitor **4**; Table 1). The main differences in the experimental composition between the inhibition assay and ITC binding assay are the different concentrations of enzymes and ligands (Assay: 2 nM NS2B/NS3; ITC: 15–40 μM NS2B/NS3) and the presence of a substrate in the inhibition assay and its absence in the ITC assay. One could speculate that the conformational equilibrium between inactive open conformation and active closed conformation of the protease is affected by the substrate's presence. As the inhibitors under elucidation also bind to the active closed conformation, a shift of equilibrium by the substrate to this conformation might enhance the effective inhibitory potency.[46,47] However, this hypothesis requires further elucidation and is beyond the scope of this manuscript. The buffer-corrected thermodynamic binding profiles (Table 1, buffer correction calculation and signature plots: Supporting Information: Figure S2, measurements in different buffers: Supporting Information: Table S1) revealed that the affinity of all tested inhibitors is driven by their large exothermic binding enthalpy with the highest contribution for inhibitors **1** and **4** ($\Delta H°$ = −74.8 and −77.3 kJ·mol^{-1}) and lower binding enthalpies for inhibitors **2** and **3** ($\Delta H°$ = −53.1 and −58.5 kJ·mol^{-1}). Furthermore, all inhibitors show a positive temperature-dependent entropy term ($-T\Delta S°$). Again, the macrocyclic compound **1** and linear compound **4** share higher absolute values ($-T\Delta S°$ = 34.7 and 47.8 kJ·mol^{-1}, respectively) than inhibitors **2** and **3** ($-T\Delta S°$ = 20.4 and 26.6 kJ·mol^{-1}). Consequently, the increased affinity of inhibitor **1** compared to **2** and **3** is despite the macrocyclization not being accompanied by a favorable entropy term. Similar results of macrocyclization were reported in the literature.[37,40] Buffer ionization correction also revealed that probably no proton transfers occur for all inhibitors **1**–**4** upon complex formation (n_{proton} = 0.17, 0.30, −0.10, and 0.24, respectively, Table 1, Supporting Information: Figure S2). Inhibitors **1**–**3** share three basic centers: the guanidinium moiety and the aliphatic amines of the lysine sidechains. All of these tend to be protonated under physiological pH and this protonation seems not to change upon binding as all three moieties are involved in polar, ionic interactions (Figure 1). Inhibitor **4** contains two additional functionalities, a basic amine, and an acid. Again, binding seems not to change protonation states. In solution as well as within the complex, all mentioned moieties should be charged. Finally, isobaric heat capacity changes (ΔC_P) were determined from ITC experiments at different temperatures (Supporting Information: Table S2). Changes in heat capacity are commonly correlated with the burial of nonpolar and polar surfaces indicating desolvation

effects.[48,49] For inhibitors **1**–**3**, negative ΔC_P-values of −3.77, −2.16, and −2.88 kJ·mol^{-1}·K^{-1} were determined (Supporting Information: Figure S3). Differently, for compound **4**, no linear correlation between ΔH and temperature was observed. While the negative ΔC_P-values of **1**–**3** indicate the burial of hydrophobic areas, the absolute values are very large for pure protein–ligand interactions.[50–52] Most likely, the shifted equilibrium between open and closed conformations upon active site-directed ligand binding is also accompanied by burial of hydrophobic areas. This is also suggested by a decrease in the fraction of hydrophobic accessible surface area from 52% in the open to 48% in the closed ZIKV NS2B/NS3 conformation and seems to be a common feature for flaviviral proteases (Supporting Information: Table S3). In the closed (active) conformation, the NS2B cofactor is tightly wrapped around NS3 while being more solvent exposed in the open conformation. Hence, the largest ΔC_P of compound **1** compared to **2** and **3** could be interpreted in different ways, like the stronger burial of hydrophobic patches of the ligand or a more efficient shift of the conformational equilibrium toward the closed conformation. However, ITC alone cannot finally answer that question and several layered effects may contribute to the observed results.[53–55]

2.3 | Molecular dynamics

To further elucidate this binding behavior, molecular dynamic simulations (MDs) were conducted starting from the ZIKV NS2B/NS3-**1** complex structure (PDB-ID 6Y3B).[27] For inhibitors **2**–**4**, starting from the same structure, ligands were modified and energetically minimized within the binding site prior to MD as at that time no complex structures of NS2B/NS3 with those ligands were available in the PDB. Throughout the simulations of the NS2B/NS3–inhibitor complexes, it was observed that inhibitors **2** and **3** retained some partial flexibility even in the protease-bound state, while the macrocycle **1** and inhibitor **4** (**4** likely due to an intramolecular ionic interaction between the charged termini) are highly rigid (Figure 2a). For inhibitors **2** and **3**, the acetylated 3-aminomethylphenylacetyl group was found to be more flexible in the bound state, whereas the P2 and P3 Lys sidechains stayed in a stable orientation. Tracking the rigid and flexible parts of the ligands over the MD trajectory (Figure 2b), for inhibitors **1** and **4**, both RMSD traces (rigid part only and whole molecule) stayed below 2.2 Å over 10 ns of simulation time, whereas the RMSD at the flexible parts of compounds **2** and **3** drastically increases throughout the simulation compared to the RMSD only of the rigid parts. To further elucidate the ligand flexibility based on MD, the internal torsion energies of the bound and unbound states were investigated. In analogy to the *SPAM* method,[56] which is used to estimate binding thermodynamics of water molecules from MD simulations, the distributions of ligand torsion energies were computed (Figure 2c). One should be aware that the transfer of this method from water to a more complex system neglects many parts of molecular recognition and might not be used as a quantitative metric. For a qualitative interpretation of torsion energy distribution and its implications on ligand flexibility, however, it might give insight into torsional degrees of freedom and their changes upon binding. In our variant of this method, the energies of all torsion angles are calculated throughout the simulations of ligands in the bound and unbound state. The torsion energies were found to follow a Gaussian distribution for all ligands, which is also found for the interaction energies of water molecules in the original *SPAM* method. This normal distribution is a prerequisite for further elucidation. The difference in the mean value of torsion energies between bound and unbound ligands indicates how favorable the bound conformation is in terms of torsion strain. A shift to higher values indicates that the bound conformation might be less favorable than other conformations found in the solution. The broadness of the distribution (indicated by the standard deviation of the Gaussian fitting curve) might be a surrogate for the degrees of torsional freedom in bound and unbound states and subsequently on its changes upon binding. A sharper distribution of ligand torsion energies in the bound state compared to the free ligand could indicate rigidification upon complex formation. This was observed for both inhibitors **1** and **4**, but for compounds **2** and **3**, no such pronounced effect was noticed with a similarly sharp distribution of torsion energies in the bound and unbound state. This observation might indicate one reason for the more favorable temperature-dependent entropy term ($-T\Delta S°$) of inhibitors **2** and **3**, which retain more torsional degrees of freedom in the complex with ZIKV NS2B/NS3 compared to inhibitors **1** and **4**. Further, for compound **1**, the highest shift of the average torsion energy toward higher values was observed, indicating that the conformation trapped by macrocyclization might not be optimal and structural reorganization is still required upon binding.

Notably, changes in the conformational entropy of the protein could easily dominate binding entropies. Interestingly, while some rigidification of NS2B/NS3 upon ligand binding was observed when comparing apo- and ligand-bound structures, especially within the binding site, there were no obvious differences observed between the ligands. Both structural changes (indicated by residual C_α-RMSD) and backbone dynamics (C_α-root-mean-square fluctuation, RMSF) in complex with the different ligands were broadly similar (Supporting Information: Table S4, Figure S4a,b). Subsequently, order parameters (S^2) for both backbone and sidechains were investigated (Supporting Information: Table S5, Figure S4c,d). S^2 are associated with the conformational entropy of residues[57] with $S^2 = 1$ indicating low flexibility and $S^2 = 0$ indicating high flexibility.[58,59] Backbone S^2 are fairly in line with the RMSF with only Gly159 and Ser160 being a little more flexible in the NS2B/NS3-**2** complex. Sidechain S^2 showed overall more differences but no obvious trends that discriminate inhibitors **1** and **4** binding from inhibitors **2** and **3**. However, even though differences in S^2 for certain residues both within the binding site and more distant were observed, the corresponding configurational entropies of the residues seem to approximately cancel each other out. Therefore, the analysis of protein residue dynamics remains inconclusive, but slight differences might not hold the explanation for the observed thermodynamic binding profiles from the ITC experiments.

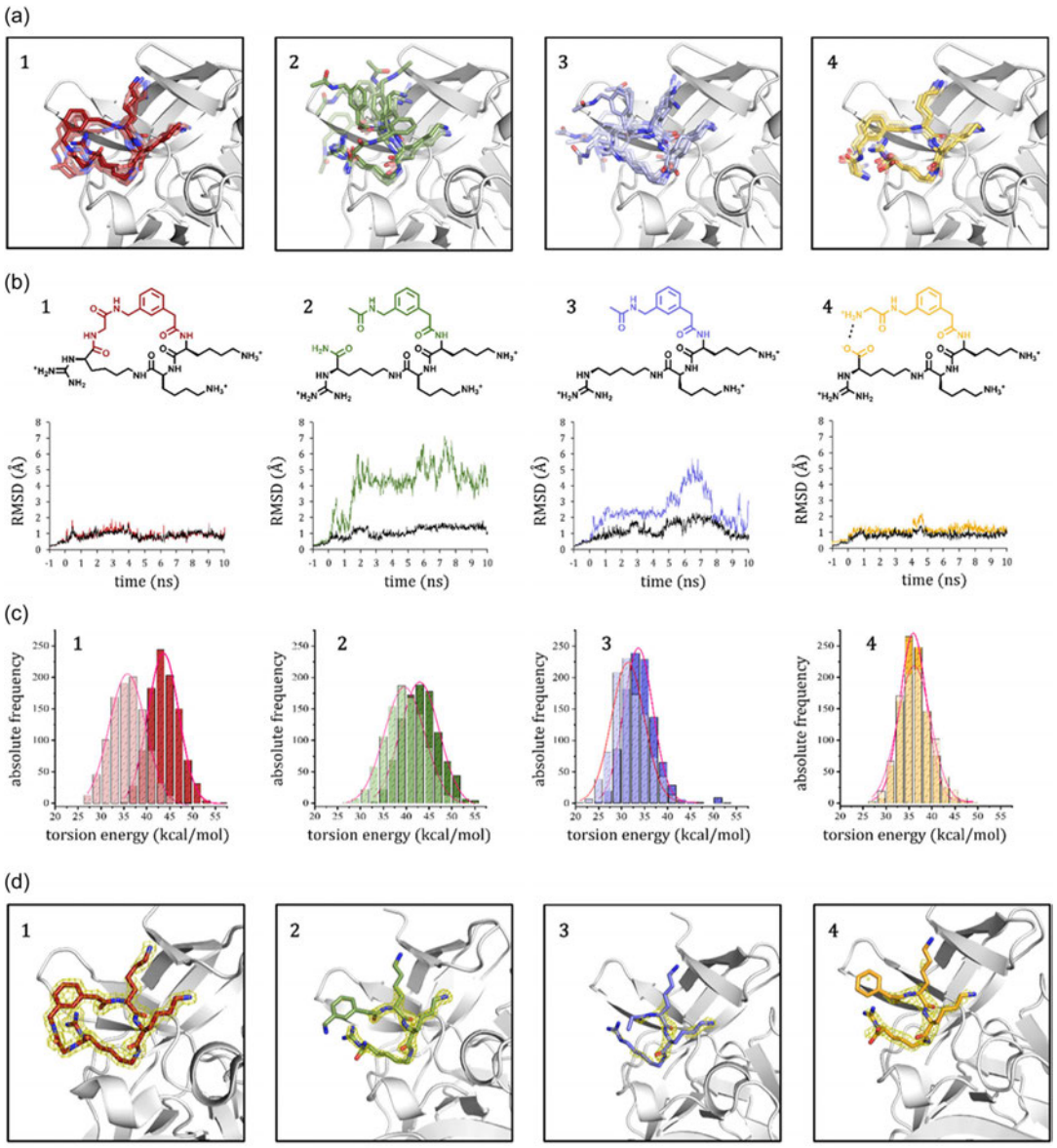


FIGURE 2 Investigating ligand flexibilities of **1** (red), **2** (green), **3** (blue), and **4** (yellow) in complex with NS2B/NS3 (white). (a) Ligand overlay of six MD frames (after 0, 2, 4, 6, 8, and 10 ns). The transparency reduces with elapsed time. For a clear view, only the starting structure of the enzyme is shown. (b) 2D structures of inhibitors **1**–**4** (top) and heavy atom root mean square deviation (RMSD) traces over 10 ns simulation time (bottom). The flexible parts of inhibitors **2** and **3** and corresponding atoms of inhibitors **1** and **4** are colored. RMSD traces are shown for the entire molecule (colored) and the more rigid part only (black). (c) Calculated ligand torsion energy distributions in the unbound (transparent) and the complexed (opaque) state. (d) Crystallographic binding modes of all inhibitors **1**–**4** in complex with ZIKV NS2B/NS3 (PDB-IDs: 6Y3B, 7ZLD, 7ZLC, and 7ZMI) with the $2F_o–F_c$ simulated annealing omit maps (yellow) are shown at $\sigma = 2.0$ for inhibitors **1** and **2**, $\sigma = 1.5$ for inhibitor **3**, and $\sigma = 1.0$ for inhibitor **4**.

2.4 | X-ray crystallography

These findings were further supported by crystal structure analysis. By the time, co-crystal structures of inhibitors **2**–**4** in complex with NS2B/NS3 could be obtained (Supporting Information: Table S6). All ligands revealed a highly similar binding mode in line with the starting structures from MD simulations. Intriguingly, lower electron densities were observed for the ligand atoms, which should remain highly flexible according to the MDs (Figure 2d). This was true when investigating the simulated annealing omit maps at $\sigma = 2$ for compounds **1** and **2** and $\sigma = 1.5$ for compound **3**. For the low-affinity ligand **4**, however, the map is shown for $\sigma = 1$. This compound is only resolved poorly, possibly due to its lower overall binding affinity and the lower crystal structure resolution (Supporting Information: Table S6). However, this torsion-centric interpretation neglects the manifold other effects of ligand binding on the thermodynamic binding profiles. Especially desolvation effects and target rigidification may have a strong impact on entropy as well. The binding modes are highly conserved between the different ligands resembling the observed interactions of the NS2B/NS3-**1** complex (Figures 1 and 2d). The low electron density of the *N*-terminal amide of inhibitors **2** and **3** and the aromatic system of **3** indicates the flexibility of these moieties. Notably, rather than being involved in direct protein–ligand interactions, these parts are oriented toward the solvent and "hovering above" the guanidinium group of the ligands (Figure 3). On the one hand, this can stabilize the ligand's binding mode via intramolecular cation–π interaction. On the other hand, the higher flexibility of that moiety can also indicate its reduced desolvation as hinted by the lower absolute ΔC_p-values of **2** and **3** compared to **1** (Table 1). Finally, **4** is the only ligand containing an acidic, negatively charged moiety reducing its potency as the active site has a highly negative potential (Supporting Information: Figure S5). Furthermore, strong interactions often result in a loss of conformational degrees of freedom for both ligands and proteins and subsequently are accompanied by an enthalpy–entropy compensation.[53,60] Therefore, B-factors as a surrogate for residue flexibility of the NS2B/NS3–inhibitor complexes were further analyzed. While raw B-factors are highly influenced by resolution, crystal packing, or refinement methods used, B-factors were normalized with the BANΔIT webserver for comparability between the different structures.[16] This analysis of normalized B′-factors indicated that some binding site residues in the complexes with inhibitors **2**–**4** retain higher flexibility (high B′-factors) while binding-site residues are slightly more rigid (lower B′-factor) for the NS2B/NS3-**1** complex. This holds true in general, but especially for NS3 residues 129–132 and Gly159 (S1 pocket), Val154, Val155, Ser160, and Tyr161 (close to the aromatic linker) and residues 83–86 of the S3 pocket formed by NS2B (Figure 3, Supporting Information: Figure S4E, Supporting Information: Table S7). All these residues are located around the termini of the acyclic ligands and the linking position of macrocycle **1**. Hence, the tight binding of inhibitor **1** might also result in a rigidification of the protein and subsequently an accompanied entropic penalty. While protein dynamics can be influenced by ligand binding in regions distant from the binding site, B'-factors showed some minor differences there as well. However, residues with the highest B′-differences are primarily located at the termini or in loop regions where flexibility is usually higher and larger deviations between different structures are observed regularly (e.g., NS2B residues 64–68 and NS3 residues 29–32, Supporting Information: Figure S4e).[16] Therefore, it can be presumed that the highly similar binding modes of **1**–**4** result in similar global effects for protein dynamics.

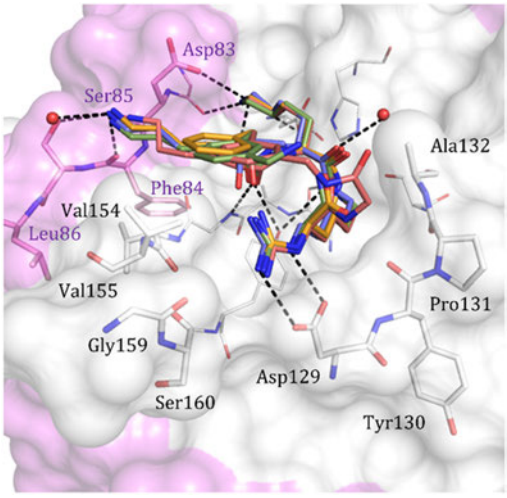


FIGURE 3 Overlay of binding modes of compounds **1**–**4** in complex with the ZIKV NS2B (magenta)/NS3 (white) protease (PDB-IDs: 6Y3B, 7ZLD, 7ZLC, and 7ZMI). Inhibitors are depicted as sticks with colored carbon atoms (**1**: salmon, **2**: green, **3**: blue, **4**: yellow). Polar interactions are shown as black dashed lines. For a clear view, only residues from the NS2B/NS3-**3** complex are shown and only those residues with differences in B'-factors depending on which ligand is bound (as described in the main text) are labeled.

3 | CONCLUSION

Cyclization and macrocyclization are common strategies in drug design and optimization. One key aspect is the generally accepted hypothesis that cyclic ligands trapped in the bioactive conformation require less structural preorganization and rigidification upon binding and hence are accompanied by a lower entropic penalty. The protonation-corrected thermodynamic binding profiles from ITC experiments of inhibitors **1**–**4** binding to the ZIKV NS2B/NS3 protease (Table 1, Supporting Information: Figure S2) revealed that this common explanation for the higher affinity of a macrocyclic compound was not valid for inhibitor **1** in comparison to inhibitors **2** and **3** and that the higher affinity of inhibitor **1** cannot be attributed to a beneficial $-T\Delta S°$ term. Both inhibitors **1** and **4** have a higher temperature-dependent binding entropy of 34.9 and 47.8 $kJ \cdot mol^{-1}$, whereas the linear compounds **2** and **3** share a less

disturbing $-T\Delta S°$ of 20.4 and 26.6 kJ·mol^{-1}, respectively. To further investigate these counterintuitive results, MD simulations and co-crystallization experiments were performed. The ligand RMSD traces and the torsion energy distributions derived from MD simulations indicated that inhibitors **2** and **3** retain some flexibility upon binding, whereas inhibitors **1** and **4** experience stronger rigidification upon complex formation. This observation was further supported by the crystallographically determined NS2B/NS3 complex structures of inhibitors **2**–**4**. Ligands **2** and **3** retain some flexibility in the complex as indicated by a lower electron density for the same atoms that appeared to be more flexible in the MDs. B'-factor analysis revealed that besides the ligand, also the protease binding site in the complex with macrocyclic ligand **1** is more rigid. While this ligand has a high binding enthalpy of $\Delta H° = -75.1$ kJ·mol^{-1}, the accompanied entropic penalty is rather large with $-T\Delta S° = 34.9$ kJ·mol^{-1}. As a word of caution, the thermodynamics of small molecules binding to proteins is way more complex and this torsion-centric evaluation neglects many additional underlying effects including slight differences in direct interactions and (de-) solvation. Additionally, large changes in heat capacity indicated that the observed effects might be shadowed by a conformational change as commonly described for the NS2B/NS3 protease. Hence, the interpretation is limited by the assumption that the high structural similarity within the described ligands results in similar direct protein–ligand interactions (Figures 2d and 3) and an equal effect on the conformational equilibrium. Subsequently, the study highlights the limitations in our understanding of how to link binding thermodynamics to structural features even for well-described systems covering ITC, crystallographic, and MD information. The interpretation can only describe a qualitative torsion effect rather than yielding quantitative energy terms to be directly correlated with ITC results. Nevertheless, the study demonstrates the possibilities of MD simulations in combination with structural biology to support the explanation of thermodynamic binding profiles for macrocycles. The implementation of additional model systems to identify common features of cyclic and acyclic ligand pairs might hold the potential to assist decision-making in prospective macrocycle design in the future.

4 | EXPERIMENTAL

4.1 | Chemistry

4.1.1 | Reagents

All reagents and solvents were purchased from Sigma-Aldrich Chemie GmbH, Thermo Fisher Scientific Inc., and Carl Roth GmbH + Co. KG.

4.1.2 | Synthesis

The synthesis and analytical characterization of inhibitors **1**–**4** and the substrate PhAc-LKKR-AMC are described in previous publications.[27,35] The InChI codes of inhibitors **1**–**4**, together with some biological activity data, are provided as the Supporting Information.

4.2 | Recombinant protein expression and purification

The bivalently expressed ZIKV protease construct NS2B/NS3 (bZiPro) for enzyme inhibition assays and ITC experiments was expressed and purified, as described previously.[26,61] Briefly, the vector (pETDuet) containing bZiPro (#86846, www.addgene.com) was transformed into competent *Escherichia coli* BL21 Gold (DE3) cells (Agilent Technologies) and grown in the LB medium containing 100 mg·L^{-1} ampicillin at 37°C until an optical density (OD_{600}) of 0.8 was reached. Overexpression was induced overnight by the addition of 1 mM isopropyl-β-D-thiogalactoside (IPTG) at 18°C for 20 h. After harvesting by centrifugation at 9000 rpm at 4°C for 15 min, cells were flash-frozen in liquid nitrogen and stored at −80°C until further use. For purification of NS2B/NS3, the cell pellet was resuspended in buffer A1 (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 0.1% (v/v) $Triton_{X\text{-}100}$, RNase, DNase, lysozyme, 1 mM DTT) and lysed by sonication (Sonoplus HD 2200). The lysate was cleared by centrifugation (20,000 rpm at 4°C for 1 h). The supernatant was subjected to immobilized metal affinity chromatography (IMAC) on a HisTrap HP 5 ml column (Cytiva Europe GmbH). After washing with buffer B1 (20 mM Tris-HCl, pH 8, 300 mM NaCl, 20 mM imidazole), the protein was eluted in a linear gradient to 250 mM imidazole. The His_6-tag was removed by thrombin protease cleavage during a dialysis step against buffer C1 (20 mM Tris-HCl, pH 8, 150 mM NaCl) overnight (4°C). After reverse IMAC to remove the His_6-tagged thrombin protease (kindly provided by Prof. Dr. Ute Hellmich, Institute of Organic Chemistry & Macromolecular Chemistry, Friedrich Schiller University Jena), NS2B/NS3 was further purified using a size exclusion chromatography step with a HiLoad 16/600 Superdex 75 column (GE Healthcare) in buffer C1. Collected fractions containing NS2B/NS3 were concentrated using Vivaspin 10 MWCO spin concentrators (Sartorius AG), flash-frozen in liquid nitrogen, and stored at −80°C. Protein identity and purity were confirmed by SDS-PAGE with Coomassie blue staining.

bZiPro for crystallographic experiments was expressed and purified in a similar manner with minor differences: Cell pellets were lysed with buffer A2 (20 mM sodium phosphate, pH 8.0, 500 mM NaCl, 2 mM β-mercaptoethanol, 10 mM imidazole, and 5% (v/v) glycerol). Nickel nitrilotriacetic acid (Ni-NTA) chromatography was performed using a HisTrap FF crude column (Cytiva Europe GmbH). After washing with buffer containing 20–30 mM imidazole, the protein was eluted with buffer B2 containing 150 mM imidazole. The His_6-tag was removed by the addition of thrombin, followed by dialysis of the mixture against buffer C2 (20 mM Na-HEPES, pH 7.5, 150 mM NaCl, 5% [v/v] glycerol, and 2 mM dithiothreitol [DTT]) for 20 h. After dialysis, the protein was further purified by size-exclusion chromatography using a HiLoad 26/600 Superdex 200 pg column (GE Healthcare).

4.3 | Enzyme inhibition assay

Fluorometric enzyme activity assays were performed on a Spark® (Tecan) microplate reader using white flat-bottom 96-well

microtiter plates (Greiner Bio-One). The substrate PhAc-Leu-Lys-Lys-Arg-7-amido-4-methylcoumarin (PhAc-LKKR-AMC) resembles the preferred cleavage site and releases fluorogenic AMC after cleavage. Assay buffer (180 µl, 20 mM TRIS [pH 8.5], 10% [v/v] Glycerol, 0.01% [v/v] $Triton_{X\text{-}100}$, 2 mM DTT) was supplemented with 10 µl of aqueous inhibitor solution in a half logarithmic dilution series (final concentrations ranging from >25·IC_{50} to <0.35·IC_{50}) and 5 µl of NS2B/NS3 to give a final protein concentration of 2 nM. The reaction was started by the addition of 5 µl of substrate solution resulting in a 10 µM final concentration. Fluorescence was monitored for 10 min at 25°C with excitation at 380 nm and detection at 460 nm. For the determination of the substrate K_M-value, 10 µl of H_2O instead of the inhibitor solution was added. The PhAc-LKKR-AMC substrate was used in eight different concentrations (100, 50, 25, 10, 5, 2.5, 1.0, and 0 µM; Supporting Information: Figure S6). IC_{50}-values and enzyme kinetic parameters were calculated from technical triplicates of each concentration (one dilution series of inhibitor, one batch of enzyme, and three wells containing the same reaction) using GRAFIT (Version 5.0.13; Erithacus Software Limited). K_i values for inhibitors were calculated from IC_{50} and the substrate K_M (3.31 ± 0.27 µM) using the Cheng–Prusoff equation.[62]

4.4 | Isothermal titration calorimetry

ITC experiments in technical triplicates (at least 3 ITC experiments of the same inhibitor and enzyme batches, although several enzyme expressions and purifications were conducted to cover all experiments presented) were performed on a MicroCal PEAQ-ITC automated system (Malvern Instruments). Experiments were performed in assay buffer (20 mM TRIS [pH 8.5], 10% [v/v] Glycerol, 0.01% [v/v] $Triton_{X\text{-}100}$, 2 mM DTT) supplemented with 5% (v/v) DMSO to better match the conditions of the enzyme inhibition assay. 300–600 µM of the ligands **1**–**3** was titrated to 15–20 µM NS2B/NS3 if not stated elsewhere (Supporting Information: Figure S7). Since the K_d value for inhibitor **4** binding to NS2B/NS3 was too low to give a sufficient slope at the inflection point, a low-c titration[63] was performed with a 20-fold molar excess of inhibitor **4** in the syringe resulting in 30–50 µM NS2B/NS3 and 1–2 mM inhibitor **4** if not stated elsewhere (Supporting Information: Figure S7). For those measurements, the stoichiometry was set to 1.0. Control experiments (buffer vs. buffer, buffer vs. titrant, titrand vs. buffer) were subtracted from the raw data to correct offset. Experiments were performed at 25°C, the stirring speed was set to 750 rpm, and the reference power to 41.9 µW. Nineteen injections à 2 µl were added to the reaction cell with a duration of 4 s and a spacing time of 150 s. Data integration and evaluation were performed using the MicroCal PEAQ-ITC analysis software (Vers. 1.21, Malvern Panalytical Ltd). Corrections for the buffer ionization enthalpies resulting from proton transitions upon NS2B/NS3 binding were performed in assay buffer containing HEPES or BICINE instead of TRIS. Determination of changes in isobaric heat capacity upon binding (ΔC_P) was performed in HEPES buffers at two additional temperatures (15°C and 35°C).

4.5 | MD simulations

MD simulations were performed starting from the ZIKV NS2B/NS3-**1** complex structure (PDB-ID 6Y3B).[27] An apo structure was generated by the removal of the ligand, while complexes with the ligands **2**–**4** were generated by manipulation (removal or addition of atoms) of the reference ligand **1** within the complex and subsequent energy minimization of the newly introduced ligand atoms within MOE[64] using the Merck Molecular Force Field (MMFF94).[65] The ligands were parameterized using the Generalized Amber Force Field (GAFF2[66]) with AM1-BCC[67] charges within antechamber[68] of the AmberTools20.[69] Complex structures were subsequently built with tleap[69] including crystallographic solvent molecules. After initial relaxation over 200 time steps with sander, counter ions (Na^+ for the protein–ligand complexes, Cl^- for ligand simulations) were added and a TIP3P[70] waterbox exceeding the structure by 10.0 Å was built with tleap. MDs were performed using NAMD2.14[71] and the AMBER forcefield (ff14SB[72]). The system was equilibrated over 1 ns by heating from 100 to 300 K over 500 ps and releasing harmonic constraints on the protein and ligand atoms over the following 500 ps in a constant volume box. MD production runs were performed over 10 ns with an NPT ensemble using periodic boundary conditions and a van der Waals cut-off of 14.0 Å with time steps of 2 fs allowing rigid bond lengths. For torsion energy analysis, ligand simulations without protein were conducted under the same conditions. Trajectories were written every ps and concatenated to include every 10th frame with catdcd 4.0 to result in 1000 frames per 10 ns prior analysis in VMD-1.9.3.[73] Order parameters S^2 were calculated using the isotropic reorientational eigenmode dynamics (iRED) approach within cpptraj[74] of the AmberTools20.[69] Figures were created with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).

4.6 | Crystallization and structure determination

The bZiPro/inhibitor complexes (molar ratio 1:12) were incubated for 1 h on ice at a protein concentration of 40 mg/ml. Two microliters of the bZiPro/inhibitor mixture was added to 2 µl of the reservoir solution and incubated at 18°C in a hanging drop vapor diffusion experiment. The reservoir solution consisted of 0.1 M sodium acetate pH 4.6, 0.2 M ammonium sulfate, and 16%–19% PEG 2000. The PEG 2000 concentration was optimized separately for each inhibitor. Crystals appeared after 1 day and were cryoprotected using 30% PEG 2000 before being flash-frozen in liquid nitrogen. For inhibitor **2**, macroseeding was performed to obtain sufficiently diffracting crystals. Crystals of bZiPro in complex with inhibitor **2** were transferred into a fresh drop consisting of 2 µl

of the bZiPro/inhibitor mixture and 2 µl of the reservoir solution for 1 day. The new reservoir solution contained 2% less PEG 2000 compared to the initial reservoir solution.

Diffraction intensities were collected at BESSY MX beamline 14.1. Data processing and scaling were performed with the XDS program package.[75] Structure determination was done by molecular replacement with PHASER MR using the bZiPro structure with the PDB code 5GPI as a model.[76] Simulated annealing was performed within the initial refinement with PHENIX to remove potential bias from the search model.[77,78] The structures were subjected to alternating rounds of manual rebuild using Coot[79] to fit amino acids into σ-weighted $2F_o–F_c$ and $F_o–F_c$ electron density maps and the PHENIX refine program (five cycles). For the calculation of R_{free}, 5% of all data were randomly chosen and were not considered during refinement. Isotropic B-factor refinement with TLS parameters was used. Water and inhibitor molecules were located in the electron density and gradually added to the model. Multiple conformations were built if the minor populated conformation showed at least 20% occupancy. Coordinates and restraints for the inhibitors were generated with the Grade Web server.[80] The structures were deposited in the Protein Data Bank with accession codes 7ZLD (**2**), 7ZLC (**3**), and 7ZMI (**4**). The final data collection and refinement statistics are summarized in Supporting Information: Table S7.

B-factors were normalized with the BANΔIT-webserver[16] to analyze differences in protein binding-site flexibility in complex with the ligands **1**–**4** using the modified z-score method for C_α atoms of the protein backbone.

ACKNOWLEDGMENTS

The authors thank the Group of Prof. Dahai Luo for sharing the pETDuet-1 vector containing NS2B/NS3 (bZiPro) via Addgene (Watertown). They thank Anna Riede for excessive purifications of NS2B/NS3 and Prof. Dr. Tanja Schirmeister for access to facilities and instrumentation as well as scientific discussion on the project. The authors thank M. ed. Thomas Steinfurth for language proofreading of the manuscript. Parts of this research were conducted using the supercomputer Mogon and/or advisory services offered by Johannes Gutenberg University Mainz (hpc.uni-mainz.de), which is a member of the AHRP (Alliance for High Performance Computing in Rhineland Palatinate, www.ahrp.info) and the Gauss Alliance e.V. The authors gratefully acknowledge the computing time granted on the supercomputer Mogon at Johannes Gutenberg University Mainz (hpc. uni-mainz. de). The authors gratefully acknowledge the provision of synchrotron beamline at the Helmholtz-Zentrum Berlin (HZB, BL 14.1) and thank the scientific beamline staff at BESSY II for providing them with outstanding support during data collection. The authors thank the German Research Council (DFG) for supporting this work through the grant "DFG-Großgeräteantrag ITC (INST 247/921-1 FUGG)." The authors are also grateful to Helmholtz-Zentrum Berlin (HZB) for financial support for travel costs. Torsten Steinmetzer obtained funding from the LOEWE Center DRUID (Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious Diseases). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Christian Kersten http://orcid.org/0000-0001-9976-7639

REFERENCES

[1] G. S. Campos, A. C. Bandeira, S. I. Sardi, *Emerg. Infect. Dis.* **2015**, *21*, 1885.
[2] G. Kuno, *Mol. Detect. Hum. Viral Pathog.* **2016**, *503*, 313.
[3] M. Hennessey, M. Fischer, J. E. Staples, *MMWR. Morb. Mortal. Wkly. Rep.* **2016**, *65*, 1.
[4] A. N. Hazin, A. Poretti, D. Di Cavalcanti Souza Cruz, M. Tenorio, A. van der Linden, L. J. Pena, C. Brito, L. Gil, D. de Barros Miranda-Filho, E. Marques, C. M. Turchi Martelli, J. Alves, T. A. Huisman, *N. Engl. J. Med.* **2016**, *374*, 2193.
[5] E. D'Ortenzio, S. Matheron, X. de Lamballerie, B. Hubert, G. Piorkowski, M. Maquart, D. Descamps, F. Damond, Y. Yazdanpanah, I. Leparc-Goffart, *N. Engl. J. Med.* **2016**, *374*, 2195.
[6] J. F. W. Chan, G. K. Y. Choi, C. C. Y. Yip, V. C. C. Cheng, K. Y. Yuen, *J. Infect.* **2016**, *72*, 507.
[7] 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, *N. Engl. J. Med.* **2016**, *374*, 951.
[8] M. R. Duffy, T.-H. Chen, W. T. Hancock, A. M. Powers, J. L. Kool, R. 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, *N. Engl. J. Med.* **2009**, *360*, 2536.
[9] 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, *Emerg. Infect. Dis.* **2017**, *23*, 669.
[10] V. M. Cao-Lormeau, A. Blake, S. Mons, S. Lastère, C. Roche, J. Vanhomwegen, T. Dub, L. Baudouin, A. Teissier, P. Larre, A. L. Vial, C. Decam, V. Choumet, S. K. Halstead, H. J. Willison, L. Musset, J. C. Manuguerra, P. Despres, E. Fournier, H. P. Mallet, D. Musso, A. Fontanet, J. Neil, F. Ghawché, *Lancet* **2016**, *387*, 1531.
[11] B. D. Lindenbach, H.-J. Theil, C. M. Rice, *Fields Virol.* **2007**, *5*, 1101.
[12] S. Voss, C. Nitsche, *Bioorg. Med. Chem. Lett.* **2020**, *30*, 126965.
[13] É. E. da Silva Rodrigues, H. Maus, S. J. Hammerschmidt, A. Ruggieri, E. C. dos Santos, Ê. J. Bassi, L. Anderson, P. G. V. Aquino, J. X. de Araújo-Júnior, F. Wei, X. Liu, P. Zhan, T. Schirmeister, E. F. da Silva-Júnior, *Human Viruses: Diseases, Treatments and Vaccines* (Ed: S. I. Ahmad), Springer International Publishing, **2021**, 233.
[14] 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.
[15] H. Maus, F. Barthels, S. J. Hammerschmidt, K. Kopp, B. Millies, A. Gellert, A. Ruggieri, T. Schirmeister, *Bioorg. Med. Chem.* **2021**, *47*, 116392.
[16] F. Barthels, T. Schirmeister, C. Kersten, *Mol. Inform.* **2021**, *40*, 1.
[17] O. Carugo, *Acta Crystallogr. D Struct. Biol.* **2022**, *78*, 69.
[18] A. Sandner, T. Hüfner-Wulsdorf, A. Heine, T. Steinmetzer, G. Klebe, *J. Med. Chem.* **2019**, *62*, 9753.
[19] W. W. Phoo, Z. Zhang, M. Wirawan, E. J. C. Chew, A. B. L. Chew, J. Kouretova, T. Steinmetzer, D. Luo, *Antiviral Res.* **2018**, *160*, 17.
[20] 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, *Nat. Commun.* **2016**, *7*, 13410.

[21] 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.
[22] 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.
[23] M. E. Hill, M. Yildiz, J. A. Hardy, *Biochemistry* **2019**, *58*, 776.
[24] 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.
[25] E. F. da Silva-Júnior, J. X. de Araújo-Júnior, *Bioorg. Med. Chem.* **2019**, *27*, 3963.
[26] S. Huber, N. J. Braun, L. C. Schmacke, J. P. Quek, R. Murra, D. Bender, E. Hildt, D. Luo, A. Heine, T. Steinmetzer, *J. Med. Chem.* **2022**, *65*, 6555.
[27] 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.
[28] C. Nitsche, H. Onagi, J. P. Quek, G. Otting, D. Luo, T. Huber, *Org. Lett.* **2019**, *21*, 4709.
[29] A. Poulsen, C. Kang, T. Keller, *Curr. Pharm. Des.* **2014**, *20*, 3422.
[30] M. A. M. Behnam, D. Graf, R. Bartenschlager, D. P. Zlotos, C. D. Klein, *J. Med. Chem.* **2015**, *58*, 9354.
[31] H. A. Lim, J. Joy, J. Hill, C. San Brian Chia, *Eur. J. Med. Chem.* **2011**, *46*, 3130.
[32] M. J. Stoermer, K. J. Chappell, S. Liebscher, C. M. Jensen, C. H. Gan, P. K. Gupta, W. J. Xu, P. R. Young, D. P. Fairlie, *J. Med. Chem.* **2008**, *51*, 5714.
[33] J. J. H. Chu, R. C. H. Lee, M. J. Y. Ang, W.-L. Wang, H. A. Lim, J. L. K. Wee, J. Joy, J. Hill, C. S. B. Chia, *Antiviral Res.* **2015**, *118*, 68.
[34] H. M. Berman, *Nucleic Acids Res.* **2000**, *28*, 235.
[35] M. Z. Hammamy, C. Haase, M. Hammami, R. Hilgenfeld, T. Steinmetzer, *ChemMedChem* **2013**, *8*, 231.
[36] E. H. Rühmann, M. Rupp, M. Betz, A. Heine, G. Klebe, *ChemMedChem* **2016**, *11*, 309.
[37] J. E. DeLorbe, J. H. Clements, M. G. Teresk, A. P. Benfield, H. R. Plake, L. E. Millspaugh, S. F. Martin, *J. Am. Chem. Soc.* **2009**, *131*, 16758.
[38] A. P. Benfield, M. G. Teresk, H. R. Plake, J. E. DeLorbe, L. E. Millspaugh, S. F. Martin, *Angew. Chem. Int. Ed.* **2006**, *45*, 6830.
[39] B. Wienen-Schmidt, H. R. A. Jonker, T. Wulsdorf, H. D. Gerber, K. Saxena, D. Kudlinzki, S. Sreeramulu, G. Parigi, C. Luchinat, A. Heine, H. Schwalbe, G. Klebe, *J. Med. Chem.* **2018**, *61*, 5922.
[40] J. E. Delorbe, J. H. Clements, B. B. Whiddon, S. F. Martin, *ACS Med. Chem. Lett.* **2010**, *1*, 448.
[41] B. P. Morgan, D. R. Holland, B. W. Matthews, P. A. Bartlett, *J. Am. Chem. Soc.* **1994**, *116*, 3251.
[42] R. Derda, M. R. Jafari, *Prot. Pept. Lett.* **2019**, *25*, 1051.
[43] G. Ruiz-Gómez, J. D. A. Tyndall, B. Pfeiffer, G. Abbenante, D. P. Fairlie, *Chem. Rev.* **2010**, *110*, PR1.
[44] J. J. Li, D. D. Holsworth, L. Y. Hu, *Chemtracts* **2003**, *16*, 439.
[45] J. R. Frost, C. C. G. Scully, A. K. Yudin, *Nat. Chem.* **2016**, *8*, 1105.
[46] D. Yang, A. Singh, H. Wu, R. Kroe-Barrett, *Anal. Biochem.* **2016**, *508*, 78.
[47] C. Velours, M. Aumont-Nicaise, S. Uebel, P. England, A. Velazquez-Campoy, D. Stroebel, G. Bec, P. Soule, C. Quétard, C. Ebel, A. Roussel, J. B. Charbonnier, P. F. Varela, *Eur. Biophys. J.* **2021**, *50*, 313.
[48] A. Cooper, *Biophys. Chem.* **2005**, *115*, 89.
[49] A. Velazquez-Campoy, N. Markova, *Malvern Panalytical: Knowledge Center* **2015**, 1.
[50] S. Vega, O. Abian, A. Velazquez-Campoy, *Biochim. Biophys. Acta.* **2016**, *1860*, 868.
[51] G. A. Holdgate, A. Tunnicliffe, W. H. J. Ward, S. A. Weston, G. Rosenbrock, P. T. Barth, I. W. F. Taylor, R. A. Pauptit, D. Timms, *Biochemistry* **1997**, *36*, 9663.
[52] E. Freire, J. Gómez, *J. Mol. Biol.* **1995**, *252*, 337.
[53] A. Cooper, C. M. Johnson, J. H. Lakey, M. Nöllmann, *Biophys. Chem.* **2001**, *93*, 215.
[54] A. Cooper, *J. Phys. Chem. Lett.* **2010**, *1*, 3298.
[55] D. L. Cameron, J. Jakus, S. R. Pauleta, G. W. Pettigrew, A. Cooper, *J. Phys. Chem. B* **2010**, *114*, 16228.
[56] G. Cui, J. M. Swails, E. S. Manas, *J. Chem. Theory Comput.* **2013**, *9*, 5539.
[57] D. W. Li, R. Brüschweiler, *J. Am. Chem. Soc.* **2009**, *131*, 7226.
[58] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4546.
[59] H. Stöckmann, A. Bronowska, N. R. Syme, G. S. Thompson, A. P. Kalverda, S. L. Warriner, S. W. Homans, *J. Am. Chem. Soc.* **2008**, *130*, 12420.
[60] J. D. Dunitz, *Chem. Biol.* **1995**, *2*, 709.
[61] Z. Zhang, Y. Li, Y. R. Loh, W. W. Phoo, A. W. Hung, C. Kang, D. Luo, *Science* **2016**, *354*, 1597.
[62] C. Yung-Chi, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099.
[63] Y.-L. Zhang, Z.-Y. Zhang, *Anal. Biochem.* **1998**, *261*, 139.
[64] I. Tubert-Brohman, W. Sherman, M. Repasky, T. Beuming, *J. Chem. Inf. Model.* **2013**, *53*, 1689.
[65] T.a Halgren, *J. Comput. Chem.* **1996**, *17*, 490.
[66] D. Vassetti, M. Pagliai, P. Procacci, *J. Chem. Theory Comput.* **2019**, *15*, 1983.
[67] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2002**, *23*, 1623.
[68] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **2006**, *25*, 247.
[69] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668.
[70] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
[71] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781.
[72] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696.
[73] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33.
[74] D. R. Roe, T. E. Cheatham, *J. Chem. Theory Comput.* **2013**, *9*, 3084.
[75] W. Kabsch, *Int. Tables Crystallogr.* **2006**, *F*, 218.
[76] A. J. McCoy, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *63*, 32.
[77] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213.
[78] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68*, 352.
[79] P. Emsley, K. Cowtan, *Acta Crystallogr. D. Biol. Crystallogr.* **2004**, *60*, 2126.
[80] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein, G. Bricogne, Cambridge, United Kingdom, Global Phasing Ltd. **2011**. https://www.globalphasing.com

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: S. J. Hammerschmidt, S. Huber, N. J. Braun, M. Lander, T. Steinmetzer, C. Kersten, *Arch. Pharm.* **2022**, e2200518. https://doi.org/10.1002/ardp.202200518

Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogues.

Stefan J. Hammerschmidt[1], Simon Huber[2], Niklas J. Braun[2], Marc Lander[1], Torsten Steinmetzer[2], Christian Kersten[1,*]

1 Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany
2 Institute of Pharmaceutical Chemistry, Philipps-University, Marbacher Weg 6, 35032 Marburg, Germany

*Correspondence:
Dr. Christian Kersten. Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany
Email: kerstec@uni-mainz.de

CONTENT

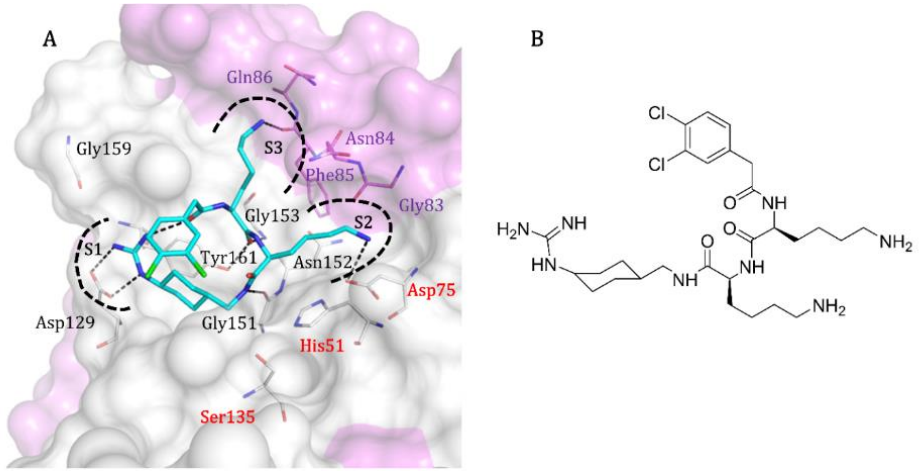


FIGURE S1 Structure and binding mode of a linear reference compound in complex with the West-Nile NS2B/NS3 protease (PDB: 2YOL). **A** The inhibitor is depicted in cyan as sticks. For a clear view, only amino acids forming polar interactions in Figure 1 (black dashed lines) are shown as lines with violet (NS2B) or white carbon atoms (NS3). Amino acids forming the catalytic triad are labelled in red. Figure was created using *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC*. The overall binding mode is similar to the one of inhibitor **1** in complex with ZIKV NS2B/NS3 (Figure 1). In this structure, however, no interacting water molecules are resolved probably due to the low resolution of 3.2 Å. **B** 2D structure of the linear inhibitor used for cocrystallization shown in panel A.

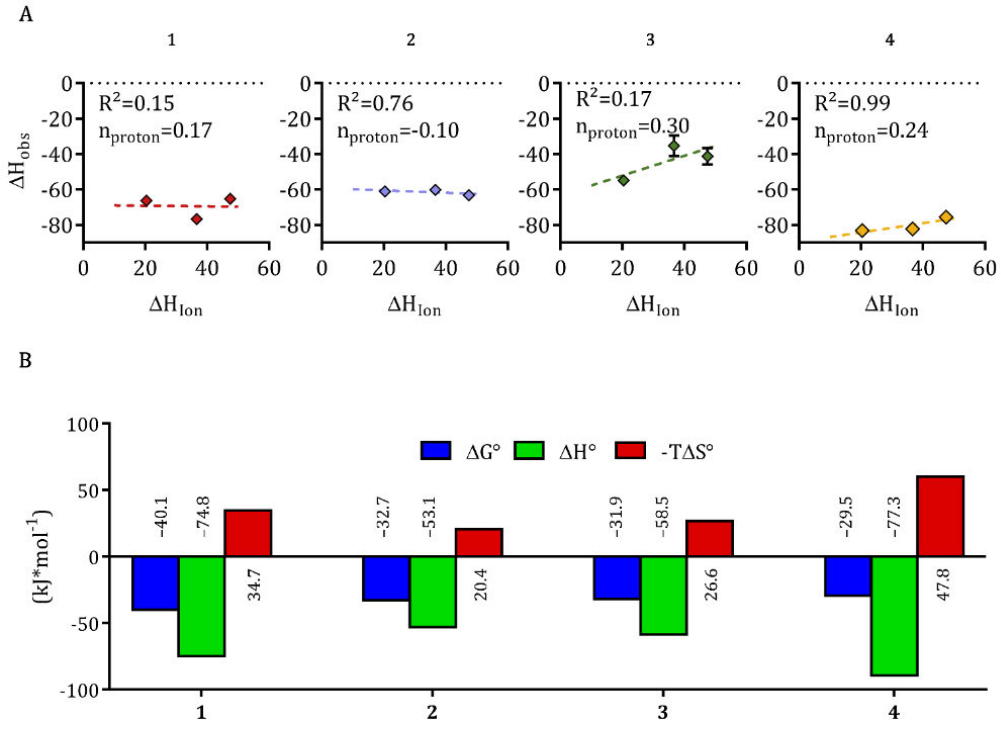


FIGURE S2 Buffer ionization enthalpy correction. ITC experiments for each inhibitor were performed in HEPES, BICINE and TRIS buffered solutions. **A** The plot of observed ΔH (ΔH_{obs}) against the ionization enthalpy of each buffer reveals the number of transferred protons (n_{proton}) as the slope of the regression line. Whereas for **1** and **2**, some fluctuations in ΔH° were observed, there is a good regression for **3** and **4**. Altogether there are no integer proton transfers (n_{proton} ≤0.3) detectable. **B** The signature plots of the buffer ionization corrected binding events show the contributions of binding enthalpy (ΔH°, green) and binding entropy (–TΔS°, red) to overall ΔG° (blue). Graphs were generated using GraphPad Prism version 7.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

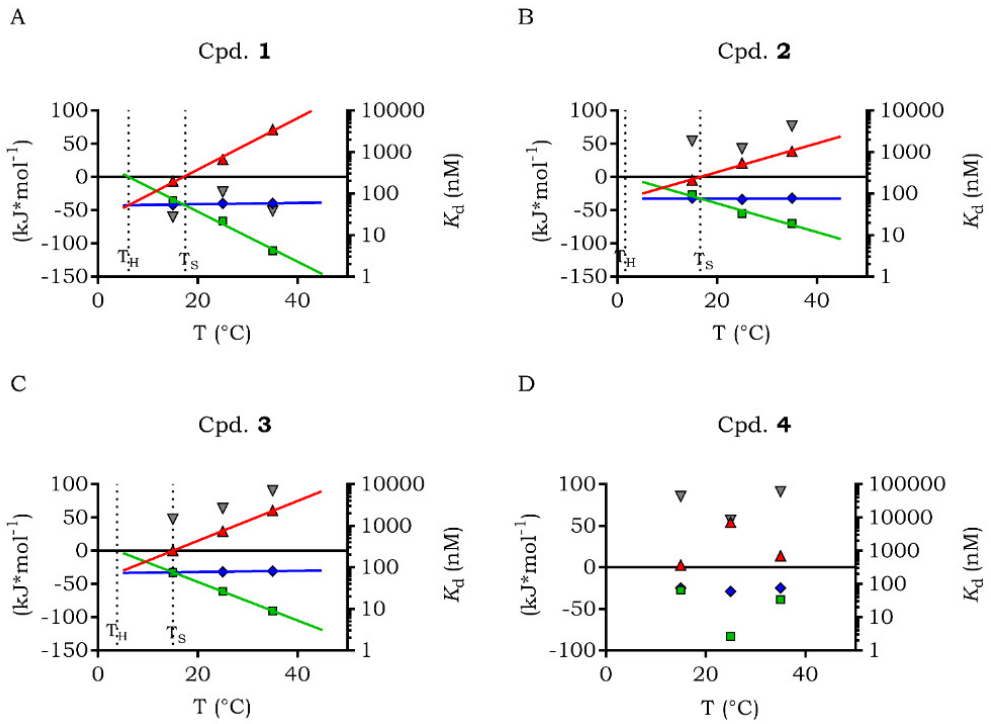


FIGURE S3 Temperature dependency plots of inhibitors **1**–**4**. ITC experiments for each inhibitor were performed in HEPES at 15 °C, 25 °C and 35 °C to calculate ΔC_P as the slope of the ΔH regression (green). T_H and T_S are the temperatures with ΔH_{obs} or $-T\Delta S_{obs}$ (red) = 0. ΔG is depicted in blue, K_d values are given in gray. GraphPad Prism version 7.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. **A-D** Temperature dependency plots of compounds **1**–**4**. Since **4** did not show a linear correlation between ΔH and T, ΔC_P was not derived.

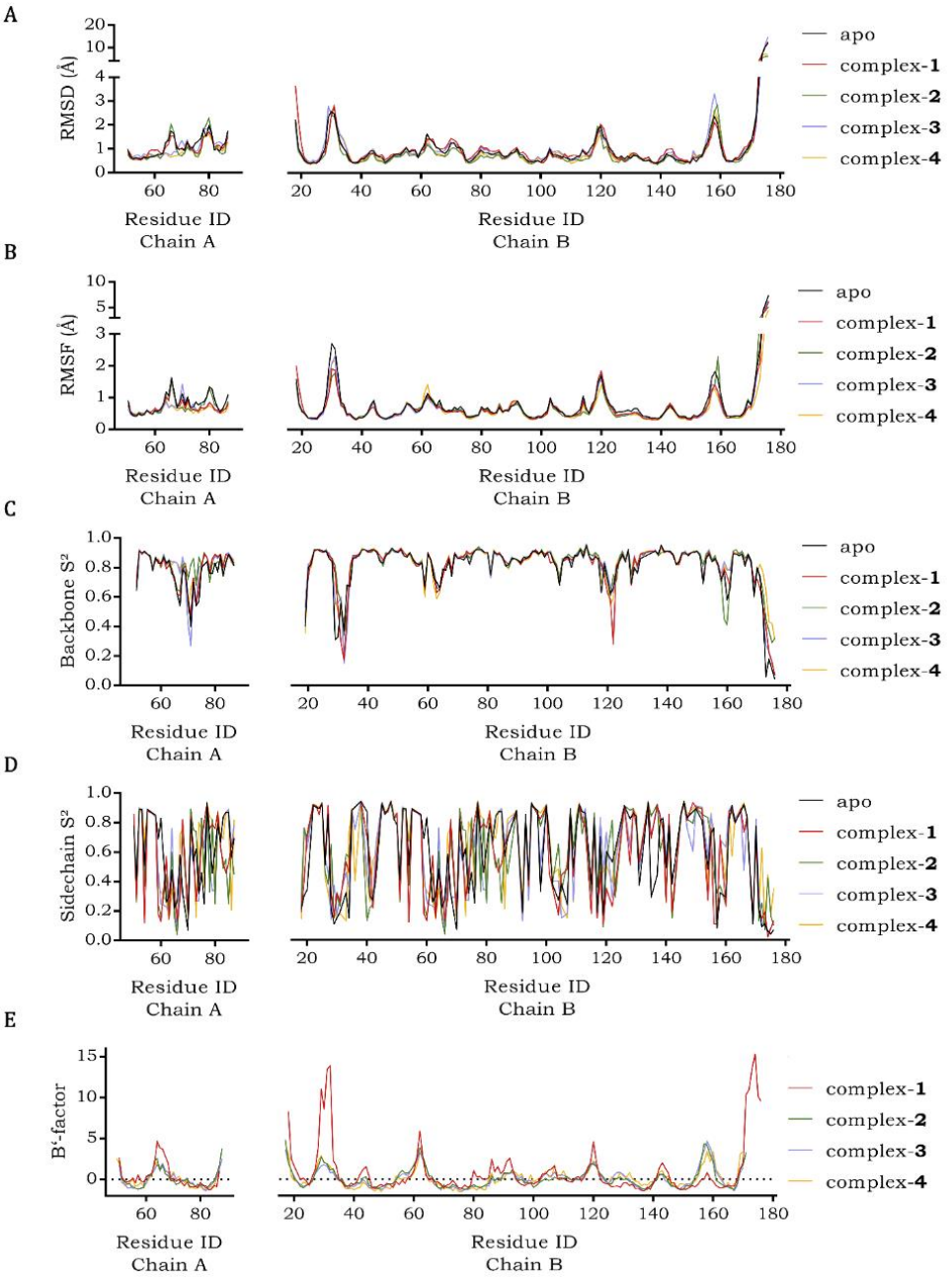


FIGURE S4 Analysis of protein dynamics per residue of 10 ns MD simulations and crystal structure analysis. A) MD-derived root-mean-square deviation (RMSD) values. B) MD-derived root-mean-square fluctuation (RMSF) values. C) MD-derived backbone order parameters (S^2). D) MD-derived sidechain S^2. E) Normalized order parameters (B') from crystal structure analysis.

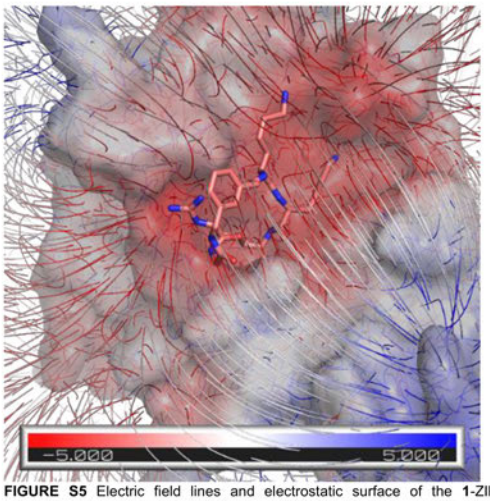
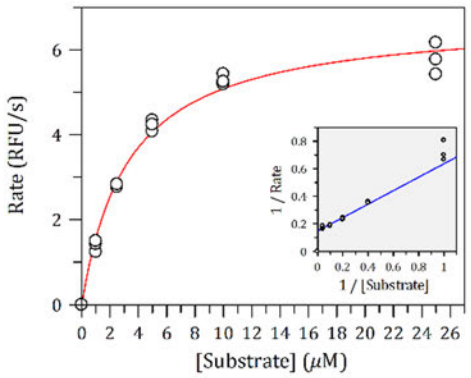


FIGURE S5 Electric field lines and electrostatic surface of the **1**-ZIKV NS2B/NS3 (bZiPro) complex. The electrostatic surface (red to blue from –5.000 to 5.000 k_BT/e_c) of the protease demonstrates the highly negatively charged substrate binding site of ZIKV NS2B/NS3 protease, exemplarily shown for the complex structure with inhibitor **1** (shown with sticks and carbon atoms in pink, PDB-ID: 6Y3B). Calculated with the Adaptive Poisson-Boltzmann Solver (APBS[1]) plugin within PyMOL.



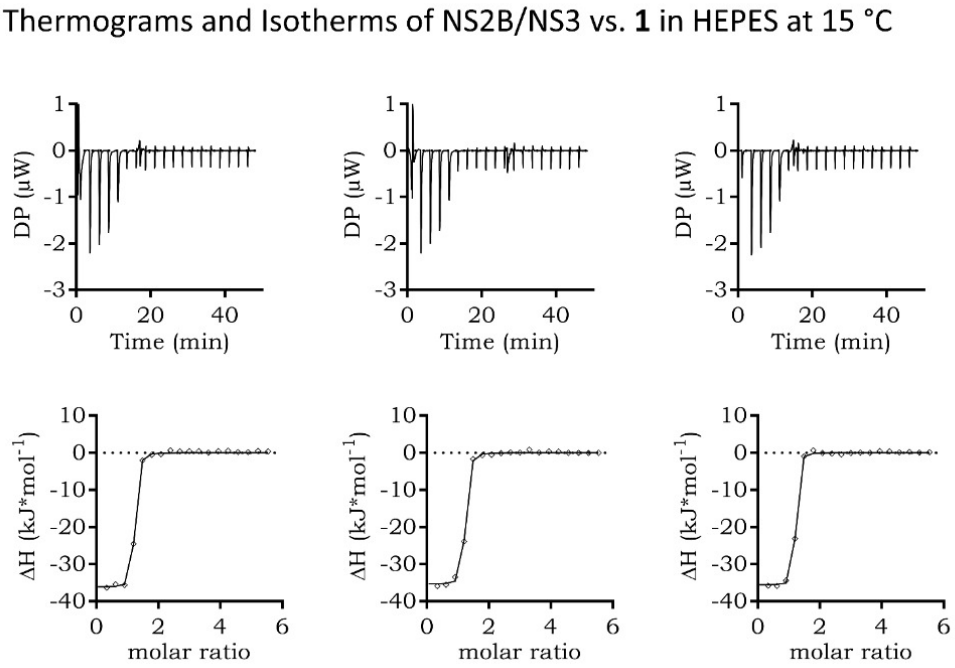
Parameter	Value	Std. Error
Vmax	6.7619	0.1767
Km	3.3074	0.2735

FIGURE S6 K_M-determination of the ZIKV NS2B/NS3 (bZiPro)-catalyzed cleavage of substrate PhAc-LKKR-AMC. In the gray colored graph, a linearized Lineweaver-Burk plot is given. Figure generated with GRAPHIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, UK).

K_M values were obtained by fitting data to the Michaelis-Menten equation (1)

$$V = \frac{V_{max}[S]}{K_M+[S]} \qquad (1)$$

with v (ΔF/min) as the substrate hydrolysis rate V_{max} as the maximum slope of the dose−response curve and the substrate concentration [S].[2] Since the high substrate concentrations of 50 and 100 µM displayed substrate inhibition of the enzyme, similar to what was reported for the closely related West-Nile virus NS2B/NS3 protease, those values were neglected in the K_M-determinations.[3]



HEPES_bZiPro_**1**_15deg_1
Model: One Set of Sites
[Cell] (M) = 15.0e-6
[Syr] (M) = 450e-6
N (sites) = 1.12 ± 4.6e-3
KD (M) = 30.2e-9 ± 6.67e-9
ΔH (kJ/mol) = -36.2 ± 0.248
Offset (kJ/mol) = 0
ΔG (kJ/mol) = -41.5
-TΔS (kJ/mol) = -5.37

HEPES_bZiPro_**1**_15deg_2
Model: One Set of Sites
[Cell] (M) = 15.0e-6
[Syr] (M) = 450e-6
N (sites) = 1.12 ± 5.6e-3
KD (M) = 33.8e-9 ± 8.54e-9
ΔH (kJ/mol) = -35.4 ± 0.295
Offset (kJ/mol) = 0
ΔG (kJ/mol) = -41.3
-TΔS (kJ/mol) = -5.85

HEPES_bZiPro_**1**_15deg_3
Model: One Set of Sites
[Cell] (M) = 15.0e-6
[Syr] (M) = 450e-6
N (sites) = 1.11 ± 3.9e-3
KD (M) = 17.2e-9 ± 5.94e-9
ΔH (kJ/mol) = -35.6 ± 0.222
Offset (kJ/mol) = 0
ΔG (kJ/mol) = -42.9
-TΔS (kJ/mol) = -7.29

FIGURE S7A Thermograms and isotherms of NS2B/NS3 vs. **1** in HEPES at 15 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

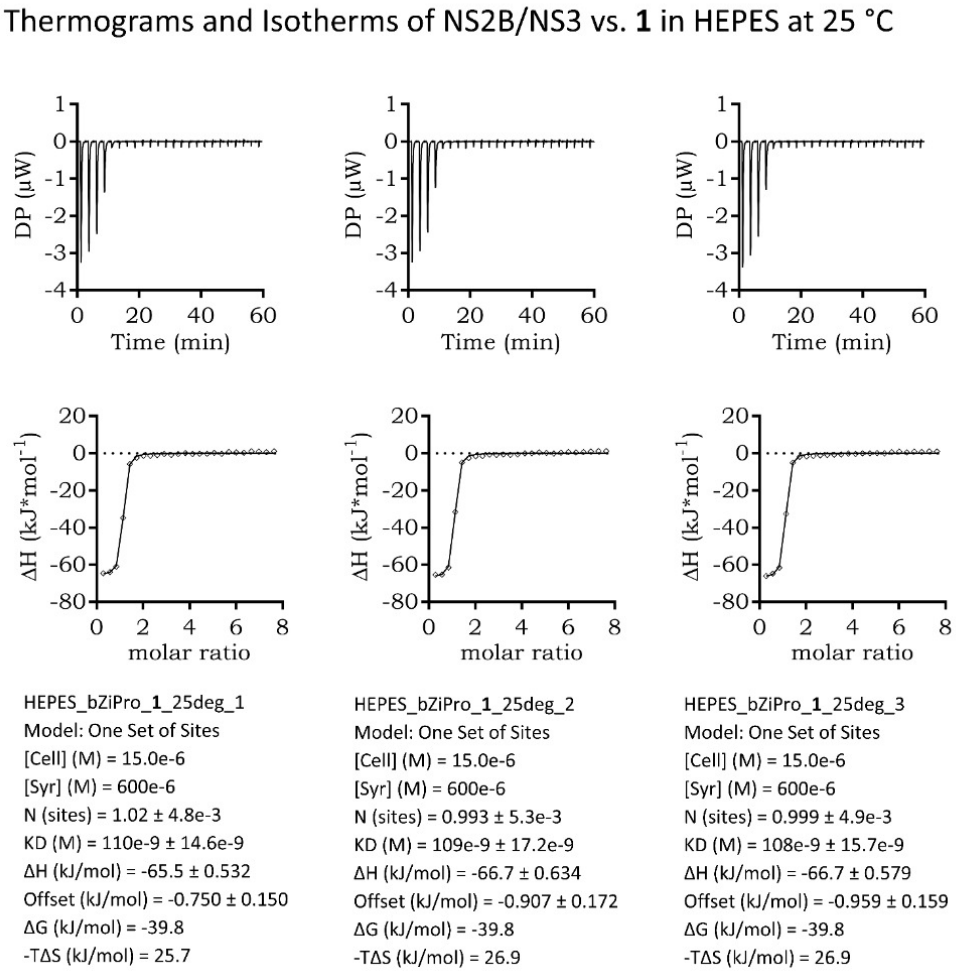


FIGURE S7B Thermograms and isotherms of NS2B/NS3 vs. **1** in HEPES at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

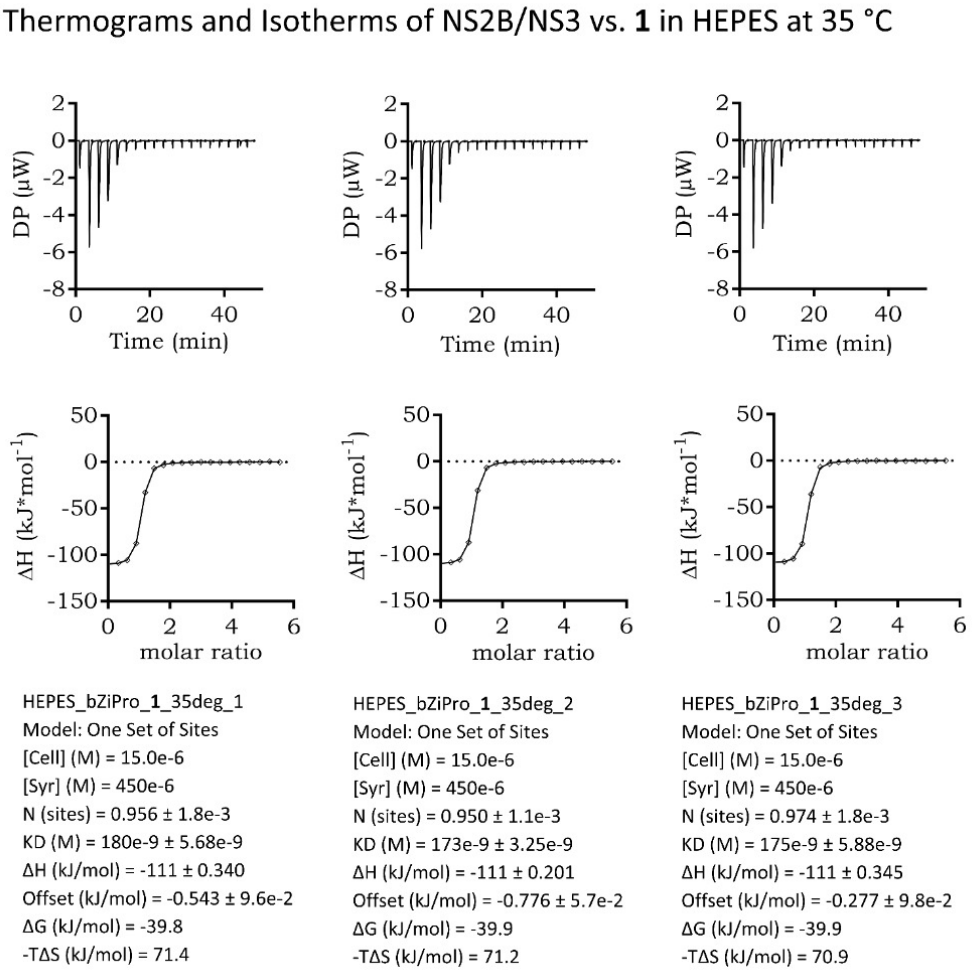


FIGURE S7C Thermograms and isotherms of NS2B/NS3 vs. **1** in HEPES at 35 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

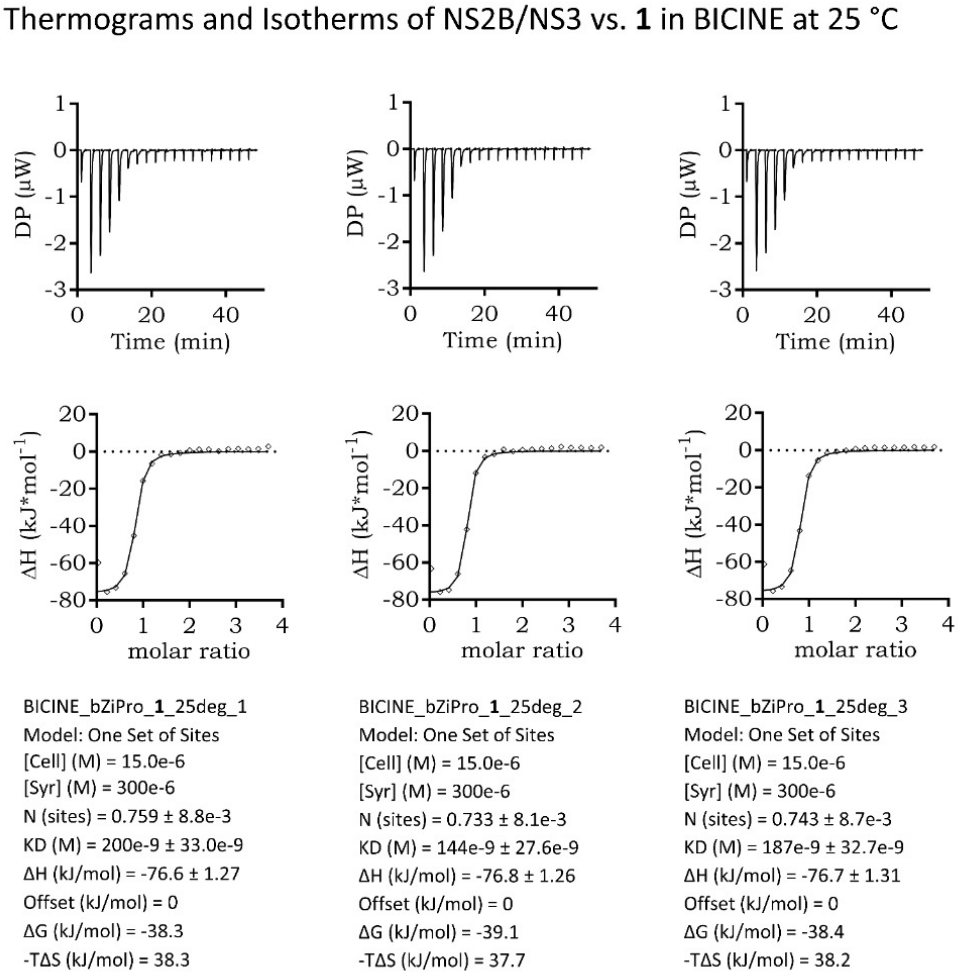


FIGURE S7D Thermograms and isotherms of NS2B/NS3 vs. **1** in BICINE at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

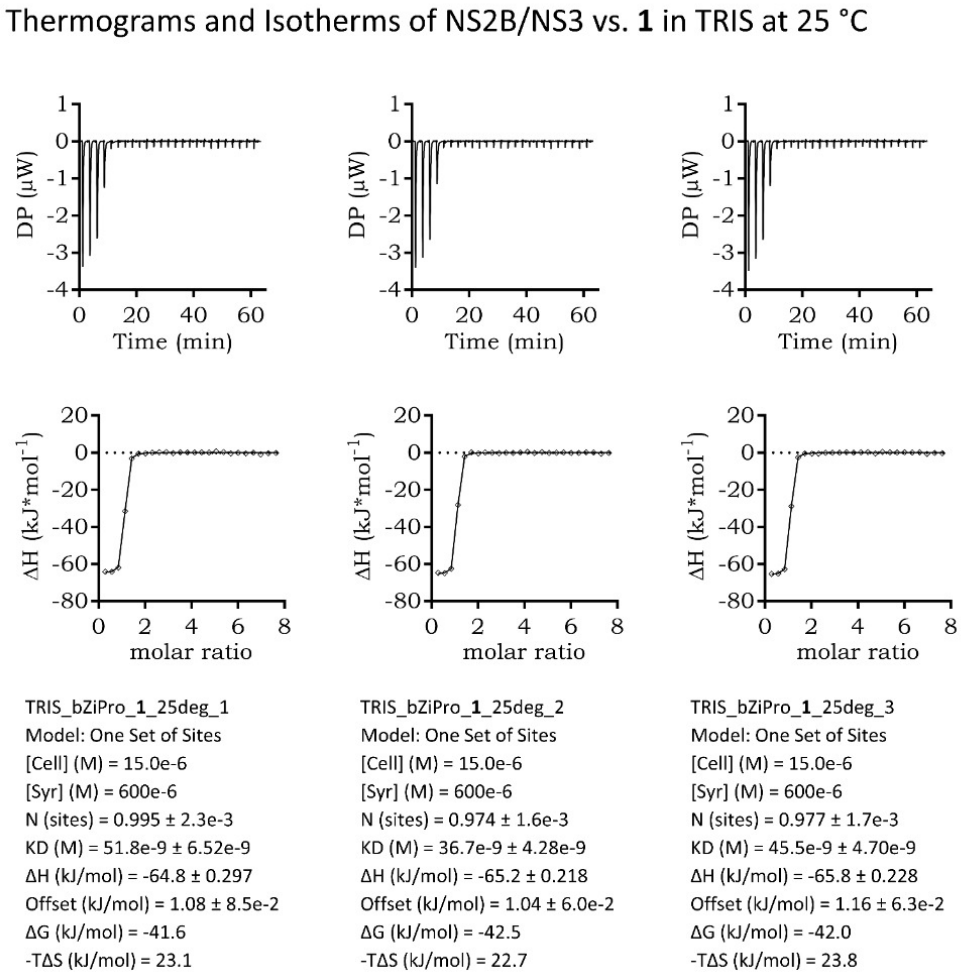


FIGURE S7E Thermograms and isotherms of NS2B/NS3 vs. **1** in TRIS at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

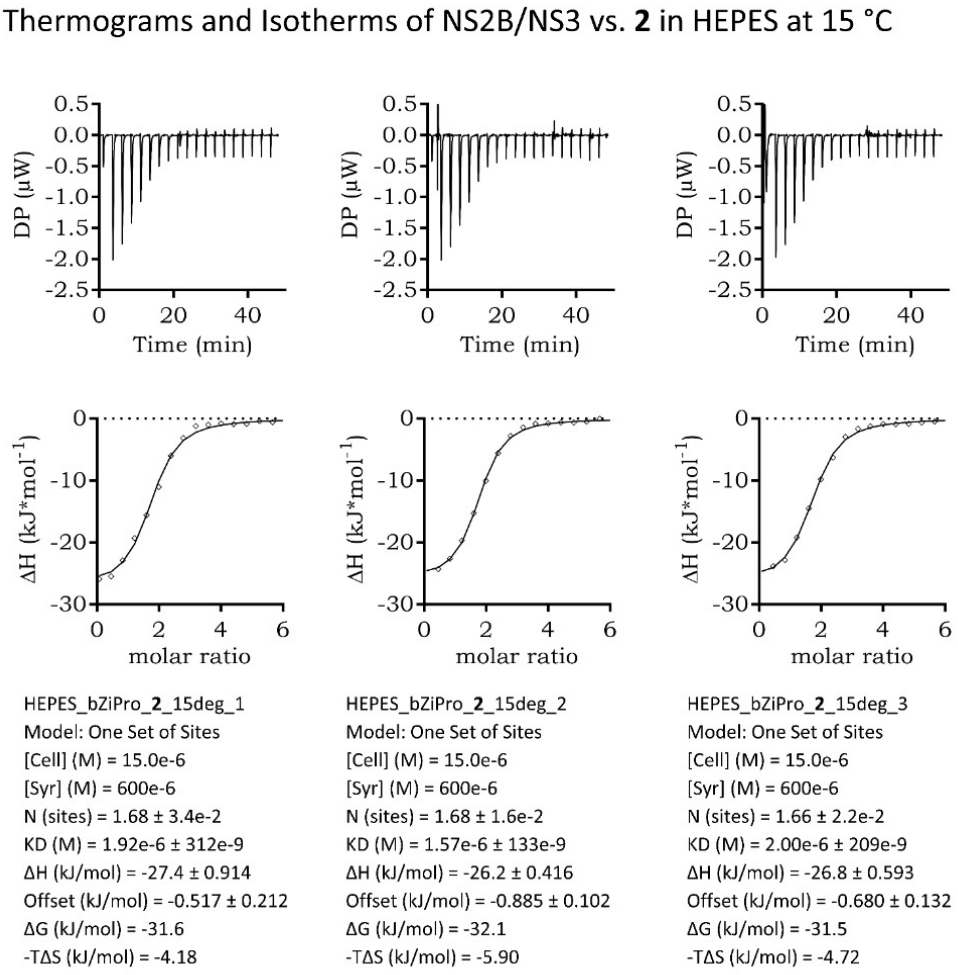


FIGURE S7F Thermograms and isotherms of NS2B/NS3 vs. **2** in HEPES at 15 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

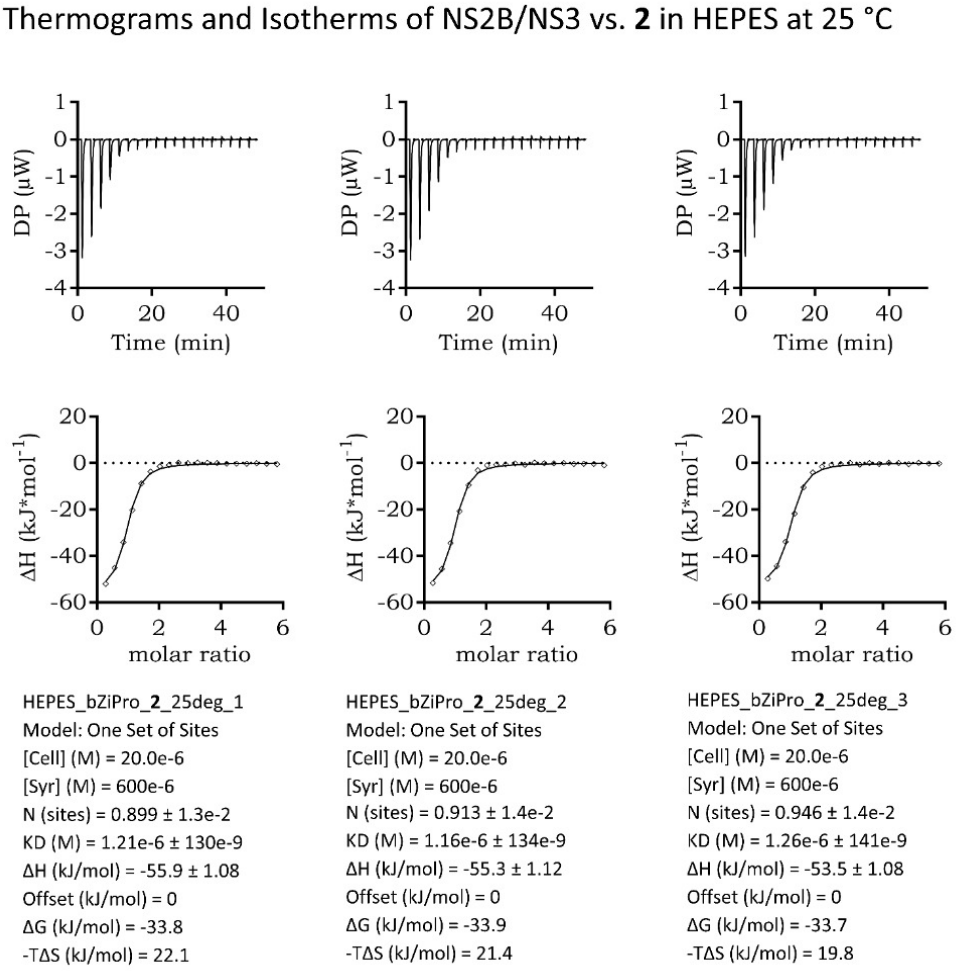


FIGURE S7G Thermograms and isotherms of NS2B/NS3 vs. **2** in HEPES at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

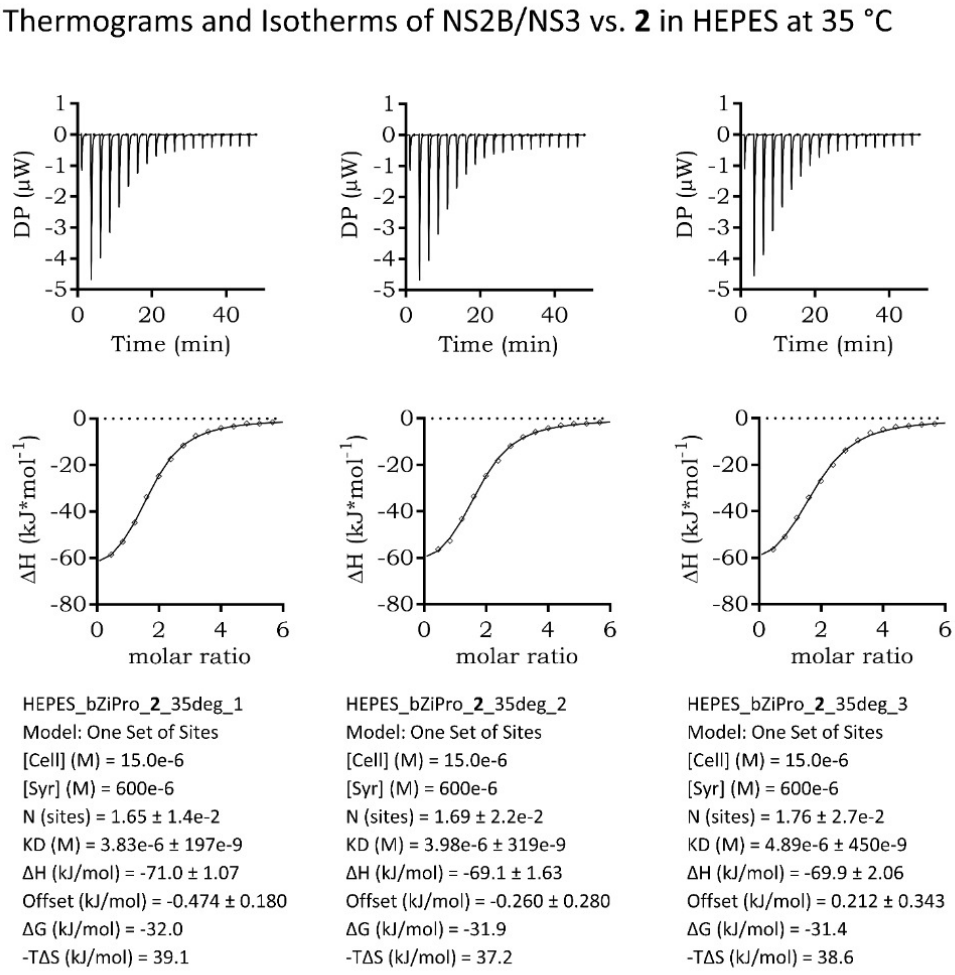


FIGURE S7H Thermograms and isotherms of NS2B/NS3 vs. **2** in HEPES at 35 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

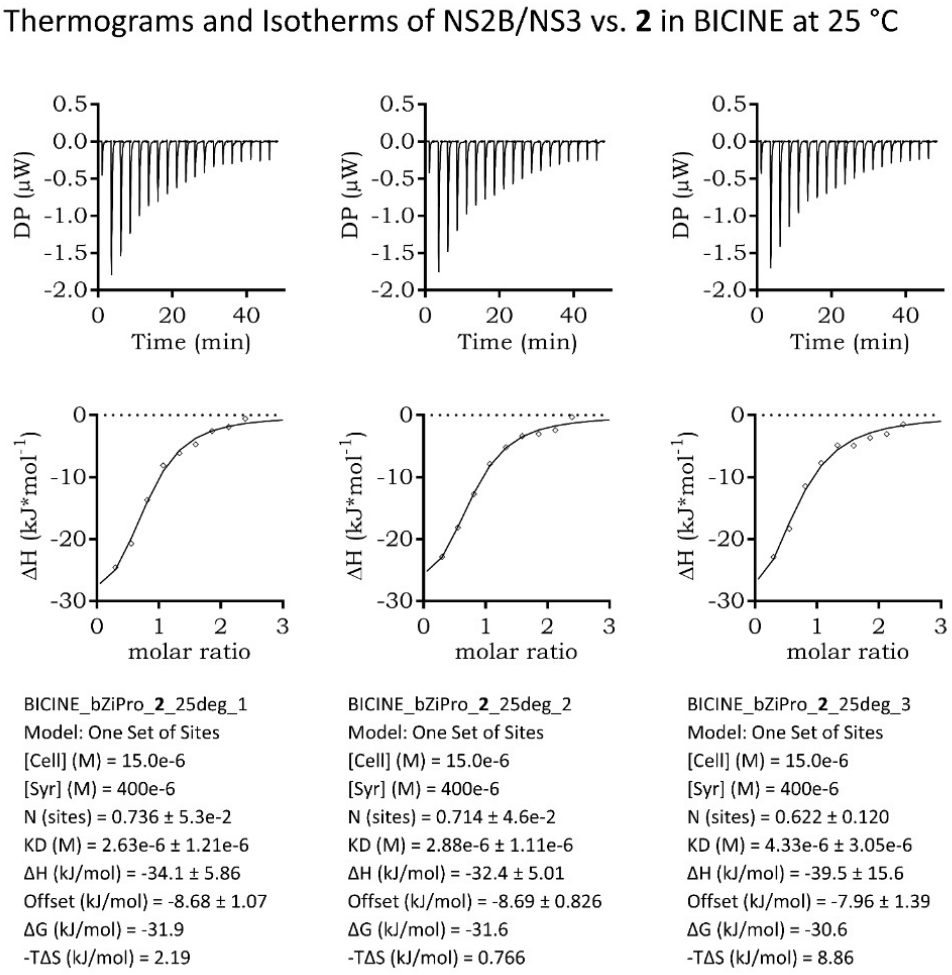


FIGURE S7I Thermograms and isotherms of NS2B/NS3 vs. **2** in BICINE at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

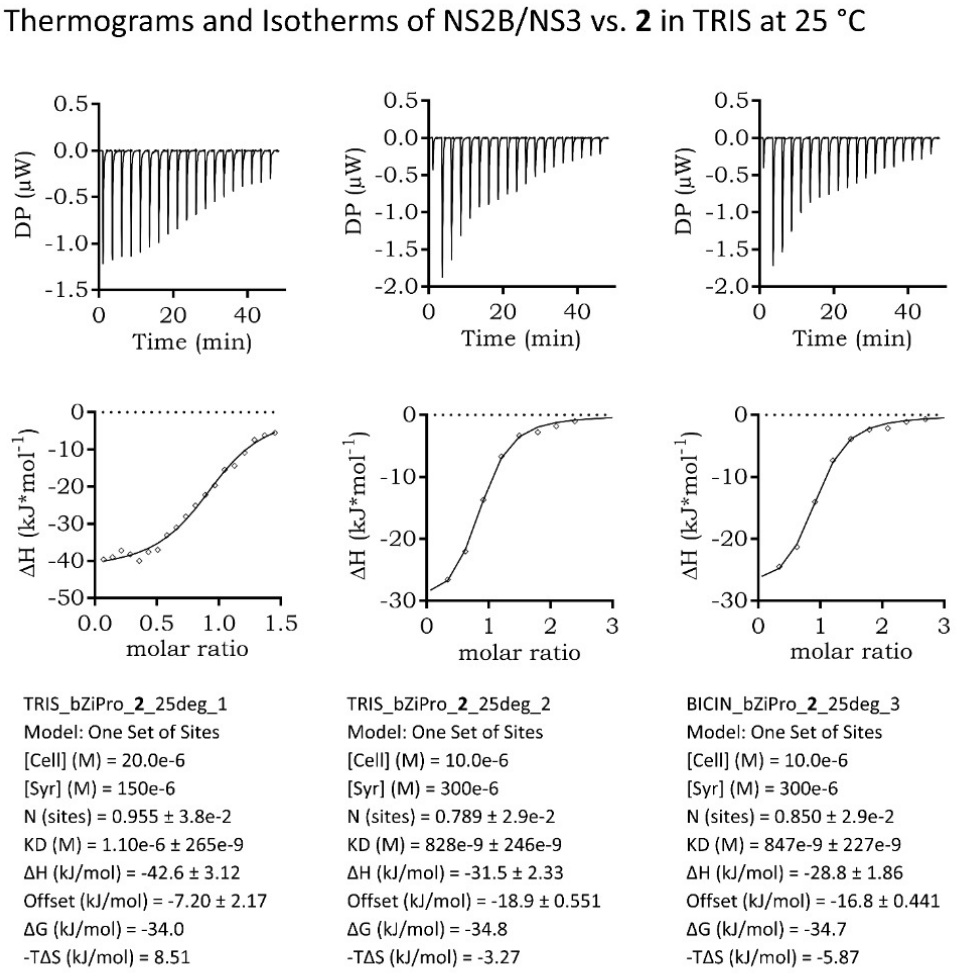


FIGURE S7J Thermograms and isotherms of NS2B/NS3 vs. **2** in TRIS at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

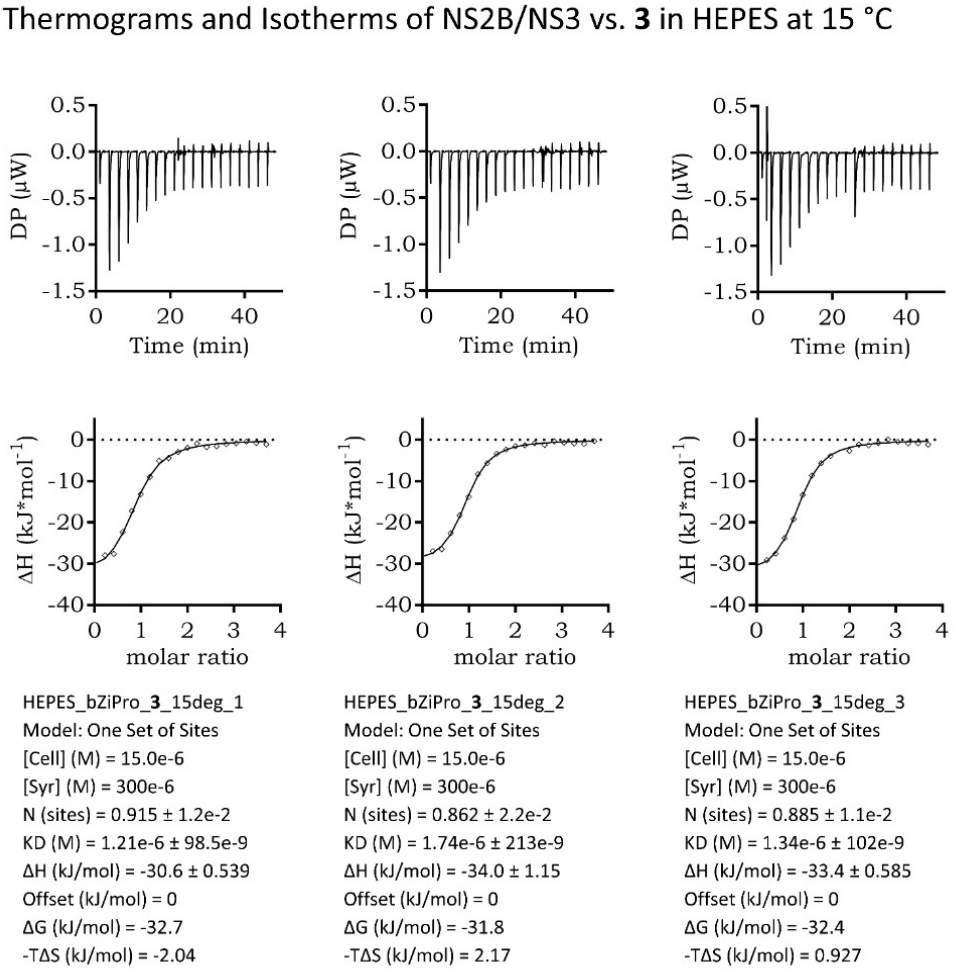


FIGURE S7K Thermograms and isotherms of NS2B/NS3 vs. **3** in HEPES at 15 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

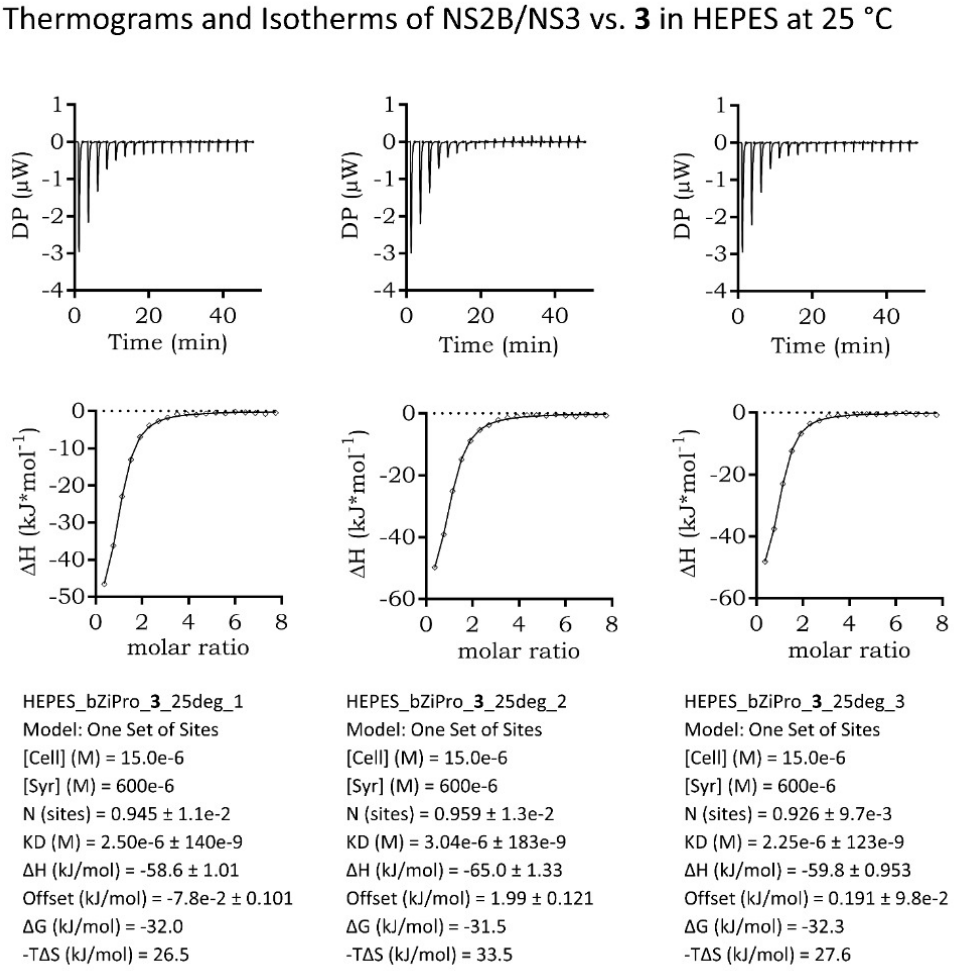


FIGURE S7L Thermograms and isotherms of NS2B/NS3 vs. **3** in HEPES at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

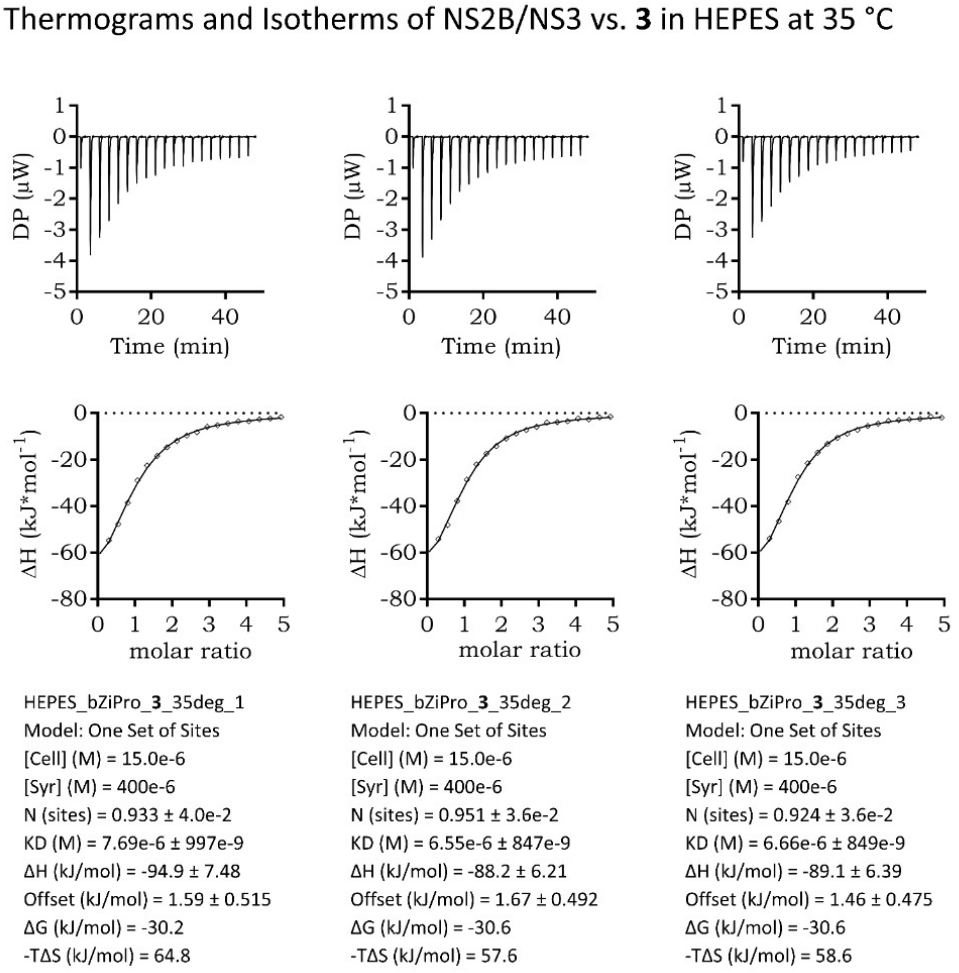


FIGURE S7M Thermograms and isotherms of NS2B/NS3 vs. **3** in HEPES at 35 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

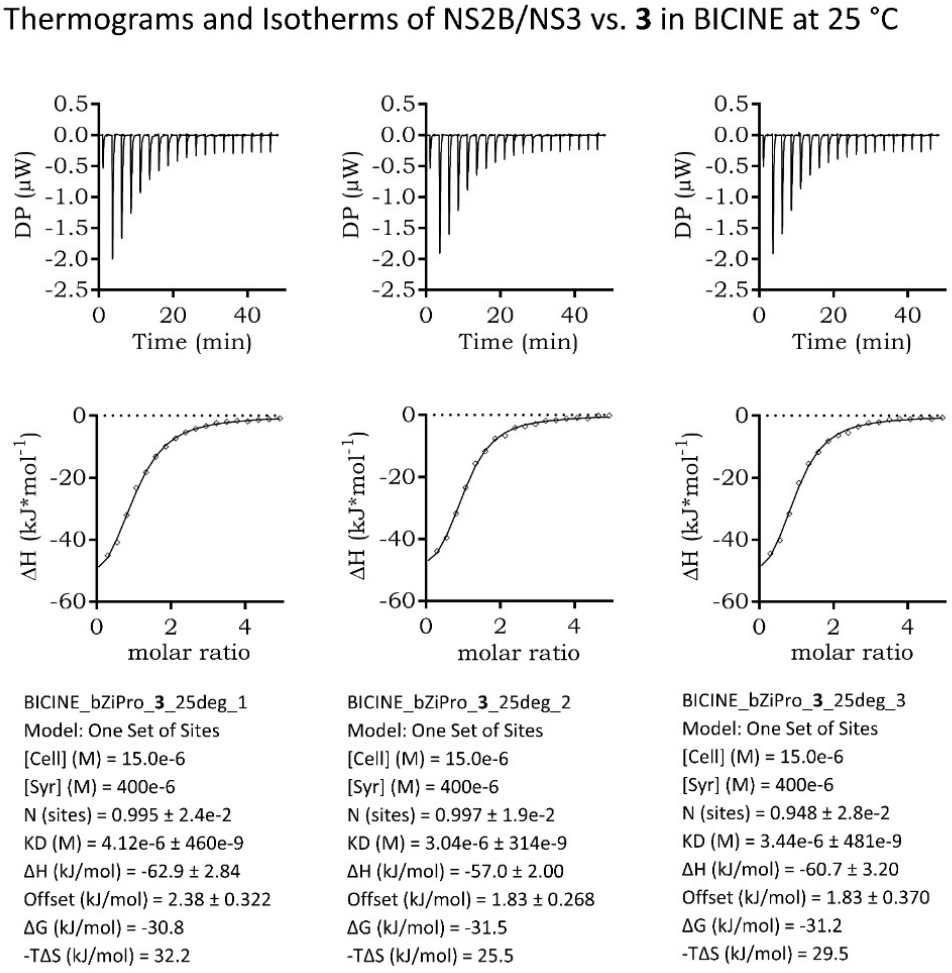


FIGURE S7N Thermograms and isotherms of NS2B/NS3 vs. **3** in BICINE at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

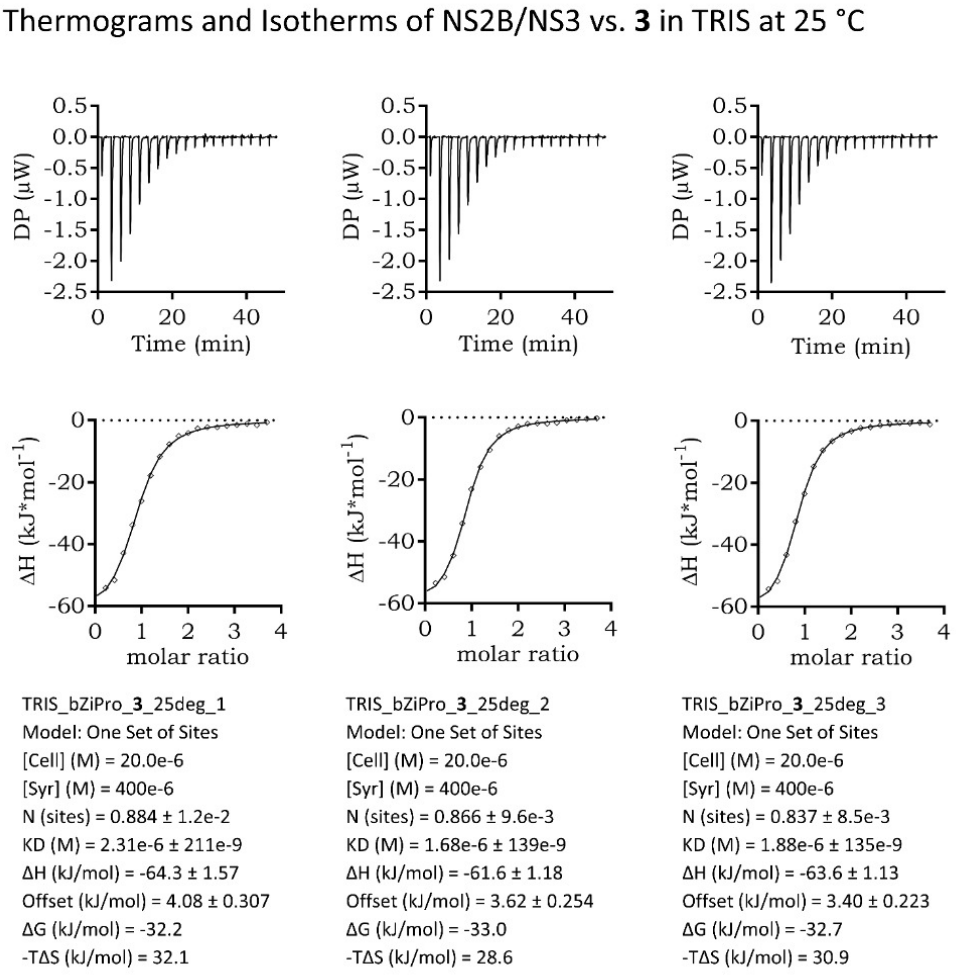


FIGURE S7O Thermograms and isotherms of NS2B/NS3 vs. **3** in TRIS at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

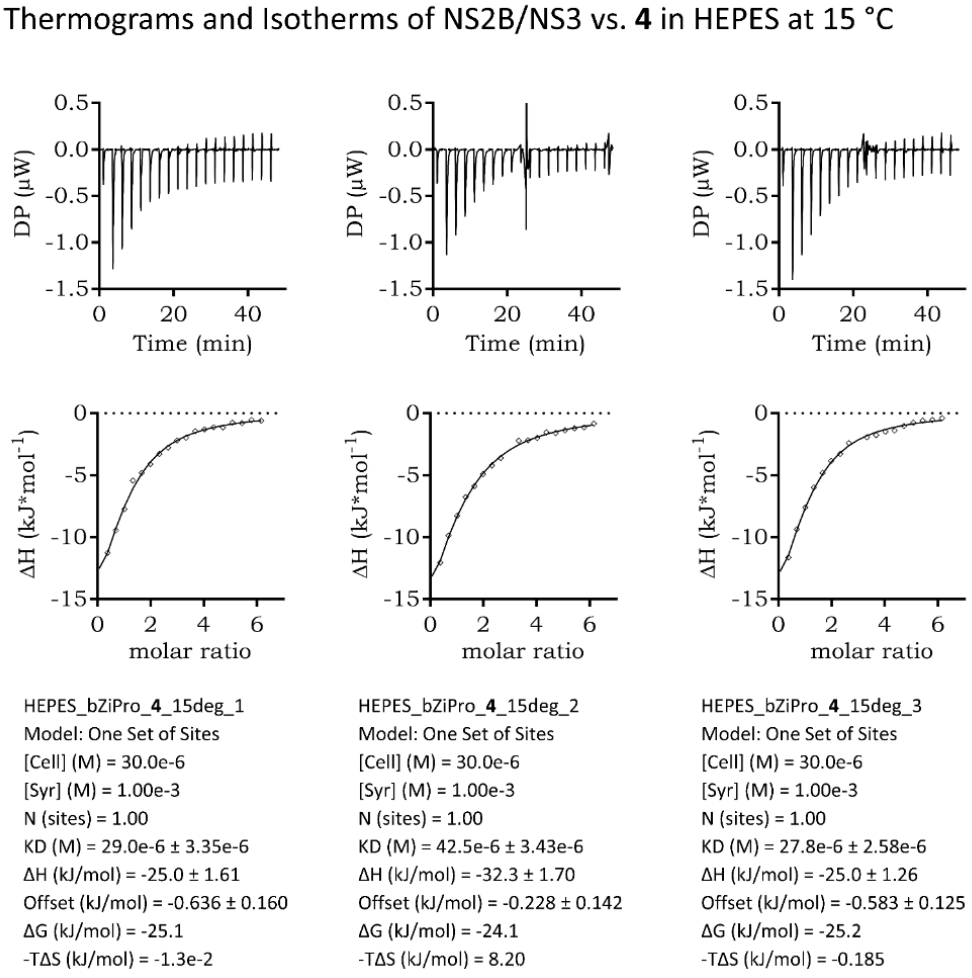


FIGURE S7P Thermograms and isotherms of NS2B/NS3 vs. **4** in HEPES at 15 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

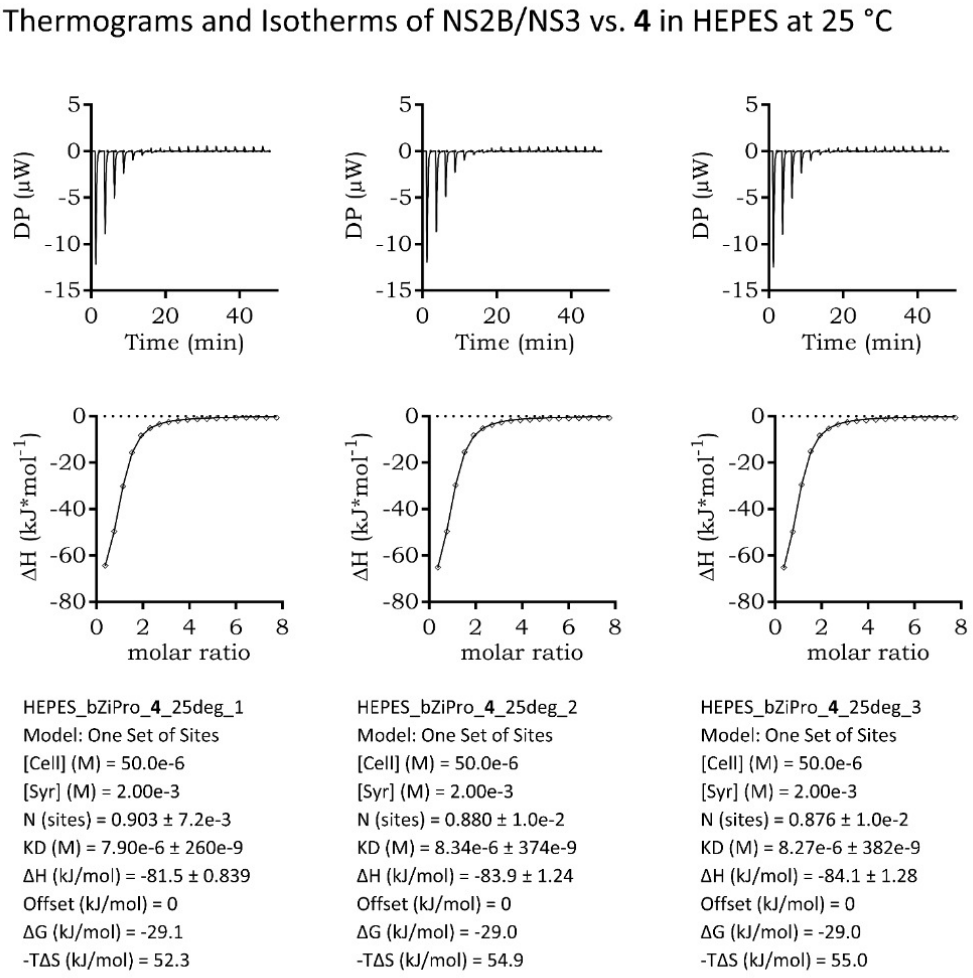


FIGURE S7Q Thermograms and isotherms of NS2B/NS3 vs. **4** in HEPES at 25 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

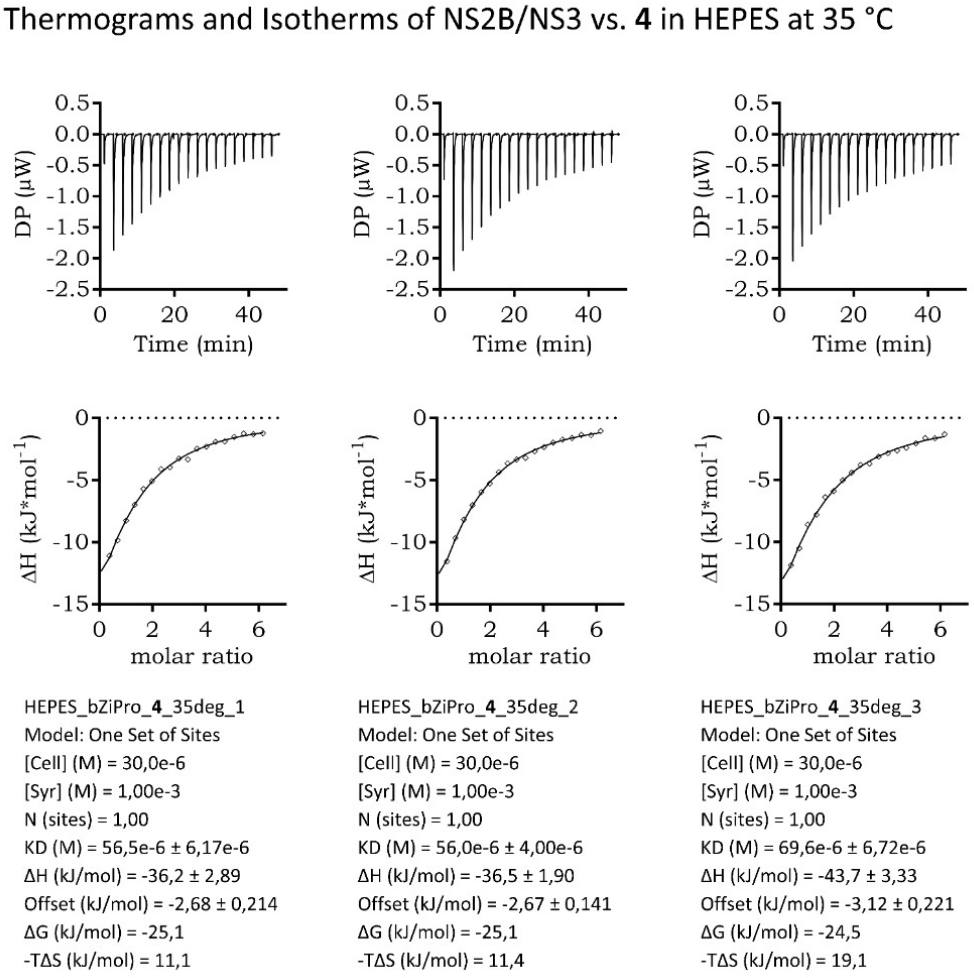


FIGURE S7R Thermograms and isotherms of NS2B/NS3 vs. **4** in HEPES at 35 °C. Isotherms are shown with subtracted offset and baseline. Control experiments are already included.

Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogues.

Stefan J. Hammerschmidt[1], Simon Huber[2], Niklas J. Braun[2], Marc Lander[1], Torsten Steinmetzer[2], Christian Kersten[1,*]

1 Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany
2 Institute of Pharmaceutical Chemistry, Philipps-University, Marbacher Weg 6, 35032 Marburg, Germany

*Correspondence:
Dr. Christian Kersten. Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany
Email: kerstec@uni-mainz.de

CONTENT

Table S1. Inhibition constants obtained by fluorometric enzyme activity assay and thermodynamic binding profiles determined by ITC. Thermodynamic values are given in the respective buffer or as buffer ionization corrected values (corrected), respectively. All experiments were performed in technical triplicates.

Cpd.	K_i (nM)	ITC buffer	K_d (nM)	ΔG° (kJ·mol^{-1})	ΔH° (kJ·mol^{-1})	–TΔS° (kJ·mol^{-1})
1	1.98 ± 0.15	HEPES[a]	109 ± 9	–39.8 ± 0.0	–66.3 ± 0.3	26.5 ± 0.6
		BICINE[b]	177 ± 18	–38.6 ± 0.4	–76.7 ± 0.7	38.1 ± 0.3
		TRIS[c]	44.7 ± 3.0	–42.0 ± 0.4	–65.3 ± 0.2	23.2 ± 0.7
		corrected	110 ± 7	–40.1 ± 1.7	–74.8	34.7
2	14.8 ± 1.32	HEPES	1210 ± 80	–33.8 ± 0.1	–54.9 ± 0.6	21.1 ± 1.0
		BICINE	3280 ± 1150	–31.4 ± 0.6	–35.3 ± 5.8	3.94 ± 3.53
		TRIS	1860 ± 629	–32.8 ± 0.8	–41.3 ± 4.6	8.41 ± 1.59
		corrected	2120 ± 440	–32.7 ± 0.3	–53.1	20.4
3	158 ± 14	HEPES	2600 ± 90	–31.9 ± 0.3	–61.1 ± 0.1	29.2 ± 3.1
		BICINE	3530 ± 240	–31.2 ± 0.3	–60.2 ± 1.6	29.1 ± 2.8
		TRIS	1960 ± 100	–32.6 ± 0.3	–63.2 ± 0.8	30.5 ± 1.5
		corrected	2700 ± 90	–31.9 ± 0.2	–58.5	26.6
4	611 ± 72	HEPES	8170 ± 200	–29.0 ± 0.0	–83.2 ± 0.7	54.1 ± 1.2
		BICINE	13600 ± 700	–27.8 ± 0.1	–82.3 ± 2.7	54.5 ± 1.6
		TRIS	2890 ± 130	–31.6 ± 0.1	–75.6 ± 0.6	43.9 ± 0.7
		corrected	8220 ± 250	–29.5 ± 0.1	–77.3	47.8

[a] buffer ionization enthalpy (ΔH_{ion}) HEPES: 20.4 kJ·mol^{-1}, [b] ΔH_{ion} BICINE: 28.4 kJ·mol^{-1}, [c] ΔH_{ion} TRIS: 47.5 kJ·mol^{-1}.[1]

Table S2. Thermodynamic binding profiles of compounds 1–4 at different temperatures in HEPES. ΔCP values are given for compounds 1–3. For inhibitor 4, ΔC_P could not be determined (n.d.).

Cpd.	Temperature (°C)	K_d (nM)	ΔG (kJ·mol^{-1})	ΔH_{obs} (kJ·mol^{-1})	$-T\Delta S_{obs}$ (kJ·mol^{-1})	ΔC_P (kJ·mol^{-1}·K^{-1})	T_H (°C)	T_S (°C)
1	15	27.1 ± 4.1	–41.9 ± 0.7	–35.7 ± 0.1	–6.17 ± 0.82			
	25	109 ± 9	–39.8 ± 0.0	–66.3 ± 0.3	26.5 ± 0.6	–3.77 ± 0.41	6.1	17.5
	35	38.2 ± 3.1	–39.9 ± 0.0	–111 ± 0	71.2 ± 0.2			
2	15	1830 ± 130	–31.7 ± 0.3	–26.8 ± 0.4	–4.93 ± 0.72			
	25	1210 ± 80	–33.8 ± 0.1	–54.9 ± 0.6	21.1 ± 1.0	–2.16 ± 0.38	1.6	16.6
	35	4230 ± 195	–31.8 ± 0.3	–70.0 ± 0.9	38.3 ± 0.8			
3	15	1430 ± 80	–32.3 ± 0.4	–32.7 ± 0.5	0.35 ± 1.77			
	25	2600 ± 90	–31.9 ± 0.3	–61.1 ± 0.1	29.2 ± 3.1	–2.90 ± 0.03	3.8	15
	35	6970 ± 520	–30.5 ± 0.2	–90.7 ± 3.9	60.3 ± 3.2			
4	15	43100 ± 1800	–24.8 ± 0.5	–27.4 ± 0.9	2.67 ± 3.91			
	25	8170 ± 200	–29.0 ± 0.0	–83.2 ± 0.7	54.1 ± 1.2	n.d.	n.d.	n.d.
	35	60700 ± 3300	–24.9 ± 0.3	–38.8 ± 1.6	13.9 ± 3.7			

Table S3. Analysis of the accessible surface area (ASA, calculated with MOE 2020.09) for flaviviral *open* and *closed* NS2B/NS3 conformations.

Protease	conformation	PDB-ID	ASA	hydrophobic ASA	polar ASA	fraction of hydrophobic ASA
Zika virus	closed	6A3B	11284.908	5424.3081	5860.6001	0.48
Zika virus	open	5GXJ	8724.4502	4498.5605	4225.8901	0.52
Dengue virus	closed	3U1I	11464.108	4872.2217	6591.8872	0.42
Dengue virus	open	2FOM	10166.805	5051.7329	5115.0723	0.50
West Nile virus	closed	5IDK	10824.537	5091.7651	5732.7715	0.47
West Nile virus	open	2GGV	11586.871	6070.897	5515.9741	0.52

Table S4. MD per residue analysis of Cα root-mean-square deviation compared to the starting structures (RMSD) and root-mean-square fluctuation over time (RMSF). Chain A = NS2B, chain B = NS3. Binding-site residues are highlighted in green.

			RMSD [Å]					RMSF [Å]				
residue	amino acid	chain	apo	complex -1	complex -2	complex -3	complex -4	apo	complex -1	complex -2	complex -3	complex -4
50	ASP	A	0.99	0.95	0.78	0.81	0.88	0.92	0.77	0.76	0.65	0.75
51	MET	A	0.67	0.80	0.62	0.63	0.59	0.51	0.56	0.55	0.47	0.50
52	TYR	A	0.59	0.55	0.58	0.66	0.56	0.43	0.47	0.49	0.41	0.42
53	ILE	A	0.63	0.56	0.61	0.73	0.56	0.41	0.42	0.46	0.41	0.41
54	GLU	A	0.61	0.55	0.59	0.74	0.57	0.49	0.52	0.54	0.51	0.49
55	ARG	A	0.58	0.54	0.57	0.65	0.52	0.44	0.46	0.51	0.44	0.46
56	ALA	A	0.67	0.62	0.63	0.79	0.70	0.49	0.47	0.51	0.47	0.47
57	GLY	A	0.61	0.68	0.54	0.72	0.60	0.61	0.61	0.60	0.58	0.58
58	ASP	A	0.66	0.91	0.63	0.80	0.77	0.52	0.54	0.52	0.52	0.50
59	ILE	A	0.66	0.96	0.72	0.72	0.77	0.49	0.58	0.50	0.50	0.47
60	THR	A	0.69	0.94	0.69	0.73	0.67	0.60	0.67	0.59	0.56	0.58
61	TRP	A	0.77	0.96	0.76	0.69	0.70	0.57	0.59	0.53	0.51	0.53
62	GLU	A	0.76	0.83	0.59	0.61	0.59	0.65	0.61	0.54	0.53	0.56
63	LYS	A	0.81	0.79	0.74	0.73	0.77	0.79	0.77	0.76	0.71	0.76
64	ASP	A	1.27	1.08	1.24	0.83	0.83	1.18	0.99	1.04	0.86	0.90
65	ALA	A	1.34	1.18	1.42	0.78	0.71	1.05	0.94	0.90	0.68	0.69
66	GLU	A	1.74	1.54	2.04	0.73	0.65	1.62	1.61	1.53	0.76	0.78
67	VAL	A	1.64	1.53	1.71	0.88	0.70	1.14	1.14	1.05	0.72	0.67
68	THR	A	1.04	1.16	0.97	0.87	0.70	0.69	0.73	0.68	0.67	0.59
69	GLY	A	0.98	0.96	0.74	1.15	0.77	0.86	0.74	0.67	0.85	0.64
70	ASN	A	1.00	1.03	0.71	1.35	0.86	1.12	0.94	0.69	1.42	0.72
71	SER	A	0.93	0.99	0.83	1.12	0.80	0.77	0.66	0.56	0.88	0.56
72	PRO	A	1.29	1.30	1.08	1.37	1.23	1.03	0.92	0.67	1.01	0.93
73	ARG	A	1.00	1.12	0.93	1.03	1.10	0.70	0.67	0.59	0.70	0.60
74	LEU	A	1.05	0.98	0.95	1.10	0.93	0.70	0.65	0.62	0.79	0.51
75	ASP	A	1.23	0.83	0.70	0.87	0.82	0.86	0.61	0.60	0.64	0.60
76	VAL	A	1.32	0.82	0.66	0.94	0.76	0.77	0.57	0.59	0.59	0.52
77	ALA	A	1.43	1.01	0.90	1.27	0.90	0.84	0.70	0.78	0.68	0.61
78	LEU	A	1.81	1.51	1.75	1.87	1.48	0.87	0.66	0.78	0.67	0.62
79	ASP	A	1.53	1.56	2.00	1.83	1.52	1.01	0.70	0.99	0.71	0.68
80	GLU	A	1.98	1.71	2.29	1.86	1.60	1.34	0.84	1.30	0.86	0.75
81	SER	A	1.67	1.52	1.69	1.57	1.39	1.27	0.77	1.07	0.77	0.70
82	GLY	A	0.97	1.08	1.00	1.18	1.06	0.94	0.61	0.77	0.66	0.60
83	ASP	A	0.98	1.14	1.07	1.26	1.09	0.68	0.61	0.71	0.55	0.54
84	PHE	A	0.90	1.13	0.84	1.29	1.03	0.57	0.52	0.56	0.53	0.50
85	SER	A	1.13	1.01	0.78	1.21	0.90	0.66	0.61	0.62	0.58	0.53
86	LEU	A	1.23	1.13	0.84	1.20	1.03	0.85	0.64	0.66	0.66	0.56
87	VAL	A	1.75	1.54	1.73	1.45	1.29	1.10	0.81	0.88	0.82	0.75
18	THR	B	2.23	3.65	2.10	1.90	2.15	1.59	2.01	1.48	1.44	1.43

19	THR	B	1.04	2.26	0.96	0.90	0.88	0.82	1.31	0.79	0.75	0.74
20	ASP	B	0.76	1.21	0.72	0.83	0.69	0.56	0.78	0.60	0.56	0.58
21	GLY	B	0.52	0.60	0.50	0.66	0.49	0.49	0.47	0.52	0.50	0.49
22	VAL	B	0.42	0.45	0.39	0.51	0.39	0.36	0.37	0.37	0.36	0.35
23	TYR	B	0.39	0.38	0.39	0.45	0.35	0.35	0.36	0.38	0.32	0.34
24	ARG	B	0.45	0.40	0.45	0.53	0.41	0.35	0.36	0.37	0.32	0.33
25	VAL	B	0.40	0.43	0.40	0.48	0.40	0.34	0.36	0.36	0.31	0.32
26	MET	B	0.49	0.55	0.51	0.61	0.53	0.45	0.50	0.46	0.42	0.45
27	THR	B	0.75	0.82	0.62	0.92	0.92	0.56	0.57	0.57	0.48	0.57
28	ARG	B	1.08	0.88	0.78	1.38	1.37	1.02	0.77	0.76	0.77	0.90
29	ARG	B	2.34	1.42	1.48	2.79	2.75	2.00	1.16	1.19	1.69	1.46
30	LEU	B	2.56	2.37	2.40	2.44	2.52	2.69	1.90	1.66	2.08	2.10
31	LEU	B	2.47	2.76	2.37	2.82	2.84	2.53	1.86	1.76	2.29	2.16
32	GLY	B	1.82	1.87	1.75	1.96	1.61	1.76	1.16	1.16	1.32	1.30
33	SER	B	1.30	1.22	0.97	1.58	1.60	0.86	0.79	0.79	0.74	0.81
34	THR	B	1.06	1.04	0.77	1.37	1.44	0.78	0.70	0.66	0.67	0.69
35	GLN	B	0.87	0.89	0.76	1.06	1.11	0.49	0.45	0.46	0.43	0.44
36	VAL	B	0.55	0.65	0.50	0.67	0.73	0.50	0.46	0.45	0.45	0.45
37	GLY	B	0.43	0.44	0.43	0.44	0.43	0.40	0.37	0.36	0.35	0.35
38	VAL	B	0.42	0.43	0.43	0.37	0.36	0.32	0.31	0.31	0.30	0.29
39	GLY	B	0.44	0.48	0.43	0.44	0.39	0.38	0.37	0.40	0.37	0.35
40	VAL	B	0.57	0.63	0.47	0.60	0.48	0.37	0.36	0.42	0.38	0.36
41	MET	B	0.58	0.64	0.48	0.70	0.51	0.42	0.42	0.44	0.42	0.40
42	GLN	B	0.70	0.76	0.54	0.78	0.60	0.54	0.52	0.53	0.50	0.50
43	GLU	B	0.90	0.82	0.76	0.91	0.79	0.77	0.66	0.74	0.68	0.71
44	GLY	B	0.93	0.79	0.84	0.99	0.87	0.91	0.71	0.90	0.86	0.87
45	VAL	B	0.73	0.73	0.54	0.78	0.62	0.54	0.47	0.49	0.48	0.49
46	PHE	B	0.66	0.71	0.49	0.66	0.54	0.42	0.40	0.41	0.40	0.39
47	HIS	B	0.65	0.76	0.49	0.65	0.60	0.37	0.37	0.37	0.33	0.35
48	THR	B	0.44	0.55	0.38	0.46	0.43	0.33	0.34	0.35	0.32	0.32
49	MET	B	0.53	0.62	0.45	0.54	0.51	0.37	0.36	0.39	0.34	0.34
50	TRP	B	0.65	0.75	0.56	0.65	0.60	0.41	0.40	0.43	0.37	0.37
51	HIS	B	0.74	0.78	0.69	0.68	0.69	0.50	0.47	0.48	0.42	0.43
52	VAL	B	0.78	0.78	0.69	0.65	0.70	0.48	0.46	0.48	0.43	0.45
53	THR	B	0.75	0.81	0.67	0.70	0.71	0.49	0.51	0.51	0.45	0.46
54	LYS	B	0.86	0.93	0.76	0.87	0.88	0.62	0.62	0.64	0.58	0.60
55	GLY	B	1.05	1.06	0.90	1.06	1.08	0.82	0.82	0.82	0.79	0.86
56	ALA	B	0.87	0.90	0.72	0.80	0.80	0.79	0.75	0.75	0.73	0.76
57	ALA	B	0.91	0.97	0.73	0.79	0.77	0.68	0.67	0.68	0.62	0.64
58	LEU	B	0.94	0.96	0.76	0.81	0.77	0.56	0.64	0.60	0.51	0.56
59	ARG	B	0.65	0.69	0.63	0.76	0.66	0.57	0.59	0.67	0.56	0.61
60	SER	B	0.95	0.92	0.78	0.92	0.92	0.68	0.68	0.70	0.63	0.71
61	GLY	B	1.13	0.97	0.88	1.07	1.19	0.94	0.92	0.88	0.85	1.18
62	GLU	B	1.63	1.41	1.17	1.24	1.44	1.13	1.05	0.98	0.99	1.41
63	GLY	B	1.40	1.39	1.07	1.15	1.17	0.97	0.93	1.02	0.86	0.97
64	ARG	B	1.29	1.40	0.95	0.93	0.96	0.83	0.83	0.90	0.77	0.85

65	LEU	B	1.07	1.06	0.75	0.89	0.81	0.75	0.68	0.70	0.65	0.69
66	ASP	B	1.01	1.02	0.73	0.89	0.79	0.92	0.73	0.81	0.69	0.71
67	PRO	B	0.87	1.00	0.72	0.88	0.74	0.65	0.61	0.67	0.61	0.56
68	TYR	B	0.81	1.03	0.68	0.93	0.73	0.57	0.56	0.59	0.55	0.49
69	TRP	B	0.92	1.20	0.78	1.11	0.96	0.60	0.56	0.59	0.58	0.53
70	GLY	B	1.23	1.43	1.01	1.29	1.23	0.59	0.55	0.59	0.57	0.56
71	ASP	B	1.25	1.37	1.04	1.23	1.23	0.70	0.57	0.61	0.63	0.56
72	VAL	B	1.04	1.27	0.97	1.16	1.17	0.63	0.55	0.59	0.60	0.57
73	LYS	B	1.07	1.02	0.85	1.03	0.95	0.72	0.58	0.60	0.64	0.57
74	GLN	B	0.96	0.79	0.77	0.83	0.74	0.62	0.54	0.57	0.57	0.48
75	ASP	B	0.54	0.53	0.46	0.52	0.44	0.41	0.39	0.42	0.38	0.36
76	LEU	B	0.45	0.58	0.41	0.55	0.45	0.39	0.38	0.42	0.38	0.34
77	VAL	B	0.51	0.66	0.42	0.59	0.48	0.40	0.40	0.43	0.39	0.35
78	SER	B	0.65	0.80	0.50	0.71	0.58	0.43	0.43	0.45	0.42	0.39
79	TYR	B	0.78	0.91	0.58	0.79	0.66	0.46	0.45	0.45	0.43	0.43
80	CYS	B	0.96	1.13	0.74	1.08	0.89	0.71	0.68	0.64	0.63	0.64
81	GLY	B	0.84	1.03	0.64	0.96	0.76	0.63	0.56	0.58	0.57	0.54
82	PRO	B	0.85	1.03	0.69	0.98	0.76	0.68	0.60	0.67	0.64	0.55
83	TRP	B	0.78	0.91	0.64	0.81	0.62	0.56	0.50	0.53	0.51	0.47
84	LYS	B	0.77	0.89	0.61	0.76	0.63	0.52	0.49	0.50	0.49	0.44
85	LEU	B	0.81	0.89	0.74	0.77	0.70	0.53	0.55	0.56	0.52	0.49
86	ASP	B	0.87	0.95	0.77	0.76	0.74	0.68	0.74	0.81	0.69	0.71
87	ALA	B	0.72	0.76	0.60	0.64	0.58	0.55	0.58	0.58	0.55	0.51
88	ALA	B	0.68	0.62	0.60	0.58	0.59	0.61	0.61	0.61	0.59	0.59
89	TRP	B	0.71	0.71	0.67	0.63	0.62	0.58	0.63	0.57	0.58	0.59
90	ASP	B	0.83	0.86	0.77	0.76	0.75	0.73	0.85	0.73	0.74	0.78
91	GLY	B	0.84	0.97	0.85	0.85	0.83	0.79	0.91	0.76	0.77	0.86
92	LEU	B	1.00	1.06	0.89	0.89	0.92	0.89	0.90	0.80	0.82	0.90
93	SER	B	0.77	0.80	0.66	0.66	0.64	0.72	0.74	0.67	0.68	0.70
94	GLU	B	0.74	0.86	0.66	0.67	0.59	0.54	0.53	0.50	0.49	0.46
95	VAL	B	0.58	0.77	0.55	0.56	0.46	0.44	0.45	0.41	0.44	0.40
96	GLN	B	0.48	0.57	0.41	0.47	0.40	0.39	0.36	0.35	0.35	0.32
97	LEU	B	0.44	0.48	0.37	0.53	0.46	0.38	0.33	0.32	0.35	0.32
98	LEU	B	0.49	0.48	0.44	0.54	0.48	0.38	0.34	0.34	0.34	0.32
99	ALA	B	0.51	0.50	0.51	0.68	0.62	0.42	0.38	0.36	0.36	0.37
100	VAL	B	0.46	0.44	0.45	0.58	0.54	0.40	0.37	0.36	0.36	0.36
101	PRO	B	0.55	0.44	0.45	0.60	0.55	0.49	0.44	0.43	0.43	0.42
102	PRO	B	0.68	0.58	0.60	0.70	0.67	0.58	0.50	0.54	0.52	0.49
103	GLY	B	0.99	0.89	0.88	1.05	1.01	0.99	0.92	0.94	0.89	0.88
104	GLU	B	0.75	0.70	0.65	0.86	0.79	0.73	0.73	0.67	0.66	0.64
105	ARG	B	0.77	0.67	0.70	0.89	0.82	0.66	0.72	0.60	0.68	0.62
106	ALA	B	0.72	0.63	0.70	0.92	0.86	0.57	0.61	0.53	0.60	0.54
107	LYS	B	0.67	0.59	0.63	0.80	0.76	0.55	0.54	0.44	0.53	0.54
108	ASN	B	0.65	0.79	0.53	0.73	0.66	0.44	0.43	0.40	0.42	0.42
109	ILE	B	0.60	0.66	0.47	0.68	0.59	0.43	0.40	0.40	0.43	0.39
110	GLN	B	0.56	0.71	0.55	0.50	0.39	0.42	0.42	0.41	0.42	0.35

111	THR	B	0.59	0.74	0.56	0.53	0.42	0.47	0.44	0.42	0.46	0.36
112	LEU	B	0.66	0.72	0.57	0.56	0.50	0.54	0.49	0.45	0.45	0.42
113	PRO	B	0.70	0.68	0.51	0.63	0.51	0.63	0.53	0.44	0.53	0.48
114	GLY	B	0.94	0.94	0.52	1.00	0.71	1.05	0.88	0.57	0.92	0.79
115	ILE	B	0.76	0.86	0.59	0.81	0.69	0.64	0.64	0.50	0.61	0.54
116	PHE	B	0.87	0.68	0.59	0.70	0.67	0.51	0.53	0.47	0.53	0.47
117	LYS	B	1.09	0.78	0.72	0.87	0.89	0.80	0.66	0.62	0.70	0.62
118	THR	B	1.26	1.16	1.31	1.19	1.07	0.92	0.98	0.97	0.97	0.81
119	LYS	B	1.79	1.62	1.93	1.61	1.35	1.47	1.58	1.54	1.40	1.11
120	ASP	B	1.96	2.04	1.66	1.78	1.62	1.70	1.83	1.65	1.58	1.38
121	GLY	B	1.49	1.86	1.01	1.42	1.45	1.22	1.37	1.09	1.21	1.22
122	ASP	B	1.38	1.44	1.10	1.32	1.25	0.98	0.90	0.83	0.96	0.86
123	ILE	B	1.05	0.90	0.78	0.97	0.80	0.77	0.74	0.65	0.77	0.62
124	GLY	B	0.90	0.78	0.50	0.88	0.70	0.70	0.60	0.48	0.65	0.51
125	ALA	B	0.68	0.64	0.46	0.66	0.55	0.53	0.44	0.39	0.45	0.40
126	VAL	B	0.62	0.54	0.44	0.54	0.46	0.56	0.46	0.40	0.45	0.41
127	ALA	B	0.71	0.63	0.57	0.61	0.57	0.55	0.48	0.43	0.47	0.43
128	LEU	B	0.58	0.50	0.45	0.49	0.47	0.58	0.49	0.43	0.47	0.46
129	ASP	B	0.69	0.58	0.51	0.60	0.59	0.61	0.49	0.47	0.48	0.47
130	TYR	B	0.61	0.76	0.65	0.60	0.83	0.57	0.51	0.51	0.48	0.45
131	PRO	B	0.72	0.84	0.78	0.75	0.81	0.66	0.53	0.54	0.52	0.48
132	ALA	B	0.72	0.79	0.71	0.72	0.70	0.66	0.52	0.50	0.55	0.47
133	GLY	B	0.66	0.66	0.59	0.62	0.63	0.53	0.46	0.44	0.47	0.43
134	THR	B	0.53	0.59	0.56	0.54	0.55	0.42	0.37	0.36	0.36	0.34
135	SER	B	0.50	0.60	0.59	0.50	0.51	0.41	0.35	0.34	0.34	0.33
136	GLY	B	0.53	0.64	0.67	0.55	0.52	0.39	0.35	0.35	0.35	0.33
137	SER	B	0.47	0.46	0.50	0.42	0.40	0.36	0.32	0.31	0.32	0.30
138	PRO	B	0.42	0.41	0.39	0.41	0.37	0.36	0.31	0.32	0.32	0.31
139	ILE	B	0.41	0.46	0.35	0.45	0.40	0.35	0.32	0.32	0.31	0.30
140	LEU	B	0.48	0.68	0.44	0.60	0.51	0.40	0.43	0.37	0.40	0.37
141	ASP	B	0.56	0.73	0.49	0.61	0.52	0.53	0.52	0.49	0.51	0.50
142	LYS	B	0.79	0.97	0.68	0.82	0.68	0.63	0.70	0.63	0.72	0.62
143	CYS	B	0.72	0.99	0.72	0.84	0.72	0.75	0.82	0.75	0.79	0.68
144	GLY	B	0.66	1.00	0.73	0.86	0.76	0.64	0.71	0.67	0.67	0.63
145	ARG	B	0.56	0.73	0.50	0.62	0.50	0.53	0.53	0.53	0.52	0.46
146	VAL	B	0.52	0.60	0.45	0.52	0.44	0.42	0.39	0.40	0.38	0.36
147	ILE	B	0.51	0.58	0.39	0.48	0.42	0.41	0.38	0.37	0.36	0.33
148	GLY	B	0.46	0.51	0.38	0.46	0.40	0.41	0.37	0.38	0.38	0.35
149	LEU	B	0.47	0.52	0.46	0.47	0.43	0.35	0.33	0.31	0.33	0.31
150	TYR	B	0.43	0.41	0.45	0.37	0.37	0.36	0.32	0.30	0.31	0.30
151	GLY	B	0.61	0.68	0.67	0.92	0.82	0.45	0.40	0.39	0.39	0.38
152	ASN	B	0.47	0.62	0.61	0.70	0.66	0.41	0.37	0.38	0.36	0.36
153	GLY	B	0.63	0.64	0.74	0.68	0.76	0.47	0.45	0.46	0.40	0.43
154	VAL	B	0.72	0.67	0.76	0.77	0.68	0.56	0.49	0.49	0.45	0.48
155	VAL	B	1.06	0.95	1.14	1.17	0.98	0.80	0.61	0.66	0.60	0.54

156	ILE	B	1.34	1.14	1.30	1.69	1.28	1.04	0.84	0.83	0.84	0.66
157	LYS	B	2.00	1.61	1.49	2.64	1.94	1.64	1.24	1.20	1.41	1.20
158	ASN	B	2.36	2.11	2.55	3.31	2.61	1.81	1.38	1.42	1.38	1.27
159	GLY	B	2.09	1.90	2.83	2.77	2.24	1.58	1.19	2.28	1.18	1.02
160	SER	B	1.25	1.37	1.87	1.84	1.61	1.09	0.88	0.96	0.76	0.75
161	TYR	B	0.71	0.96	0.86	0.89	0.83	0.66	0.55	0.50	0.50	0.50
162	VAL	B	0.48	0.54	0.54	0.52	0.48	0.43	0.37	0.37	0.37	0.36
163	SER	B	0.48	0.49	0.52	0.44	0.47	0.41	0.38	0.36	0.36	0.36
164	ALA	B	0.49	0.50	0.47	0.46	0.45	0.42	0.41	0.39	0.39	0.37
165	ILE	B	0.59	0.63	0.44	0.59	0.51	0.44	0.43	0.38	0.42	0.36
166	THR	B	0.54	0.63	0.45	0.62	0.54	0.42	0.43	0.38	0.44	0.36
167	GLN	B	0.65	0.83	0.54	0.75	0.65	0.46	0.42	0.42	0.43	0.38
168	GLY	B	0.83	1.06	0.80	0.98	0.77	0.59	0.52	0.64	0.54	0.48
169	LYS	B	1.00	1.11	0.88	1.03	0.84	0.90	0.69	0.95	0.75	0.71
170	ARG	B	1.21	1.34	1.14	1.32	1.06	0.77	0.71	0.80	0.69	0.62
171	GLU	B	1.63	1.34	1.68	1.26	1.28	1.15	1.29	1.14	0.99	0.79
172	GLU	B	2.33	2.18	2.42	1.92	2.01	1.92	1.77	2.00	1.65	1.16
173	GLU	B	3.97	4.29	4.32	3.58	4.63	2.48	2.06	3.03	2.10	1.56
174	THR	B	7.13	7.38	5.39	7.24	6.78	4.28	3.67	3.82	3.86	2.60
175	PRO	B	10.30	10.23	6.07	11.29	7.30	5.88	4.69	4.49	5.27	3.35
176	VAL	B	12.05	12.69	6.21	14.86	6.80	7.42	6.15	5.14	6.41	4.49

Table S5. Order parameter (S^2) analysis of the backbone (peptide N-H bond vector or C_α-C_β for proline) and sidechains (terminal bond vectors, not available for Gly, Ala and Trp). Chain A = NS2B, chain B = NS3. Binding-site residues are highlighted in green.

			backbone S^2					sidechain S^2				
residue	amino acid	chain	apo	complex-1	complex-2	complex-3	complex-4	apo	complex-1	complex-2	complex-3	complex-4
50	ASP	A						0.78	0.86	0.84	0.86	0.84
51	MET	A	0.64	0.70	0.67	0.74	0.71	0.41	0.28	0.26	0.28	0.33
52	TYR	A	0.91	0.92	0.91	0.91	0.91	0.89	0.89	0.89	0.90	0.87
53	ILE	A	0.90	0.89	0.90	0.90	0.90	0.89	0.89	0.88	0.88	0.88
54	GLU	A	0.91	0.91	0.91	0.90	0.91	0.51	0.12	0.37	0.44	0.67
55	ARG	A	0.89	0.89	0.89	0.89	0.89	0.89	0.88	0.89	0.89	0.89
56	ALA	A	0.88	0.87	0.88	0.87	0.88					
57	GLY	A	0.78	0.79	0.80	0.79	0.80					
58	ASP	A	0.87	0.85	0.85	0.87	0.85	0.85	0.85	0.85	0.84	0.84
59	ILE	A	0.83	0.85	0.85	0.84	0.86	0.69	0.18	0.24	0.47	0.37
60	THR	A	0.83	0.81	0.80	0.82	0.81	0.83	0.13	0.16	0.25	0.33
61	TRP	A	0.87	0.86	0.88	0.89	0.89					
62	GLU	A	0.83	0.79	0.82	0.82	0.82	0.26	0.57	0.45	0.23	0.27
63	LYS	A	0.85	0.82	0.84	0.85	0.85	0.30	0.15	0.15	0.19	0.14
64	ASP	A	0.76	0.78	0.79	0.84	0.82	0.48	0.45	0.60	0.76	0.35
65	ALA	A	0.73	0.71	0.73	0.84	0.84	0.22	0.35	0.15	0.19	0.20
66	GLU	A	0.62	0.62	0.65	0.83	0.83	0.27	0.13	0.04	0.09	0.11
67	VAL	A	0.54	0.61	0.64	0.76	0.79	0.58	0.29	0.37	0.55	0.72
68	THR	A	0.79	0.79	0.82	0.87	0.86	0.32	0.82	0.72	0.30	0.76
69	GLY	A	0.76	0.81	0.83	0.73	0.81					
70	ASN	A	0.60	0.57	0.72	0.42	0.53	0.07	0.19	0.16	0.28	0.14
71	SER	A	0.40	0.48	0.82	0.27	0.66	0.82	0.84	0.89	0.63	0.73
72	PRO	A	0.70	0.73	0.86	0.67	0.73	0.54	0.62	0.75	0.53	0.66
73	ARG	A	0.54	0.54	0.68	0.54	0.57	0.62	0.27	0.74	0.27	0.16
74	LEU	A	0.57	0.67	0.87	0.80	0.69	0.32	0.43	0.38	0.27	0.80
75	ASP	A	0.80	0.80	0.81	0.82	0.83	0.46	0.66	0.83	0.78	0.75
76	VAL	A	0.82	0.89	0.90	0.88	0.89	0.25	0.81	0.46	0.63	0.36
77	ALA	A	0.84	0.87	0.87	0.87	0.87	0.93	0.93	0.94	0.93	0.94
78	LEU	A	0.76	0.76	0.70	0.80	0.81	0.70	0.76	0.33	0.49	0.84
79	ASP	A	0.76	0.86	0.83	0.83	0.86	0.46	0.79	0.82	0.34	0.82
80	GLU	A	0.83	0.89	0.82	0.88	0.88	0.62	0.76	0.25	0.34	0.36
81	SER	A	0.77	0.87	0.84	0.86	0.88	0.79	0.87	0.83	0.81	0.88
82	GLY	A	0.88	0.89	0.84	0.89	0.89	0.67	0.64	0.65	0.66	0.72
83	ASP	A	0.75	0.80	0.74	0.81	0.83	0.84	0.53	0.49	0.84	0.69
84	PHE	A	0.83	0.89	0.84	0.89	0.89	0.87	0.46	0.55	0.87	0.87
85	SER	A	0.89	0.88	0.89	0.90	0.90	0.87	0.52	0.25	0.89	0.88
86	LEU	A	0.86	0.85	0.84	0.87	0.87	0.57	0.64	0.61	0.65	0.21
87	VAL	A	0.81	0.83	0.81	0.83	0.85	0.69	0.70	0.45	0.78	0.82
18	THR	B						0.21	0.19	0.22	0.28	0.14
19	THR	B	0.40	0.45	0.40	0.53	0.35	0.30	0.40	0.76	0.75	0.71
20	ASP	B	0.80	0.73	0.80	0.82	0.79	0.34	0.72	0.69	0.62	0.70
21	GLY	B	0.82	0.85	0.83	0.84	0.80					

70	GLY	B	0.86	0.86	0.86	0.87	0.86	0.07	0.19	0.16	0.28	0.14
71	ASP	B	0.86	0.87	0.87	0.87	0.88	0.82	0.84	0.89	0.63	0.73
72	VAL	B	0.85	0.87	0.88	0.88	0.87	0.54	0.62	0.75	0.53	0.66
73	LYS	B	0.88	0.91	0.90	0.89	0.91	0.62	0.27	0.74	0.27	0.16
74	GLN	B	0.82	0.89	0.86	0.86	0.89	0.32	0.43	0.38	0.27	0.80
75	ASP	B	0.89	0.90	0.89	0.90	0.91	0.46	0.66	0.83	0.78	0.75
76	LEU	B	0.92	0.92	0.92	0.92	0.92	0.25	0.81	0.46	0.63	0.36
77	VAL	B	0.92	0.92	0.94	0.93	0.93	0.93	0.93	0.94	0.93	0.94
78	SER	B	0.92	0.92	0.91	0.92	0.92	0.70	0.76	0.33	0.49	0.84
79	TYR	B	0.90	0.90	0.90	0.90	0.89	0.46	0.79	0.82	0.34	0.82
80	CYS	B	0.90	0.90	0.90	0.90	0.90	0.62	0.76	0.25	0.34	0.36
81	GLY	B	0.75	0.75	0.75	0.73	0.74	0.79	0.87	0.83	0.81	0.88
82	PRO	B	0.91	0.91	0.90	0.90	0.92	0.67	0.64	0.65	0.66	0.72
83	TRP	B	0.89	0.89	0.89	0.88	0.89	0.84	0.53	0.49	0.84	0.69
84	LYS	B	0.89	0.90	0.90	0.89	0.90	0.87	0.46	0.55	0.87	0.87
85	LEU	B	0.87	0.86	0.86	0.86	0.86	0.87	0.52	0.25	0.89	0.88
86	ASP	B	0.81	0.84	0.81	0.82	0.81	0.57	0.64	0.61	0.65	0.21
87	ALA	B	0.82	0.82	0.80	0.81	0.79	0.69	0.70	0.45	0.78	0.82
88	ALA	B	0.86	0.86	0.85	0.87	0.87					
89	TRP	B	0.86	0.86	0.86	0.86	0.86					
90	ASP	B	0.85	0.85	0.85	0.85	0.84	0.88	0.86	0.88	0.87	0.87
91	GLY	B	0.83	0.82	0.84	0.83	0.84					
92	LEU	B	0.87	0.85	0.87	0.85	0.85	0.19	0.20	0.18	0.13	0.26
93	SER	B	0.86	0.86	0.87	0.87	0.85	0.85	0.84	0.81	0.84	0.84
94	GLU	B	0.88	0.87	0.87	0.88	0.86	0.41	0.37	0.36	0.27	0.36
95	VAL	B	0.91	0.90	0.91	0.91	0.90	0.93	0.91	0.92	0.91	0.91
96	GLN	B	0.91	0.92	0.90	0.92	0.91	0.89	0.89	0.90	0.89	0.89
97	LEU	B	0.89	0.90	0.89	0.89	0.88	0.88	0.85	0.88	0.86	0.88
98	LEU	B	0.91	0.90	0.92	0.91	0.91	0.33	0.82	0.77	0.49	0.86
99	ALA	B	0.87	0.89	0.89	0.88	0.88					
100	VAL	B	0.90	0.90	0.90	0.90	0.91	0.91	0.91	0.91	0.92	0.92
101	PRO	B	0.90	0.83	0.90	0.91	0.91	0.68	0.45	0.70	0.70	0.72
102	PRO	B	0.84	0.84	0.85	0.84	0.86	0.38	0.36	0.38	0.40	0.45
103	GLY	B	0.80	0.81	0.82	0.82	0.82					
104	GLU	B	0.68	0.68	0.73	0.68	0.70	0.27	0.17	0.65	0.40	0.48
105	ARG	B	0.84	0.84	0.86	0.85	0.85	0.38	0.44	0.33	0.15	0.23
106	ALA	B	0.87	0.85	0.86	0.86	0.87					
107	LYS	B	0.87	0.87	0.89	0.85	0.86	0.24	0.48	0.53	0.18	0.15
108	ASN	B	0.80	0.85	0.87	0.83	0.80	0.89	0.88	0.89	0.89	0.88
109	ILE	B	0.88	0.87	0.87	0.88	0.88	0.49	0.43	0.62	0.27	0.26
110	GLN	B	0.88	0.88	0.89	0.90	0.90	0.75	0.75	0.82	0.82	0.84
111	THR	B	0.92	0.92	0.93	0.92	0.93	0.92	0.92	0.92	0.90	0.93
112	LEU	B	0.87	0.87	0.90	0.86	0.88	0.87	0.82	0.72	0.80	0.72
113	PRO	B	0.95	0.94	0.95	0.96	0.94	0.86	0.76	0.85	0.89	0.82
114	GLY	B	0.87	0.87	0.89	0.87	0.89					
115	ILE	B	0.85	0.82	0.88	0.82	0.88	0.38	0.18	0.33	0.30	0.19
116	PHE	B	0.87	0.88	0.90	0.90	0.88	0.83	0.81	0.86	0.82	0.85
117	LYS	B	0.86	0.85	0.88	0.88	0.87	0.19	0.15	0.36	0.18	0.16

22	VAL	B	0.92	0.91	0.91	0.91	0.91	0.91	0.92	0.91	0.91	0.91			
23	TYR	B	0.92	0.92	0.92	0.92	0.92	0.90	0.91	0.91	0.90	0.91			
24	ARG	B	0.92	0.91	0.91	0.91	0.92	0.89	0.85	0.89	0.89	0.90			
25	VAL	B	0.91	0.90	0.90	0.91	0.90	0.92	0.93	0.93	0.93	0.93			
26	MET	B	0.91	0.91	0.92	0.91	0.91	0.51	0.54	0.46	0.56	0.42			
27	THR	B	0.86	0.88	0.88	0.90	0.88	0.85	0.91	0.38	0.46	0.32			
28	ARG	B	0.84	0.86	0.86	0.84	0.84	0.21	0.33	0.65	0.14	0.20			
29	ARG	B	0.31	0.76	0.81	0.55	0.59	0.11	0.16	0.15	0.37	0.19			
30	LEU	B	0.33	0.53	0.63	0.56	0.50	0.15	0.32	0.30	0.17	0.15			
31	LEU	B	0.56	0.29	0.57	0.60	0.53	0.19	0.26	0.23	0.29	0.18			
32	GLY	B	0.37	0.17	0.34	0.15	0.22								
33	SER	B	0.68	0.48	0.51	0.38	0.37	0.32	0.53	0.58	0.46	0.13			
34	THR	B	0.67	0.80	0.77	0.79	0.82	0.15	0.46	0.56	0.36	0.73			
35	GLN	B	0.86	0.88	0.88	0.87	0.86	0.88	0.88	0.88	0.88	0.89			
36	VAL	B	0.87	0.88	0.88	0.86	0.87	0.90	0.90	0.90	0.75	0.40			
37	GLY	B	0.88	0.89	0.89	0.89	0.88								
38	VAL	B	0.92	0.92	0.92	0.92	0.93	0.94	0.94	0.94	0.94	0.94			
39	GLY	B	0.92	0.92	0.92	0.92	0.92								
40	VAL	B	0.91	0.90	0.91	0.90	0.90	0.87	0.70	0.13	0.47	0.42			
41	MET	B	0.92	0.91	0.90	0.92	0.91	0.30	0.56	0.22	0.29	0.26			
42	GLN	B	0.92	0.92	0.92	0.92	0.93	0.27	0.22	0.24	0.26	0.54			
43	GLU	B	0.90	0.91	0.90	0.91	0.90	0.42	0.51	0.50	0.65	0.58			
44	GLY	B	0.84	0.88	0.85	0.84	0.85								
45	VAL	B	0.83	0.85	0.81	0.85	0.83	0.93	0.93	0.93	0.93	0.93			
46	PHE	B	0.88	0.89	0.90	0.88	0.89	0.85	0.87	0.86	0.85	0.86			
47	HIS	B	0.90	0.90	0.89	0.90	0.90	0.88	0.87	0.87	0.88	0.89			
48	THR	B	0.93	0.93	0.92	0.92	0.93	0.94	0.94	0.94	0.94	0.94			
49	MET	B	0.91	0.91	0.91	0.90	0.91	0.82	0.83	0.82	0.82	0.85			
50	TRP	B	0.90	0.91	0.91	0.90	0.91	0.78	0.86	0.84	0.86	0.84			
51	HIS	B	0.92	0.93	0.93	0.93	0.93	0.41	0.28	0.26	0.28	0.33			
52	VAL	B	0.88	0.90	0.90	0.90	0.90	0.89	0.89	0.89	0.90	0.87			
53	THR	B	0.87	0.88	0.88	0.88	0.87	0.89	0.89	0.88	0.88	0.88			
54	LYS	B	0.90	0.90	0.90	0.90	0.89	0.51	0.12	0.37	0.44	0.67			
55	GLY	B	0.85	0.85	0.86	0.85	0.83	0.89	0.88	0.89	0.89	0.89			
56	ALA	B	0.83	0.81	0.83	0.82	0.82								
57	ALA	B	0.82	0.81	0.82	0.83	0.83								
58	LEU	B	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.85	0.84	0.84			
59	ARG	B	0.67	0.65	0.67	0.66	0.60	0.69	0.18	0.24	0.47	0.37			
60	SER	B	0.89	0.90	0.90	0.90	0.90	0.83	0.13	0.16	0.25	0.33			
61	GLY	B	0.85	0.85	0.85	0.85	0.83								
62	GLU	B	0.73	0.75	0.77	0.79	0.68	0.26	0.57	0.45	0.23	0.27			
63	GLY	B	0.70	0.63	0.66	0.67	0.59	0.30	0.15	0.15	0.19	0.14			
64	ARG	B	0.66	0.65	0.65	0.64	0.63	0.48	0.45	0.60	0.76	0.35			
65	LEU	B	0.79	0.74	0.76	0.78	0.78	0.22	0.35	0.15	0.19	0.20			
66	ASP	B	0.78	0.82	0.82	0.82	0.82	0.27	0.13	0.04	0.09	0.11			
67	PRO	B	0.85	0.87	0.86	0.87	0.88	0.58	0.29	0.37	0.55	0.72			
68	TYR	B	0.74	0.78	0.78	0.78	0.78	0.32	0.82	0.72	0.30	0.76			
69	TRP	B	0.89	0.90	0.90	0.90	0.90								

“

3.2. Improving Binding Entropy by Introducing Higher Ligand Symmetry

3.2.1. Context, Project Summary, and Own Contribution

ITC has established its position in rational drug design as an indispensable method for the thermodynamic characterization of binding events. Using computational methods such as forcefield-based scoring functions, the enthalpic contribution to binding affinity can be predicted as a sum of single interactions with constantly improving accuracy.[227,228] More sophisticated entropic contributions such as rotational preorganization, rigidification, and (de-)solvation effects are regularly simplified or omitted.[227–229] Thus, the prediction of entropic effects demands time- and budget-consuming MD simulations and experimental verification. In this work, we shed light on the rotational preorganization of MT-SP1 inhibitors using molecular docking and ITC experiments. Therefore, we designed a set of trivalent inhibitors and introduced ligand symmetry to reduce the number of energetically distinguishable binding modes. This project was derived from a previous work of the group of [redacted]. They developed symmetric inhibitors of trypsin-like proteases, including MT-SP1.[230] The improved inhibitory activity was attributed to the higher ligand symmetry, entropically benefiting inhibitor binding. This hypothesis, however, was not confirmed by comparison of a suitable set of inhibitors and determining their thermodynamic binding profiles.[230] The binding site of MT-SP1 is particularly well suited for experiments concerning ligand symmetry as it has a highly negative surface potential and tolerates three basic residues in the deep S1 and the shallow S2 and S3/4 pockets.[88,89,230] These features will likely enable multiple rotational binding modes of di- or tribasic inhibitors as long as a basic residue is placed into the S1 pocket (Figure 18A).[74,230–232] Therefore, [redacted] kindly provided us with two inhibitors (cpds **8** and **9**) of his work (Figure 18B), which should differ in the proportion of advantageous binding modes from all rotating binding modes (Figure 18C).[230]

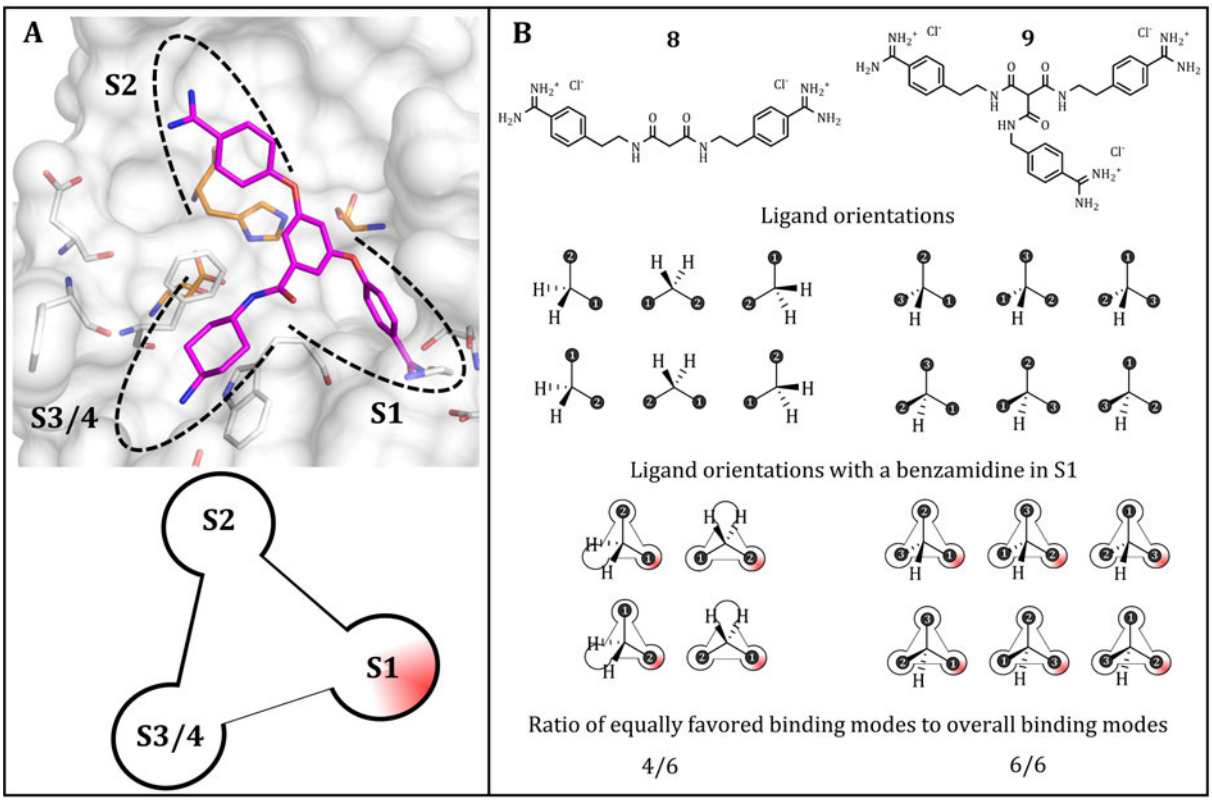


Figure 18. *The symmetric binding site of MT-SP1 and rotational preorganization of ligands **8** and **9**.* (**A**) Top: The tribasic inhibitor (*N*-(4-aminocyclohexyl)-3,5-bis(4-carbamimidoylphenoxy)benzamide) (pink) in complex with MT-SP1 (PDB-ID: 4O9V).[91] The protein surface is shown in white, and the catalytic triad is shown as orange sticks. For a clear view, only essential amino acids forming the binding pockets are depicted as sticks. Binding sub-pockets are indicated according to Schechter&Berger.[43] Figure created using PyMOL.[44] Bottom: schematic depiction of the binding site with highlighted S1 pocket as more critical for substrate recognition. (**B**) 2D-projection of the inhibitors **8** and **9** and their respective schematic depiction[230] in all possible orientations (120°-steps rotated and flipped) of cpds **8** (left) and **9** (right). From the equally preferred orientations with a benzamidine moiety residing in the S1 pocket, the ratio of equally favored binding modes to overall binding modes can be calculated as a proxy for the statistical benefit in binding entropy.

[redacted] supported us with an MT-SP1 plasmid to establish the purification of recombinant MT-SP1 from *E. coli,* including solubilization from inclusion bodies, refolding, and subsequent autoactivation, eventually yielding highly active and pure protease. First, enzyme inhibition assays and ITC experiments showed a more beneficial binding entropy of the branched trivalent inhibitor **9** than the linear inhibitor **8**. This proof-of-concept encouraged us to design a set of four trivalent inhibitors (cpds **10**–**13**), harboring amine and amidine residue permutations and lacking the central sp^3 hybridized carbon, thus allowing us to conclude the influences of rotational preorganization (Figure 19).

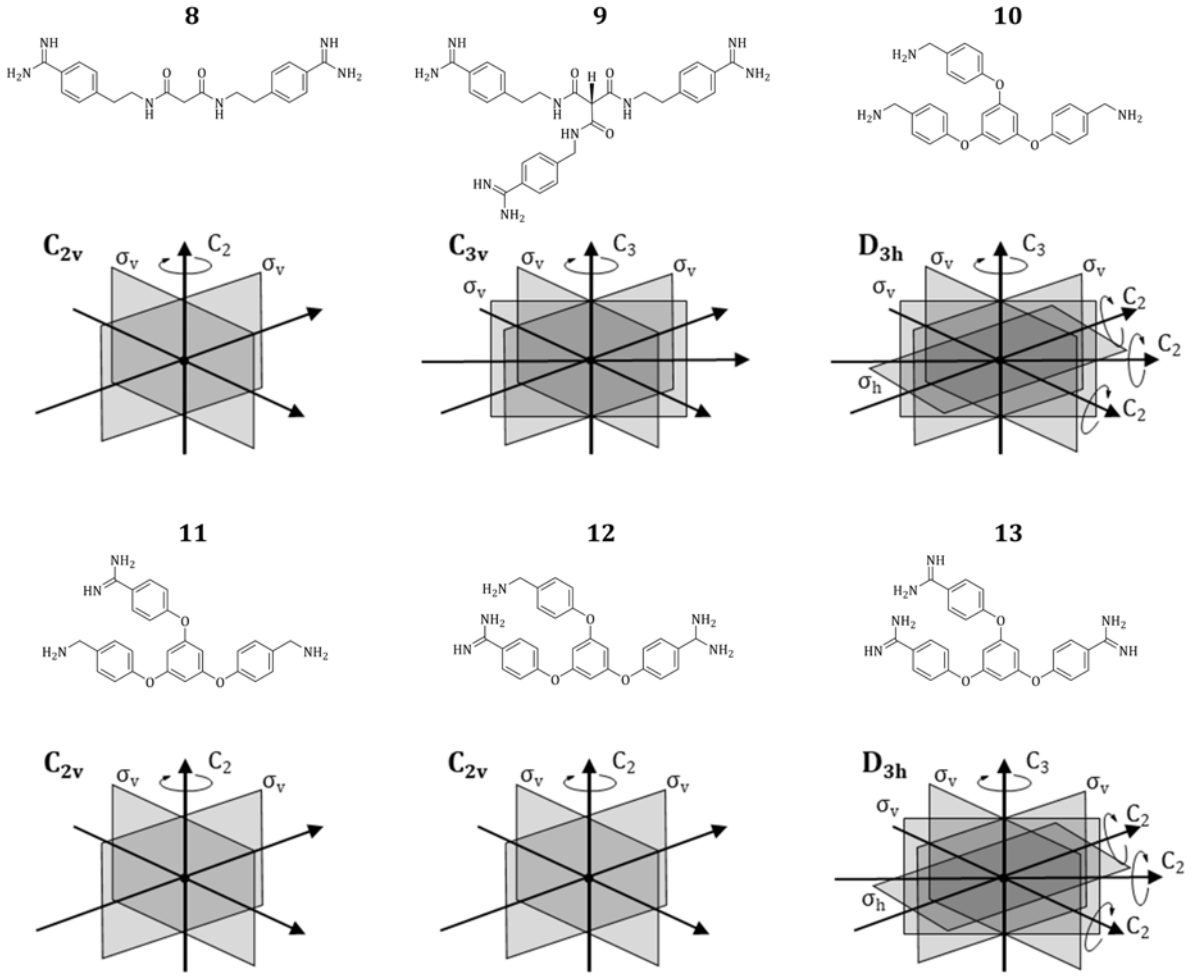


Figure 19. *2D-Structure and Symmetry point groups of the cpds **8–13***. Cps **8** and **9** were provided by [redacted] as reference compounds. Since they harbor a stereocenter, we further improved ligand symmetry by designing cpds **10** and **13** with a phloroglucinol joint.

Molecular docking studies indicated that an amidine residue is always preferred in the S1 pocket, limiting the number of beneficial binding modes out of all possible rotational orientations. The proportion of beneficial modes could be correlated with higher inhibitory potencies of these inhibitors in enzyme inhibition assays. Since the affinities were also in line with the number of amidine residues, we could not allocate them to the entropic benefits. Therefore, ITC experiments were employed to determine the contribution of binding entropy to the observed affinities and binding modes. Indeed, with a higher ratio of beneficial binding modes, the entropic penalty of inhibitors **10–13** was reduced, although these results were not pronounced. However, calculations using a mathematical model that describes the additivity of binding affinities of multiple binding modes were in good accordance with our observations. Altogether, we could show that introducing ligand symmetry can be a viable strategy to reduce the number of potential binding modes, hence achieving a superior binding entropy.

(2) **Hammerschmidt, S. J.**[#] [redacted] Improving binding entropic by ligand symmetry – a case study with human matriptase. *Manuscript submitted to RSC Medicinal Chemistry.*

Own contribution: Establishment of enzyme expression and purification, establishment of matriptase inhibition assays, establishment of ITC experiments, writing of the original draft & editing of the manuscript.

Contributions from other authors: Enzyme expression and purification, matriptase inhibition assays, ITC experiments, synthesis of inhibitors, molecular docking studies, writing parts of the chapter "Synthesis of trivalent Inhibitors", writing of the experimental section "Chemistry" and "Molecular Docking Studies" & editing of the manuscript.

Manuscript submitted to *RSC Medicinal Chemistry* (impact factor: 3.47).

3.2.2. Publication

The following manuscript, from page 89 to page 128, contains unpublished data.

Improving binding entropy by ligand symmetry – a case study with human matriptase.

Stefan J. Hammerschmidt[a,†], [redacted]

[a] Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany
[b] Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany
† Authors contributed equally.
* E-Mail: [redacted]

Abstract

Understanding different contributions to the binding entropy of ligands is of utmost interest to better predict affinity and the thermodynamic binding profiles of protein-ligand interactions and to develop new strategies for ligand optimization. To these means, the largely neglected effects of introducing higher ligand symmetry, thereby reducing the number of energetically distinguishable binding modes on binding entropy using the human matriptase as a model system, was investigated. A set of new trivalent phloroglucinol-based inhibitors that address the roughly symmetric binding site of the enzyme was designed, synthesized, and subjected to isothermal titration calorimetry. These highly symmetric ligands that can adopt multiple indistinguishable binding modes exhibited high entropy-driven affinity in line with affinity-change predictions.

INTRODUCTION

Protein-ligand binding thermodynamics.

Isothermal titration calorimetry (ITC) has become an indispensable method to characterize binding thermodynamics in rational drug design. In a single ITC experiment, the enthalpic (ΔH) and temperature-dependent entropic (–TΔS) contributions to the overall binding free energy (ΔG) can be quantified. The determined thermodynamic binding profile is composed of several energetic superimposed contributions. Thus, attributing the observed thermodynamics to the respective underlying events is a highly complex procedure.[1] The ΔH term reflects not only direct interactions between the isolated binding partners but also the breaking and formation of hydrogen bonds in desolvation and resolvation events,[2] and proton transfers.[3] Hydrophobic interactions,[4] conformational flexibility, and changes in rotational, translational, and torsional degrees of freedom influenced –TΔS.[5] Several studies thoroughly investigated those contributions to binding entropy, predominantly in terms of intrinsic rigidification of the binding partners.[6,7] Although occasionally appreciated as beneficial for binding affinity[8,9] one so far underrepresented aspect is the impact of ligand symmetry on binding thermodynamics by enabling multiple binding modes.

The human matriptase.

The human matriptase belongs to the family of human type II transmembrane serine proteases (TMPRSS) and is known under several aliases (TMPRSS14; suppressor of tumorgenicity 14 (ST14); prostamin; membrane-type serine proteinase 1 (MT-SP1)). Along with other TMPRSS, MT-SP1 shares a characteristically conserved *N*-terminal transmembrane region, a variable stem region, and a *C*-terminal trypsin-like serine protease region.[10] MT-SP1 was first identified as a protease, highly expressed in breast cancer cells, and found to be involved in matrix degradation. Hence it is also associated with tumor invasion and metastasis.[11,12] Several substrates of MT-SP1 were found to be implicated with tumorigeneses, such as hepatocyte growth factor (HGF) and urokinase-type plasminogen activator (uPA).[13] Therefore, inhibitors for MT-SP1 may represent potential anticancer agents.[14,15] Besides that, MT-SP1 plays critical roles in epithelial barrier formation, hair follicle development, and thymic functions.[16] Furthermore, it could be targeted to modify osteoarthritis as it increases metalloproteinase activity.[17] The MT-SP1 catalyzes proteolysis with its catalytic triad, formed by the amino acids (aa) His57, Asp102, and Ser195 (Figure 1). The negatively charged Asp189 at the bottom of the S1 pocket determines the primary substrate specificity for basic aa. In the S2 pocket, the size of the preferred aa is determined by the

proximal residue Phe99, sterically hindering larger substrate residues. The fused S3/4 subsite is rather shallow and forms an open distal binding area, with Trp215 forming the bottom. Phe97 closes the distal part of the S3/4 pocket.[18,19]

Inhibitors of MT-SP1.

Besides substrate-analog ligands, for MT-SP1, several mono- to tri-basic small molecules were identified as potent inhibitors.[14,20,21] The attraction of multi-cationic ligands is enforced by the highly negative surface potential of the MT-SP1 active site (Figure S1A). As known for several trypsin-like proteases, benzamidine moieties were successfully used to address the S1 subsite as arginine mimetics.[22,23] Exemplarily, this can be observed in the MT-SP1-ligand complex structure of the tri-basic inhibitor (*N*-(4-aminocyclohexyl)-3,5-bis(4-carbamimidoylphenoxy)benzamide, **15**) (Figures 1, S1B, RCSB protein data bank (PDB) ID: 4O9V).[24] The inhibitor features one benzamide moiety, deeply burrowed in the S1 pocket, while another benzamidine is oriented towards the more solvent-exposed S2 subsite. The central aromatic ring connects the benzamidines *via* meta-substituted ether bonds and one amide to a *N*-4-aminocyclohexyl moiety. The benzamidine in the S1 pocket forms several polar interactions with the sidechains of Asp189 and Ser190 and the backbone oxygens of Ser190 and Gly219 (Figure 1). An interaction with the catalytic Ser195 stabilizes the phenolic ether oxygen. The second benzamidine moiety reaches the S2 pocket without forming any direct interactions but resides adjacent to Asp60A/B and Asp96. In the S3/4 pocket, Gln175 forms a hydrogen bond with the terminal amine of the cyclohexyl residue.[14,15,22,25,26]

Ligand symmetry.

Knowledge of the correct binding mode is an essential prerequisite for rational structure-based inhibitor design. When crystal structures are not directly accessible, molecular docking studies can support decision-making to develop a binding mode hypothesis.

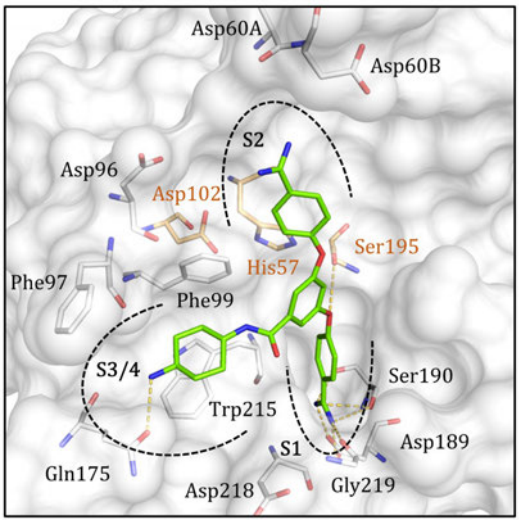


Figure 1. The binding pocket of the cocrystal structure of MT-SP1 (white) with the tribasic inhibitor **15** (green) with polar interactions as yellow dashed lines (PDB ID: 4O9V).[24] Selected aa are shown as sticks and labeled according to the chymotrypsinogen numbering (catalytic triad: orange sticks). Sub-pockets are indicated with black dashed lines according to Schechter&Berger.[27] Figure was created using PyMOL.[28]

In the case of multibasic inhibitors of MT-SP1, however, it is challenging to assign basic residues to their correct binding subsites.[29] This becomes even more challenging when the binding scores of the predicted binding poses are highly similar. Since even X-ray structures of MT-SP1-bisbenzamidine complexes indicate that multibasic inhibitors can bind in different orientations,[22,24] Furtmann *et al.*[29] introduced higher symmetry to bi- and tribasic ligands to reduce the number of energetically distinguishable binding modes and to ease the selection of the correct one. The bivalent inhibitor **1** and the trivalent inhibitor **2** reduce the number of distinguishable binding modes from six to two for compound (cpd) **1** (assuming S1 binding) and to one or two for cpd 2, depending on whether there is a preferred orientation of the central proton (Table 1).[29] In the binding events of a protein P and a ligand L, the K_d is defined by the concentrations of the free binding partners and the complex concentration (Equation 1).

$$K_d = \frac{[P]\cdot[L]}{[PL]} = \frac{1}{K_a} \quad (1)$$

In the case of multiple (n) binding modes, the concentrations of each complex can be summed up.[8] Hence, the apparent association constant $K_{a,app}$ (the reciprocal value of the apparent dissociation

constant $K_{d,app}$) can be described as the sum of the incremental association constants $K_{a,inc}$ of each binding mode (Equation 2).

$$K_{d,app} = \frac{[P] \cdot [L]}{[PL]_1 + [PL]_2 + \cdots + [PL]_n} = \frac{1}{K_{a,app}} = \frac{1}{\sum_{i=1}^{n} K_{a,inc,n}} \quad (2)$$

Subsequently, the observed affinity described by the $K_{a,app}$-value linearly scales with the number of identical (same $K_{a,inc}$) binding modes. Due to the relationship between $K_{a,app}$ and ΔG (Equation 3), with R as the universal gas constant and T as the absolute temperature in K, doubling the number of identical binding modes, for example, always provides a change of –1.7 kJ·mol^{-1} of binding free energy (ΔΔG, Equation 4 with same $K_{a,inc}$) at 25 °C.

$$\Delta G = -RT \cdot \ln(K_a) \quad (3)$$

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 \quad (4)$$

$$= -RT \cdot \ln(K_{a,inc}) - \left(-RT \cdot \ln(2 \cdot K_{a,inc})\right)$$

$$= -RT \cdot \ln\left(\frac{2}{1}\right) = -1.7 \text{ kJ} \cdot \text{mol}^{-1}$$

Additionally, as one macrostate (the formation of the protein-ligand complex [PL]) can be realized differently, the gain in binding affinity should appear as a more favorable binding entropy.

RESULTS AND DISCUSSION

Trivalent MT-SP1 inhibitor design.

In this work, we follow up on this idea and analyze the thermodynamic effects of higher symmetry in MT-SP1 ligands. Therefore, a set of trivalent inhibitors was designed to yield different numbers of identical binding modes. In theory, non-symmetric dibasic ligands can accommodate six different possible binding modes within the MT-SP1 binding site. For symmetric linear inhibitors like cpd **1**, each binding mode is identical to its 180° rotated orientation resulting in three possible solutions. Since a benzamidine residue can be considered deeply burrowed in the S1 pocket, for ligand **1**, instead of three, only two binding modes are relevant (occupation of S2 and S3/4 with a free S1 is less likely). In the case of the trivalent MT-SP1 binding pocket, the branched inhibitors can rotate in approximately 120° steps to place one residue in the S1, one in S2, and the last in S3/4. For inhibitors with three different substituents, it further must be considered that the quaternary central carbon atom allows for two distinct orientations, one with the hydrogen towards the protease and one with the hydrogen towards the solvent. This gives six possible binding modes for each potential enantiomer which is reduced to only two for inhibitor **2**, which shares three identical substituents. To achieve even higher ligand symmetry than cpd **2**, the central moiety was exchanged by phloroglucinol to remove the central sp^3 carbon, thereby adding a further symmetry plane (Table S1). By decorating the phloroglucinol with either *p*-benzamidine or *p*-benzylamine substituents (cpds **3**–**6**), we obtained the highly symmetric cpds **3** and **6** in the D_{3h} point group and the less symmetric cpds **4** and **5** in the C_{2v} point group (Table S1). However, assuming benzamidine is preferred over benzylamine in the S1 pocket, the number of potential energetic different binding modes was reduced to three or fewer. Molecular docking studies (Table S2) of these cpds suggested that binding modes are well aligned with that of cpd **15**. For cpd **1**, two orientations with similar binding scores can be observed, with one benzamidine moiety deeply bound in the S1 pocket and the second benzamidine either in the S2 or S3/4, supporting the assumption that one residue needs to be placed in the S1 pocket. For cpd **2**, only conformations with the central proton pointing to the protein were obtained, indicating a preferred orientation of this inhibitor and only 3 of 6 equally preferred binding modes. In the docking studies with the inhibitors **3**–**6,** a preference for benzamidine substituents was observed, especially in the S1 sub-pocket, as reported previously (PDB IDs: 1EAX, 4R0I).[19,24,30] This results in six identical binding modes for inhibitor **3** harboring three benzylamine moieties and two favorable out of six possible binding modes for cpd **4** with the benzamidine in S1 while benzylamines occupy S2 and S3/4. For cpd **5**, two different binding modes with highly similar docking scores were observed in which the benzylamine occupies either the S3/4 or the S2 pocket while the benzamidines address the respective other sub-pocket and S1 (Figure 2B). As for cpd **3**, all six possible binding modes are identical for cpd **6** with three amidines (Figure 2A). This highlights the difficulty of selecting the true binding mode for asymmetric inhibitors with several possible binding modes or eventually

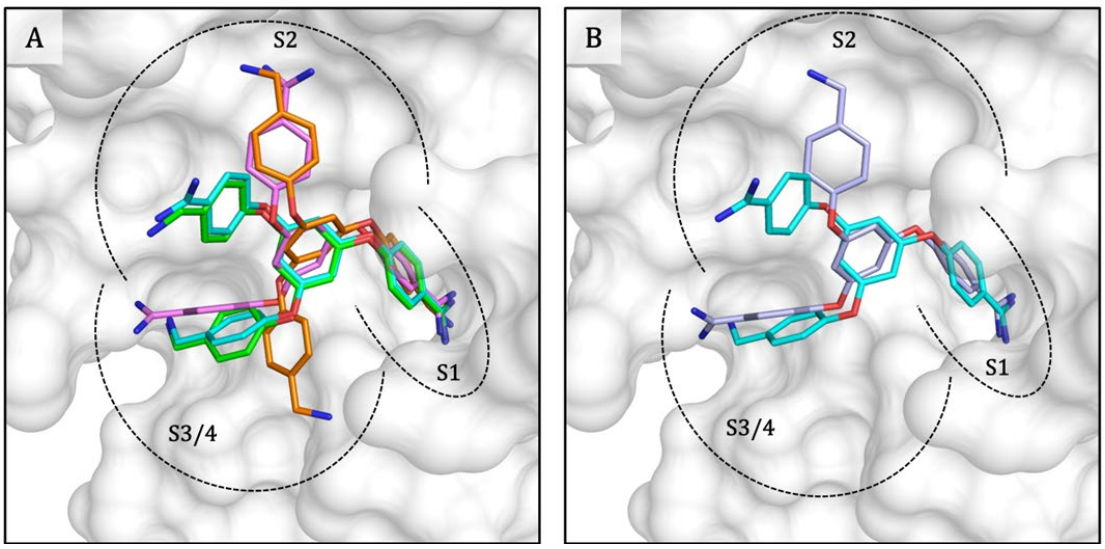
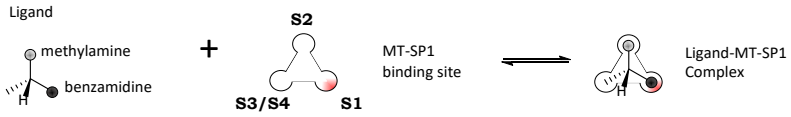
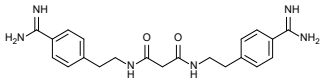
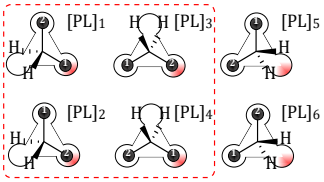
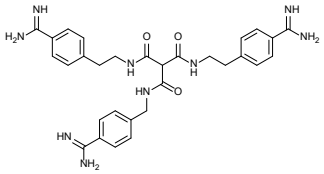
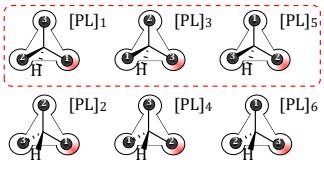
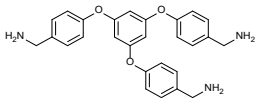
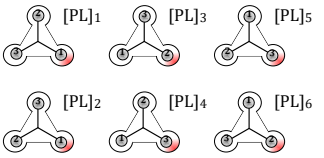
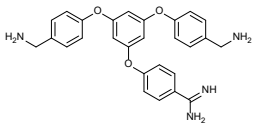
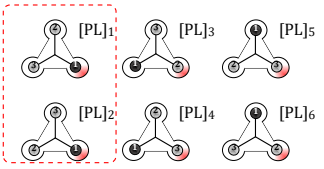
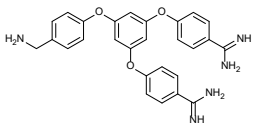
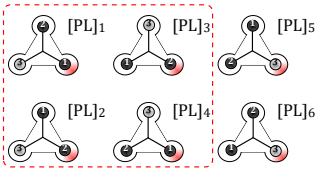
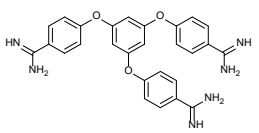
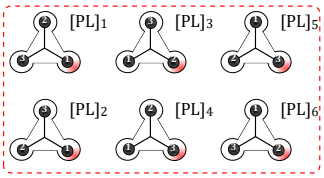


Figure 2. Binding poses of **3**–**6** predicted by molecular docking against MT-SP1 (PDB ID: 4O9V).[24] **A** The top scored binding poses of cpds **3**–**6** (**3**: pink sticks, **4:** green sticks, **5**: cyan sticks, **6**: orange sticks) into the active site of MT-SP1 (white surface). Amine functionalities are colored purple, amidines are colored blue. Binding sub-pockets are indicated according to Schechter&Berger[27] with black dashed lines. **B** The two predicted binding poses of cpd **5** with one amidine in S1 and the second amidine either in the S2 (cyan; FlexX-score –42.7 kJ·mol^{-1}, rank 1) or in S3/4 binding pocket (pale blue; FlexX-score –41.9 kJ·mol^{-1}, rank 15). Generally, within S2, orientations of the basic moieties towards Asp60B or Asp96 were observed with similar scores.

indicates multiple different but energetically rather similar binding modes for cpds **1** and **5** (Table 1). While ligand binding is generally accompanied by an entropic penalty upon ligand preorganization and the loss of rotational, translational, and torsional degrees of freedom, multiple binding modes with identical interactions should reduce that penalty and result in a more favorable temperature-dependent entropy term (–TΔS). Based on these considerations, the thermodynamic binding profiles should differ for tri-basic modified phloroglucinol derivatives carrying *p*-benzylamine or *p*-benzamidine sidechains (**3**–**6**) following their number of similar possible binding solutions. However, these effects may be superimposed by the influence of the methylamine to amidine exchange. Inhibitor **6** might display the most favorable binding entropy (ratio of preferred orientations 6/6), followed by cpds **5** (4/6) and **4** (2/6). Cpd **3** was used to evaluate the effects of amine to amidine replacement in the S1 pocket.

Table 1. Binding modes of inhibitors **1**–**6**. The importance of the amidine placement in the S1 pocket is highlighted in red. Graphical depictions are modified after Furtmann *et al.*[29]



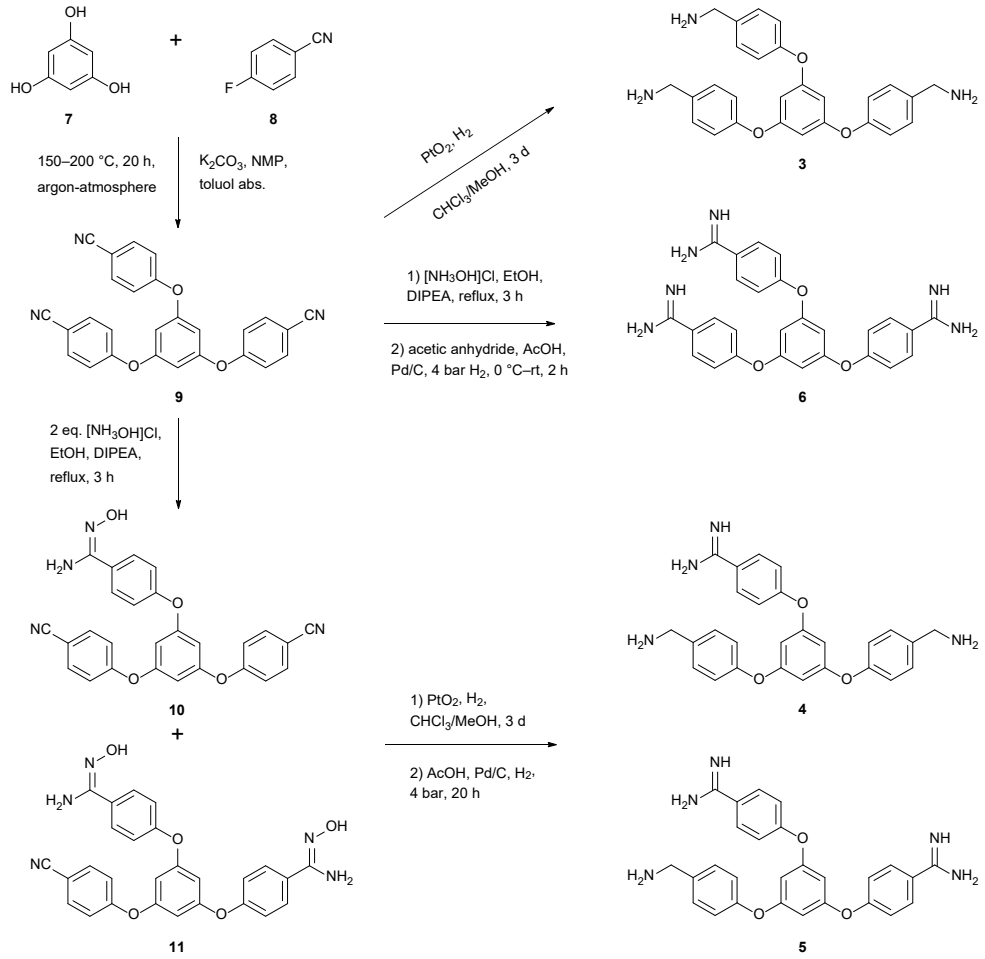
Cpd	Structure	Theoretical ligand orientations	Relationship of Ka-values	Ratio of preferred orientations per possible orientations
1			Ka,1 = Ka,2 ≈ Ka,3 = Ka,4 << Ka,5 = Ka,6	0.33 (2/6) or 0.67 (4/6)a
2			Ka,1 = Ka,3 = Ka,5 ≈ Ka,2 = Ka,4 = Ka,6	0.5 (3/6) or 1.0 (6/6)b
3			Ka,1 = Ka,2 = Ka,3 = Ka,4 = Ka,5 = Ka,6	1.0 (6/6)
4			Ka,1 = Ka,2 << Ka,3 = Ka,4 ≈ Ka,5 = Ka,6	0.33 (2/6)
5			Ka,1 = Ka,2 ≈ Ka,3 = Ka,4 << Ka,5 = Ka,6	0.33 (2/6) or 0.67 (4/6)a
6			Ka,1 = Ka,2 = Ka,3 = Ka,4 = Ka,5 = Ka,6	1.0 (6/6)

a: depending on whether amidine is preferred over methylamine in S2 or S3/4 or exhibits similar affinity; b: depending on whether there is a preferred orientation of the central proton.

Synthesis.

Symmetrically and asymmetrically branched trivalent inhibitors **3**–**6** were synthesized in a three- to four-step reaction starting from phloroglucinol (Scheme 1).

Phloroglucinol **7** reacted with *p*-fluorobenzonitrile **8** in a base-catalyzed substitution reaction. After column chromatography, 4,4',4''-(benzene-1,3,5-triyltris(oxy))tribenzonitrile **9** was obtained in moderate yield. The symmetrical inhibitor with three amine substituents **3** was obtained by reduction of **9** with platinum(IV) oxide. Purification by HPLC afforded ((benzene-1,3,5-triyltris(oxy))tris(benzene-4,1-diyl))trimethanamine **3**. The corresponding cpd **6** with three amidine substituents was obtained from cpd **9** in a two-step reaction. By reacting **9** with an excess of hydroxylammonium chloride in a base-catalyzed reaction, the corresponding *N*'-hydroxybenzimidamide substituted cpd was obtained as an intermediate. Reduction at 4 bar in acetic acid with palladium on activated carbon and acetic anhydride, resulting in acetylation of the hydroxyl group, gave the corresponding triple benzimidamide substituted cpd **6**. 4,4',4''-(benzene-1,3,5-triyltris(oxy))tribenzimidamide **6** was obtained by purification by HPLC. The asymmetric inhibitors **4** and **5** were obtained by slightly modifying the previously described reaction procedure. Cpd **9** was reacted with 2 eq. hydroxylammonium chloride in a base-catalyzed reaction. The resulting mixture of cpds **10** and **11** was separated by HPLC.



Scheme 1. Synthetic route to obtain trivalent inhibitors containing either three amines (cpd **3**), three amides (cpd **6**), or the respective permutations (cpds **4**, **5**).

Cpds **10** and **11** were first reduced with platinum(IV) oxide, selectively reducing the nitrile to the amine.

The intermediate product obtained in each case was then reduced in acetic acid with palladium on activated carbon at 4 bar. Just as in the reaction of **9** to **6**, the hydroxyl group is first acetylated by acetic acid. The target cpds **4** and **5** were obtained by purification by several subsequent HPLC runs.

Inhibitory activities.

As described previously, inhibitor **2**, carrying trivalent benzamidines, is more potent than cpd **1** with only two benzamidines (Table 2).[29] The first benzamidine residue increases the potency more than introducing the second or third benzamidine. This indicates a specific subsite, S1, that strongly favors benzamidine over benzylamine, whereas the other two subsites appear less discerning.[19,30] To exclude different protonation states of the ligand substituents contributing to differences in inhibitory potency, pk_a-values of all cpds were calculated (Figure S3). According to the predictions, amidines are more basic (pk_a-values 11.2–12.2) than methylamines (pk_a-values 8.8–9.7), but the dominant protomers of all ligands carry a formal charge of +3 under assay conditions (pH = 8). Increasing the pH of the inhibition assay conditions to 8.5 or 9 did not increase the K_i of cpd **6**, while the K_i of cpd **3** slightly increased, hinting at de-protonation of that cpd resulting in weaker binding (Table S3). More acidic conditions of pH 7.5 and 7.0 reduced the protease activity drastically (Table S3, Figure S4).

Thermodynamic binding profiles.

ITC experiments were performed to quantify the enthalpic and entropic contributions to ΔG of inhibitors **1**–**6** (Table 2, Figures 3, S5). ITC measurements for the reported cpds 1 and 2 were conducted in tris(hydroxymethyl)aminomethane (TRIS) buffer for experiment establishment. The obtained K_d values are in excellent accordance with the reported K_i values.[29] A comparison of the thermodynamic binding profile of **1** and **2** revealed a loss in binding enthalpy upon introduction of the third benzamidine residue, whereas the binding entropy is highly increased (ΔH: –14.0 *vs*. –6.77 kJ·mol^{-1}; –TΔS: –19.4 *vs*. –30.1 kJ·mol^{-1} for **1** and **2**, respectively). The large hydration shell of amidines, which is (partly) stripped upon binding, could lead to a positive ΔH, and a negative –TΔS contribution could explain this.[31] However, introducing an additional branch into cpd **1** to yield cpd **2** comes with several altered interactions that do not allow ascribing the observed entropic effects to the higher ligand symmetry. Next, inhibitors **3**–**6** were subjected to ITC experiments. These experiments were performed in TRIS and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer to correct for proton transfer since they display sufficiently different ionization enthalpies of 20.4 and 47.5 kJ·mol^{-1} (Table S4).[32] Upon binding of cpds **3**–**6** to MT-SP1, no protons were transferred (number of transferred protons (n_{proton}) = –0.03, –0.06, –0.09, –0.08, respectively). This supports the hypothesis of the inhibition assay and pk_a-prediction that all ligands are present in the fully protonated state under assay conditions. Consequently, differences in affinity do not arise from different ligand charges (Figure S3). The obtained dissociation constants (K_d) also align with the results from the inhibition assay and lie within the expected fluctuation between different methods.[33,34] Noteworthy, the trend in increasing affinities for inhibitors with successive more benzamidine moieties is consistent with K_d values of 10.2, 2.52, 1.20, and 0.637 μM for cpds **3**–**6**, respectively. We achieved potent MT-SP1 binders that are mainly entropy-driven, whereas cpds **1** and **2** showed heat release upon binding. It can be hypothesized that the ligand symmetry in terms of three basic residues increases the number of potential binding modes and hence favors entropy. Since inhibitors **3**–**6** have shorter substituents and less rotatable bonds than cpds **1** and **2**, this can probably explain the thermodynamic differences of better binding entropy and worse binding enthalpy. Cpds **1** and **2** with more flexible bonds can be organized to form better polar interactions with the MT-SP1 sub-pockets but undergo rigidification upon binding. In contrast, smaller, more rigid cpds **3**–**6** are entropically favored but cannot form optimal polar interactions, resulting in an enthalpy-entropy compensation.[35]

Table 2. Binding constants and thermodynamic binding profiles of inhibitors **1**–**6**.

Cpd	K_i (μM)	K_d (μM)	ΔG° (kJ·mol^{-1})	ΔH° (kJ·mol^{-1})	−TΔS° (kJ·mol^{-1})	n_{proton}
1	4.23 ± 0.15[a]	1.45 ± 0.19	−33.4 ± 0.2	−14.0 ± 0.5[b]	−19.4 ± 0.4[b]	n.d.
2	0.393 ± 0.004[a]	0.392 ± 0.121	−36.9 ± 1.2	−6.8 ± 2.6[b]	−30.1 ± 1.3[b]	n.d.
3	18.4 ± 2.0	10.2 ± 1.2	−28.9 ± 0.9	5.2 ± 0.2	−34.1 ± 0.6	−0.04
4	0.775 ± 0.115	2.52 ± 0.32	−32.1 ± 0.3	8.5 ± 0.2	−40.6 ± 0.7	−0.06
5	0.459 ± 0.051	1.20 ± 0.14	−33.9 ± 0.4	8.3 ± 0.1	−42.2 ± 0.4	−0.09
6	0.398 ± 0.043	0.637 ± 0.078	−35.6 ± 0.5	8.1 ± 0.1	−43.7 ± 0.5	−0.08

a: Ref [29]. b: not buffer corrected values determined in ITC_{TRIS} buffer. n.d.: not determined.

In line with the affinities, the buffer ionization-corrected thermodynamic binding profile of cpd **3** differs from cpds **4**–**6**, which are very similar. Interestingly, all phloroglucinol inhibitors are characterized by a positive ΔH compensated by a highly negative −TΔS. Cpd **3** has the least positive binding enthalpy but also a less favorable binding entropy than inhibitors **4**–**6** (ΔH°: 5.2 vs. 8.1 to 8.5 kJ·mol^{-1}; −TΔS°: −34.1 *vs.* −40.6 to −43.7 kJ·mol^{-1}). Most likely, this is caused by a methylamine residue placed in the S1 pocket (cpd **3**) instead of a benzamidine (**4**–**6**). According to the number of equal binding modes, cpd **3** should similarly benefit from a high number of identical binding modes like cpd **6** (Table 1). However, this effect is masked by the impact of placing a benzamidine instead of a methylamine residue in the S1 pocket. The observed differences could emanate from the larger hydration shell of benzamidines stripped upon binding to the narrow S1 subsite or from the displacement of a loosely bound water, as reported recently for the closely related S1 pocket of trypsin.[36,37] Both effects could explain why benzamidines are favored over benzylamines in S1, thereby leading to a beneficial binding entropy of cpds **4**–**6** compared to cpd **3**. Therefore, only inhibitors **4**–**6** can be used for direct comparisons to assess the impact of ligand symmetry on binding thermodynamics. For cpds **3** and **6**, exhibiting six identical binding modes (Table 1), the incremental affinities ($K_{d,inc}$) of all 6 binding modes $K_{d,1}$ to $K_{d,6}$ are equal. Hence, $K_{d,inc}$ of each binding mode can be calculated by multiplying the $K_{d,app}$ with six (Equation 2). This gives a $K_{d,inc}$ of 61.2 μM for each of the six binding modes for cpd **3** and 3.82 μM for those of cpd **6**, respectively (Table S5). With a factor of ~16 between both incremental affinities, the contributions of binding modes of cpds **4** and **5** with a methylamine binding into the S1 pocket can be neglected if more affine binding modes with benzamidines in the S1 pocket are possible. Assuming that S2 and S3/4 do not strongly discriminate benzamidines from benzylamines, the affinities of cpds **4** and **5** can hence be estimated from $K_{d,inc}$ of cpd **6** with the respective number of identical binding modes of cpds **4** (two identical binding modes) and **5** (four identical binding modes) with an benzamidine in S1. Consequently, the difference of binding free energy from **4** to **5** and from **5** to **6** for the exchange of methylamine to amidine substituents can also be calculated (Equation 4, Table S5). This results in $\Delta\Delta G_{(calc)\mathbf{4\rightarrow5}}$ of −1.7 kJ·mol^{-1} and $\Delta\Delta G_{(calc)\mathbf{5\rightarrow6}}$ of −1.0 kJ·mol^{-1}, while $\Delta\Delta G_{(ITC)\mathbf{4\rightarrow5}}$ of −1.8 kJ·mol^{-1} and $\Delta\Delta G_{(ITC)\mathbf{5\rightarrow6}}$ of −1.7 kJ·mol^{-1} were observed in ITC experiments. The prediction for the cpd **4** to **5** conversion is in excellent agreement with the experiment. However, the difference between calculation and experiment for the cpd **5** to **6** exchange might point out that the assumption of identical affinity for amidine and methylamine in S2 and S3/4, at least for one of the sites, is not ultimately correct.

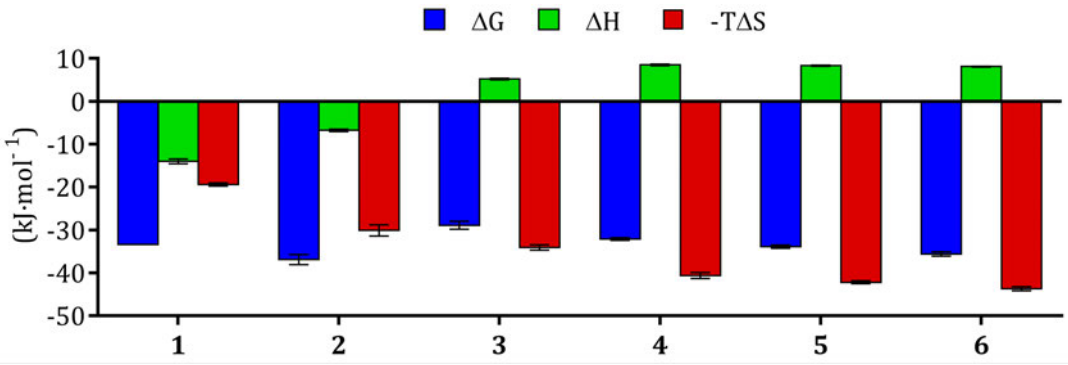


Figure 3. Signature plots of inhibitors **1**–**6** displaying the buffer ionization corrected enthalpic (green) and temperature-dependent entropic (red) contributions to the overall binding energy (blue). Gaussian progressed errors of all measurements in HEPES and TRIS buffers are indicated as black bars.

For cpds **4**–**6**, ΔH° stays relatively constant (8.0, 8.3, 8.1 kJ·mol^{-1}), indicating low selectivity between methylamine and amidine moieties of the S2 and the S3/4 binding sites arising from enthalpic interactions. The differences in ΔG° (–32.1, –33.9, –35.6 kJ·mol^{-1}) are mainly due to effects in –TΔS° (–40.6, –42.2, –43.7 kJ·mol^{-1}) for cpds **4**–**6**, respectively. This trend aligns with the higher number of possible solutions to generate energetically equal complexes of these cpds while placing a benzamidine moiety into the S1 pocket. Whereas cpd **4** only has two possibilities to adopt a conformation with benzamidine in the S1 pocket out of six binding modes (2/6), and cpd **5** has four out of six similar preferred binding modes (4/6 if S2 and S3/4 are less demanding). Cpd **6** has six similar binding modes (6/6), leading to the lowest entropic penalty (Tables 1, 2, Figure 3). Remarkably, the entropic benefits are similar to the calculated ΔΔG values of –1.7 kJ·mol^{-1} for the ligands **4** to **5** and –1.0 kJ·mol^{-1} for the ligands **5** to **6,** only taking the effects of the increased number of indistinguishable binding modes into account (Table S5). Although not pronounced, the relatively good accordance of the calculated ΔΔG for the statistic effect of multiple binding modes with the gain in binding entropy (Δ–TΔS, Table S5) supports the hypothesis of an entropic benefit of high ligand symmetry. The slight difference between the calculated $\Delta\Delta G_{(calc)\mathbf{5\rightarrow6}}$ (–1.0 kJ·mol^{-1}) and the measured $\Delta\Delta G_{(ITC)\mathbf{5\rightarrow6}}$ (–1.7 kJ·mol^{-1}) leaves space for diverse interpretations and indicates diminutive discrimination of amidine from methylamine residues in the S2 and S3/4 pockets.

However, it is to be noted that other superimposed effects arising from the methylamine-to-amidine exchange could also explain the thermodynamic observations. (i) Binding in the S2 and S3/4 sub-pockets leads to a higher partial desolvation of benzamidines than benzylamines, probably leading to a more beneficial binding entropy of ligands with more benzamidine moieties. (ii) The benzylamine residues contain one additional rotatable bond compared to benzamidines. If these get rigidified in the complex state, this could also lead to an entropic penalty. (iii) In case the S2 and S3/4 sub-pockets do not discriminate methylamine from amidine moieties, the observed subtle effects could indeed result from the higher number of identical binding modes for cpds **4**–**6** (cpd **4**: 2/6; cpd **5**: 4/6; cpd **6**: 6/6, Tables 1 and 2). (iv) Increased binding affinity of the ligands usually leads to the higher rigidification of the target protein. However, it is unlikely that these combined effects would lead to observed gains in binding entropy similar to the calculated values. It must be stated that the influence of ligand symmetry cannot be investigated without any superimposed effects caused by ligand modifications. Due to changes in molecular moieties and, subsequently, the thermodynamic binding profiles of the described cpds, these interpretations cannot be undisputedly dissected.

CONCLUSIONS

Protein-ligand binding thermodynamics are still not fully understood. The results highlight the complexity of energetics of molecular interactions, where multiple effects are superimposed. Using a series of new MT-SP1 ligands, the rarely investigated influence of higher ligand symmetry on thermodynamic binding profiles was elucidated.

Therefore, a new set of four phloroglucinol-based trivalent inhibitors harboring three amine functionalities (**3**) that are successively replaced by amidines (**4**–**6**) was designed. As earlier studies showed, introducing more benzamidines increased the inhibitory activity against MT-SP1.[19,24,30] With high symmetry, remarkably entropy-driven potent inhibitors of MT-SP1 were obtained.

Interestingly, all designed inhibitors exhibited an enthalpic penalty despite their ionic nature with three positive charges. Based on the equal number of identical binding modes, the binding profile of cpd **3** should similarly entropically benefit from symmetry like cpd **6**. However, the binding profile of cpd **3** differed from those of cpds **4**–**6** with a lower enthalpic penalty and less beneficial binding entropy. This observation can be attributed to the placement of a methylamine moiety in the S1 pocket for cpd **3** and a benzamidine residue for cpds **4**–**6**, as reported previously.[19,24,30] Our results indicate low selectivity of the S2 and S3/4 sub-pockets between benzylamine and benzamidine residues. Although not pronounced, we could observe the trend of a more beneficial binding entropy for ligands with a higher number of solutions for identically preferred binding modes to the overall number of possible binding modes, which is in good accordance with the estimated values from calculations.

Nevertheless, we are aware that various superimposed entropic effects could easily dominate our observations. Altogether, this work shows that introducing higher symmetry into ligands can be a promising strategy to yield highly entropic-driven MT-SP1 binders. With a linear correlation between K_a and the number of binding modes (Equation 2), this strategy improves affinity by symmetry. However, only moderate improvements can be yielded as it does not scale linearly with ΔG, which would be required to overcome activity cliffs.[8,38] Further studies are needed to elucidate whether this strategy can be generally applied using targets harboring multiple structurally similar sub-pockets to allow for symmetric ligand design, as is the case in MT-SP1. To our knowledge, this is the first study to employ ITC to quantify the entropic impact of higher symmetry in ligands.

Experimental section

Reagents.

All reagents and solvents were purchased from Sigma-Aldrich Chemie GmbH (Munich, Germany), Thermo Fisher Scientific Inc. (Waltham, Massachusetts; USA), and Carl Roth GmbH + Co. KG (Karlsruhe, Germany) if not stated elsewhere.

Synthesis.

Synthesis and chemical analytics of cpds **1** and **2** as control substances and for assay development were described previously.[29] Cpds **3**–**6** were obtained in a purity >95% by the following synthetic routes (Scheme 1).

General procedure for synthesis of cpd 9. Synthesis of 4,4',4''-(Benzene-1,3,5-triyltris(oxy))tribenzonitrile (cpd **9**) was performed modified according to Roy *et al.*[39] Phloroglucinol (cpd **7**, 2.10 g, 16.65 mmol, 1 equivalent (eq.), *p*-fluorobenzonitrile (cpd **8**, 5.73 g, 47.31 mmol, 2.84 mmol) and potassiumcarbonate (9.33 g, 67.51 mmol, 4.1 eq.) were dissolved in toluol abs. (20 mL) under argon atmosphere. *N*-methylpyrrolidinone (20 mL) was added to the reaction mixture, which was then heated in a Dean-Stark apparatus for 3 h at 150 °C. Toluene was removed under reduced pressure, and the remaining reaction mixture was heated to 200 °C for 20 h under an argon atmosphere. The cooled reaction mixture was poured into water, forming a brown precipitate. The pH of the mixture was adjusted to a pH of 3 with 1 M HCl. The precipitate was removed by filtration and dried in a vacuum. The crude product was purified by column chromatography with dichlormethane (DCM) to give cpd **9** as a colorless solid. (2.05 g, 4.77 mmol, yield: 30%). m_p = 165 °C, R_f = 0.6 (CH/EtOAc 2:1). ^{1}H NMR (300 MHz, Chloroform-*d*) δ = 7.68–7.62 (m, 6H), 7.11–7.05 (m, 6H), 6.57 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ = 159.94, 157.98, 134.52, 119.04, 118.45, 107.59, 107.30.

General procedure for synthesis of cpd 3. For the synthesis of ((Benzene-1,3,5-triyltris(oxy))tris(benzene-4,1-diyl))trimethanamine (cpd **3**) 4,4',4''-(Benzene-1,3,5-triyltris(oxy))tribenzonitrile (cpd **9**, 0.10 g, 0.23 mmol, 1 eq.) was dissolved in chloroform (1 mL). Platinum(IV)oxide (0.01 g, 10%) and methanol (10 mL) were added. Cpd **9** was reduced with H_2 at atmospheric pressure for 3 d. The catalyst was removed by filtration through Celite, and the residue was purified by HPLC. 0.03 g, 0.07 mmol,

yield: 30%. m_p: decomp. > 300 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 8.54 (s, 6H), 7.55 (d, *J* = 8.6 Hz, 6H), 7.12 (d, *J* = 8.6 Hz, 6H), 6.31 (s, 3H), 3.98 (s, 6H). ^{13}C NMR (75 MHz, DMSO) δ = 159.00, 155.51, 131.13, 130.04, 119.21, 103.03, 41.48. ^{1}H NMR (400 MHz, DMSO-d_6) δ = 8.54 (s, 6H), 7.55 (d, *J* = 8.5 Hz, 6H), 7.13 (d, *J* = 8.5 Hz, 6H), 6.31 (s, 3H), 3.98 (s, 6H). ^{13}C NMR (101 MHz, DMSO) δ = 159.03, 155.53, 131.17, 131.03, 130.07, 119.25, 103.03, 41.49. m/z for $[C_{27}H_{27}N_3O_3+H]^+$ calc. 442.2, found 442.1.

General procedure for synthesis of cpd 6. Synthesis of 4,4',4''-(benzene-1,3,5-triyltris(oxy))tribenzimidamide (cpd **6**) was performed modified according to Furtmann *et al.*[29] 4,4',4''-(Benzene-1,3,5-triyltris(oxy))tribenzonitrile (cpd **9**, 0.10 g, 0.23 mmol, 1 eq.) and hydroxylammoniumchloride (0.19 g, 2.76 mmol, 12 eq.) were dissolved in ethanol (20 mL). DIPEA (0.45 mL, 2.76 mmol, 12 eq.) was added dropwise, and the reaction mixture was heated to reflux for 3 h. The solvent was removed under reduced pressure, and the residue was dissolved in glacial acid (5 mL) under ice cooling. Acetic anhydride (0.18 mL, 1.84 mmol, 8 eq.) and palladium on carbon (0.01 g) were added. The reaction mixture was hydrated at 4 bar for 3 h. The catalyst was removed by filtration through Celite, and the residue was purified by HPLC. 0.03 g, 0.06 mmol, yield: 27%. m_p = 250 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 9.28 (d, *J* = 11.1 Hz, 9H), 7.91–7.84 (m, 6H), 7.35–7.29 (m, 6H), 6.73 (s, 3H). ^{13}C NMR (75 MHz, DMSO) δ = 164.78, 160.47, 157.64, 130.63, 123.20, 118.32, 107.31, 40.35, 40.08, 39.80, 39.52, 39.24, 38.96, 38.69. ^{1}H NMR (400 MHz, DMSO-d_6) δ = 9.29 (d, *J* = 17.9 Hz, 9H), 7.87 (d, *J* = 8.9 Hz, 6H), 7.33 (d, *J* = 8.9 Hz, 6H), 6.74 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ = 164.80, 160.51, 157.65, 130.66, 123.22, 118.33, 107.38. m/z for $[C_{27}H_{24}N_6O_3+H]^+$ calc. 481.2, found 481.2.

General procedure for synthesis of cpds 10 and 11. Synthesis of (*E*)-4-(3,5-bis(4-cyanophenoxy)phenoxy)-*N*'-hydroxybenzimidamide (cpd **10**) and (*Z*)-4-(3-(4-cyanophenoxy)-5-(4-((*E*)-*N*'-hydroxycarbamimidoyl)phenoxy)phenoxy)-*N*'-hydroxybenzimidamide (cpd **11**) was performed modified according to Furtmann *et al.*[29] 4,4',4''-(Benzene-1,3,5-triyltris(oxy))tribenzonitrile (cpd **9**, 0.40 g, 0.92 mmol, 1 eq.) and hydroxylammoniumchloride (0.15 g, 1.84 mmol, 2 eq.) were dissolved in ethanol (20 mL). DIPEA (0.32 mL, 1.84 mmol, 2 eq.) was added dropwise, and the reaction mixture was heated to reflux for 3 h. The solvent was removed under reduced pressure, and the residue was purified by HPLC. Cpd **10** was obtained next to cpd **11**.

Analytical data of cpd **10**

0.16 g, 0.35 mmol, yield: 38%. m_p = 90 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 11.15 (s, 1H), 8.71 (s, 2H), 7.88–7.83 (m, 4H), 7.78–7.71 (m, 2H), 7.30–7.23 (m, 6H), 6.77 (t, *J* = 2.1 Hz, 1H), 6.73 (d, *J* = 2.1 Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ = 160.10, 157.99, 157.26, 134.79, 130.02, 118.86, 118.69, 118.65, 107.65, 106.09, 40.36.

Analytical data of cpd **11**

0.11 g, 0.22 mmol, yield: 24%. m_p = 55 °C. ^{1}H NMR (300 MHz, DMSO-d_6) δ = 11.25 (s, 2H), 8.63 (s, 4H), 7.91–7.83 (m, 2H), 7.80–7.73 (m, 4H), 7.34–7.23 (m, 6H), 6.73–6.64 (m, 3H). ^{13}C NMR (75 MHz, DMSO) δ = 160.08, 159.53, 157.95, 157.14, 134.76, 129.94, 118.77, 118.60, 107.17, 106.04, 40.35, 40.08, 39.80, 39.52, 39.24, 38.96, 38.68.

General procedure for synthesis of cpd 4. Synthesis of 4-(3,5-bis(4-(aminomethyl)phenoxy)phenoxy)benzimidamide (cpd **4**) was performed modified following Huang *et al.*[40] (*E*)-4-(3,5-bis(4-cyanophenoxy)phenoxy)-*N*'-hydroxybenzimidamide (cpd **10**, 0.1 g, 0.22 mmol) was dissolved in methanol and platinum(IV) oxide (0.01 g) was added. The reaction mixture was reduced with hydrogen at atmospheric pressure for 2 d. The catalyst was separated by filtration over Celite. Excess solvent was removed, and the residue was dissolved in glacial acetic acid. Palladium was added on activated carbon (0.01 g) and reduced at 4 bar with hydrogen. The catalyst was removed by filtration over Celite, and the residue was purified by several HPLC runs. 0.02 g, 0.04 mmol, yield: 20%. m_p = 255 °C. ^{1}H NMR (600 MHz, DMSO-d_6) δ = 9.38 (s, 1H), 9.15 (s, 2H), 8.46 (s, 4H), 7.90 (d, *J* = 8.9 Hz, 2H), 7.59–7.52 (m, 4H), 7.26 (d, *J* = 8.9 Hz, 2H), 7.19–7.13 (m, 4H), 6.45 (d, *J* = 2.2 Hz, 2H), 6.42 (t, *J* = 2.2 Hz, 1H), 3.99 (s, 4H). ^{13}C NMR (151 MHz, DMSO) δ = 164.66, 160.63, 159.27, 159.07, 157.19, 155.52, 131.23, 130.76, 130.71, 130.20, 122.81, 119.37, 118.12, 117.93, 104.52, 104.38, 41.54. m/z for $[C_{27}H_{26}N_4O_3+H]^+$ calc. 455.2, found 455.2.

General procedure for synthesis of cpd 5. Synthesis of 4,4'-((5-(4-(aminomethyl)phenoxy)-1,3-phenylene)bis(oxy))dibenzimidamide (cpd **5**) was performed modified following Huang *et al.*[40]

(Z)-4-(3-(4-cyanophenoxy)-5-(4-((E)-*N*'-hydroxycarbamimidoyl)phenoxy)phenoxy)-*N*'-hydroxybenzimidamide (cpd **11**, 0.1 g, 0.20 mmol) was dissolved in methanol and platinum(IV) oxide (0.01 g) was added. The reaction mixture was reduced with hydrogen at atmospheric pressure for 2 d. The catalyst was separated by filtration over Celite, the excess solvent was removed, and the residue was dissolved in glacial acetic acid. Palladium was added on activated carbon (0.01 g) and reduced at 4 bar with hydrogen. The catalyst was removed by filtration over Celite, and the residue was purified by several HPLC runs. 0.04 g, 0.06 mmol, yield: 28%. m_p = 85 °C. ^{1}H NMR (600 MHz, DMSO-d_6) δ = 9.37 (s, 2H), 9.28 (s, 4H), 8.27 (s, 2H), 7.87–7.85 (m, 4H), 7.54–7.47 (m, 2H), 7.31–7.25 (m, 4H), 7.23–7.18 (m, 2H), 6.64–6.57 (m, 3H), 4.02 (s, 2H). ^{13}C NMR (151 MHz, DMSO) δ = 164.86, 164.83, 160.58, 159.45, 158.83, 158.62, 157.47, 155.56, 131.13, 130.68, 130.65, 130.26, 123.18, 119.52, 118.27, 118.16, 118.14, 106.11, 105.92, 41.67. m/z for $[C_{27}H_{25}N_5O_3+H]^+$ calc. 468.2, found 468.2.

Molecular Docking studies.

A crystal structure of MT-SP1 in complex with an inhibitor containing two benzamidine residues is freely available in the PDB[41] (PDB ID 4O9V).[24] Molecular docking was performed with LeadIT (LeadIT version 2.3.2; BioSolveIT GmbH, Sankt Augustin, Germany, 2017, www.biosolveit.de/LeadIT).[42] For binding site definition, the crystallographic reference ligand **15** (*N*-(4-aminocyclohexyl)-3,5-bis(4-carbamimidoylphenoxy)benzamide)[24] was used. Residues within 6.5 Å around the ligand and one water molecule forming at least two polar interactions (water-1004) were included. The final binding site consisted of His57, Cys58, Tr59, Ile60, Asp60A, Asp60B, Tyr60G, Asp96, Phe97, Thr98, Phe99, Tyr146, Gln174, Gln175, Met180, Val183, Asp189, Ser190, Cys191, Gln192, Gly193, Asp194, Ser195, Val213, Ser214, Trp215, Gly216, Asp217, Gly219, Cys220, Ala221, Lys224, Pro225, Gly226, Val227 and Tyr228. Prior to docking, ligands were protonated and energy minimized using the Merck Molecular Force Field (MMFF94x)[43,44] within the Molecular Operating Environment (Molecular Operating Environment (MOE); 2019.0102; Chemical Computing Group ULC: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2019. https://www.chemcomp.com/index.htm).[45] The docking was performed using LeadIT under default parameters with ligand binding driven by entropy and enthalpy (hybrid approach). A redocking with the reference ligand cpd 15 was performed to validate the receptor setup. The top pose has a score of –53.9 kJ·mol^{-1} and a root mean square deviation (RMSD)-Value of 1.65 Å (Figure S6A). Binder *vs.* non-binder discrimination was validated with ligands taken from the ChEMBL database[46] with a pChEMBL value of 6 or higher (corresponding to a concentration at half-maximal inhibition (IC_{50}), K_i, K_d, etc. of 1 µM or lower). 100 diverse ligands were selected using the RDKit (RDKit 2020_03_4 (Q1 2020): Open-source cheminformatics. https://www.rdkit.org[47]) diversity picker within KNIME (Version 4.2.1, KNIME AG, Talacker 50, 8001 Zurich, Switzerland, https://www.knime.com/)[48] and 1170 decoys with matched physicochemical properties were generated using the Rapid Decoy Retriever (RADER).[49] The resulting receiver operating characteristic (ROC)-curve has an area under the curve (AUC)-value of 0.91 (Figure S6B). Figures were created using PyMOL.[28]

Recombinant protein expression, refolding, and purification.

The recombinant human matriptase was mainly expressed as described previously.[14,17] Briefly, the pQE-30 vector containing the zymogen of the catalytic domain of MT-SP1 (pro-MT-SP1; uniport: Q9Y5Y6 aa 596–855) harboring a hexahistidine (His_6) tag at the N-terminus of the pro-peptide, was transformed into competent *Escherichia coli* (*E. coli*) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). 10 L main cultures were grown in LB medium containing 100 µg/mL ampicillin at 37 °C and 160 rpm until they reached an optical density (OD_{600}) of ~0.8. Overexpression was induced by adding 1 mM isopropyl-*β*-D-thiogalactopyranosid (IPTG) for ~16 h at 20 °C. Cells were harvested by centrifugation (9000 rpm at 4 °C for 15 min). Protein purification was performed on an ÄKTA start protein purification system (GE Healthcare, Chicago, IL, USA) fast protein liquid chromatography (FPLC). Lysis was performed with cell pellets, resuspended in 250 mL lysis buffer A (50 mM TRIS·HCl pH 8.0, 10% (v/v) glycerol, 300 mM NaCl, 0.1% (v/v) $Triton_{X-100}$, RNase, DNase, lysozyme, 1 mM dithiothreitol (DTT)) by thorough sonication (Sonoplus HD 2200; Bandelin, Berlin, Germany). Cell debris and inclusions bodies were collected by centrifugation (20000 rpm at 4 °C for 1 h), and the supernatant was discarded. The inclusion bodies were then solubilized in 500 mL denaturing buffer B (50 mM

TRIS·HCl pH 8.0, 5% (v/v) glycerol, 6 M urea, 20 mM imidazole) by stirring overnight at 4 °C. The insoluble impurities were removed by another centrifugation step (20000 rpm at 4 °C for 1 h), and the supernatant was subjected to immobilized metal affinity chromatography (IMAC) on a HisTrap HP 5 ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany) containing a nickel nitrilotriacetic acid (Ni-NTA) Sepharose resin. After washing with 5 column volumes (CV) of wash buffer C (50 mM TRIS·HCl pH 8.0, 6 M urea, 20 mM imidazole), pro-MT-SP1 was eluted with a linear gradient of elution buffer D (50 mM TRIS·HCl pH 8.0, 6 M urea, 200 mM imidazole). The pooled fractions were combined and refolded by two-step dialysis against 2 L of dialysis buffer E (50 mM TRIS·HCl pH 9.0, 1 mM β-ME, 3 M urea) and 2 L of dialysis buffer F (50 mM TRIS·HCl pH 9.0, 1 mM β-ME) at 4 °C for > 8 h each, using a ZelluTrans dialysis tube (Carl Roth, Karlsruhe, Germany) with a molecular weight cut-off (MWCO) of 12–14 kDa. Finally, the MT-SP1 was subjected to anion exchange chromatography (AEX) on a HiTrap 5 mL column (Cytiva Europe GmbH). After washing with 5 CV of buffer F, a linear gradient of buffer F and AEX elution buffer G (50 mM TRIS·HCl pH 9.0, 1 mM β-ME, 1 M NaCl) was applied to elute MT-SP1 (iso-electric points (PI) calculated with ExPASy ProtParam[50]: Zymogen: 6.21; MT-SP1: 5.78; His_6-pro-peptide: 7.94). During dialysis, the 3.5 kDa His_6-pro-peptide is cleaved *via* trans-activation (Figure S7).[51–53] Eluted fractions were combined and concentrated to ~50 µM using Amicon® Ultra-15 Centrifugal Filters, Ultracel, 10 kDa MWCO, (Millipore, Billerica, MA, USA). After that, MT-SP1 was aliquoted, flash-frozen in liquid nitrogen, and stored at –80 °C until further use. The yield was approximately 0.4 mg of active MT-SP1 per 1 L of *E. coli* culture.

Enzyme inhibition assays.

Inhibitory activity against human matriptase was determined using a fluorescence-based enzyme activity assay. The increasing fluorescence over time was recorded in white flat-bottom 96-well microtiter plates (Greiner bio-one, Kremsmünster, Austria) on a TECAN Spark 10M (Agilent Technologies). Each well contained a total volume of 200 µL composed of 180 µL reaction buffer (50 mM TRIS·HCl pH 8.0, 150 mM NaCl, 5 mM $CaCl_2$, and 0.01% (v/v) $Triton_{X\text{-}100}$), 5 µL of MT-SP1 in reaction buffer (final concentration 2.5 nM), and 10 µL of one concentration of a half logarithmic or 1:1 dilution series of the inhibitors in DMSO. The reaction was initiated without incubation by adding 5 µL of the Boc-Leu-Arg-Arg-AMC substrate (Bachem AG, Bubendorf, Switzerland) in DMSO to a final concentration of 85 µM (or 100 µM for pH 8.0 measurement of **3** and **6**) and thoroughly mixed. Fluorescence of technical triplicates was recorded in 30 s intervals for 10 min at 25 °C (λ_{ex} 380 nm; λ_{em} 460 nm; slit 20 nm). IC_{50} values were calculated by fitting to the four-parameter equation using GRAFIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, West Sussex, UK; Figure S2).[54] K_i values for inhibitors were calculated from IC_{50}- and Michaelis-Menten constants (K_M) using the Cheng-Prusoff equation.[55]w

pH-dependency.

Investigations of the influences of the inhibitor's protonation states were conducted like the enzyme inhibition studies but with reaction buffers adjusted to pH 7.0–9.0 in 0.5 increments.

K_M determination.

The Michaelis-Menten constants K_M of the Boc-Leu-Arg-Arg-AMC substrate to MT-SP1 was determined at different pH values using pure DMSO instead of inhibitor solutions and with 8 different final substrate concentrations (800, 400, 200, 100, 50, 25, 12.5, 6.25 µM, Figure S4). The K_M was calculated using GRAFIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, West Sussex, UK).[54] K_M values were obtained by fitting to the Michaelis-Menten equation (Equation 5)

$$v = (V_{max}\cdot[S])/(K_M+[S]) \quad (5)$$

with v (ΔF/min) as the substrate hydrolysis rate V_{max} as the maximum slope of the dose-response curve, and the substrate (S) concentration.

Isothermal Titration Calorimetry.

ITC experiments were performed in technical triplicates on a MicroCal PEAQ-ITC automated system (Malvern Instruments, Worcestershire, UK). Experiments were conducted in ITC_{TRIS} buffer (50 mM TRIS·HCl pH 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 1 mM β-ME) and ITC_{HEPES} buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 5 mM $CaCl_2$, 1 mM β-ME). 200–500 µM solutions of the inhibitors (except for **3**, of which 2000 µM were used to perform a low-C titration) were titrated in 18 injection à 2 µL (after one pre-injection of 0.4 µL) with a duration of 4 s and a spacing time of 150 s to 15–50 µM of MT-SP1 in the reaction cell (see Table S6 for exact concentrations). Control experiments were performed for each titration set and subtracted (buffer *vs*. titrant, titrant *vs*. buffer) from or summed

(buffer *vs*. buffer) to the raw data to correct for unspecific dilution events. All experiments were performed at 25 °C with a stirring speed of 750 rpm and a reference power of 41.9 µW. Data integration and evaluation was performed using the MicroCal PEAQ-ITC analysis software (Version 1.21, Malvern Panalytical Ltd, Worcestershire, UK). Correction of proton transfers upon ligand binding was performed using *ITCcalc* (https://itccalc.uni-mainz.de/).[56] Figures were created using GraphPad Prism 7.0.4.[57]

Author Contributions

SJH: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. HM: Formal analysis, Investigation, Methodology, Writing – review & editing. ACW: Formal analysis, Investigation, Software, Validation, Writing – review & editing. MG: Conceptualization, Writing – review & editing. CK: Conceptualization, Project administration, Software, Supervision, Writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

[redacted] for sharing the pQE-30 vector containing human matriptase and [redacted] for sharing his expertise in MT-SP1 preparation. We thank [redacted] for access to facilities and instrumentation and highly appreciate the scientific discussions on the project. We thank [redacted] for support in the lab and [redacted] for assistance in establishing synthetic routes. The German Research Foundation (DFG-Großgeräteantrag INST 247/921-1 FUGG) funding of the ITC instrument is gratefully acknowledged.

References

(1) Pantsar, T.; Poso, A. Binding Affinity via Docking: Fact and Fiction. *Molecules* **2018**, *23* (8), 1899. https://doi.org/10.3390/molecules23081899.

(2) Baum, B.; Muley, L.; Smolinski, M.; Heine, A.; Hangauer, D.; Klebe, G. Non-Additivity of Functional Group Contributions in Protein–Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry. *J. Mol. Biol.* **2010**, *397* (4), 1042–1054. https://doi.org/10.1016/j.jmb.2010.02.007.

(3) Pierce, M. M.; Raman, C. S.; Nall, B. T. Isothermal Titration Calorimetry of Protein–Protein Interactions. *Methods* **1999**, *19* (2), 213–221. https://doi.org/10.1006/meth.1999.0852.

(4) Biela, A.; Nasief, N. N.; Betz, M.; Heine, A.; Hangauer, D.; Klebe, G. Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin. *Angew. Chemie Int. Ed.* **2013**, *52* (6), 1822–1828. https://doi.org/10.1002/anie.201208561.

(5) Geschwindner, S.; Ulander, J.; Johansson, P. Ligand Binding Thermodynamics in Drug Discovery: Still a Hot Tip? *J. Med. Chem.* **2015**, *58* (16), 6321–6335. https://doi.org/10.1021/jm501511f.

(6) Watanabe, M.; Nakamura-Nakayama, M.; Fujihara, M.; Kawasaki, M.; Nakano, S.; Kakuta, H. Increased Molecular Flexibility Widens the Gap between K i and K d Values in Screening for Retinoid X Receptor Modulators. *ACS Med. Chem. Lett.* **2022**, *13* (2), 211–217. https://doi.org/10.1021/acsmedchemlett.1c00575.

(7) Verteramo, M. L.; Stenström, O.; Ignjatović, M. M.; Caldararu, O.; Olsson, M. A.; Manzoni, F.; Leffler, H.; Oksanen, E.; Logan, D. T.; Nilsson, U. J.; Ryde, U.; Akke, M. Interplay between Conformational Entropy and Solvation Entropy in Protein–Ligand Binding. *J. Am. Chem. Soc.* **2019**, *141* (5), 2012–2026. https://doi.org/10.1021/jacs.8b11099.

(8) Feinberg, H.; Castelli, R.; Drickamer, K.; Seeberger, P. H.; Weis, W. I. Multiple Modes of Binding Enhance the Affinity of DC-SIGN for High Mannose N-Linked Glycans Found on Viral Glycoproteins. *J. Biol. Chem.* **2007**, *282* (6), 4202–4209. https://doi.org/10.1074/jbc.M609689200.

(9) Lauria, A.; Terenzi, A.; Bartolotta, R.; Bonsignore, R.; Perricone, U.; Tutone, M.; Martorana, A.; Barone, G.; Almerico, A. Does Ligand Symmetry Play a Role in the Stabilization of DNA G-Quadruplex Host-Guest Complexes? *Curr. Med. Chem.* **2014**, *21* (23), 2665–2690. https://doi.org/10.2174/0929867321666140217155156.

(10) Szabo, R.; Bugge, T. H. Membrane-Anchored Serine Proteases as Regulators of Epithelial Function. *Biochem. Soc. Trans.* **2020**, *48* (2), 517–528. https://doi.org/10.1042/BST20190675.

(11) Lin, C.-Y.; Anders, J.; Johnson, M.; Sang, Q. A.; Dickson, R. B. Molecular Cloning of CDNA for Matriptase, a Matrix-Degrading Serine Protease with Trypsin-like Activity. *J. Biol. Chem.* **1999**, *274* (26), 18231–18236. https://doi.org/10.1074/jbc.274.26.18231.

(12) Ding, K.-F. Effect of SNC19/ST14 Gene Overexpression on Invasion of Colorectal Cancer Cells. *World J. Gastroenterol.* **2005**, *11* (36), 5651. https://doi.org/10.3748/wjg.v11.i36.5651.

(13) List, K. Matriptase: A Culprit in Cancer? *Futur. Oncol.* **2009**, *5* (1), 97–104. https://doi.org/10.2217/14796694.5.1.97.

(14) Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M. E.; Bode, W.; Stürzebecher, J. Secondary Amides of Sulfonylated 3-Amidinophenylalanine.

New Potent and Selective Inhibitors of Matriptase. *J. Med. Chem.* **2006**, *49* (14), 4116–4126. https://doi.org/10.1021/jm051272l.

(15) Hammami, M.; Rühmann, E.; Maurer, E.; Heine, A.; Gütschow, M.; Klebe, G.; Steinmetzer, T. New 3-Amidinophenylalanine-Derived Inhibitors of Matriptase. *Medchemcomm* **2012**, *3* (7), 807. https://doi.org/10.1039/c2md20074k.

(16) List, K.; Haudenschild, C. C.; Szabo, R.; Chen, W.; Wahl, S. M.; Swaim, W.; Engelholm, L. H.; Behrendt, N.; Bugge, T. H. Matriptase/MT-SP1 Is Required for Postnatal Survival, Epidermal Barrier Function, Hair Follicle Development, and Thymic Homeostasis. *Oncogene* **2002**, *21* (23), 3765–3779. https://doi.org/10.1038/sj.onc.1205502.

(17) Wilkinson, D. J.; Habgood, A.; Lamb, H. K.; Thompson, P.; Hawkins, A. R.; Désilets, A.; Leduc, R.; Steinmetzer, T.; Hammami, M.; Lee, M. S.; Craik, C. S.; Watson, S.; Lin, H.; Milner, J. M.; Rowan, A. D. Matriptase Induction of Metalloproteinase-Dependent Aggrecanolysis In Vitro and In Vivo: Promotion of Osteoarthritic Cartilage Damage by Multiple Mechanisms. *Arthritis Rheumatol.* **2017**, *69* (8), 1601–1611. https://doi.org/10.1002/art.40133.

(18) Sisay, M. T.; Steinmetzer, T.; Stirnberg, M.; Maurer, E.; Hammami, M.; Bajorath, J.; Gütschow, M. Identification of the First Low-Molecular-Weight Inhibitors of Matriptase-2. *J. Med. Chem.* **2010**, *53* (15), 5523–5535. https://doi.org/10.1021/jm100183e.

(19) Friedrich, R.; Fuentes-Prior, P.; Ong, E.; Coombs, G.; Hunter, M.; Oehler, R.; Pierson, D.; Gonzalez, R.; Huber, R.; Bode, W.; Madison, E. L. Catalytic Domain Structures of MT-SP1/Matriptase, a Matrix-Degrading Transmembrane Serine Proteinase. *J. Biol. Chem.* **2002**, *277* (3), 2160–2168. https://doi.org/10.1074/jbc.M109830200.

(20) Béliveau, F.; Tarkar, A.; Dion, S. P.; Désilets, A.; Ghinet, M. G.; Boudreault, P.-L.; St-Georges, C.; Marsault, É.; Paone, D.; Collins, J.; Macphee, C. H.; Campobasso, N.; Groy, A.; Cottom, J.; Ouellette, M.; Pope, A. J.; Leduc, R. Discovery and Development of TMPRSS6 Inhibitors Modulating Hepcidin Levels in Human Hepatocytes. *Cell Chem. Biol.* **2019**, *26* (11), 1559-1572.e9. https://doi.org/10.1016/j.chembiol.2019.09.004.

(21) Pilgram, O.; Keils, A.; Benary, G. E.; Müller, J.; Merkl, S.; Ngaha, S.; Huber, S.; Chevillard, F.; Harbig, A.; Magdolen, V.; Heine, A.; Böttcher-Friebertshäuser, E.; Steinmetzer, T. Improving the Selectivity of 3-Amidinophenylalanine-Derived Matriptase Inhibitors. *Eur. J. Med. Chem.* **2022**, *238*, 114437. https://doi.org/10.1016/j.ejmech.2022.114437.

(22) Goswami, R.; Mukherjee, S.; Wohlfahrt, G.; Ghadiyaram, C.; Nagaraj, J.; Chandra, B. R.; Sistla, R. K.; Satyam, L. K.; Samiulla, D. S.; Moilanen, A.; Subramanya, H. S.; Ramachandra, M. Discovery of Pyridyl Bis(Oxy)Dibenzimidamide Derivatives as Selective Matriptase Inhibitors. *ACS Med. Chem. Lett.* **2013**, *4* (12), 1152–1157. https://doi.org/10.1021/ml400213v.

(23) Beckmann, A.-M.; Gilberg, E.; Gattner, S.; Huang, T. L.; Vanden Eynde, J. J.; Mayence, A.; Bajorath, J.; Stirnberg, M.; Gütschow, M. Evaluation of Bisbenzamidines as Inhibitors for Matriptase-2. *Bioorg. Med. Chem. Lett.* **2016**, *26* (15), 3741–3745. https://doi.org/10.1016/j.bmcl.2016.05.071.

(24) Goswami, R.; Mukherjee, S.; Ghadiyaram, C.; Wohlfahrt, G.; Sistla, R. K.; Nagaraj, J.; Satyam, L. K.; Subbarao, K.; Palakurthy, R. K.; Gopinath, S.; Krishnamurthy, N. R.; Ikonen, T.; Moilanen, A.; Subramanya, H. S.; Kallio, P.; Ramachandra, M. Structure-Guided Discovery of 1,3,5 Tri-Substituted Benzenes as Potent and Selective Matriptase Inhibitors Exhibiting in Vivo Antitumor Efficacy. *Bioorg. Med. Chem.* **2014**, *22* (12), 3187–3203. https://doi.org/10.1016/j.bmc.2014.04.013.

(25) Enyedy, I. J.; Lee, S.; Kuo, A. H.; Dickson, R. B.; Lin, C.; Wang, S. Structure-Based Approach for the Discovery of Bis-Benzamidines as Novel Inhibitors of Matriptase. *J. Med. Chem.* **2001**, *44* (9), 1349–1355. https://doi.org/10.1021/jm000395x.

(26) Colombo, É.; Désilets, A.; Duchêne, D.; Chagnon, F.; Najmanovich, R.; Leduc, R.; Marsault, E. Design and Synthesis of Potent, Selective Inhibitors of Matriptase. *ACS Med. Chem. Lett.* **2012**, *3* (7), 530–534. https://doi.org/10.1021/ml3000534.

(27) Harper, E.; Berger, A. On the Size of the Active Site in Proteases: Pronase. *Biochem. Biophys. Res. Commun.* **1972**, *46* (5), 1956–1960. https://doi.org/10.1016/0006-291X(72)90076-9.

(28) The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography*. Schrödinger 2002, pp 82–92. http://www.pymol.org.

(29) Furtmann, N.; Häußler, D.; Scheidt, T.; Stirnberg, M.; Steinmetzer, T.; Bajorath, J.; Gütschow, M. Limiting the Number of Potential Binding Modes by Introducing Symmetry into Ligands: Structure-Based Design of Inhibitors for Trypsin-Like Serine Proteases. *Chem. - A Eur. J.* **2016**, *22* (2), 610–625. https://doi.org/10.1002/chem.201503534.

(30) Goswami, R.; Wohlfahrt, G.; Mukherjee, S.; Ghadiyaram, C.; Nagaraj, J.; Satyam, L. K.; Subbarao, K.; Gopinath, S.; Krishnamurthy, N. R.; Subramanya, H. S.; Ramachandra, M. Discovery of O-(3-Carbamimidoylphenyl)-l-Serine Amides as Matriptase Inhibitors Using a Fragment-Linking Approach. *Bioorg. Med. Chem. Lett.* **2015**, *25* (3), 616–620. https://doi.org/10.1016/j.bmcl.2014.12.008.

(31) Shen, Y.; Liu, L.; Zheng, Q.; Zhao, X.; Han, Y.; Guo, Q.; Wang, Y. Quantitative Insights into Tightly and Loosely Bound Water in Hydration Shells of Amino Acids. *Soft Matter* **2021**, *17* (44), 10080–10089. https://doi.org/10.1039/D1SM01234G.

(32) Goldberg, R. N.; Kishore, N.; Lennen, R. M. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data* **2002**, *31* (2), 231–370. https://doi.org/10.1063/1.1416902.

(33) Yang, D.; Singh, A.; Wu, H.; Kroe-Barrett, R. Comparison of Biosensor Platforms in the Evaluation of High Affinity Antibody-Antigen Binding Kinetics. *Anal. Biochem.* **2016**, *508*, 78–96. https://doi.org/10.1016/j.ab.2016.06.024.

(34) Velours, C.; Aumont-Nicaise, M.; Uebel, S.; England, P.; Velazquez-Campoy, A.; Stroebel, D.; Bec, G.; Soule, P.; Quétard, C.; Ebel, C.; Roussel, A.; Charbonnier, J.-B.; Varela, P. F. Macromolecular Interactions in Vitro, Comparing Classical and Novel Approaches. *Eur. Biophys. J.* **2021**, *50* (3–4), 313–330. https://doi.org/10.1007/s00249-021-01517-5.

(35) Dunitz, J. D. Win Some, Lose Some: Enthalpy-Entropy Compensation in Weak Intermolecular Interactions. *Chem. Biol.* **1995**, *2* (11), 709–712. https://doi.org/10.1016/1074-5521(95)90097-7.

(36) Schiebel, J.; Gaspari, R.; Wulsdorf, T.; Ngo, K.; Sohn, C.; Schrader, T. E.; Cavalli, A.; Ostermann, A.; Heine, A.; Klebe, G. Intriguing Role of Water in Protein-Ligand Binding Studied by Neutron Crystallography on Trypsin Complexes. *Nat. Commun.* **2018**, *9* (1), 3559. https://doi.org/10.1038/s41467-018-05769-2.

(37) Ansari, N.; Rizzi, V.; Parrinello, M. Water Regulates the Residence Time of Benzamidine in Trypsin. *Nat. Commun.* **2022**, *13* (1), 5438. https://doi.org/10.1038/s41467-022-33104-3.

(38) Stumpfe, D.; Hu, H.; Bajorath, J. Evolving Concept of Activity Cliffs. *ACS Omega* **2019**, *4* (11), 14360–14368. https://doi.org/10.1021/acsomega.9b02221.

(39) Roy, S.; Biradha, K. Coordination Polymers of Silver(I) with the Flexible Tritopic Ligand 1,3,5-Tri(4-Cyanophenoxy)Benzene: Guest Inclusion and Luminescent Properties. *Aust. J. Chem.* **2013**, *66* (4), 436. https://doi.org/10.1071/CH12476.

(40) Huang, Y.; Yu, M.; Sun, N.; Tang, T.; Yu, F.; Song, X.; Xie, Q.; Fu, W.; Shao, L.; Wang, Y. Discovery of Carbazole Carboxamides as Novel RORγt Inverse Agonists. *Eur. J. Med. Chem.* **2018**, *148*, 465–476. https://doi.org/10.1016/j.ejmech.2018.02.050.

(41) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(42) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261* (3), 470–489. https://doi.org/10.1006/jmbi.1996.0477.

(43) Halgren, T. a. Merck Molecular Force Field. II. MMFF94 van Der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* **1996**, *17* (5–6), 520–552. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<520::AID-JCC2>3.0.CO;2-W.

(44) Halgren, T. A. 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.

(45) Vinet, L.; Zhedanov, A. A "Missing" Family of Classical Orthogonal Polynomials. *J. Chem. Inf. Model.* **2010**, *53* (9), 1689–1699. https://doi.org/10.1088/1751-8113/44/8/085201.

(46) Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. ChEMBL: Towards Direct Deposition of Bioassay Data. *Nucleic Acids Res.* **2019**, *47* (D1), D930–D940. https://doi.org/10.1093/nar/gky1075.

(47) Landrum, G.; Tosco, P.; Kelley, B.; sriniker; gedeck; Vianello, R.; NadineSchneider; Ric; Dalke, A.; N, D.; Cole, B.; Kawashima, E.; Turk, S.; Swain, M.; AlexanderSavelyev; Vaucher, A.; Cosgrove, D.; Wójcikowski, M.; Probst, D.; Godin, G.; Pahl, A.; Berenger, F.; JLVarjo; strets123; JP; DoliathGavid; Sforna, G.; Jensen, J. H.; Fuller, P.; O'Boyle, N.; guillaume godin; Pahl, A.; Berenger, F.; JLVarjo; strets123; JP; DoliathGavid; Sforna, G.; Jensen, J. H.; Fuller, P.; O'Boyle, N. Rdkit/Rdkit: 2020_03_4 (Q1 2020) Release. Zenodo July 3, 2020. https://doi.org/10.5281/ZENODO.3929204.

(48) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Studies in Classification, Data Analysis, and Knowledge Organization*; Springer, 2008; pp 319–326. https://doi.org/10.1007/978-3-540-78246-9_38.

(49) Wang, L.; Pang, X.; Li, Y.; Zhang, Z.; Tan, W. RADER: A RApid DEcoy Retriever to Facilitate Decoy Based Assessment of Virtual Screening. *Bioinformatics* **2017**, *33* (8), 1235–1237. https://doi.org/10.1093/bioinformatics/btw783.

(50) Godfrey, W. L.; Rudd, C. J.; Iyer, S.; Recktenwald, D. Purification of Cellular and Organelle Populations by Fluorescence-Activated Cell Sorting for Proteome Analysis. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, 2005; pp 67–78. https://doi.org/10.1385/1-59259-890-0:067.

(51) Oberst, M. D.; Williams, C. A.; Dickson, R. B.; Johnson, M. D.; Lin, C.-Y. The Activation of Matriptase Requires Its Noncatalytic Domains, Serine Protease Domain, and Its Cognate Inhibitor. *J. Biol. Chem.* **2003**, *278* (29), 26773–26779. https://doi.org/10.1074/jbc.M304282200.

(52) Kojima, K.; Inouye, K. Activation of Matriptase Zymogen. *J. Biochem.* **2011**, *150* (2), 123–125. https://doi.org/10.1093/jb/mvr075.

(53) Miyake, Y.; Yasumoto, M.; Tsuzuki, S.; Fushiki, T.; Inouye, K. Activation of a Membrane-Bound Serine Protease Matriptase on the Cell Surface. *J. Biochem.* **2009**, *146* (2), 273–282. https://doi.org/10.1093/jb/mvp066.

(54) Robin J. Leatherbarrow. GraFit 6. Erithacus Software Limited: East Grinstead, West Sussex, U.K. 2007. http://erithacus.com/grafit/.

(55) Yung-Chi, C.; Prusoff, W. H. Relationship between the Inhibition Constant (KI) and the Concentration of Inhibitor Which Causes 50 per Cent Inhibition (I50) of an Enzymatic Reaction. *Biochem. Pharmacol.* **1973**, *22* (23), 3099–3108. https://doi.org/10.1016/0006-2952(73)90196-2.

(56) Hammerschmidt, S. J.; Barthels, F.; Kersten, C. Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc (Unpublished/Submitted Manuscript). **2023**.

(57) GraphPad Prism. GraphPad Software: San Diego, California USA 2014. www.graphpad.com.

Improving binding entropy by ligand symmetry – a case study with human matriptase.

Stefan J. Hammerschmidt[a,†], [redacted]

[a] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany*
[b] *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany*
† Authors contributed equally.
* E-Mail: [redacted]

Content

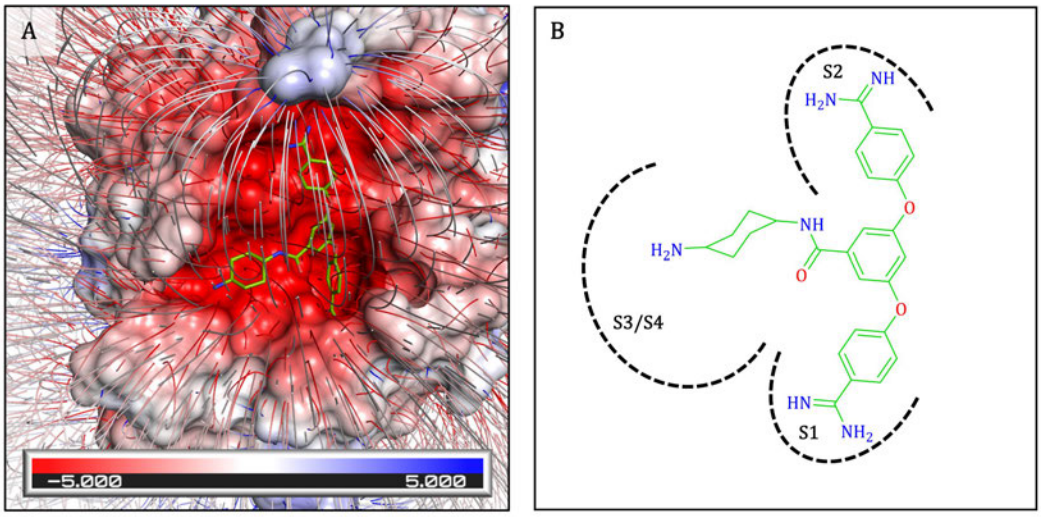


Figure S1. Electrostatic surface of the MT-SP1 binding site and 2D structure of cpd **15**. **A** The electrostatic surface and electric field lines (red to blue from −5.000 to 5.000 kBT/ec) of the protease are exemplarily shown for the complex structure (PDB ID: 4O9V)[1] with inhibitor **15** (*N*-(4-aminocyclohexyl)-3,5-bis(4-carbamimidoylphenoxy)benzamide) depicted as sticks with green carbon atoms. Calculated with the Adaptive Poisson-Boltzmann Solver (APBS)[2] plugin within PyMOL.[3] **B** 2D structure of inhibitor **15** with indicated binding sites.

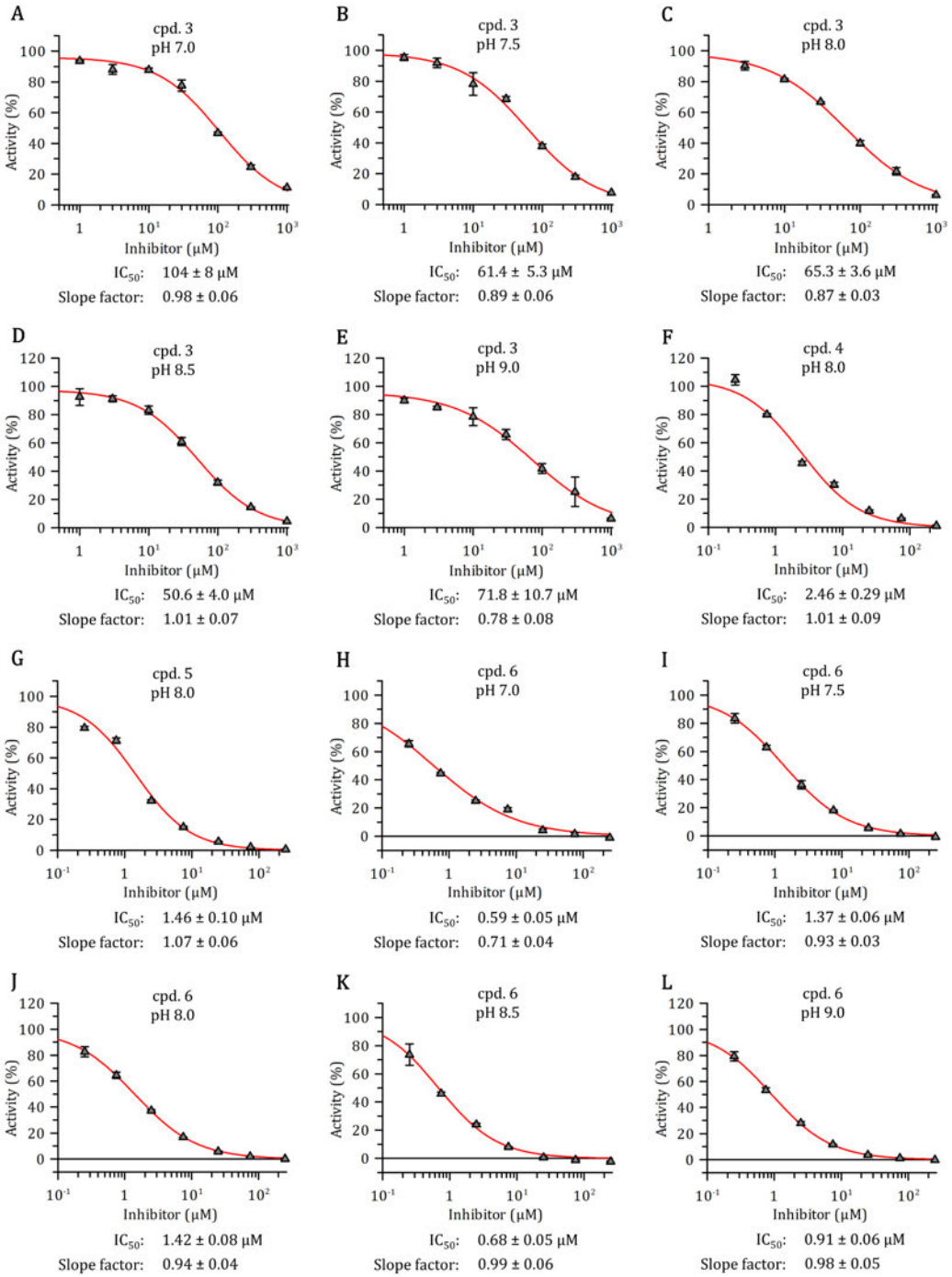


Figure S2. IC_{50} fits of enzyme inhibition assay data. **A–E** Cpd **3** at pH 7.0–9.0. **F**, **G** Cpds **4** and **5** at pH 8.0. **H–L** Cpd **6** at pH 7.0–9.0. Errors of technical triplicates are indicated as bars. Figure generated with GRAPHIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, UK).[4]

Protonation states. To evaluate if the protonation states of the moieties determine the different affinities of amines and amidines, we calculated the pka values of ligands 1–6 (Figure S3A). Through the symmetry of the ligands, pka values for the same functional groups can be exchanged and should be understood as those of successive protonation events. The pka values range from 8.78 to 9.74 for amides and 11.19 to 12.15 for the protonated states of the amidines. Hence, at the experimental pH of 8.0, the great majority should be fully protonated, as even for cpd 3 with the weakest basicity, more than 80% is fully protonated (Figure S3B–F).

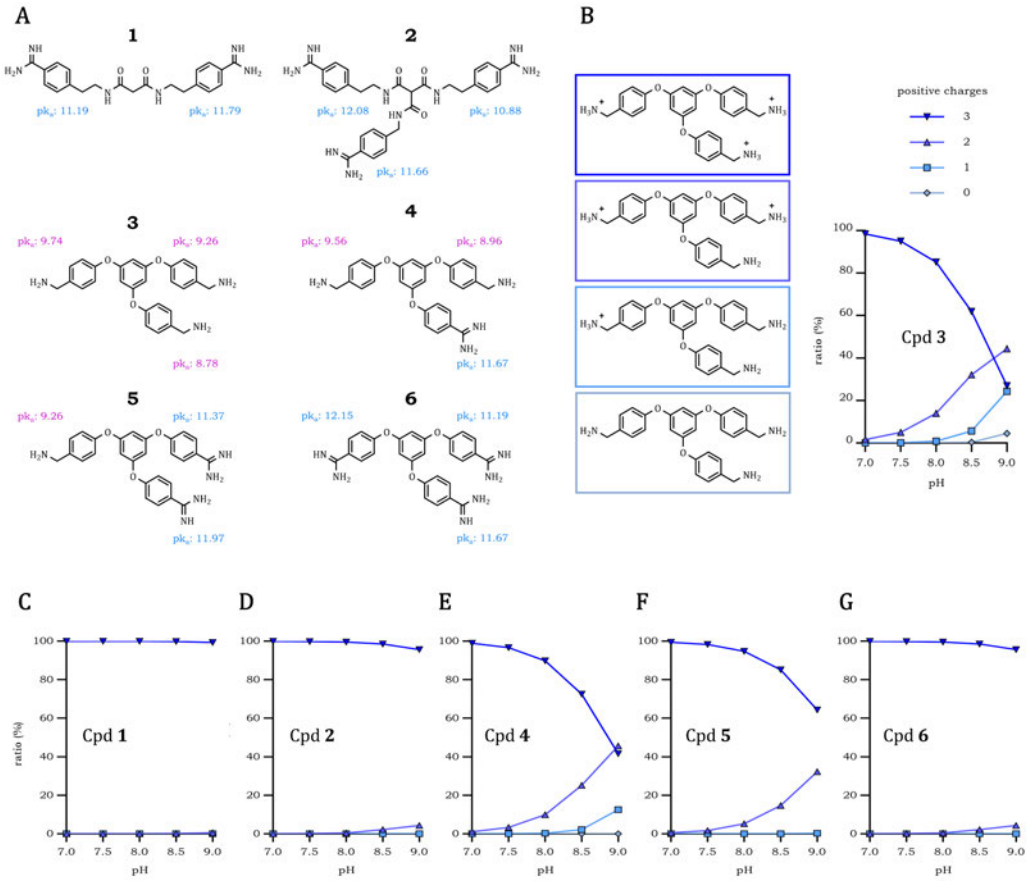


Figure S3. Evaluation of pk_a values amines and amidines used in this study. Calculated with MarvinSketch.[5] **A** pk_a of amins (purple) and amidines (blue) of inhibitors **1**–**6**. **B** Protonation states of cpd **3** and their occurrences at relevant pH values of 7.0–9.0, predicted by MarvinSketch. From dark blue to pale blue: fully protonated species to not protonated state. **C**–**G** Protonation states of cpds 1, 2, 4, 5, and 6, respectively, with only two possible positions for protonation being present in inhibitor **1** and inhibitors **2**–**6** featuring three basic sites for protonation.

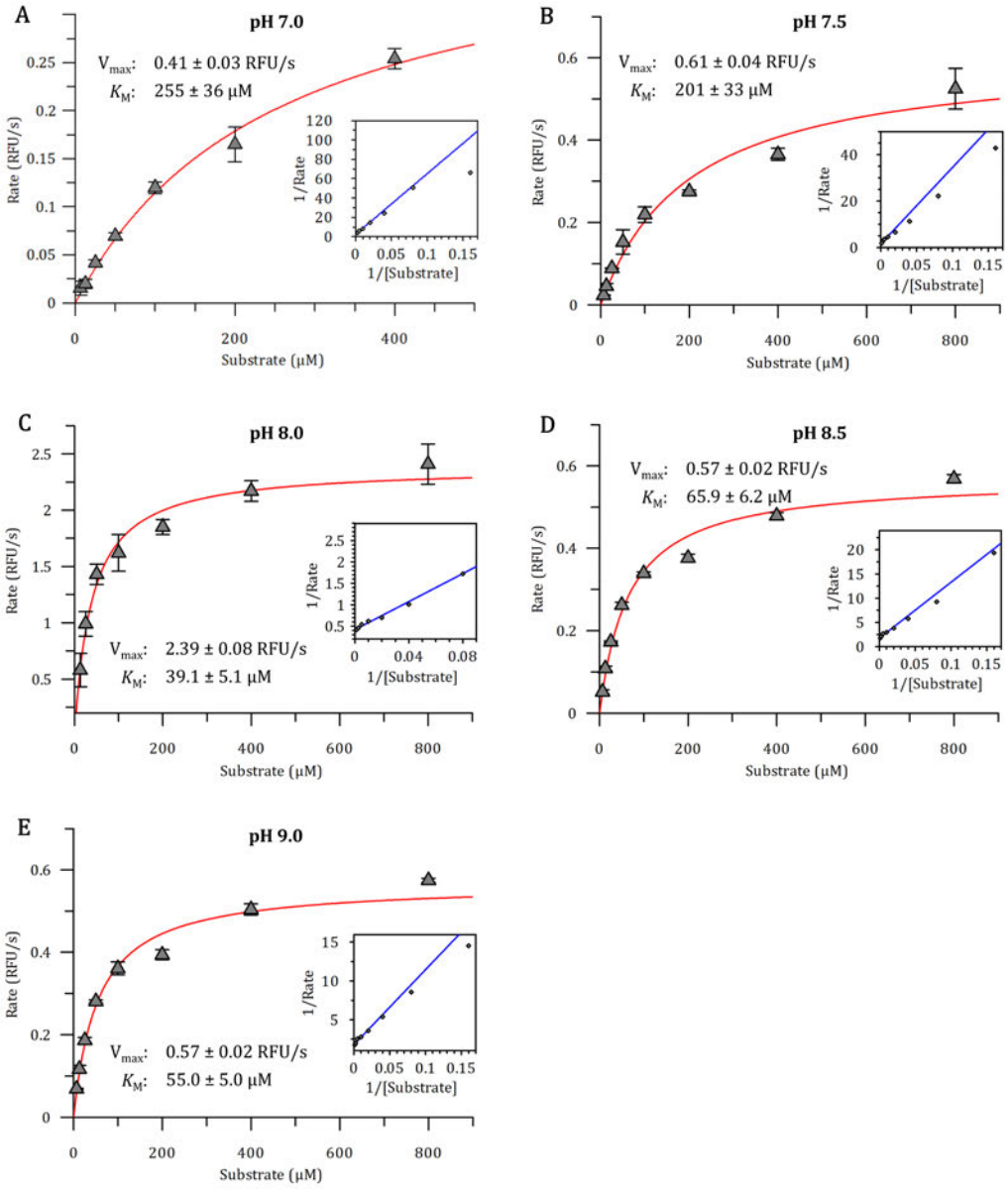


Figure S4. K_M-determinations of the MT-SP1 cleavage of Boc-LRR-AMC at different pH-Values. **A** pH 7.0; **B** pH 7.5; **C** pH 8.0; **D** pH 8.5; **E** pH 9.0. In the smaller secondary graph, the respective linearized Lineweaver-Burk plot is given. Figure generated with GRAPHIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, UK).[4]

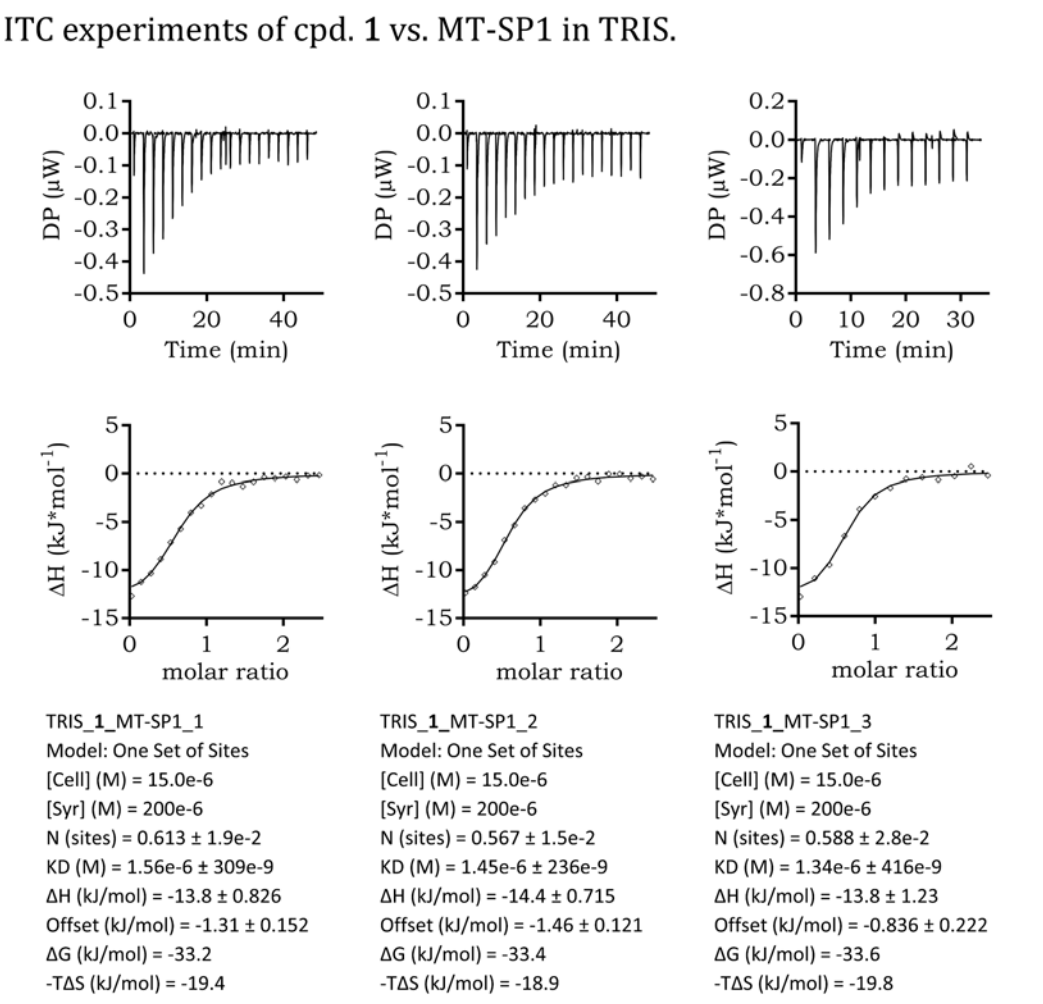


Figure S5A. Thermograms and isotherms of cpd **1** vs. MT-SP1 in ITC_{TRIS} buffer.

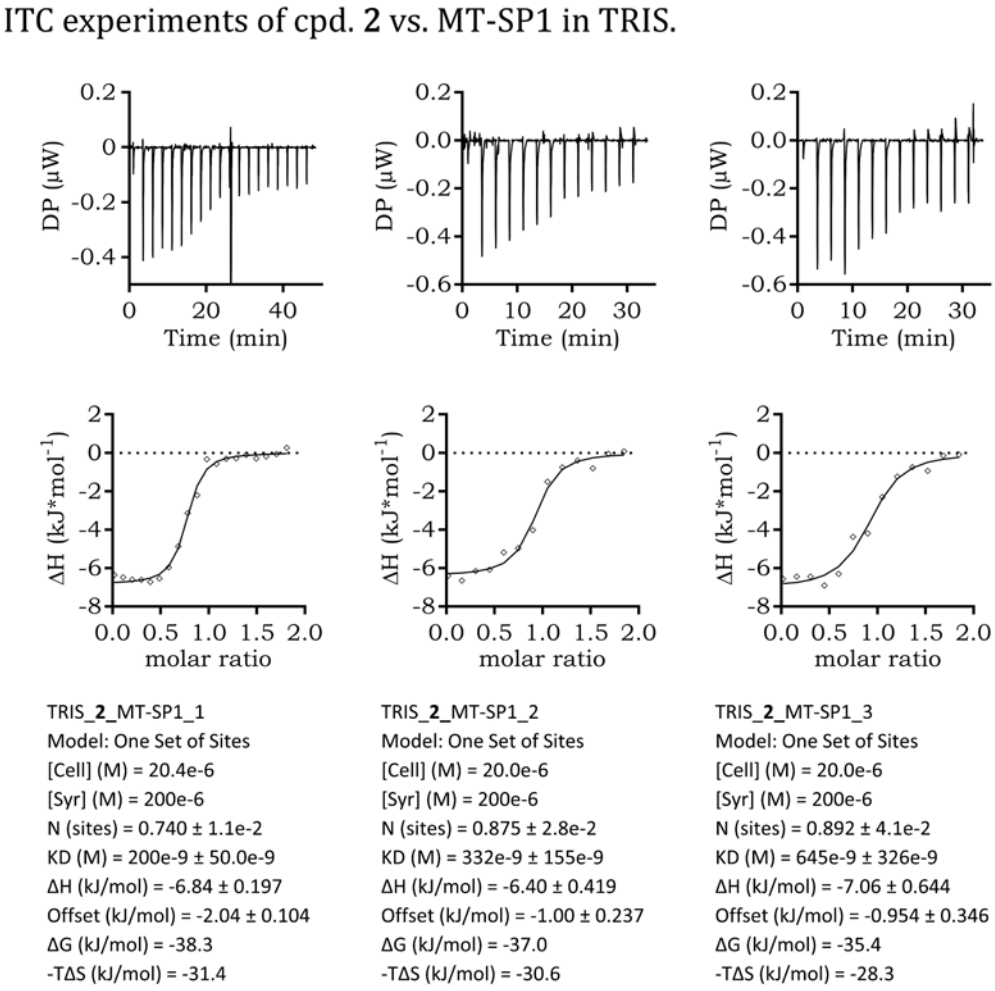


Figure S5B. Thermograms and isotherms of cpd **2** vs. MT-SP1 in ITC_{TRIS} buffer.

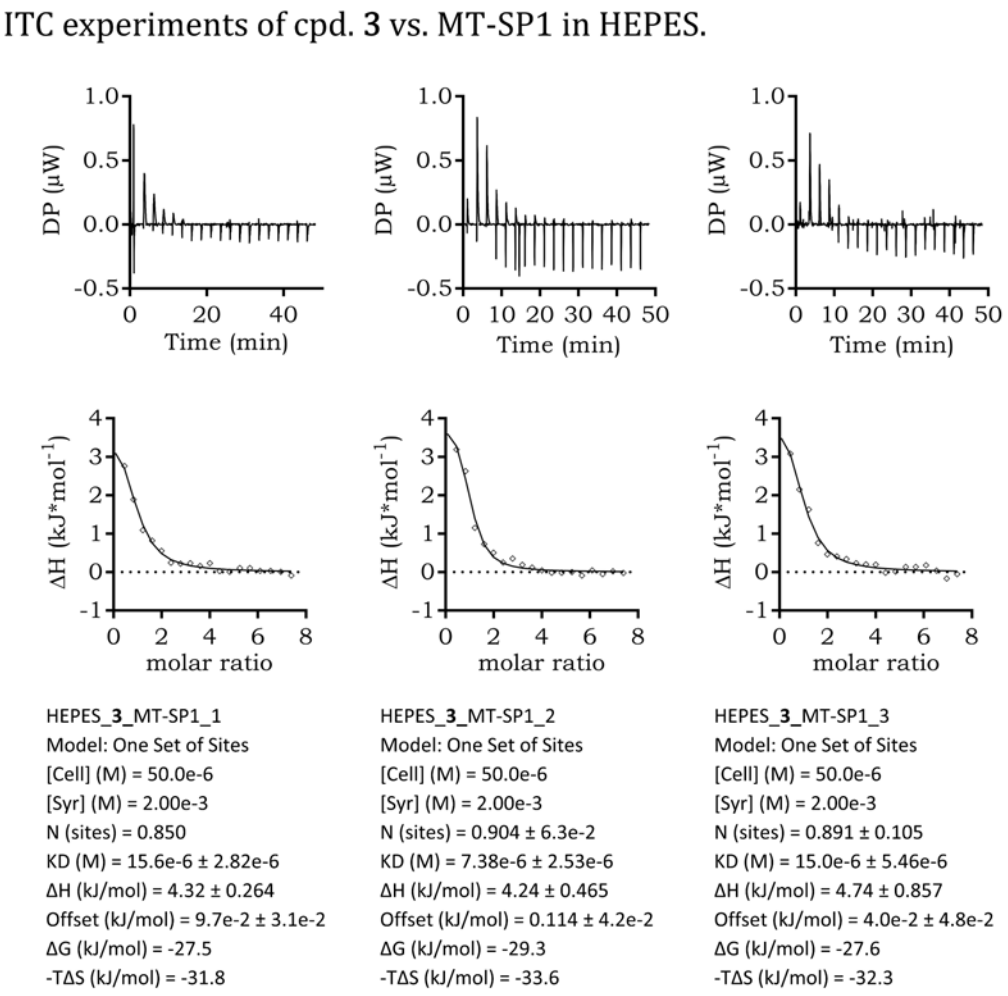


Figure S5C. Thermograms and isotherms of cpd **3** vs. MT-SP1 in ITC_{HEPES} buffer.

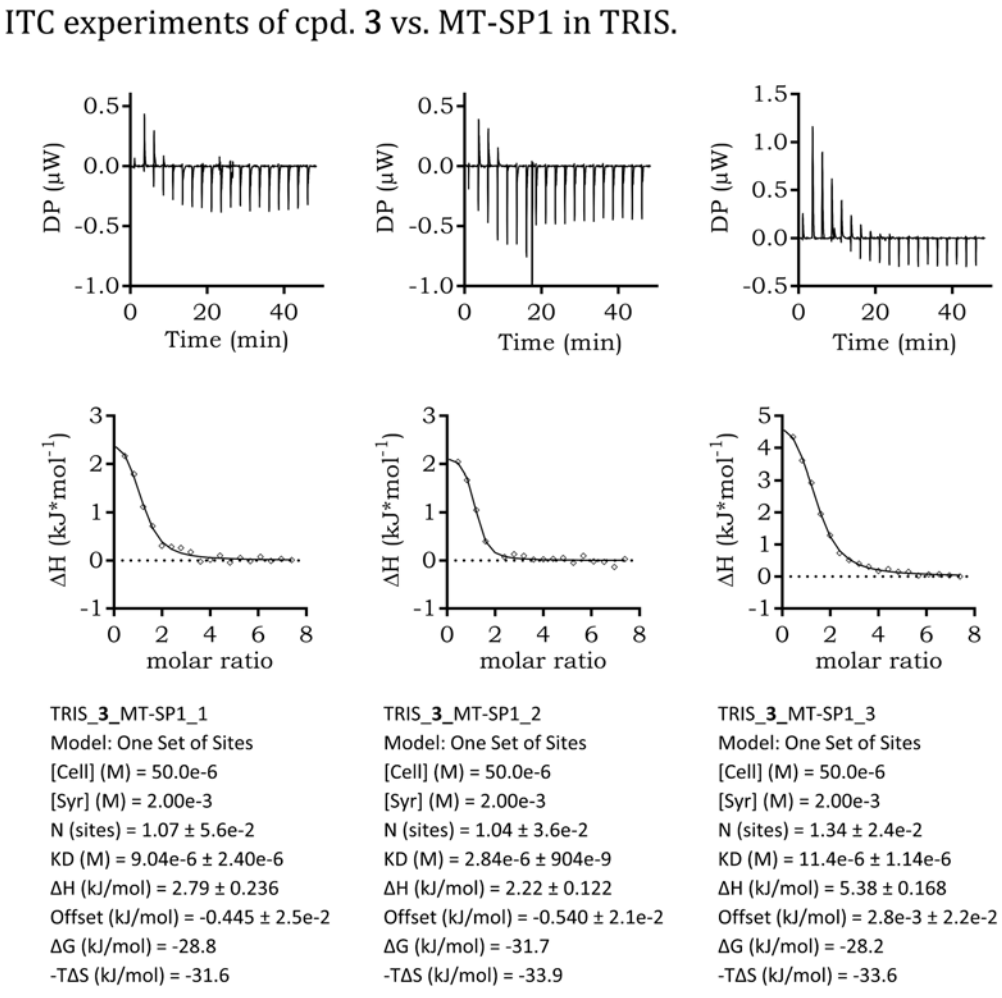


Figure S5D. Thermograms and isotherms of cpd **3** vs. MT-SP1 in ITC_{TRIS} buffer.

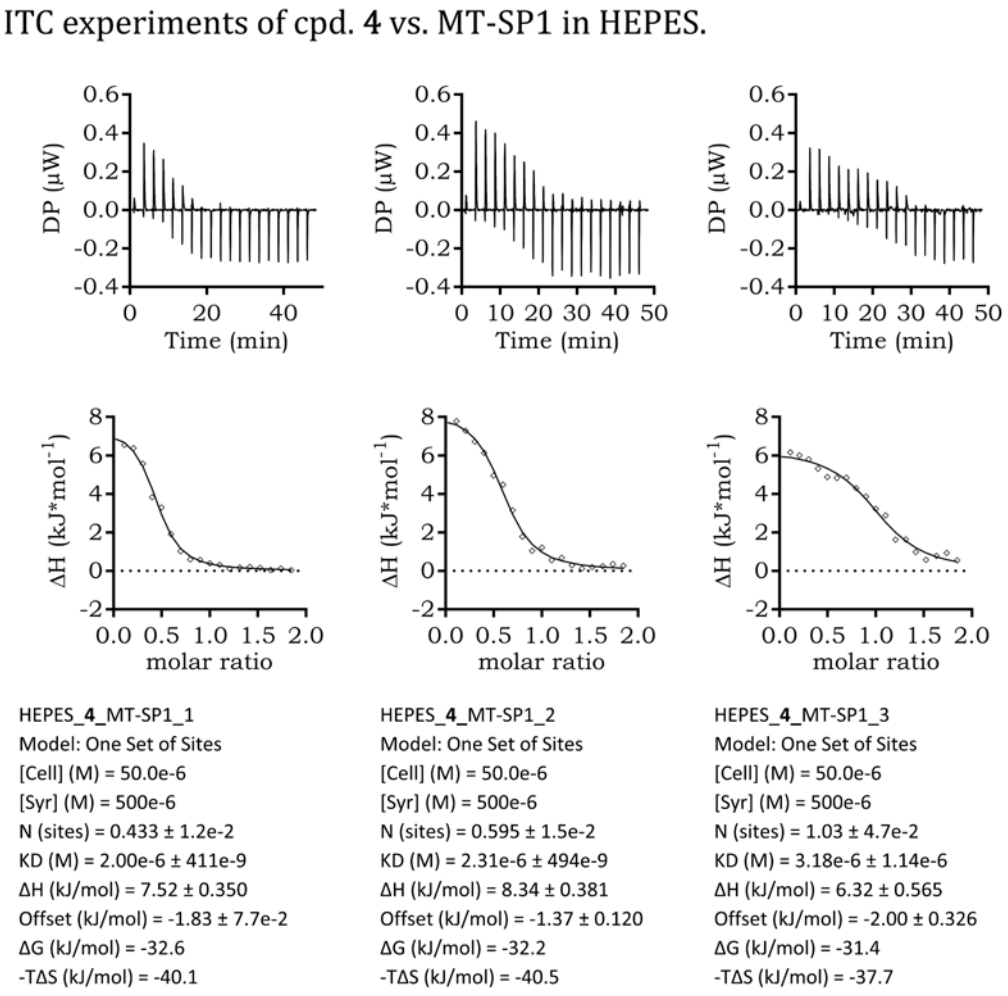


Figure S5E. Thermograms and isotherms of cpd **4** vs. MT-SP1 in ITC_{HEPES} buffer.

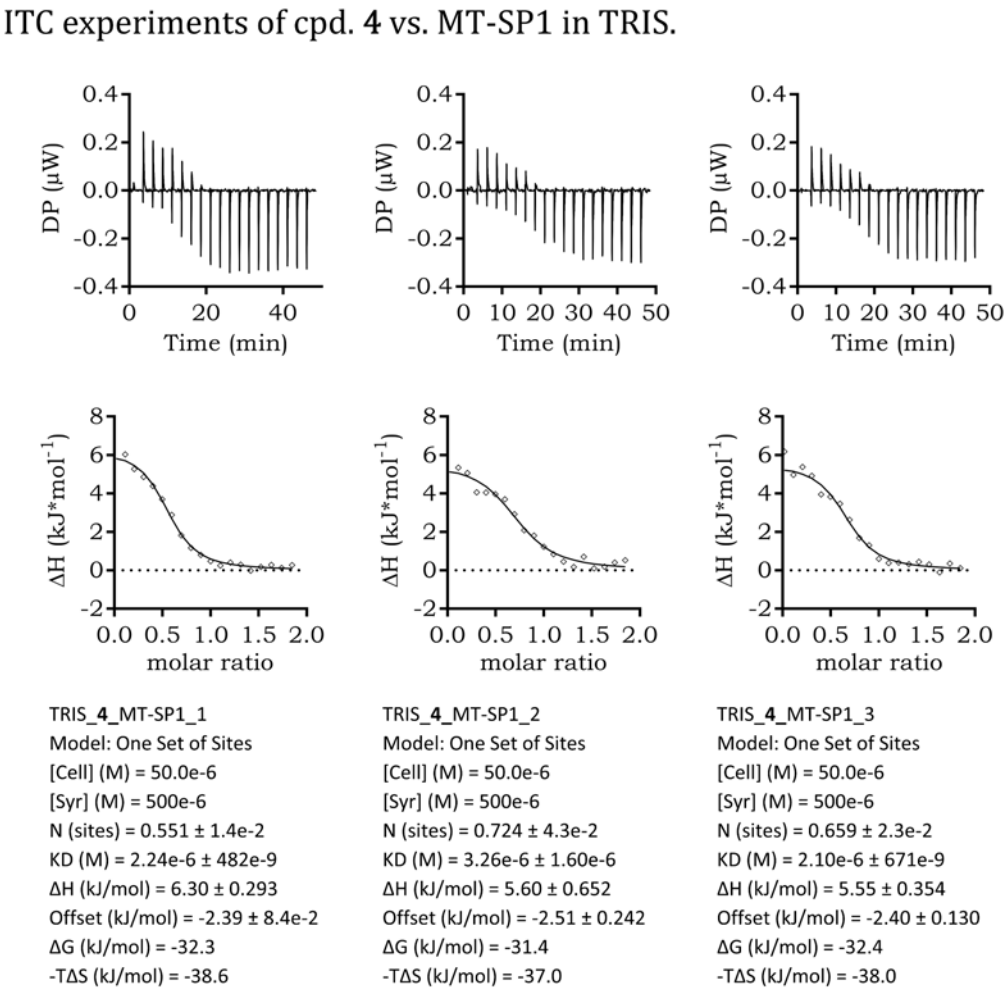


Figure S5F. Thermograms and isotherms of cpd **4** vs. MT-SP1 in ITC_{TRIS} buffer.

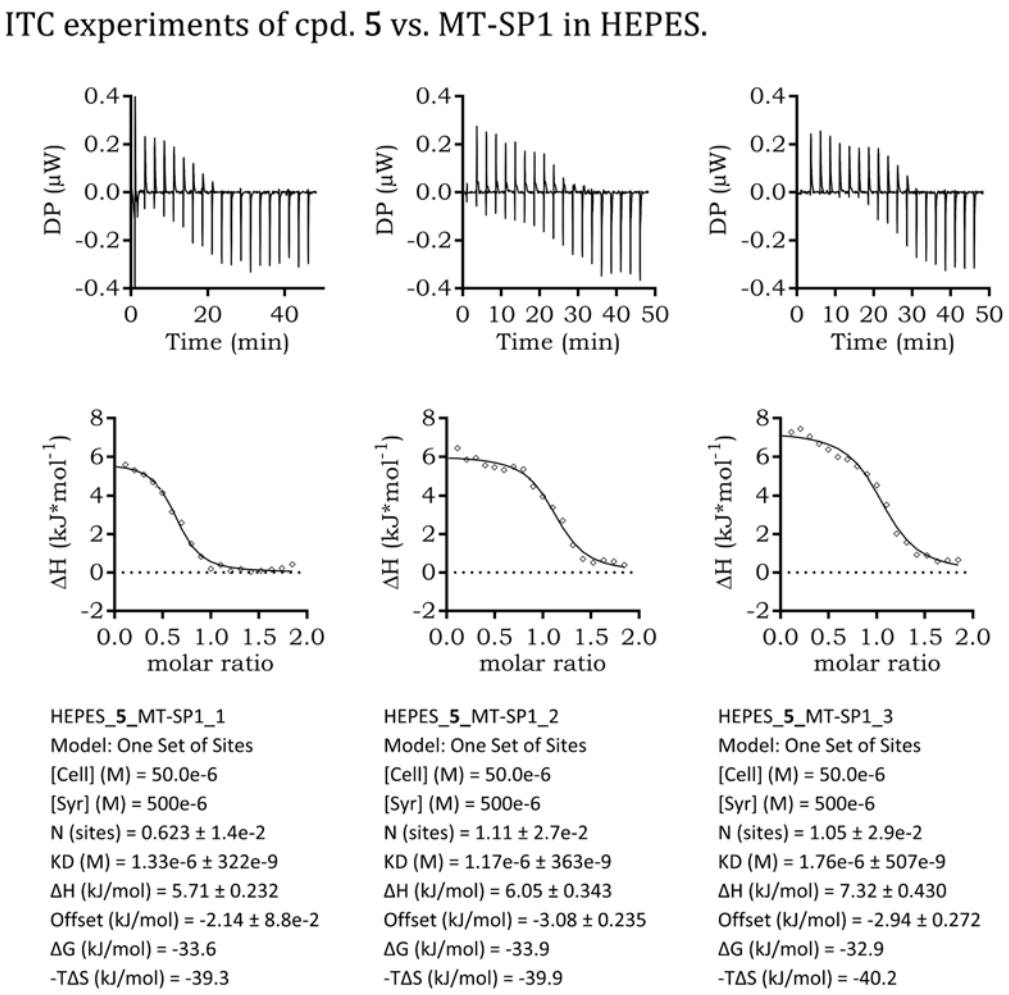


Figure S5G. Thermograms and isotherms of cpd **5** vs. MT-SP1 in ITC_{HEPES} buffer.

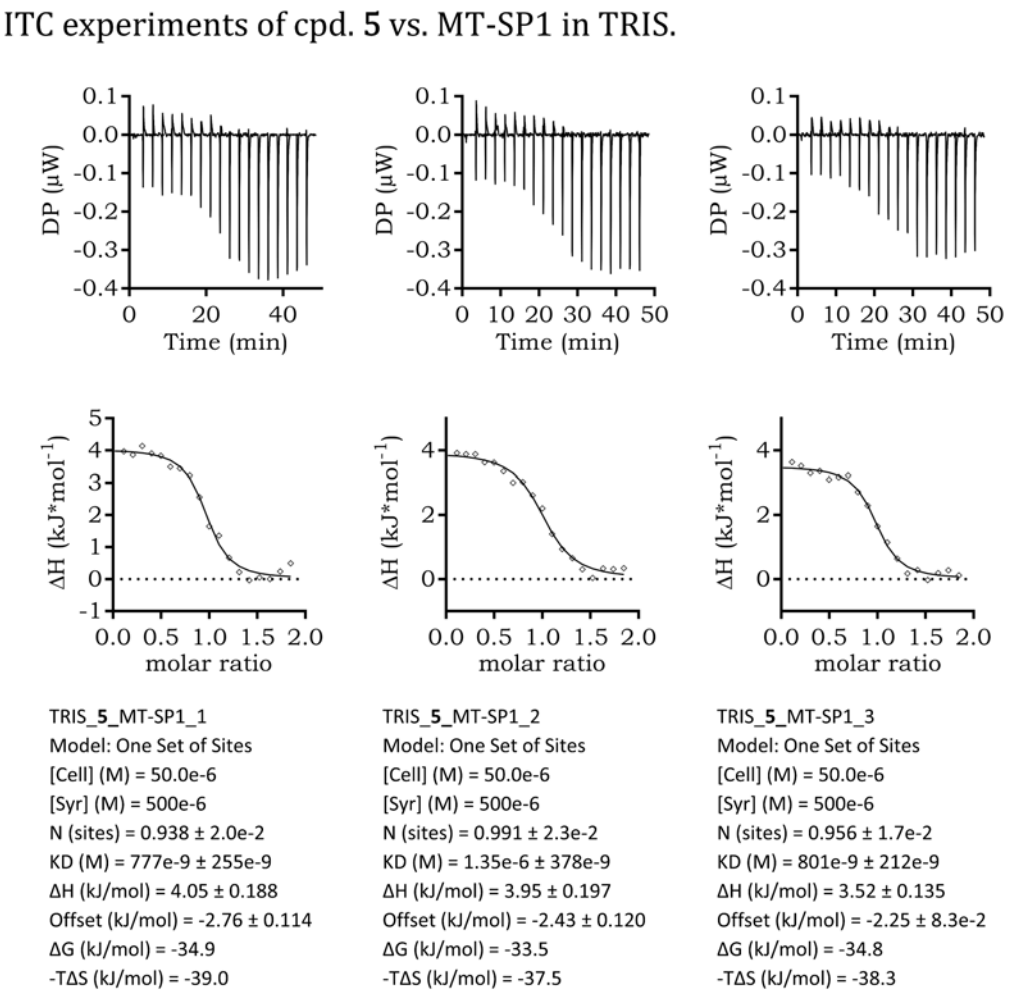


Figure S5H. Thermograms and isotherms of cpd **5** vs. MT-SP1 in ITC_{TRIS} buffer.

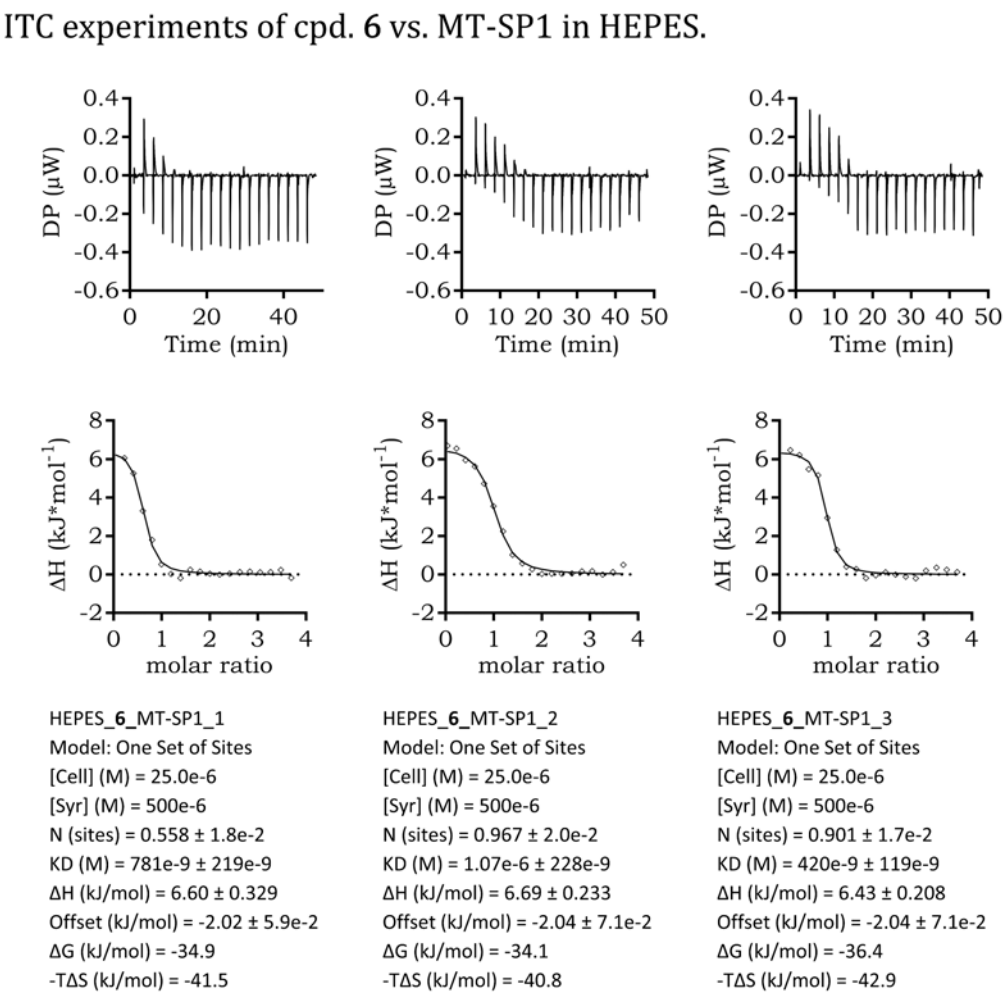


Figure S5I. Thermograms and isotherms of cpd **6** vs. MT-SP1 in ITC_{HEPES} buffer.

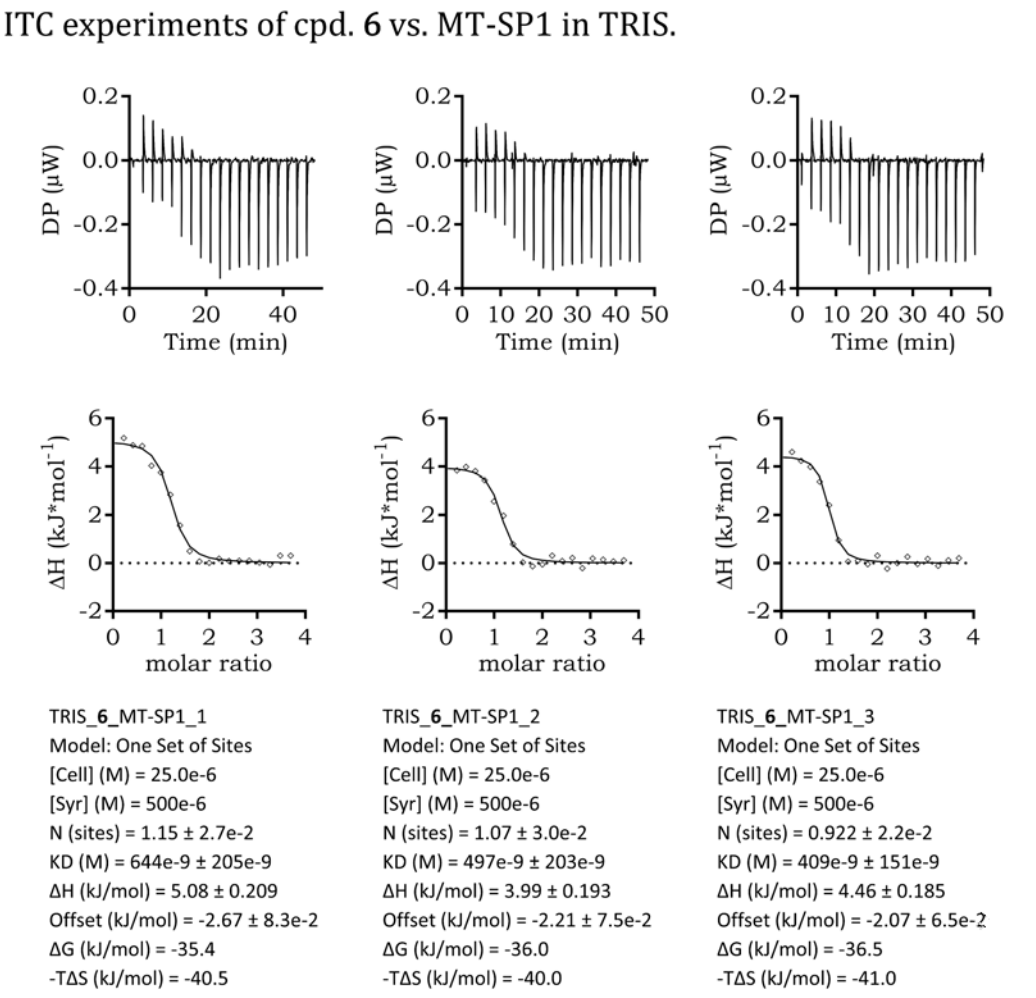


Figure S5J. Thermograms and isotherms of cpd **6** vs. MT-SP1 in ITC_{TRIS} buffer.

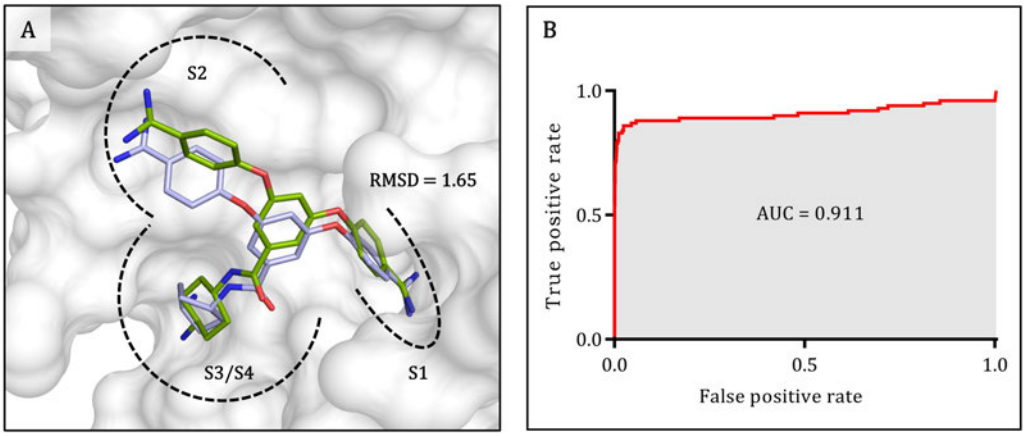


Figure S6. Docking validation. **A** redocking of cpd **15** (PDB ID 4O9V)[1] reveals an RMSD of 1.65 Å (X-ray: green sticks, redocked pose: pale blue sticks, FlexX score –53.9 kJ·mol^{-1}). **B** ROC-curves of the 100 binders and 1170 decoys for validation of the molecular docking studies.

Figure S7. SDS-PAGE analysis of the MT-SP1 purification and autoactivation process. Precision Plus Protein™ Dual Xtra Prestained Protein Standards (BioRad, Hercules, CA, USA) was used as a molecular weight marker. Protein bands were stained with Coomassie brilliant blue. Lane 1 contains the MT-SP1 after Ni-NTA and subsequent dialysis, which was loaded to the AEX column. Lanes 2–9 show the respective elution fractions. Lane 10 contains a sample of the pooled and concentrated MT-SP1. In lane 1, there is mainly the zymogen, whereas autoactivation occurs during AEX. After concentrating of the eluted fractions, the active MT-SP1 is prevailing.

Table S1. Symmetry point groups of inhibitors **1**–**6**. The assignment was performed as described in Lauria *et al.*[6]

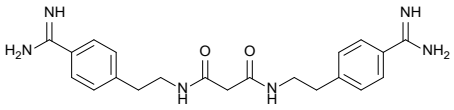
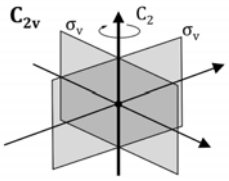
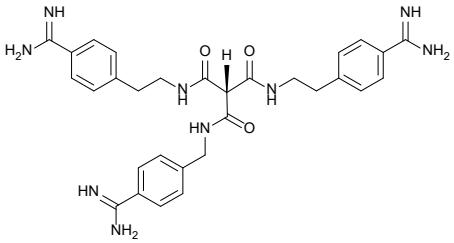
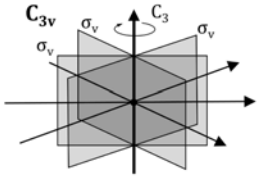
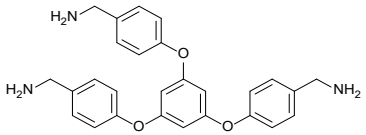
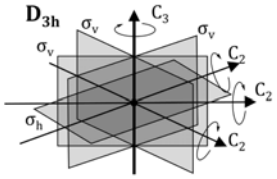
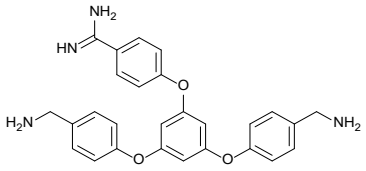
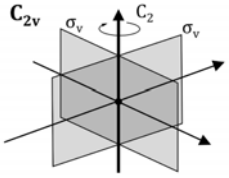
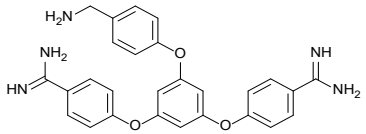
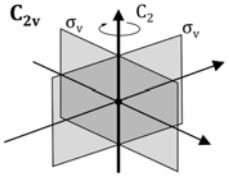
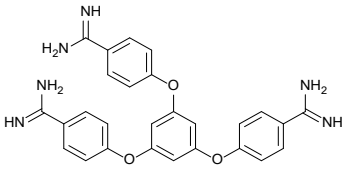
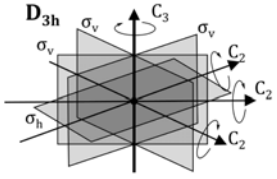
Cpd	Structure	Symmetry point group
1		
2		
3		
4		
5		
6		

Table S2. FlexX docking scores for poses resembling possible solutions. Predicted binding modes are shown in Figure 2.

Cpd	Rank	Docking score (kJ·mol^{-1})	orientation
1	1	–43.6	
	3	–42.5	
2	1	–49.0	
			(inconclusive binding mode)
	4	–46.2	
3	1	–33.4	
4	1	–40.1	
5	1	–42.7	
	15	–41.9	
6	1	–45.5	

Table S3. K_M values of Boc-LRR-AMC, and IC_{50} and K_i values of **3–6** in dependence of buffer pH.

Enzyme	K_M (μM)[a]				
	pH 7.0	pH 7.5	pH 8.0	pH 8.5	pH 9.0
MT-SP1	255 ± 37	201 ± 33	39.1 ± 5.1	65.9 ± 6.2	55.0 ± 5.0
Cpd	IC_{50} (μM)				
	pH 7.0	pH 7.5	pH 8.0	pH 8.5	pH 9.0
3	104 ± 8	61.4 ± 5.3	65.3 ± 3.6	50.6 ± 4.0	71.8 ± 10.7
4	n.d.	n.d.	2.46 ± 0.29	n.d.	n.d.
5	n.d.	n.d.	1.46 ± 0.10	n.d.	n.d.
6	0.589 ± 0.048	1.37 ± 0.06	1.42 ± 0.08	0.680 ± 0.051	0.910 ± 0.057
Cpd	K_i (μM)				
	pH 7.0	pH 7.5	pH 8.0	pH 8.5	pH 9.0
3	77.9 ± 6.7	43.1 ± 4.3	18.4 ± 2.0	22.1 ± 2.1	28.2 ± 4.5
4	n.d.	n.d.	0.775 ± 0.115	n.d.	n.d.
5	n.d.	n.d.	0.459 ± 0.051	n.d.	n.d.
6	0.440 ± 0.039	0.961 ± 0.064	0.398 ± 0.043	0.297 ± 0.027	0.358 ± 0.030

a: substrate BOC-LRR-AMC. n.d.: not determined.

Table S4. ITC results including errors from direct titrations. Experiments were performed at least in triplicates.

Ligand	buffer	N	K_d (μM)	ΔG_{obs} (kJ·mol^{-1})	ΔH_{obs} (kJ·mol^{-1})	$-T\Delta S_{obs}$ (kJ·mol^{-1})
1	TRIS	0.59 ± 0.01	1.45 ± 0.19	–33.4 ± 0.2	–14.0 ± 0.5	–19.4 ± 0.4
2	TRIS	0.84 ± 0.02	0.392 ± 0.121	–36.9 ± 1.2	–6.8 ± 0.3	–30.1 ± 1.3
3	HEPES	0.88 ± 0.06	12.7 ± 2.2	–28.1 ± 0.8	4.4 ± 0.3	–32.6 ± 0.8
	TRIS	1.15 ± 0.03	7.76 ± 0.94	–29.6 ± 1.5	3.5 ± 0.1	–33.0 ± 1.0
4	HEPES	0.69 ± 0.02	2.50 ± 0.44	–32.1 ± 0.5	7.4 ± 0.6	–39.4 ± 1.2
	TRIS	0.65 ± 0.02	2.53 ± 0.60	–32.0 ± 0.5	5.8 ± 0.3	–37.9 ± 0.6
5	HEPES	0.93 ± 0.01	1.42 ± 0.23	–33.5 ± 0.4	6.4 ± 0.2	–39.8 ± 0.4
	TRIS	0.96 ± 0.01	0.976 ± 0.168	–34.4 ± 0.6	3.8 ± 0.1	–38.3 ± 0.6
6	HEPES	0.81 ± 0.01	0.757 ± 0.113	–35.1 ± 1.0	6.6 ± 0.2	–41.7 ± 0.9
	TRIS	1.05 ± 0.02	0.517 ± 0.109	–36.0 ± 0.5	4.5 ± 0.1	–40.5 ± 0.4

Table S5. Comparison of calculated ($\Delta\Delta G_{calc}$) and experimental ($\Delta\Delta G_{ITC}$) changes in ΔG, assuming equal $K_{a,inc}$ for inhibitors **4**–**6**.

Cpd	Qualitative binding mode affinity comparison	Estimated incremental affinity $K_{a,inc}$ (μM)		$\Delta\Delta G_{(calc)}$ (kJ·mol^{-1})	$\Delta\Delta G_{(ITC)}$ (kJ·mol^{-1})	$\Delta-T\Delta S_{(ITC)}$ (kJ·mol^{-1})
3	$[PL]_1 = [PL]_2 = [PL]_3 = [PL]_4 = [PL]_5 = [PL]_6$	$K_{a,app} = K_{a,1} = K_{a,2} = K_{a,3} = K_{a,4} = K_{a,5} = K_{a,6} = 6 \cdot K_{a,inc}$				
4	$[PL]_1 = [PL]_2 >> [PL]_3 = [PL]_4 \approx [PL]_5 = [PL]_6$	$K_{a,app} = K_{a,1} = K_{a,2} = 2 \cdot K_{a,inc}$	$\Delta\Delta G_{\mathbf{4}\rightarrow\mathbf{5}}$	$-RT \cdot \ln(2 \cdot K_{a,inc}) - (-RT \cdot \ln(4 \cdot K_{a,inc})) = -RT \cdot \ln(\frac{4}{2}) = -1.7$	–1.8	–1.6
5	$[PL]_1 = [PL]_2 \approx [PL]_3 = [PL]_4 >> [PL]_5 = [PL]_6$	$K_{a,app} = K_{a,1} = K_{a,2} = K_{a,3} = K_{a,4} = 4 \cdot K_{a,inc}$				
6	$[PL]_1 = [PL]_2 = [PL]_3 = [PL]_4 = [PL]_5 = [PL]_6$	$K_{a,app} = K_{a,1} = K_{a,2} = K_{a,3} = K_{a,4} = K_{a,5} = K_{a,6} = 6 \cdot K_{a,inc}$	$\Delta\Delta G_{\mathbf{5}\rightarrow\mathbf{6}}$	$-RT \cdot \ln(4 \cdot K_{a,inc}) - (-RT \cdot \ln(6 \cdot K_{a,inc})) = -RT \cdot \ln(\frac{6}{4}) = -1.0$	–1.7	–1.5

Table S6. Used concentrations for direct titrations.

Cpd	Cell (μM)	Syringe (μM)
1	15	200
2	15	200
3	50	2000
4	50	500
5	50	500
6	25	500

REFERENCES

(1) Goswami, R.; Mukherjee, S.; Ghadiyaram, C.; Wohlfahrt, G.; Sistla, R. K.; Nagaraj, J.; Satyam, L. K.; Subbarao, K.; Palakurthy, R. K.; Gopinath, S.; Krishnamurthy, N. R.; Ikonen, T.; Moilanen, A.; Subramanya, H. S.; Kallio, P.; Ramachandra, M. Structure-Guided Discovery of 1,3,5 Tri-Substituted Benzenes as Potent and Selective Matriptase Inhibitors Exhibiting in Vivo Antitumor Efficacy. *Bioorg. Med. Chem.* **2014**, *22* (12), 3187–3203. https://doi.org/10.1016/j.bmc.2014.04.013.

(2) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; Chun, M.; Li, P.; Gohara, D. W.; Dolinsky, T.; Konecny, R.; Koes, D. R.; Nielsen, J. E.; Head-Gordon, T.; Geng, W.; Krasny, R.; Wei, G.; Holst, M. J.; McCammon, J. A.; Baker, N. A. Improvements to the <scp>APBS</Scp> Biomolecular Solvation Software Suite. *Protein Sci.* **2018**, *27* (1), 112–128. https://doi.org/10.1002/pro.3280.

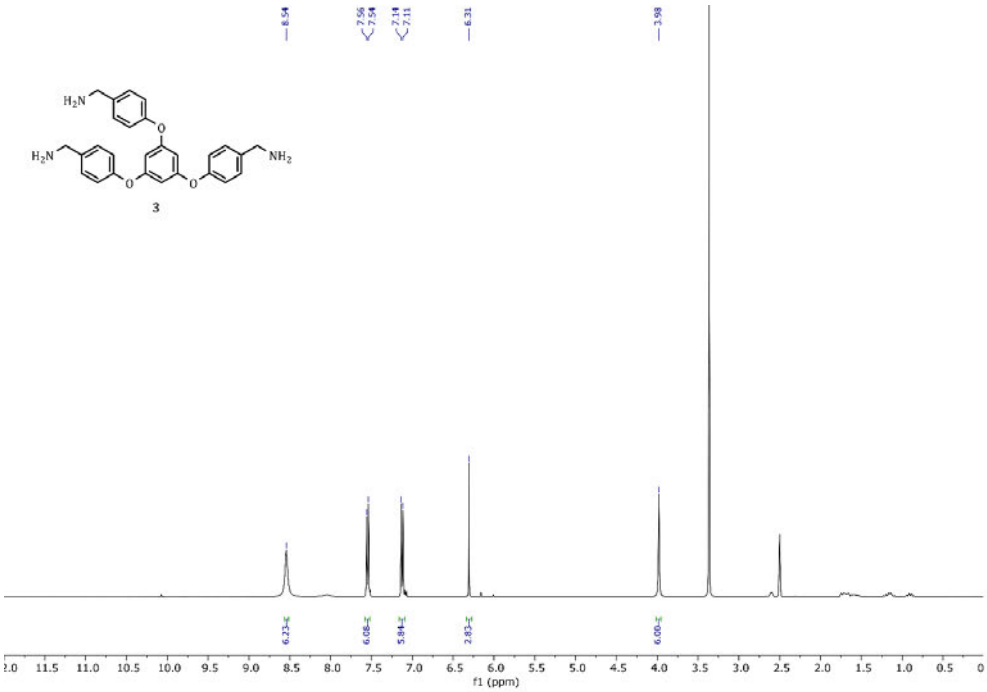
(3) The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography*. Schrödinger 2002, pp 82–92. http://www.pymol.org.

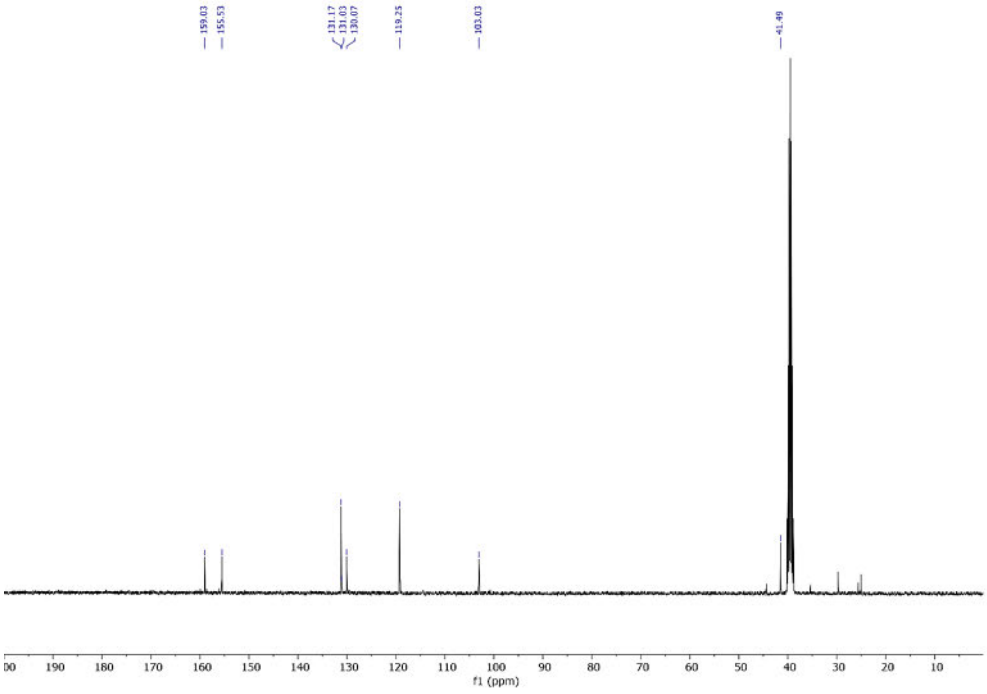
(4) Robin J. Leatherbarrow. GraFit 6. Erithacus Software Limited: East Grinstead, West Sussex, U.K. 2007. http://erithacus.com/grafit/.

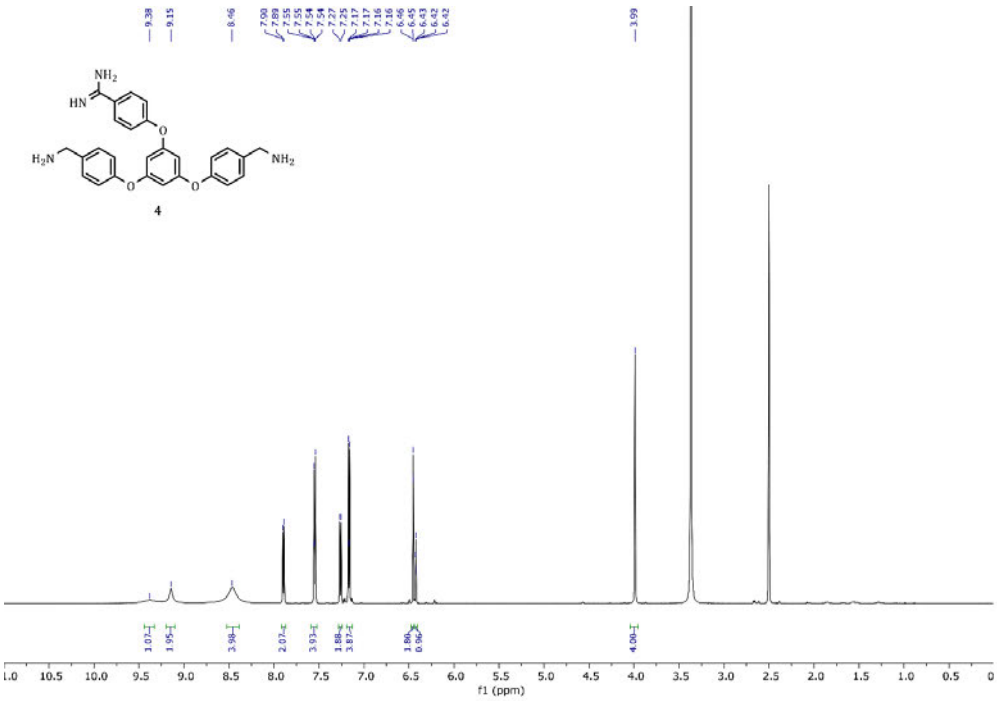
(5) Marvin 21.19. *Marvin 21.19*. ChemAxon (http://www.chemaxon.com): Budapest, Ungarn 2021. http://www.chemaxon.com.

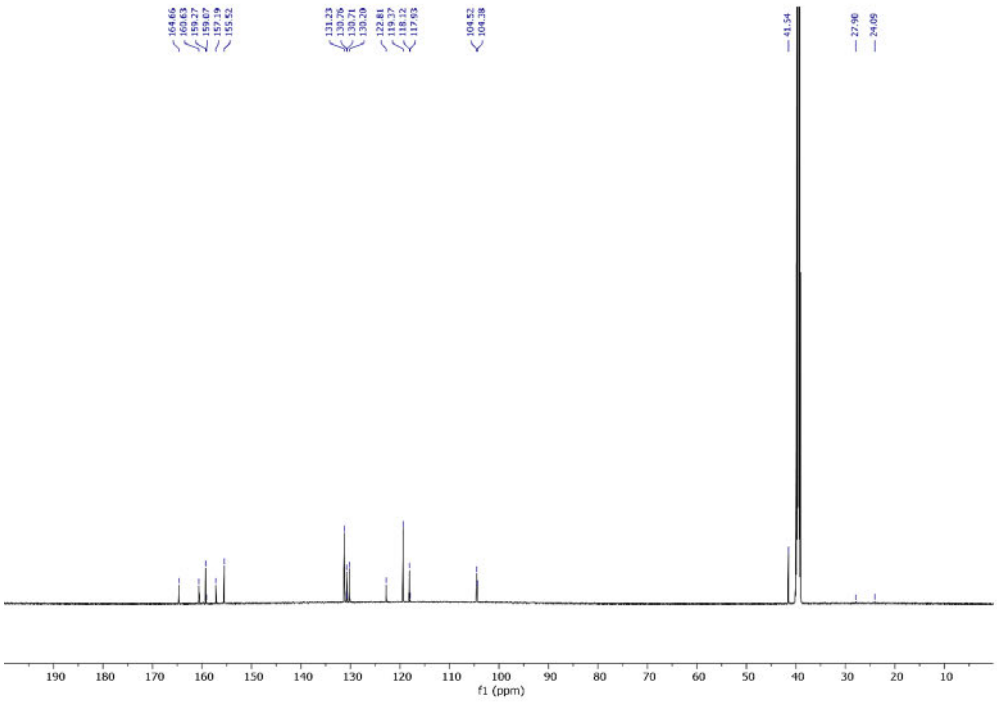
(6) Lauria, A.; Terenzi, A.; Bartolotta, R.; Bonsignore, R.; Perricone, U.; Tutone, M.; Martorana, A.; Barone, G.; Almerico, A. Does Ligand Symmetry Play a Role in the Stabilization of DNA G-Quadruplex Host-Guest Complexes? *Curr. Med. Chem.* **2014**, *21* (23), 2665–2690. https://doi.org/10.2174/0929867321666140217155156.

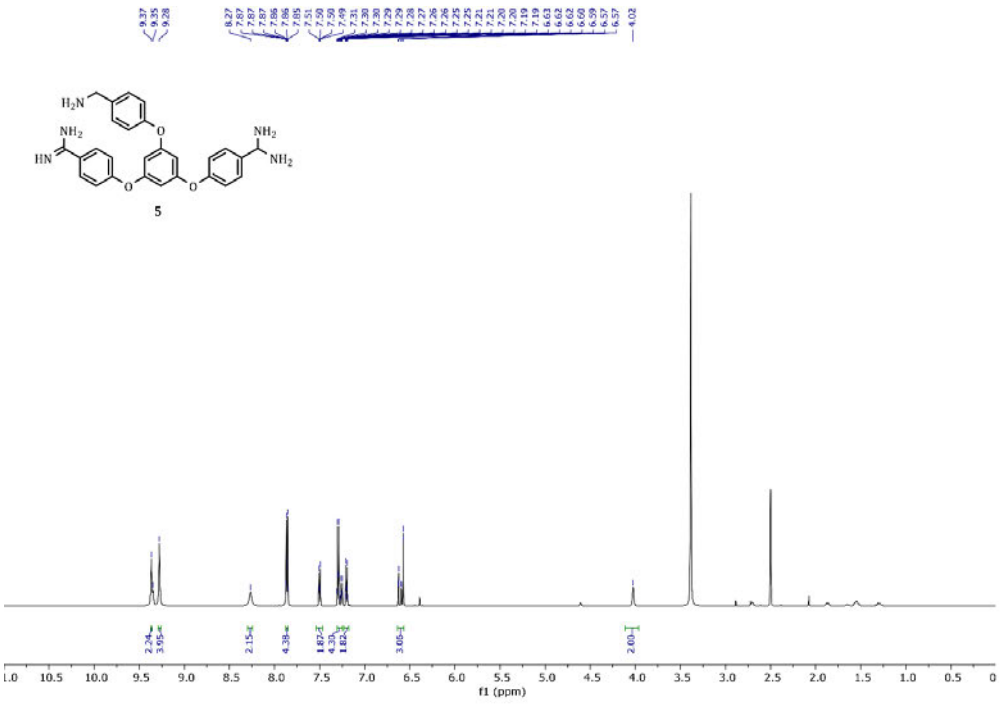
Spectral Appendix

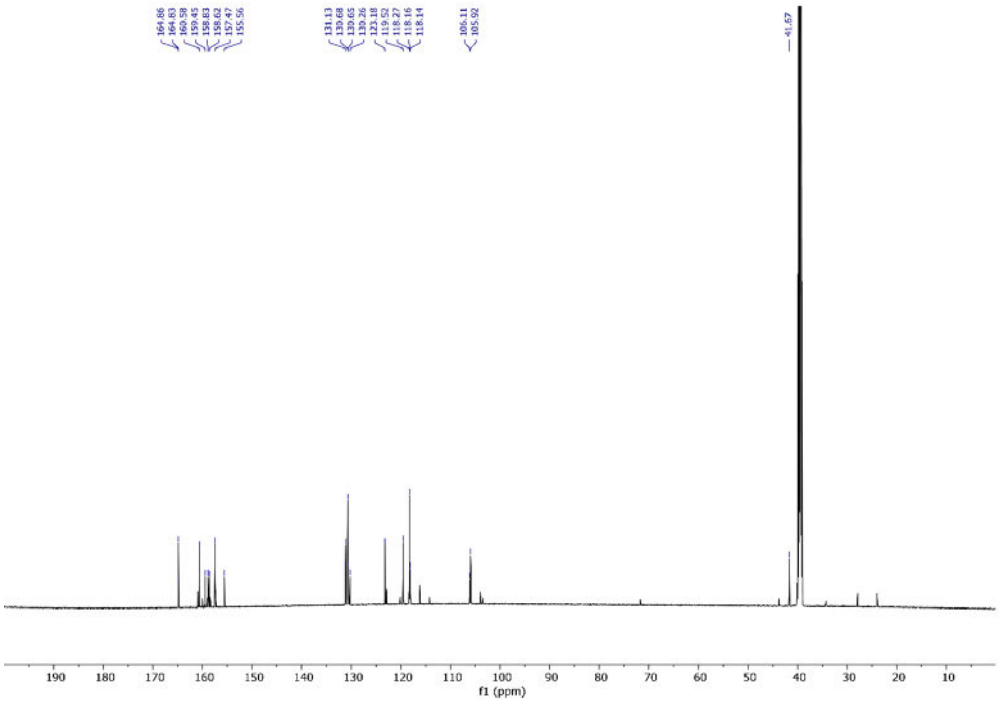


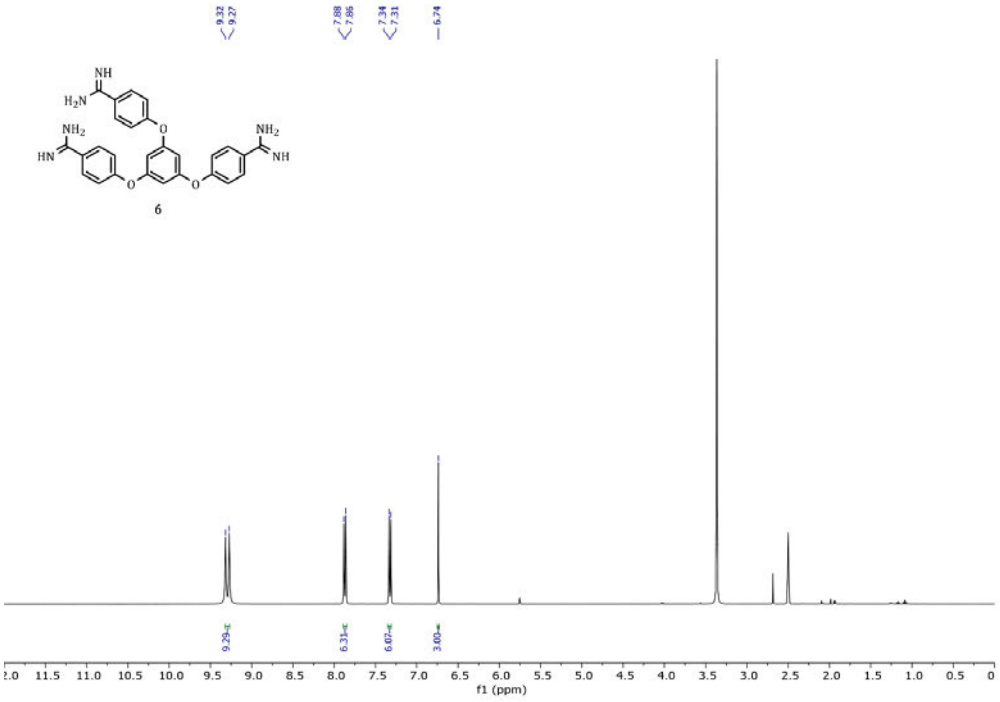


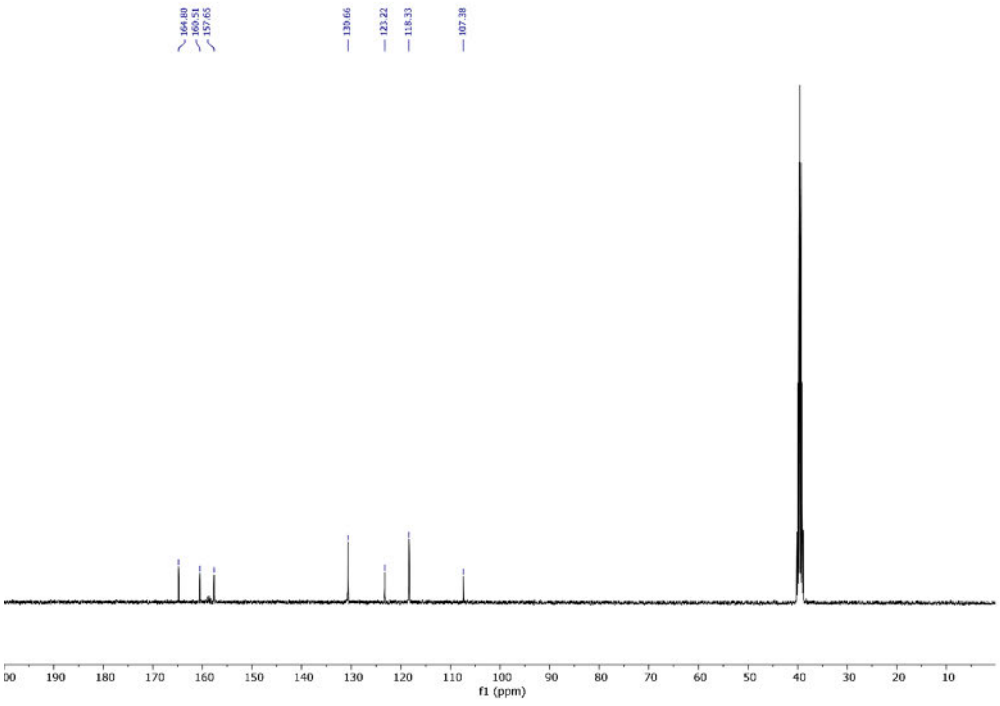

9.38
9.15
8.46
7.90
7.89
7.55
7.55
7.54
7.54
7.27
7.25
7.17
7.17
7.16
7.16
6.46
6.45
6.43
6.42
6.42
3.99
NH2
HN
O
H2N
NH2
O
O
4
1.07
1.95
3.98
2.07
3.93
1.88
3.87
1.86
6.98
4.00
f1 (ppm)


164.66
160.63
159.77
159.67
157.19
155.52
131.23
130.76
130.71
130.29
122.81
119.37
118.32
117.93
104.52
104.38
41.54
27.90
24.09
f1 (ppm)


9.37
9.35
9.28
8.27
7.87
7.87
7.87
7.86
7.86
7.85
7.51
7.50
7.50
7.49
7.31
7.30
7.30
7.29
7.29
7.28
7.27
7.26
7.26
7.25
7.21
7.21
7.20
7.20
7.19
7.19
6.63
6.62
6.62
6.60
6.59
6.57
6.57
4.02
H2N
NH2
HN
O
O
O
O
NH2
NH2
5
2.24
3.95
2.15
4.38
1.87
4.33
1.82
3.05
2.01
f1 (ppm)


164.86
164.63
160.58
159.45
158.83
158.62
157.47
155.56
131.13
130.68
130.65
130.36
123.18
119.52
118.27
118.16
118.14
106.11
105.92
41.67
f1 (ppm)


9.32
9.27
7.88
7.86
7.34
7.31
6.74
NH
H2N
NH2
HN
O
NH2
NH
O
O
6
9.29
6.31
6.07
3.00
f1 (ppm)


164.80
160.61
157.65
130.66
123.22
118.33
107.38
f1 (ppm)

3.3. Advanced Isothermal Titration Calorimetry for Medicinal Chemists with *ITCcalc*

3.3.1. Context, Project Summary, and Own Contribution

In the other parts of this work, ITC was employed to characterize thermodynamic binding profiles of ZIKV and MT-SP1 inhibitors and cation binding to SrtA. Beyond following standard procedures, several tasks demanded advanced experimental setups such as displacement titrations to determine the thermodynamics of too-high or too-low affine inhibitors, corrections for buffer ionization enthalpy and the determination of n_{proton}, the determination of ΔC_P, and consecutive titrations. Several educational papers concerning ITC are directed to undergraduate students and describe more basic experiments[165,168,169,233–238] or homebuilt calorimeters.[233–236] We noticed that no educational articles focus on reliable teaching experiments to demonstrate best practice experimental design and advanced result evaluation, covering these more sophisticated methodologies. To close this gap and to guide medicinal chemists through these issues, we present hands-on experiments in the trypsin-ligand system and an even more cost-efficient EDTA-based variant. To provide a platform guiding through the underlying calculations of the obtained experimental results, we launched the open, accessible ITC web calculator *ITCcalc* (https://itccalc.uni-mainz.de), guiding through the underlying evaluation of the obtained experimental results. This calculator combines the functionalities of a ΔC_P calculator, a buffer protonation correction tool, a displacement calculator, and a ligand descriptor calculation tool that will prove valuable for ligand development purposes.

(3) **Hammerschmidt, S. J.:** [redacted] Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc. *Manuscript submitted to J. Chem. Educ.*

Own contribution: ITC experiments and evaluation, writing of the original draft & editing of the manuscript.

Contributions from other authors: Programming of the *ITCcalc* webserver, writing & editing of the manuscript.

Manuscript submitted to *ACS chemical education* (impact factor: 3.21).

3.3.2. Publication

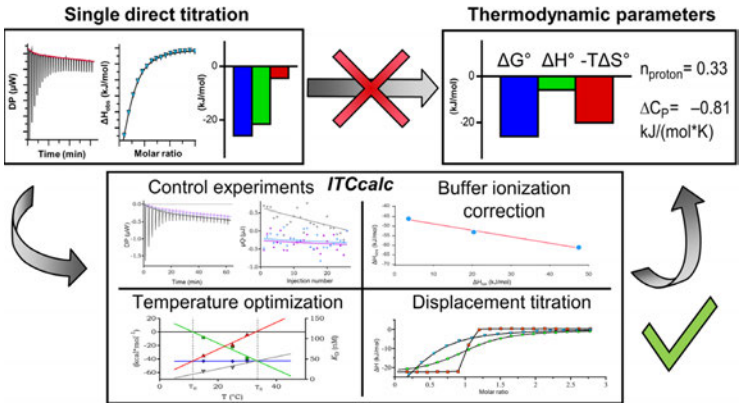
The following manuscript, from page 130 to page 173, contains unpublished data.

Advanced Isothermal Titration Calorimetry for Medicinal Chemists with *ITCcalc*

Stefan J. Hammerschmidt[a,†], [redacted]

[a.] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany*
[b.] *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany*
† Authors contributed equally.
* [redacted]

ABSTRACT



Isothermal titration calorimetry (ITC) is the method of choice for thermodynamic binding profiling. As a label-free in-solution technique, ITC experiments do not require modified or immobilized proteins or ligands and hence, can be performed under close-to-native environmental conditions. It directly follows the released heat of molecular interaction and does not need competing reporter molecules. However, advanced experimental ITC setups are frequently necessary to derive thermodynamic parameters for medicinal chemistry applications. Beyond 1:1 protein-ligand binding, complex systems include displacement titrations to determine thermodynamics for ligands with affinities either too high or too low to be derived from direct titrations, the characterization of protonation changes upon binding, or the elucidation of heat capacities.

We present hands-on experiments for graduate medicinal chemists and alike familiar with standard ITC methods, reliably highlighting these issues and their respective solutions. It can be chosen from two setups: a biochemical trypsin-ligand system or a more cost-efficient EDTA-based variant. This demonstration of best practice protocols comes together with the open, accessible ITC-web calculator *ITCcalc* (https://itccalc.uni-mainz.de), guiding through the underlying evaluation of the obtained experimental results.

KEYWORDS

Continuing Education, Biochemistry, Chemical Education Research, Laboratory Instruction, Physical Chemistry, Transfer, Hands-On Learning, Biophysical Chemistry, Calorimetry / Thermochemistry, Heat Capacity, Instrumental Methods, Stoichiometry, Thermodynamics, Medicinal Chemistry, Drug Discovery

INTRODUCTION

Theoretical background

To date, the method of choice for characterizing thermodynamic binding profiles constitutes isothermal titration calorimetry (ITC).[1] Besides sole binding measurements, a well-designed ITC experiment allows for the extraction of the binding stoichiometry (N), the dissociation constant (K_D), and changes in free binding energy or Gibbs energy (ΔG), enthalpy (ΔH), and entropy (ΔS) of a binding process. Almost every biological or chemical process is accompanied by the consumption or release of heat.[2] Consequently, ligand binding to a target molecule is mostly accompanied by an enthalpic term ≠ 0. The affinity of binding processes (the reciprocal K_D) depends on ΔG, as shown in Equation 1, with R as the universal gas constant and T as the absolute temperature in Kelvin.

$$\Delta G = R * T * \ln(K_D) \quad (1)$$

As described by Gibbs in 1873, ΔG itself is the sum of the changes in enthalpy and the negative product of T, and the change in the system's entropy (Equation 2).[3]

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

To aid in rational drug discovery and for research purposes, it is of particular interest to determine binding enthalpies using instrumental methods. However, the impact of thermodynamic data in decision-making on drug development success rates is still discussed.[4,5] On the one hand, it is compelling that thermodynamic parameters can be used to describe the contributions of modifications to the binding energy. On the other hand, some success stories were found to be rather rationalized retrospective explanations based on structural information.[6,7] Most researchers, however, try to achieve mainly enthalpy-driven ligands, as their affinity is believed to result in more specific target-ligand interactions. Entropic binders are expected to result in rather unspecific lipophilic interactions.[4] Additionally, ligand optimization is often accompanied by an improved entropy term, while binding enthalpy decreases due to an enthalpy-entropy compensation.[8]

Besides prioritization of enthalpic binders over entropic ones, additional metrics are used to validate progress in lead optimization during drug development. These relate the observed target affinity to molecular size and/or lipophilicity.[8,9] Commonly used metrics comprise:

- ligand efficiency (LE)[10,11]
- lipophilic ligand efficiency (LLE)[12]
- ligand efficiency dependent lipophilicity (LELP)[13]
- LLE adjusted for heavy atom count (LLE_{AT})[8]

The LE is calculated from Equation 3, with HA as the number of heavy atoms (non-hydrogens). The LE offers an intriguingly simple way to evaluate the potency of lead structures concerning their size.

$$LE = \Delta G / HA = \left(-2.303\left(\frac{R * T}{HA}\right)\right) * \log(K_D) \quad (3)$$

It reflects the concept that if two ligands are similar in potency, further developing the smaller one is more promising, as it offers more room for optimization before size inflation limits bioavailability.[13,14] The widespread use of LE as a hard limit (usually LE > ~0.3 (kcal*mol^{-1})/HA), resembling a K_D of 10 nM for a molecule with 38 HA corresponding to ~500 g*mol^{-1}) for compound progression raised concerns about the negative effects of replacing complex considerations with strict usage of LE.[15] It is convenient that the nature of some large binding sites, such as protein-protein binding interfaces, requires larger ligands with a lower LE that would not be identified by the application of hard LE cutoffs.[16,17] Another throwback of LE is its nonlinear scaling when used to compare ligands of significantly different sizes.

In typical lead optimization steps, initial enthalpic binding ligands with low molecular weight are grown and supplemented with bulky hydrophobic substituents. Thereby affinity and lipophilicity increase, wheres and the LLE (Equation 4) decreases along these optimization campaigns and allows for the distinction of polar and ionic interactions rather than unspecific lipophilic potency improvements. Usually, an LLE > ~5 is desirable as it reflects a molecule with a IC_{50} of 10 nM and a logP < 3.[12,18]

$$LLE = p(IC_{50}) - \log P \quad (4)$$

The main benefits of LLE are its normalization to logP, a well-known ligand property in medicinal chemistry that enables better scaling. However, a major throwback of LLE is the lack of consideration of molecular size. Alternatively, some combined metrics were presented to follow LE and LLE optimization. Attempts to combine LE with LLE to account for molecular size and lipophilicity led to the development of LELP, the division of logP and LE (Equation 5).[13] An optimal LELP for lead discovery is in the range of –10 < LELP < 10.[13]

$$LELP = \log P / LE \quad (5)$$

The LLE_{AT} constitutes LLE, corrected for the molecules' heavy atom counts, and is scaled to be comparable to LE (Equation 6).

$$LLE_{AT} = 0.111 + 1.37\,(LLE / HA) \quad (6)$$

Advanced Experimentation

Control experiments. Control experiments are a mandatory requirement for the collection of reliable ITC data. If not thoroughly performed, dilution enthalpies or dissociation reactions of agglomerates can bias the results or, in some cases, be misinterpreted as specific (low-C) binding events.[19,20] Therefore, a typical control experiment set comprises the titrations of buffer to the analyte, ligand to buffer, and buffer to buffer. The best correction method depends on the shape of the isotherms of the control experiments. This correction method should be chosen if a regression line can express the isotherm's shape. The *mean* method can be used if the regression slope is close to zero and an offset fit is sufficient. If more complex shapes occur, e.g., when a dissociation occurs, the method of choice constitutes the *point-to-point* correction, where each injection of the correction measurement is subtracted from the respective injection of the binding experiment. This method, however, increases the overall noise, as statistical errors from two measurements are summed up.

Heat capacity and temperature-dependency

Measuring the changes in ΔH at different experimental temperatures (ΔΔH) reveals the isobaric binding heat capacity (ΔC_P) upon a reaction, as can be derived from Equation 7.

$$\Delta C_P = \left(\frac{\Delta\Delta H}{\Delta T}\right)_P \tag{7}$$

Since ΔC_P is strongly connected with (de-)solvation, flexibility, and conformational changes that occur during complex formation,[21–23] the binding heat capacity enables the assessment of potential conformational changes upon ligand binding events. Assuming no conformational changes, the structural information about the free binding partners can be used to calculate the changes in the polar and non-polar accessible surface area and to assess changes in solvation conditions.[24,25] This allows the estimation of the binding heat capacity from rigid bodies. The deviation from the calculated and the experimentally determined binding heat capacity is a proxy of the possible conformational changes.[26] One can estimate a more negative ΔC_P in reactions where hydrophobic residues are buried upon ligand binding or conformational changes.[27–30] However, those assumptions and calculations should be taken with caution, as several disturbing discrepancies in these simplified perspectives were reported.[31,32]

Based on Equation 7, ΔC_P describes the difference in heat absorbed or released during a binding process upon changed temperatures in an isobaric system. In other words, it is the changed heat capacity of the system upon ligand binding. At different temperatures, ΔG remains unchanged, whereas significant differences exist in observed enthalpy (ΔH_{obs}).[21,26,27,33] Therefore, knowing ΔC_P is crucial to estimating varying temperatures' influences on observed ΔH.

Some protein-small molecule binding events are mainly driven by beneficial entropic effects. Even a high-affinity ligand can lack sufficient enthalpy signals if the binding enthalpy is close to zero. The following

experimental modifications can circumvent these issues: Fewer injections with higher injection volumes or devices with larger reaction vessels. Increasing sample volumes leads to a linear increase in measured enthalpies but also sample consumption. Adjustments in reaction temperature, however, constitute a more elegant solution. This strategy assumes that ΔG remains largely unchanged at different temperatures (Equation 8, Figure 1).[21,26,27,33] If no pleasing effect can be achieved applying the earlier methods, displacement titrations of ligands with a more extensive enthalpic term offer the indirect characterization of the desired thermodynamics.[19,34]

Based on Equation 2 and given that ΔG remains largely unchanged at different temperatures (this effect is termed "thermodynamic homeostasis"),[21,26,27,33] the thermodynamic binding profile can be manipulated by shifting the reaction temperature to yield increased ΔH values with improved signal/noise ratios. Without temperature-dependent conformational changes or reactions, ΔH and −TΔS show a linear temperature dependency.[26] In a certain range of temperatures, the binding heat capacity ΔC_P can be used to estimate the behavior of the binding affinity upon a change in temperature (Equation 8),

$$\Delta G(T_1) = \Delta H(T_0) + \Delta C_P(T_1 - T_0) - T_1\left(\Delta S(T_0) + \Delta C_P \ln\frac{T_1}{T_0}\right) \quad (8)$$

where T_0 is the reference temperature and T_1 is the altered temperature, $\Delta H(T_0)$ is ΔH determined at reference temperature T_0 and $\Delta S(T_0)$ is ΔS determined at reference temperature T_0. When the binding constant K_A and the energetic contributions of ΔH and −TΔS to ΔG are plotted against the temperature (Figure 1), the slope of the linear ΔH regression equation is given by ΔC_P (Equations 7, 8), the slope of the −TΔS is defined as $-(\Delta C_P + \Delta S)$. Since ΔC_P is usually much larger than ΔS, the slopes of the enthalpic and entropic lines are similar, resulting in a ΔG insensitive to temperature changes. The intercept of the ΔH line defines T_H at which the binding enthalpy is zero, and the binding is only dependent on −TΔS. At T_S, the entropic contribution is zero. Using *ITCcalc*, ΔC_P, T_H, and T_S can be calculated by entering ΔH_{obs} and ΔG° into the respective temperature (Figure S1).

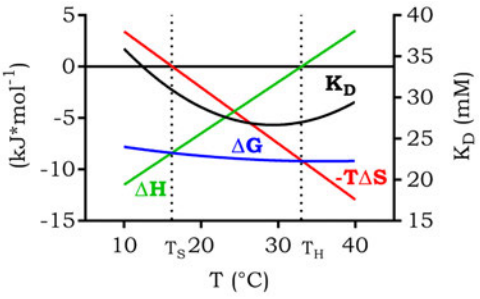


Figure 1. Schematic temperature dependency plot of the thermodynamic binding profile for a macromolecule-ligand interaction (ΔG = −9 kJ/mol; ΔH = −4 kJ/mol; −TΔS = −5 kJ/mol at 25 °C and ΔC_P = 0.5 kJ/mol/K). The ΔH plot is green−TΔS is red, and ΔG is blue.

Buffer ionization correction. The uptake or release of protons often accompanies binding events. These can occur when interactions favor another protonation state as present in the unbound state. The transferred protons are usually provided or absorbed by the buffering solvent. If not corrected, the resulting solvent ionization events display an enthalpic term that can result in misleading heat signal. Therefore, the ITC experiments must be conducted in different buffered solutions with differing ionization enthalpies. Based on the known ionization enthalpies and the binding stoichiometry, it is possible to calculate the number of protons transferred upon the binding event. Therefore, for each direct titration, ΔH_{obs} is plotted against the ionization enthalpy of the respective buffering agent. The slope of the regression line then reveals the number of transferred protons (n_{proton}), and the intercept with the x-axis gives the corrected binding enthalpy ($\Delta H°$) (Figure 1). Regarding proton transfers upon binding, the observed binding enthalpies (ΔH_{obs}) should be linearly related to the n_{proton} multiple of the respective buffer's ionization enthalpy (ΔH_{ion}). For ITC experiments with insufficient binding enthalpies and occurring proton transfers, this effect can be exploited by choosing a buffer whose ionization enthalpy increases ΔH_{obs}, leading to improved signal/noise ratios. Another utilization of this effect is adjusting pH values, enabling proton transfers.[35] It should be noted that these proton transfers are not restricted to the ligand but represent transitions of the entire analyte-ligand system. Hence, the proportions of protonated species can also lead to non-integer numbers of protons transferred. *ITCcalc* can help to select suitable buffers. Entering the desired pH gives several buffers with pk_a values in the range of pH ± 1 and the respective ΔH_{ion} values. For buffer ionization correction using *ITCcalc*, $\Delta G°$ and the obtained enthalpies ΔH_{obs} of at least two buffers can be entered to calculate a linear regression. The output values are the slope of that regression (n_{proton}) and the corrected $\Delta H°$ and $-T\Delta S°$ (Figure S2).

Displacement titrations. The C-value is an important metric to estimate the shape of a titration's Wiseman isotherm and plan experiments (Figure 2).[36] In an experiment with one binding site, where M_{tot} is the total concentration of the analyte, and K_D is the dissociation constant, C is the ratio of analyte concentration to the K_D of the analyzed binding reaction (Equation 9).[36]

$$C = M_{tot} / K_D \qquad (9)$$

Depending on the instrument, C should be between 1–1000 to deconvolute the titration curves to obtain binding affinities, with an optimum C between 10 and 100.[19,37–39] To meet these C-values, the experimental design includes adjustment of M_{tot} to the estimated K_D. In a titration with a very tight binder, C is very high. All added ligand is completely bound until the molar equivalence point is reached, the target molecule is saturated, and no more binding occurs upon further addition of more ligand. The resulting curve, therefore, displays a step function (Figure 2). No K_D can be derived from the infinite slope at the equivalence point. To satisfy the C-value rule, M_{tot} needs to be reduced. Therefore, the lower affinity limit for ITC experiments is defined by the low quantities of reactants used in these titrations, often resulting in insufficient heat signals.[34] C adopts a small value in titrations with very low-affinity binders. The obtained titration curve lacks an inflection point (Figure 2). The experiment, therefore, loses stoichiometric information and provides a highly

imprecise affinity. Hence, the upper concentration limit for direct titrations is often defined by sample consumption or poor solubility of the analytes, especially in the case of proteins.[40] To nevertheless be able to determine the thermodynamic binding profile in a low-affinity system, the binding stoichiometry must be manually set to the correct value derived from other experiments. A so-called low-C titration is performed to improve the precision of the measurement. Thereby, the ligand is titrated in higher molar excess to reach saturation of the analyte of more than 70%.[41] Therefore, binding affinities with a K_D of 10^{-9}–10^{-3} M can be assumed as the working framework for direct titrations.[19,42]

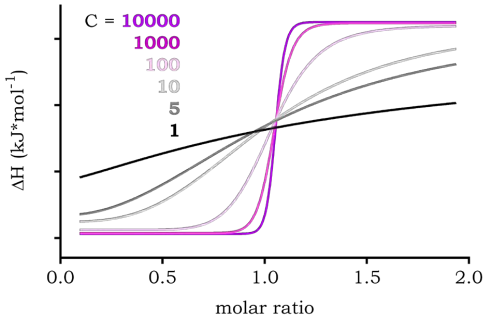


Figure 2. Wiseman isotherm for C-values ranging from 1 (black), over 5 (dark gray), 10 (pale gray), 100 (rose), 1000 (hot pink) to 10000 (violet). With lower C-values, the isotherms are getting shallower and lacking an inflection point. The information about stoichiometry gets lost. With C between 5 and 100, the curve is sigmoidal shaped and can be deconvoluted into stoichiometry, enthalpy, and binding affinity. With increasing C-values, the isotherm's shape gets steeper until the information about the affinity gets inaccurate. The stoichiometry and enthalpy, however, can still be obtained.

Displacement titrations allow expanding the affinity range of direct ITC experiments and can be utilized to reduce sample consumption. Especially in measurements to examine low-affinity binding events, sample consumption increases drastically to provide a sufficient heat signal and give a C-value between 10–100 while being limited by ligand and protein solubility simultaneously.[34] Thus, several mg of valuable protein can be consumed with only a few experiments.

In displacement experiments, a low-affinity binder is displaced by a ligand competing for the same binding site with a significantly higher affinity (at least factor 10 in K_D).[43] Due to displacement, the observed affinity (K_{Dapp}) of the high-affinity ligand is raised compared to the K_D of the direct titration experiment.[19]

Using a known strong binder, low-affinity ligands up to 20 mM[19] can be analyzed, enabling the use of ITC for fragment-based drug discovery (FBDD).[40,44,45] *Vice versa*, the high-affinity limit can be expanded to femtomolar concentrations if a high-affinity ligand (with a too-steep Wiseman isotherm, Figure 2) has to displace a lower-affinity ligand from the binding site.[46] If no effect on K_{Dapp} in the presence of the second ligand is measured, the

analytes are likely to bind to different binding sites.[19] Given the thermodynamic parameters of a known binding partner and the observed values for the displacement titration, the K_D and the thermodynamic parameters for either the unknown strong or weak binder can be calculated using Equations 10 and 11 as implemented in *ITCcalc*. According to the terminology of Zhang *et al.*, we denoted the parameters of the high-affinity ligand with the index 1 (e.g., K_{D1}, ΔH_1) and those of the low-affinity, displaced out of the binding pocket by the high-affinity ligand with the index 2. L_{2tot} is the concentration of the low-affinity ligand preincubated with the analyte in the reaction cell ligand.[19] . *ITCcalc* not only gives K_D, ΔG, ΔH, and $-T\Delta S$ of the unknown ligand but also displays the resulting signature plot (Figure S3).

$$K_{D2} = \left(\frac{K_{D1}}{K_{Dapp}} - 1\right) * \frac{1}{L_{2tot}} \tag{10}$$

$$\Delta H_2 = \left(\Delta H_1 - \Delta H_{app}\right) * \left(1 + \frac{1}{K_{D2} * L_{2tot}}\right) \tag{11}$$

Objectives

Many challenging questions in medicinal chemistry require more sophisticated experimental strategies instead of simply performing single direct titrations. Therefore, advanced ITC users can access a variety of methods to overcome limitations and improve the significance of their results. The user is strongly suggested to have some expertise in standard ITC experiments before advancing into the practices described herein. Most educational publications featuring ITC methodology are directed to undergraduate students and describe more basic experiments.[1,20,47–53] Others are focused on homebuilt calorimeters.[49–52] This article is conceptualized for more advanced ITC users amongst graduate students or late-stage undergraduates from medicinal chemistry, biochemistry, and related disciplines. It is intended to create awareness of typical hindrances and limitations in obtaining reliable data from ITC experiments for medicinal chemistry applications and to help to overcome these issues. While classical protocols for direct ITC, displacement titration ITC, and buffer ionization corrections are presented elsewhere,[1,48,50] this manuscript focuses on reproducible teaching experiments, demonstrating best practice experimental design and advanced result evaluation. As a model system, trypsin with different commercially available ligands is presented due to its biochemical/medicinal chemistry relevance and the affordable commercially available substances, leading to only 1.55 €/experiment (as mean costs per experimental set for the trypsin-based system, performed as technical triplicates including correction measurements as technical singlets, Table S1). Alternatively, an ethylenediaminetetraacetic acid (EDTA)-ion titration experimental setup, which might be readily established as a low-cost alternative for lab courses, can be obtained from the supporting information. The selected binding partners are reported to display K_D values suitable for measurements spanning from 60 fM to 22 µM (Table S2). Additional to these demonstrative experiments, we present the open-access web server *ITCcalc*

(https://itccalc.uni-mainz.de/), giving guided support through the calculations of advanced ITC experiment evaluation. Key functionalities comprise:

- A **ΔC_P calculator** allows concluding ΔC_P-based on repeated experiments at different temperatures as well as the T_S and the T_H at which the binding affinity is solely dependent on the enthalpic or entropic binding contribution, respectively. Only at temperatures $T_H < T < T_S$ is the binding both entropically and enthalpically favored (Figure S1).

- A **Buffer protonation correction tool** allows plotting ΔH_{obs} in ≥2 buffers to correct for their ionization enthalpies. It calculates the net protons transferred upon ligand binding and displays the buffer-corrected signature plot (Figure S2).

- The **displacement calculator** enables the determination of affinity (K_D) and thermodynamic binding profiles of a too-strong (very high C) or too-weak (very low C) binding ligand using a reference ligand that can be used in displacement experiments (Figure S3).

- **Ligand descriptor calculation tool**, featuring a ΔG / K_D calculator, a structure drawing panel that allows calculating the logarithm of the octanol-water coefficient (cLogP) using the Ghose and Crippen algorithm,[54] and the ligand efficiency metrics LE,[10,11] LLE[12], LELP,[13] and LLE_{AT},[8] combining lipophilicity, size, and potency to support hit-to-lead optimization decision-making (Figure S4).[55]

RESULTS AND DISCUSSION

General

A summary of all obtained data, including errors, can be found in Tables S3, S4. For clarity, all other tables omit errors. It became apparent that the stoichiometries (N) were below 1 for all titrations against trypsin, even though a 1:1 binding stoichiometry is expected. Especially for titrations with benzamidine (BA), N was 0.3–0.46. However, according to the literature, even trypsin purchased in the highest purity is only to ~70% binding competent.[56]

For all experiments, controls for each titration, consisting of a buffer to the analyte, ligand to buffer, and buffer to buffer titration, were performed and subtracted using the regression *line* method (exemplarily shown in Figures S5, S6).

Heat capacity and temperature dependency

The changed isobaric heat capacity upon ligand binding (ΔC_P) is part of a fully characterized thermodynamic binding profile. It can be obtained by a series of ITC experiments performed at varied temperatures. When ΔH is plotted against the temperature, the slope of the regression curve gives ΔC_P. ITC experiments were performed in HEPES buffer at three temperatures and calculated ΔC_P, T_S, and T_H from the ΔH_{obs} using *ITCcalc* (Figures 3, S1, Table 1).

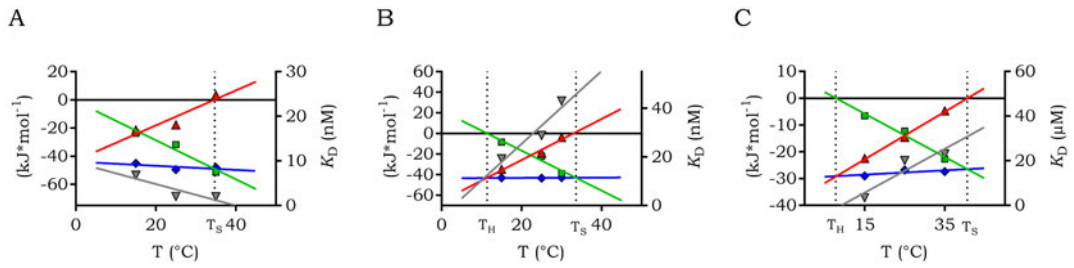


Figure 3. Temperature dependency plots of trypsin binding of aprotinin (**A**), leupeptin (**B**), and BA (**C**) at three different temperatures. The slope of the ΔH regression (green) gives the heat capacity change (ΔC_P). T_H and T_S are the temperatures at the intercept of the ΔH and the $-T\Delta S$ (red) regression lines with the ordinate. ΔG is depicted in blue, and the K_D is given in gray. The respective plots for the EDTA-based system are presented in Figure S7.

Table 1. ITC experiments at different temperatures to calculate the heat capacity of trypsin binding.

Ligand	T (°C)	ΔG (kJ*mol^{-1})	ΔH_{obs} (kJ*mol^{-1})	$-T\Delta S_{obs}$ (kJ*mol^{-1})	ΔC_P (kJ*mol^{-1}*K^{-1})	T_S (°C)	T_H (°C)
Aprotinin[a]	15	n.d.	–23.7	n.d.	–1.39	n.d.	–0.7
	25	n.d.	–31.8	n.d.			
	35	n.d.	–51.4	n.d.			
Leupeptin	15	–43.2	–8.4	–34.7	–1.94	33.6	11.4
	25	–43.4	–22.1	–18.9			
	30	–42.9	–39.0	–3.86			
BA	15	–29.0	–6.6	–22.5	–0.81	40.6	7.80
	25	–26.8	–12.3	–14.6			
	35	–27.4	–22.7	–4.7			

Experiments were performed at least in triplicates. For better readability, errors are given in Tables S3, S4. The corresponding evaluation for the EDTA-based system a presented in Table S5. a: the C-value is too high to determine K_D (ΔG and $-T\Delta S_{obs}$), n.d. not determined due to high-C titration.

This temperature dependency of ΔH can be utilized to optimize the setup for measuring the thermodynamic binding profile of leupeptin *vs*. trypsin. While at 15 °C a low binding enthalpy of –8.4 kJ*mol^{-1} was observed, higher ΔH values could be obtained at 25 °C and 30 °C (–22.1 and –39.0 kJ*mol^{-1}, respectively), which are less close to the T_H of 11.4 °C, the temperature, where no enthalpic signal could be measured (Table 1, Figure 4). Increasing ΔH_{obs} is usually accompanied by an improved signal-to-noise ratio. For all studied trypsin ligands, ΔH_{obs} further decreases with higher temperatures, whereas $-T\Delta S_{obs}$ increases, compensating each other to result in a largely temperature-independent ΔG. The negative ΔC_P values of –1.39, –1.94, and –0.81 kJ*mol^{-1}*K^{-1} for aprotinin, leupeptin, and BA indicate the burial of hydrophobic sites upon ligand binding.[27,29] It can be assumed that the highest ΔC_P is observed for BA due to its small molecular size, resulting in a relatively small binding interface.

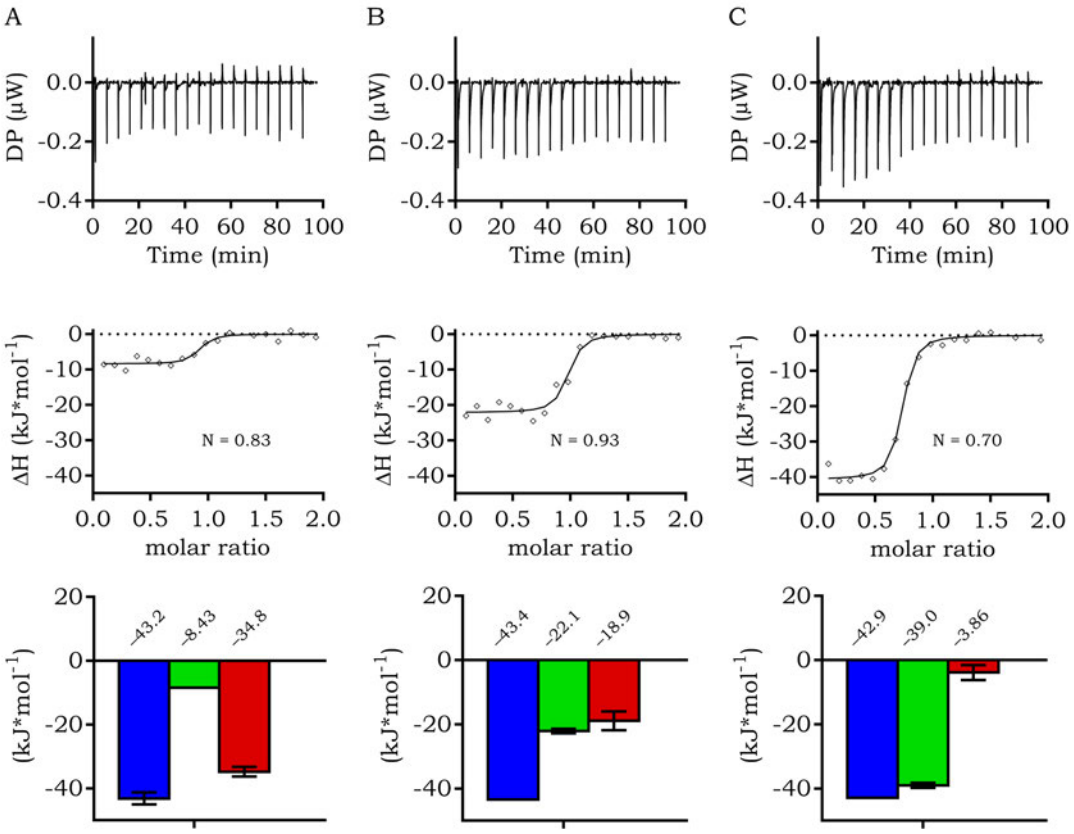


Figure 4. Exemplary curves for the signal improvement for direct leupeptin vs. trypsin titrations by adjusting the temperature. The thermograms (top) show the thermograms. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line. The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG (blue). Measurements at 15 °C (**A**), 25 °C (**B**), and 30 °C (**C**) were performed.

Buffer Ionization Correction

For BA and leupeptin binding to trypsin, direct ITC titrations were performed in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), imidazole, and tris(hydroxymethyl)aminomethane (TRIS) buffered solution at pH 7.4 (Table S6). Buffer pK_a values are within 7.4 ± 1.0 and span ΔH_{ion} from 20.4 to 47.5 kJ*mol^{-1}.[57] The buffer-corrected signature plots were calculated using *ITCcalc*. Therefore, the mean ΔG of the titrations in different buffers and the obtained ΔH_{obs} values for the respective buffers were used for calculation (Figure S2). When two or more buffers are added, *ITCcalc* calculates n_{proton}, the corrected ΔH (ΔH°), and the resulting $-T\Delta S°$. Positive values for n_{proton} represent accepted protons, and negative values are due to released protons upon binding. Since aprotinin displays a K_D below the

affinity limit for direct titrations,[58] the calculation for ΔG° and $-T\Delta S_{obs}$ are unreliable. Nevertheless, ΔH_{obs} can be obtained by the released heat from the first injections. Therefore, the proton transfer can be calculated as well. We observed no proton transitions for leupeptin ($n_{proton} = -0.05$) Figures S8, S9), the release of 0.48 protons per formed aprotinin-trypsin complex, and 0.33 protons per BA-trypsin complex, respectively (Figure 5).

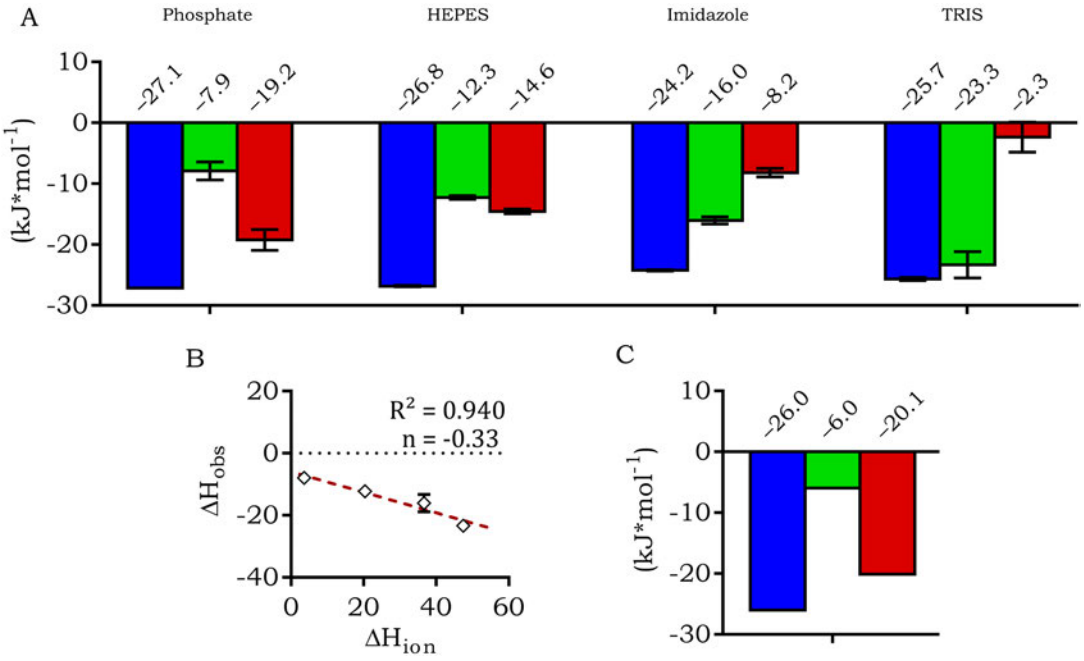


Figure 5: Buffer correction exemplarily performed for direct BA vs. trypsin titrations. **A** Thermodynamic signature plots of ITC experiments performed in phosphate, HEPES, imidazole, and TRIS buffered solution, displaying the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **B** Regression of the ΔH_{obs} vs. ΔH_{ion} of the respective buffer gives the corrected binding enthalpy ΔH° as the intercept with the ordinate. **C** Signature plot of the BA-trypsin binding after buffer ionization correction.

The result differs from previously published data for BA, where no proton transfer was described. However, these experiments were performed at a different pH of 8.0.[59] Since BA has a pk_a value of 11.8,[60] it is predominantly (>99%) in the protonated state at both pH values. Thus, this effect is probably due to altered protonation states of trypsin. For the EDTA-based system, we found 1 proton to be released upon Ca^{2+} ($n_{proton} = -0.96$) and Mg^{2+} binding ($n_{proton} = -0.95$), which is following the literature (Table S7, Figures S10–S12).[58]

Table 2. Buffer ionization correction for trypsin binding ligands.

Ligand	buffer	ΔG° (kJ*mol⁻¹)	ΔH_{obs} (kJ*mol⁻¹)	$-T\Delta S_{obs}$ (kJ*mol⁻¹)	ΔH° (kJ*mol⁻¹)	−TΔS° (kJ*mol⁻¹)	n_{proton}
Aprotinin	HEPES	n.d.	−31.8	n.d.	−21.4	n.d.	−0.48
	Imidazole		−37.0	n.d.			
	TRIS		−45.1	n.d.			
Leupeptin	HEPES	−43.1	−22.1	−18.9	−21.0	−14.3	−0.05
	Imidazole		−22.4	−17.8			
	TRIS		−23.5	−21.2			
BA	Phosphate	−26.0	−7.9	−19.2	−5.95	−20.0	−0.33
	HEPES		−12.3	−14.6			
	Imidazole		−16.0	−8.2			
	TRIS		−23.3	−2.3			

Experiments were performed at least in triplicates. For better readability, errors are not shown but can be viewed in Tables S3, S4. The corresponding evaluations for the EDTA-based system are presented in Table S7. n.d. not determined due to high-C titration.

Displacement titrations

BA is a known ligand for trypsin with a K_D of 21.7 µM.[59] To achieve an optimal C-value of 10–100,[19,39] the trypsin concentration should be above 200 µM according to Equation 9. This experimental setup would significantly increase the sample consumption and the likelihood of unspecific interactions.[61] Additionally, protein- and ligand solubility could limit the experimental leeway. A low-C experiment was performed to titrate BA vs. trypsin directly. Herefore, BA at 2.5 mM was titrated to trypsin at 100 µM to reach a stoichiometry of 5 and sufficient target saturation (Figure 6). We found BA to bind to trypsin with a K_D of 20.1 µM, which agrees with the previously published data (Table S2).[59] Alternatively, a displacement titration with a more potent ligand like leupeptin can be performed. Therefore, a solution of 250 µM BA with 10 µM trypsin was titrated with 100 µM leupeptin (Figure 6).

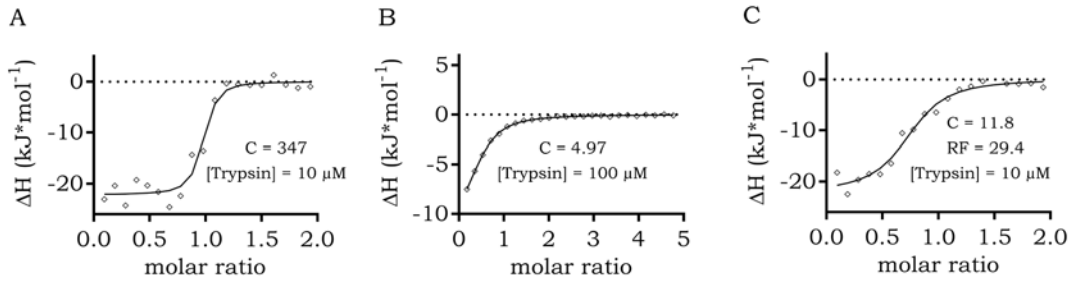


Figure 6: Exemplary thermograms for direct and displacement titrations. **A** Direct titration of the leupeptin-trypsin binding. The C-value of 347 is not in the optimal range of 10–100 and is too high to obtain a reliable K_D. **B** Direct low-C

titration of the BA-trypsin binding. The sample consumption is high, with 100 µM trypsin in the reaction cell and 2.5 mM BA in the syringe. **C** Displacement titration with 10 µM trypsin and 250 µM BA in the reaction cell. The C-value of 25 is in the desired range, and the sample consumption is reduced compared to **B**.

K_D values from direct titrations are generally accessible for leupeptin and BA. Hence, the displacement titrations can be used to calculate the thermodynamics of both BA as an unknown low-affinity binder and leupeptin as an unknown high-affinity binder using the obtained values from direct titrations with the respective other ligand (Equations 11, 12, Table 3). Based on the K_D and the ΔH values of the known ligand, the concentration of the preincubated, weaker ligand, and the parameters K_{Dapp} and ΔH_{app}, determined by the displacement titration, the desired binding parameters of the unknown ligand can be calculated using *ITCcalc* (Figure S3). This procedure was also applied for the displacement of BA with aprotinin (Table 3, Figure S13). Likewise, for the EDTA system, Ca^{2+} can be used as the high-affinity ligand and Mg^{2+} as the low-affinity ligand. The displacement experiment is accessible for both metal-ions as the unknown 'too high affinity' or 'too low affinity' ligand, respectively (Table S8, Figure S14).

Table 3. Obtained thermodynamic parameters from direct and displacement titrations.

Ligand	Titration protocol	K_{Dapp}	K_D	ΔG (kJ*mol^{-1})	ΔH (kJ*mol^{-1})	−TΔS (kJ*mol^{-1})
Leupeptin	Direct	28.8 nM		−43.4	−22.1	−18.9
	Displacement of 250 µM BA	354 nM	26.3 nM	−43.2	−29.0	−14.2
BA	Direct	20.1 µM		−26.8	−12.3	−14.6
	Displacement with leupeptin	354 nM	22.1 µM	−26.6	−6.9	−19.6
Aprotinin	Direct	n.d.		n.d.	−31.8	n.d.
	Displacement of 2500 µM BA	79.7 nM	642 pM	−52.4	−31.4	−21.1

Experiments were performed at least in triplicates. For better readability, errors are only given in Tables S3, S4. The corresponding evaluation for the EDTA-based system is presented in Table S8. n.d. not determined due to high-C titration.

The comparison of direct titrations with displacement experiments reveals good accordance in terms of affinity. The K_D values of leupeptin fit each other and the literature values[62] (direct: 28.8 nM; displ.: 26.3 nM; lit.: 40.4 nM — Tables 3, S2). The C-value is reduced from 347 to 11.8, yielding a reduction factor (RF) of 29.4. The K_D of BA, determined by the same displacement titration, is in good accordance with the literature[59] (direct: 20.0 µM; displ.: 22.1 µM; lit.: 21.7 µM — Tables 3, S2). In the case of aprotinin, the affinity could not be accurately determined in a direct titration experiment since the slope of the binding isotherm displays a steep

function. With the displacement of 2500 µM BA, the C-value was reduced to 251 (RF = 319, Figure S13), thus allowing the extraction of the K_{Dapp} of 79.7 nM. Calculations with *ITCcalc* revealed the K_D of aprotinin of 642 pM, which is still significantly higher than the literature-known value of 60 fM. This value, however, was obtained by an alternative technique, displacement equilibrium measurements, under different conditions previously (Table S8).[63]

CONCLUSION

With this work, we want to create awareness that ITC offers more experimental flexibility in binding characterization than direct titration experiments. Also, we want to demonstrate how to solve ITC-related tasks in medicinal chemistry. The desired learning achievements of these experimental setups are:

- The improvement of heat signals by exploiting different buffer ionization enthalpies (e.g., for the trypsin vs. BA titrations) and by adjusting the experimental temperature (e.g., for the trypsin vs. leupeptin titrations).
- The optimization of C-values for too-strong (leupeptin C ≈ 275) and too-low binding (BA C ≈ 5) ligands by using displacement titrations (resulting C ≈ 25).
- The correction of the thermodynamics for unspecific events, dilution enthalpies, and proton transfers by performing a variety of control measurements.
- The correction of proton transfers using experiments in several buffers exhibiting differing ionization enthalpies.

Aiming at medicinal chemists as advanced ITC users, we performed the experiments in the well-studied trypsin system.[31,59,64,65] Nevertheless, we also present low-budget teaching experiments in an EDTA-based system. For easy data processing, we established the free accessible ITC-web calculator *ITCcalc* (https://itccalc.uni-mainz.de/), providing ITC users with a convincingly simple and intuitive toolkit, assisting with underlying calculations, and giving advice for future experimental design. *ITCcalc* can also be used to assess ligands in drug development campaigns, as it calculates several standard ligand metrics, such as clogP, LE, LLE, LELP, and LLE_{AT}. The ΔG / K_D calculator can help undergraduate students to get a better intuitive understanding of the logarithmic relationship between affinity and binding energy (Equation 1, Figure S4). Thereby, the differences in ΔG between two compounds can be translated to improvements in K_D or the contributions of single modifications to the overall binding affinity. To our knowledge, we are the first to present the complete thermodynamic binding profile of the leupeptin-trypsin and the aprotinin-trypsin binding.

OUTLOOK

Besides the experiments and evaluation presented in this manuscript, it is possible to characterize the kinetic parameters (k_{on}/k_{off}) of binding by analyzing varying equilibration times near the infection point (a proprietary tool to derive binding kinetics from ITC experiments is KinITC by Affinimeter, https://www.affinimeter.com/).[66] However, analysis of binding kinetics is beyond the scope of *ITCcalc,* and this work focuses on thermodynamics. In some exceptional cases, also not covered by this work, ITC can identify different affinities of enantiomers in a single set of experiments without preceding racemate separation. For this simultaneous determination, the affinity of both enantiomers, however, needs to differ by a factor of 50–200.[67]

ABBREVIATIONS

ITC: isothermal titration calorimeter, N: binding stoichiometry, K_D: dissociation constant, ΔG: Gibbs energy, ΔH: enthalpy, ΔS: entropy, LE: ligand efficiency, LLE: ligand lipophilic efficiency, LLE_{AT}: LLE adjusted for heavy atom count, LELP: ligand efficiency dependent lipophilicity, ΔΔH: changes in ΔH at different experimental temperatures, ΔC_P: isobaric binding heat capacity, ΔH_{obs}: observed enthalpy, ΔH°: binding enthalpy, ΔH_{ion}: ionization enthalpy, n_{proton}: number of transferred protons, M_{tot}: total concentration of the analyte, observed affinity (K_{Dapp}), FBDD: fragment-based drug discovery, EDTA: ethylenediaminetetraacetic acid, cLogP: Calculated logarithmic octanol-water coefficient, BA: benzamidine, HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, TRIS: tris(hydroxymethyl)aminomethane, RF: reduction factor,

ASSOCIATED CONTENT

Supporting Information

- Supplementary Figures
    - Screenshots of evaluations performed with ITCcalc
    - Evaluations not shown in this manuscript.
- Supplementary Tables
- Experimental Section
    - Reagents
    - Implementation of ITCcalc
    - Thermograms and ITC experiments
- Hazards
- Additional References
- A guided tour to learn how to use ITCcalc and all its functions can be accessed at https://itccalc.uni-mainz.de.

AUTHOR INFORMATION

Corresponding Author

*E-mail: [redacted]

Acknowledgments

The authors thank [redacted] for access to facilities and instrumentation, as well as fruitful scientific discussion on the project. This work was supported by the German Research Council (DFG) through the grant "DFG-Großgeräteantrag ITC (INST 247/921-1 FUGG)". [redacted] received funding for digital teaching from the "Stiftung Innovation in der Hochschullehre" via Johannes Gutenberg University in the framework of "Mainzer Modelle für digital erweitertes Lehren und Lernen (ModeLL-M)".

REFERENCES

(1) Wei, C. C.; Jensen, D.; Boyle, T.; O'Brien, L. C.; De Meo, C.; Shabestary, N.; Eder, D. J. Isothermal Titration Calorimetry and Macromolecular Visualization for the Interaction of Lysozyme and Its Inhibitors. *J. Chem. Educ.* **2015**, *92* (9), 1552–1556. https://doi.org/10.1021/ed5002569.

(2) Perozzo, R.; Folkers, G.; Scapozza, L. Thermodynamics of Protein-Ligand Interactions: History, Presence, and Future Aspects. *J. Recept. Signal Transduct.* **2004**, *24* (1–2), 1–52. https://doi.org/10.1081/RRS-120037896.

(3) Gibbs, J. W. A Method of Geometrical Representation of the Thermodynamic Properties of Substances By Means of Surfaces. *Trans. Connect. Acad.* **1873**, *II*, 382–404.

(4) Geschwindner, S.; Ulander, J.; Johansson, P. Ligand Binding Thermodynamics in Drug Discovery: Still a Hot Tip? *J. Med. Chem.* **2015**, *58* (16), 6321–6335. https://doi.org/10.1021/jm501511f.

(5) Ladbury, J. E. Calorimetry as a Tool for Understanding Biomolecular Interactions and an Aid to Drug Design. *Biochem. Soc. Trans.* **2010**, *38* (4), 888–893. https://doi.org/10.1042/BST0380888.

(6) Muzammil, S.; Armstrong, A. A.; Kang, L. W.; Jakalian, A.; Bonneau, P. R.; Schmelmer, V.; Amzel, L. M.; Freire, E. Unique Thermodynamic Response of Tipranavir to Human Immunodeficiency Virus Type 1 Protease Drug Resistance Mutations. *J. Virol.* **2007**, *81* (10), 5144–5154. https://doi.org/10.1128/jvi.02706-06.

(7) Carbonell, T.; Freire, E. Binding Thermodynamics of Statins to HMG-CoA Reductase. *Biochemistry* **2005**, *44* (35), 11741–11748. https://doi.org/10.1021/bi050905v.

(8) Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The Role of Ligand Efficiency Metrics in Drug Discovery. *Nat. Rev. Drug Discov.* **2014**, *13* (2), 105–121. https://doi.org/10.1038/nrd4163.

(9) Raynolds, C. H. Ligand Efficiency Metrics: Why All the Fuss? *Futur. Med. Chem.* **2015**, *7* (11), 1363–1365. https://doi.org/10.4155/FMC.15.70.

(10) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The Maximal Affinity of Ligands. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (18), 9997–10002. https://doi.org/10.1073/pnas.96.18.9997.

(11) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discov. Today* **2004**, *9* (10), 430–431. https://doi.org/10.1016/S1359-6446(04)03069-7.

(12) Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* **2007**, *6* (11), 881–890. https://doi.org/10.1038/nrd2445.

(13) Keserü, G. M.; Makara, G. M. The Influence of Lead Discovery Strategies on the Properties of Drug Candidates. *Nat. Rev. Drug Discov.* **2009**, *8* (3), 203–212. https://doi.org/10.1038/nrd2796.

(14) Hann, M. M. Molecular Obesity, Potency and Other Addictions in Drug Discovery. *Medchemcomm* **2011**, *2* (5), 349–355. https://doi.org/10.1039/c1md00017a.

(15) Murray, C. W.; Erlanson, D. A.; Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H.; Richmond, N. J. Validity of Ligand Efficiency Metrics. *ACS Med. Chem. Lett.* **2014**, *5* (6), 616–618. https://doi.org/10.1021/ml500146d.

(16) Kenny, P. W.; Leitão, A.; Montanari, C. A. Ligand Efficiency Metrics Considered Harmful. *J. Comput. Aided. Mol. Des.* **2014**, *28* (7), 699–710. https://doi.org/10.1007/s10822-014-9757-8.

(17) Shultz, M. D. Improving the Plausibility of Success with Inefficient Metrics. *ACS Med. Chem. Lett.* **2014**, *5* (1), 2–5. https://doi.org/10.1021/ml4004638.

(18) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *46* (1997), 3–25.

(19) Zhang, Y.-L.; Zhang, Z.-Y. Low-Affinity Binding Determined by Titration Calorimetry Using a High- Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B, 1998, Vol. 261. https://doi.org/10.1006/abio.1998.2738.

(20) Moore, D. E.; Goode, D. R.; Seney, C. S.; Boatwright, J. M. Isothermal Titration Calorimetry Can Provide Critical Thinking Opportunities. *J. Chem. Educ.* **2016**, *93* (2), 304–310. https://doi.org/10.1021/acs.jchemed.5b00575.

(21) Cooper, A.; Johnson, C. M.; Lakey, J. H.; Nöllmann, M. Heat Does Not Come in Different Colours: Entropy-Enthalpy Compensation, Free Energy Windows, Quantum Confinement, Pressure Perturbation Calorimetry, Solvation and the Multiple Causes of Heat Capacity Effects in Biomolecular Interactions. *Biophys. Chem.* **2001**, *93* (2–3), 215–230. https://doi.org/10.1016/S0301-4622(01)00222-8.

(22) Cameron, D. L.; Jakus, J.; Pauleta, S. R.; Pettigrew, G. W.; Cooper, A. Pressure Perturbation Calorimetry and the Thermodynamics of Noncovalent Interactions in Water: Comparison of Protein-Protein, Protein-Ligand, and Cyclodextrin-Adamantane Complexes. *J. Phys. Chem. B* **2010**, *114* (49), 16228–16235. https://doi.org/10.1021/jp107110t.

(23) Cooper, A. Protein Heat Capacity: An Anomaly That Maybe Never Was. *J. Phys. Chem. Lett.* **2010**, *1* (22), 3298–3304. https://doi.org/10.1021/jz1012142.

(24) Sturtevant, J. M. Heat Capacity and Entropy Changes in Processes Involving Proteins. *Proc. Natl. Acad. Sci.* **1977**, *74* (6), 2236–2240. https://doi.org/10.1073/pnas.74.6.2236.

(25) Benfield, A. P.; Teresk, M. G.; Plake, H. R.; DeLorbe, J. E.; Millspaugh, L. E.; Martin, S. F. Ligand Preorganization May Be Accompanied by Entropic Penalties in Protein-Ligand Interactions. *Angew. Chemie - Int. Ed.* **2006**, *45* (41), 6830–6835. https://doi.org/10.1002/anie.200600844.

(26) Velazquez-Campoy, A.; Markova, N. Isothermal Titration Calorimetry: Theory and Practice. *Malvern Panalytical Knowl. Cent.* **2015**, No. 3, 1–19.

(27) Cooper, A. Heat Capacity Effects in Protein Folding and Ligand Binding: A Re-Evaluation of the Role of Water in Biomolecular Thermodynamics. *Biophys. Chem.* **2005**, *115* (2-3 SPEC. ISS.), 89–97. https://doi.org/10.1016/j.bpc.2004.12.011.

(28) Ward, W. H. J.; Holdgate, G. A. Isothermal Titration Calorimetry in Drug Discovery. In *Progress in Medicinal Chemistry*; 2001; Vol. 38, pp 309–376. https://doi.org/10.1016/S0079-6468(08)70097-3.

(29) Vega, S.; Abian, O.; Velazquez-Campoy, A. On the Link between Conformational Changes, Ligand Binding and Heat Capacity. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (5), 868–878. https://doi.org/10.1016/j.bbagen.2015.10.010.

(30) Hammerschmidt, S. J.; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B / NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm. (Weinheim).* **2022**, e2200518. https://doi.org/10.1002/ardp.202200518.

(31) Talhout, R.; Villa, A.; Mark, A. E.; Engberts, J. B. F. N. Understanding Binding Affinity: A Combined Isothermal Titration Calorimetry/Molecular Dynamics Study of the Binding of a Series of Hydrophobically Modified Benzamidinium Chloride Inhibitors to Trypsin. *J. Am. Chem. Soc.* **2003**, *125* (35), 10570–10579. https://doi.org/10.1021/ja034676g.

(32) Holdgate, G. A.; Tunnicliffe, A.; Ward, W. H. J.; Weston, S. A.; Rosenbrock, G.; Barth, P. T.; Taylor, I. W. F.; Pauptit, R. A.; Timms, D. The Entropic Penalty of Ordered Water Accounts for Weaker Binding of the Antibiotic Novobiocin to a Resistant Mutant of DNA Gyrase: A Thermodynamic and Crystallographic Study. *Biochemistry* **1997**, *36* (32), 9663–9673. https://doi.org/10.1021/bi970294+.

(33) Dunitz, J. D. Win Some, Lose Some: Enthalpy-Entropy Compensation in Weak Intermolecular Interactions. *Chem. Biol.* **1995**, *2* (11), 709–712. https://doi.org/10.1016/1074-5521(95)90097-7.

(34) Sigurskjold, B. W. Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry. *Anal. Biochem.* **2000**, *277* (2), 260–266. https://doi.org/10.1006/abio.1999.4402.

(35) Neeb, M.; Czodrowski, P.; Heine, A.; Barandun, L. J.; Hohn, C.; Diederich, F.; Klebe, G. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis, and p K a Calculations Trace the Locus of Charge in Ligand Binding to a TRNA-Binding Enzyme. *J. Med. Chem.* **2014**, *57* (13), 5554–5565. https://doi.org/10.1021/jm500401x.

(36) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L. N. Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.* **1989**, *179* (1), 131–137. https://doi.org/10.1016/0003-2697(89)90213-3.

(37) Thomson, J. A.; Ladbury, J. E. Part II Isothermal Titration a Tutorial. *Biocalorimetry 2 Appl. Calorim. Biol. Sci.* **2004**.

(38) Broecker, J.; Vargas, C.; Keller, S. Revisiting the Optimal c Value for Isothermal Titration Calorimetry. *Anal. Biochem.* **2011**, *418* (2), 307–309. https://doi.org/10.1016/j.ab.2011.07.027.

(39) Malvern, I. Man0573: Microcal Peaq-Itc User Manual. **2015**, 124.

(40) Rühmann, E.; Betz, M.; Fricke, M.; Heine, A.; Schäfer, M.; Klebe, G. Thermodynamic Signatures of Fragment Binding: Validation of Direct versus Displacement ITC Titrations. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850* (4), 647–656. https://doi.org/10.1016/j.bbagen.2014.12.007.

(41) Turnbull, W. B.; Daranas, A. H. On the value of c: Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Soc. Inf. Sci.* **2003**, *125* (48), 14859–14866. https://doi.org/10.1002/asi.4630310417.

(42) Ladbury, J. E.; Chowdhry, B. Z. Sensing the Heat: The Application of Isothermal Titration Calorimetry to Thermodynamic Studies of Biomolecular Interactions. *Chem. Biol.* **1996**, *3* (10), 791–801. https://doi.org/10.1016/S1074-5521(96)90063-0.

(43) Velazquez-Campoy, A.; Freire, E. Isothermal Titration Calorimetry to Determine Association Constants for High-Affinity Ligands. *Nat. Protoc.* **2006**, *1* (1), 186–191. https://doi.org/10.1038/nprot.2006.28.

(44) Rühmann, E.; Betz, M.; Heine, A.; Klebe, G. Fragment Binding Can Be Either More Enthalpy-Driven or Entropy-Driven: Crystal Structures and Residual Hydration Patterns Suggest Why. *J. Med. Chem.* **2015**, *58* (17), 6960–6971. https://doi.org/10.1021/acs.jmedchem.5b00812.

(45) Taylor, J. D.; Gilbert, P. J.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. Identification of Novel Fragment Compounds Targeted Against the PY Pocket of V-Src SH2 by Computational and NMR Screening and Thermodynamic Evaluation. *Proteins Struct. Funct. Genet.* **2007**, *67*, 981–990. https://doi.org/10.1002/prot.

(46) Thompsett, A. R.; Abdelraheim, S. R.; Daniels, M.; Brown, D. R. High Affinity Binding between Copper and Full-Length Prion Protein Identified by Two Different Techniques. *J. Biol. Chem.* **2005**, *280* (52), 42750–42758. https://doi.org/10.1074/jbc.M506521200.

(47) Zhang, Q.; Liu, J.; Nai, X.; Bao, H.; Sun, D.; Liu, M. Examination and Improvement of Undergraduate Laboratory Experiment: Thermodynamics of a Surfactant Micellization. *J. Chem. Educ.* **2020**, *97* (12), 4490–4498. https://doi.org/10.1021/acs.jchemed.0c00886.

(48) O'Brien, L. C.; Root, H. B.; Wei, C. C.; Jensen, D.; Shabestary, N.; De Meo, C.; Eder, D. J. M2+•EDTA Binding Affinities: A Modern Experiment in Thermodynamics for the Physical Chemistry Laboratory. *J. Chem. Educ.* **2015**, *92* (9), 1547–1551. https://doi.org/10.1021/acs.jchemed.5b00159.

(49) Wadsö, L.; Smith, A. L.; Shirazi, H.; Mulligan, S. R.; Hofelich, T. The Isothermal Heat Conduction Calorimeter: A Versatile Instrument for Studying Processes in Physics, Chemistry, and Biology. *J. Chem. Educ.* **2001**, *78* (8), 1080. https://doi.org/10.1021/ed078p1080.

(50) Bartle, K. D.; Osborn, P. M. A Simplified Undergraduate Calorimetry Experiment. *J. Chem. Educ.* **1973**, *50* (9), 637. https://doi.org/10.1021/ed050p637.

(51) Wadsö, L.; Li, X. A Simple Rate Law Experiment Using a Custom-Built Isothermal Heat Conduction Calorimeter. *J. Chem. Educ.* **2008**, *85* (1), 112–116. https://doi.org/10.1021/ed085p112.

(52) Wadsö, L.; Li, Y. Isothermal Titration Calorimetry in the Student Laboratory. *J. Chem. Educ.* **2011**, *88* (1), 101–105. https://doi.org/10.1021/ed100649e.

(53) Zhang, Q.; Liu, J.; Bao, H.; Nai, X.; Sun, D.; Ma, B. Determining the Ternary Phase Diagram of Benzene-Acetic Acid-Water Using Isothermal Titration Microcalorimetry to Train Upper-Level Undergraduates in Advanced Calorimetry Methods. *J. Chem. Educ.* **2020**, *97* (5), 1470–1475. https://doi.org/10.1021/acs.jchemed.0c00002.

(54) Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for Three-Dimensional-Structure-Directed Quantitative Structure-Activity Relationships. 2. Modeling Dispersive and Hydrophobic Interactions. *J. Chem. Inf. Comput. Sci.* **1987**, *27* (1), 21–35. https://doi.org//10.1021/ci00053a005.

(55) Mortenson, P. N.; Murray, C. W. Assessing the Lipophilicity of Fragments and Early Hits. *J. Comput. Aided. Mol. Des.* **2011**, *25* (7), 663–667. https://doi.org/10.1007/s10822-011-9435-z.

(56) Gopal, S. M.; Klumpers, F.; Herrmann, C.; Schäfer, L. V. Solvent Effects on Ligand Binding to a Serine Protease. *Phys. Chem. Chem. Phys.* **2017**, *19* (17), 10753–10766. https://doi.org/10.1039/c6cp07899k.

(57) Goldberg, R. N.; Kishore, N.; Lennen, R. M. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data* **2002**, *31* (2), 231–370. https://doi.org/10.1063/1.1416902.

(58) Henzl, M. T.; Larson, J. D.; Agah, S. Estimation of Parvalbumin Ca2+- and Mg2+-Binding Constants by Global Least-Squares Analysis of Isothermal Titration Calorimetry Data. *Anal. Biochem.* **2003**, *319* (2), 216–233. https://doi.org/10.1016/S0003-2697(03)00288-4.

(59) Talhout, R.; Engberts, J. B. F. N. Thermodynamic Analysis of Binding of p -Substituted Benzamidines to Trypsin. *Eur. J. Biochem.* **2001**, *268* (6), 1554–1560. https://doi.org/10.1046/j.1432-1327.2001.01991.x.

(60) Albert, A.; Goldacre, R.; Phillips, J. The Strength of Heterocyclic Bases. *J. Chem. Soc.* **1948**, No. 2240, 2240–2249. https://doi.org/10.1039/JR9480002240.

(61) Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K. High-Throughput Assays for Promiscuous Inhibitors. *Nat. Chem. Biol.* **2005**, *1* (3), 146–148. https://doi.org/10.1038/nchembio718.

(62) Kuramochi, H.; Nakata, H.; Ishii, S. Mechanism of Association of a Specific Aldehyde Inhibitor, Leupeptin, with Bovine Trypsin. *J. Biochem.* **1979**, *86* (5), 1403–1410.

(63) Vincent, J. P.; Lazdunski, M. Trypsin-Pancreatic Trypsin Inhibitor Association. Dynamics of the Interaction and Role of Disulfide Bridges. *Biochemistry* **1972**, *11* (16), 2967–2977. https://doi.org/10.1021/bi00766a007.

(64) Ngo, K.; Collins-Kautz, C.; Gerstenecker, S.; Wagner, B.; Heine, A.; Klebe, G. Protein-Induced Change in Ligand Protonation during Trypsin and Thrombin Binding: Hint on Differences in Selectivity Determinants of Both Proteins? *J. Med. Chem.* **2020**, *63* (6), 3274–3289. https://doi.org/10.1021/acs.jmedchem.9b02061.

(65) Bode, W. The Structure of Thrombin, a Chameleon-like Proteinase. *J. Thromb. Haemost.* **2005**, *3*, 2379–2388.

(66) Burnouf, D.; Ennifar, E.; Guedich, S.; Puffer, B.; Hoffmann, G.; Bec, G.; Disdier, F.; Baltzinger, M.; Dumas, P. KinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J. Am. Chem. Soc.* **2012**, *134* (1), 559–565. https://doi.org/10.1021/ja209057d.

(67) Fokkens, J.; Klebe, G. Ein Einfaches Verfahren Zur Abschätzung von Affinitätsunterschieden Bei Der Bindung von Enantiomeren an Targetproteine Ohne Vorherige Trennung Der Racemate. *Angew. Chemie* **2006**, *118* (6), 1000–1004. https://doi.org/10.1002/ange.200502302.

Advanced Isothermal Titration Calorimetry for Medicinal Chemists with *ITCcalc*

Stefan J. Hammerschmidt[a,†], [redacted]

[a] *Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Staudingerweg 5, 55128 Mainz, Germany*
[b] *Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany*
† Authors contributed equally.
* E-Mail: [redacted]

CONTENT

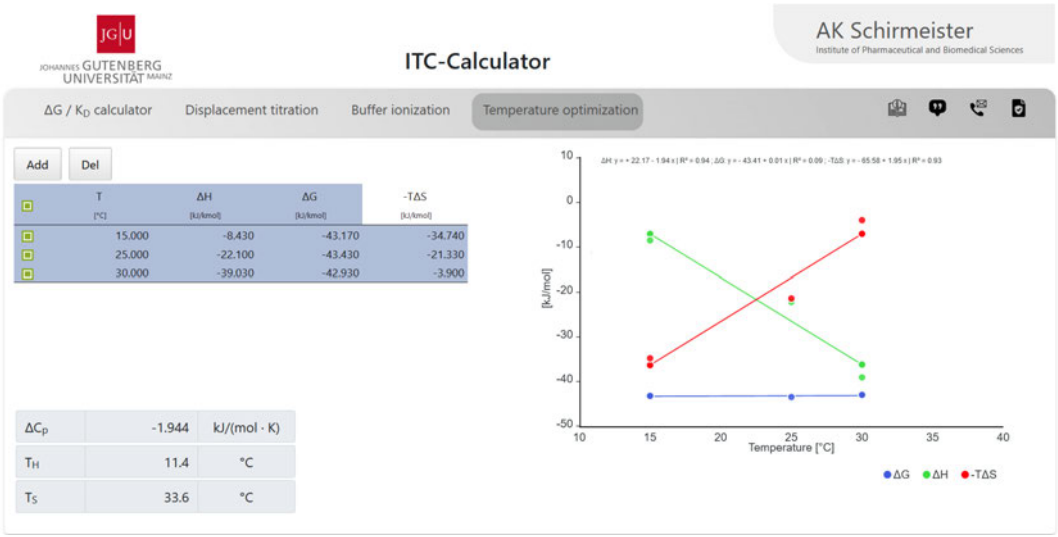


Figure S1. Screenshot of the temperature optimization tool of the *ITCcalc* webserver (https://itccalc.uni-mainz.de) shown by the example of leupeptin vs. trypsin. After entering the experimental temperature and the obtained respective thermodynamic parameters ΔH and ΔG, the tool calculates ΔC_P, T_H, and T_S.

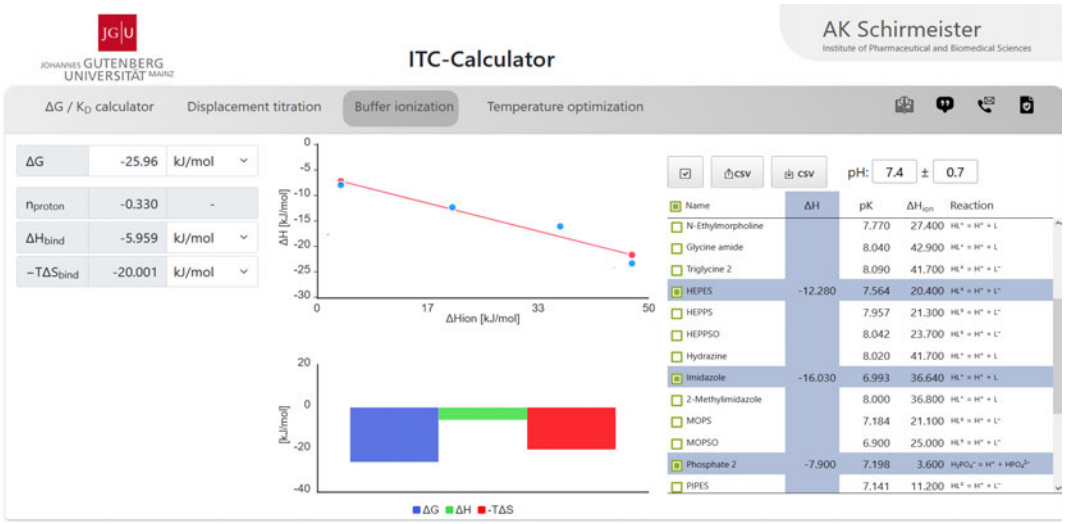


Figure S2. Screenshot of the buffer ionization tool of the *ITCcalc* webserver (https://itccalc.uni-mainz.de) shown by the example of BA vs. trypsin. After entering the mean ΔG and the ΔH_{obs} values at the boxes of the respective buffering agent, the tool calculates the number of transferred protons (n_{proton}) and gives the corrected thermodynamic profile (ΔH_{bind} and $-T\Delta S_{bind.}$).

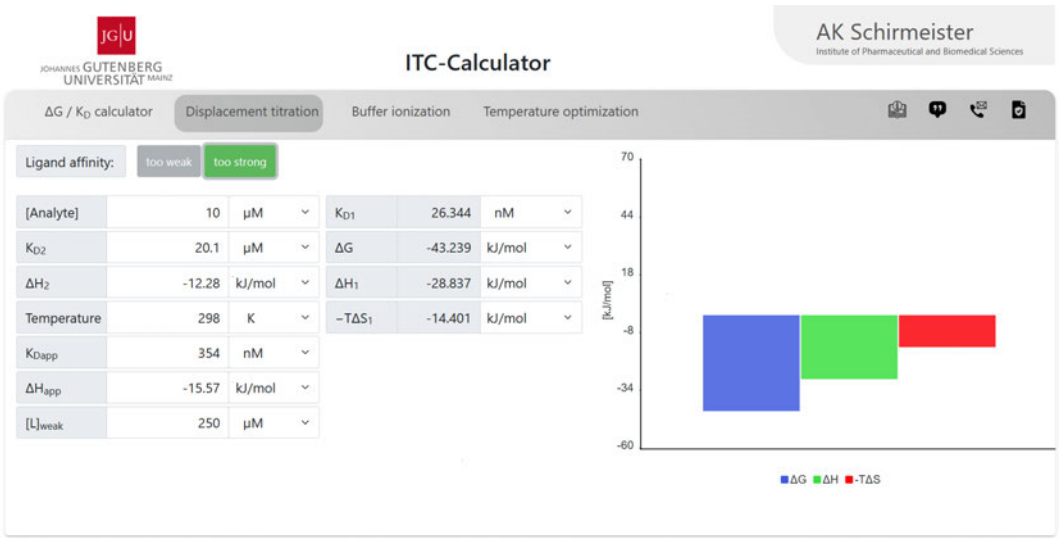


Figure S3. Screenshot of the displacement titration tool of the *ITCcalc* webserver (https://itccalc.uni-mainz.de) shown by the example of the displacement of 250 µM of BA by leupeptin. First, the user must choose if the unknown ligand is too weak or too strong for direct titrations. Several values need to be entered: The temperature, the experimental parameters analyte concentration in the reaction cell [Analyte], and the concentration of the weak ligand $[L]_{weak}$ as well as the known K_D and ΔH parameters obtained in direct titrations (K_{D2} and ΔH_2 for the weak ligand or K_{D1} and ΔH_1 for the strong ligand) and the observed parameters from the displacement experiment (K_{Dapp}, and ΔH_{app}). The displacement tool then gives the unknown ligand's calculated binding parameters (KD, ΔG, ΔH, and –TΔS).

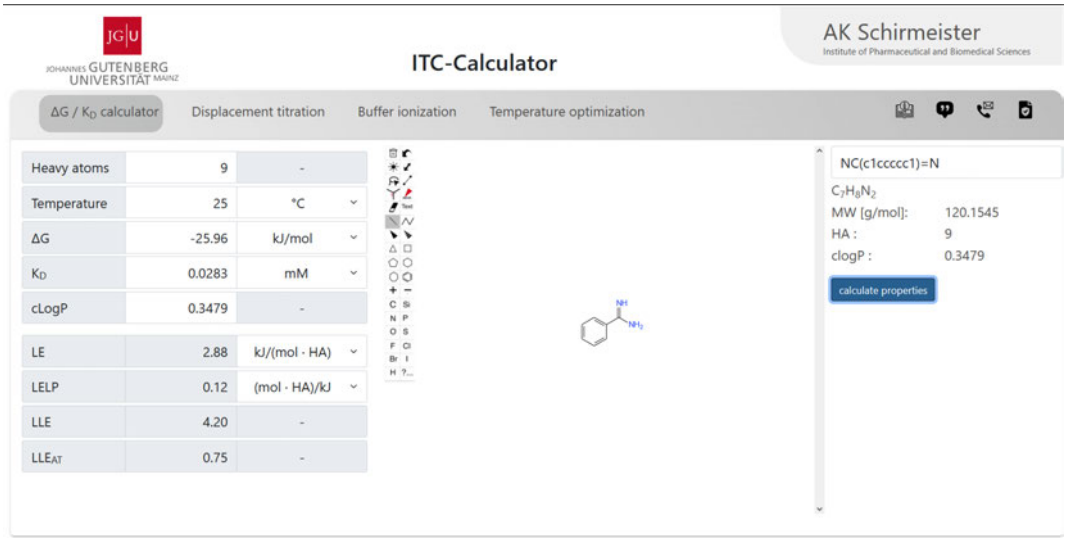


Figure S4. Screenshot of the ΔG / K_D calculator tool of the *ITCcalc* webserver (https://itccalc.uni-mainz.de) shown by the example of the binding of BA to trypsin. The structure can be drawn or entered as smiles so that the tool can compute the molecular weight (MW), the number of heavy atoms (HA), and the cLogP. Then, the ΔG or the K_D must be entered with the experimental temperature. The tool then calculates either K_D or ΔG, the LE, the LELP, the LLE, and the LLE_{AT}.

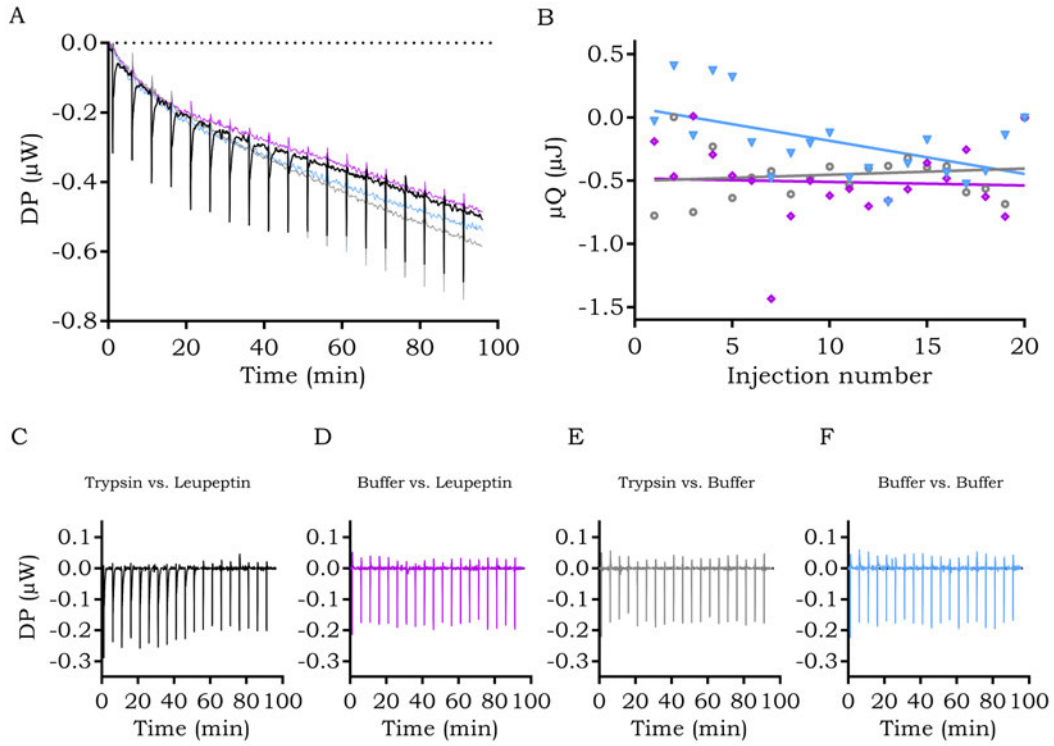
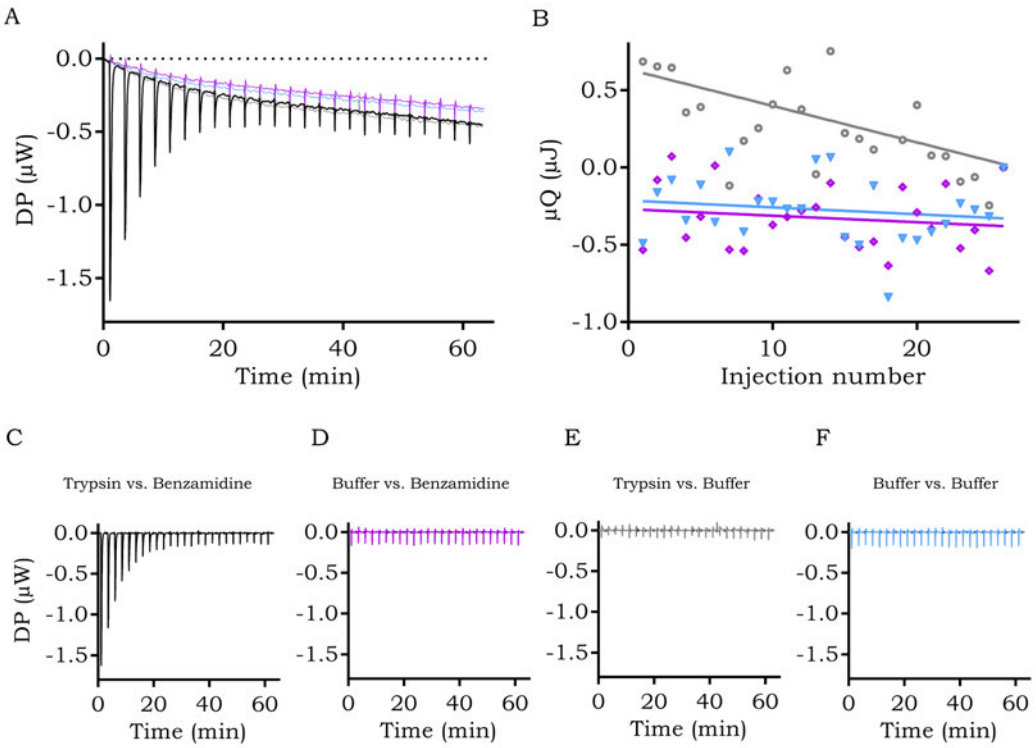


Figure S5. Performed control experiments exemplarily shown for the titration of trypsin vs. leupeptin in HEPES buffer at 25 °C. **A** Overlay of the thermograms of one trypsin vs. leupeptin titration (black) and all control experiments (buffer vs. leupeptin (pink); trypsin vs. buffer (gray); buffer vs. buffer (blue)). **B** Linear regression for correction of the control experiments. **C** Thermogram of the trypsin vs. leupeptin titration. **D** Thermogram of the buffer vs. leupeptin control experiment. **E** Thermogram of the trypsin vs. buffer control experiment. **F** Thermogram of the buffer vs. buffer control experiment.



85 Figure S6. Performed control experiments exemplarily shown for the titration of trypsin vs. BA in HEPES buffer at 25 °C. **A** Overlay of the thermograms of one trypsin vs. BA titration (black) and all control experiments (buffer vs. BA (pink); trypsin vs. buffer (gray); buffer vs. buffer (blue)). **B** Linear regression for correction of the control experiments. **C** Thermogram of the trypsin vs. BA titration. **D** Thermogram of the buffer vs. BA control experiment. **E** Thermogram of the trypsin vs. buffer control experiment. **F** Thermogram of the buffer vs. buffer control experiment.

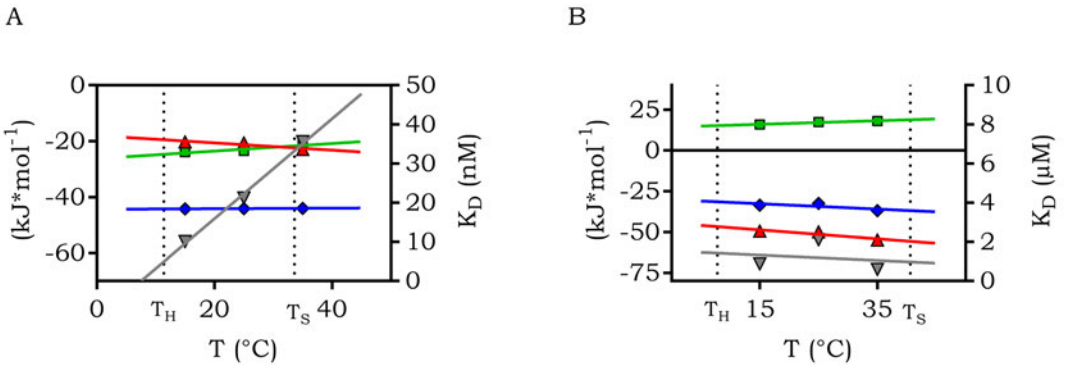
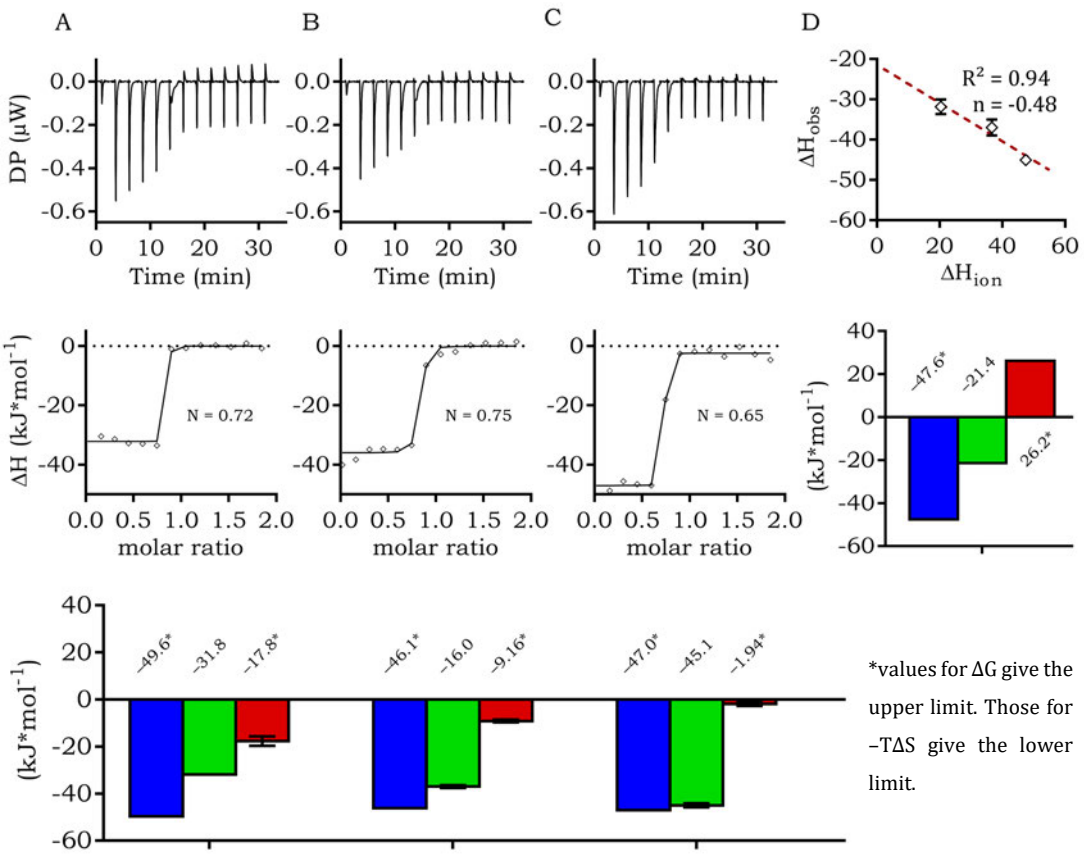


Figure S7. Temperature dependency plots of EDTA binding to Ca^{2+} (**A**) and Mg^{2+} (**B**) at three different temperatures. The slope of the ΔH regression (green) gives the heat capacity change (ΔC_P). T_H and T_S are the temperatures at the intercept of the ΔH and the –TΔS (red) regression lines with the ordinate. ΔG is depicted in blue, and the K_D is given in gray



*values for ΔG give the upper limit. Those for –TΔS give the lower limit.

Figure S8. Buffer correction exemplarily performed for direct aprotinin vs. trypsin titrations. **A**–**C** ITC experiments in HEPES (**A**), imidazole (**B**), and TRIS (**C**) buffered media. The thermograms (top) show the raw heat data. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line, calculated by the ITC software MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **D** Linear regression of ΔH_{obs} vs. ΔH_{ion} (top) and signature plot of the aprotinin-trypsin binding after buffer ionization correction (mid).

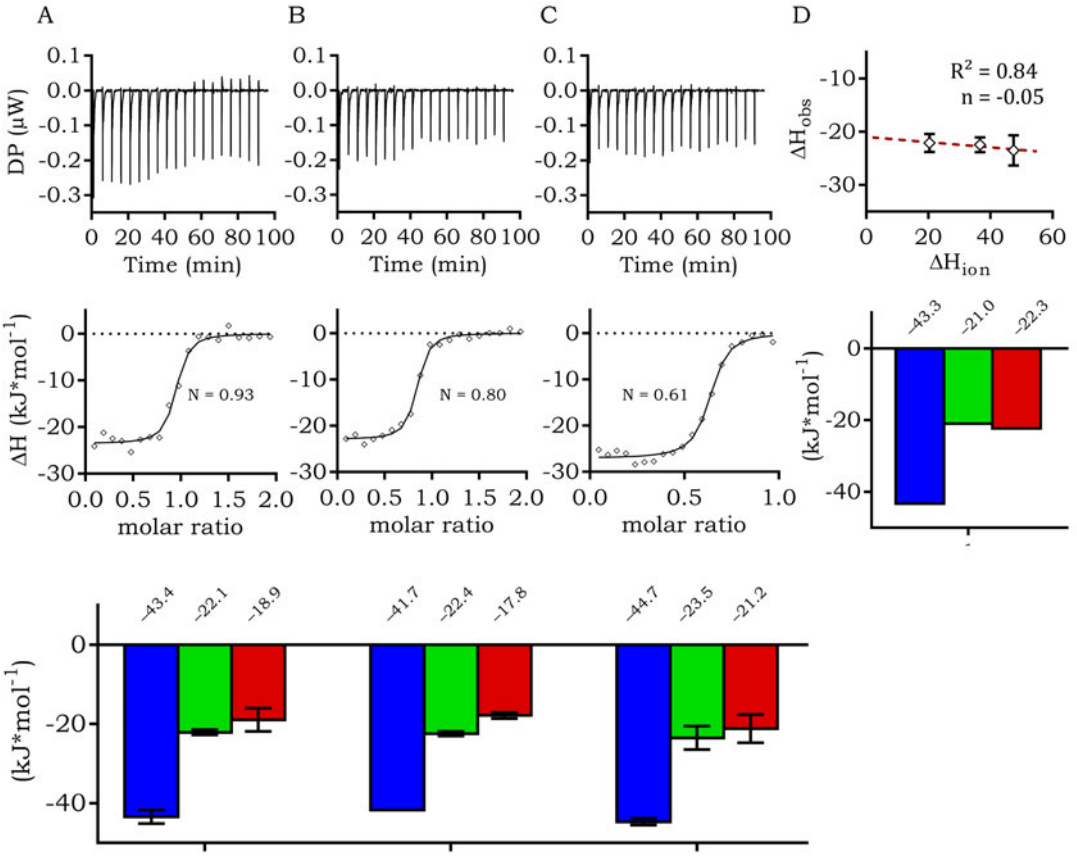


Figure S9. Buffer correction exemplarily performed for direct leupeptin vs. trypsin titrations. **A–C** ITC experiments in HEPES (**A**), imidazole (**B**), and TRIS (**C**) buffered media. The thermograms (top) show the raw heat data. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line, calculated by the ITC software MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **D** Linear regression of ΔH_{obs} vs. ΔH_{ion} (top) and signature plot of the leupeptin-trypsin binding after buffer ionization correction (mid).

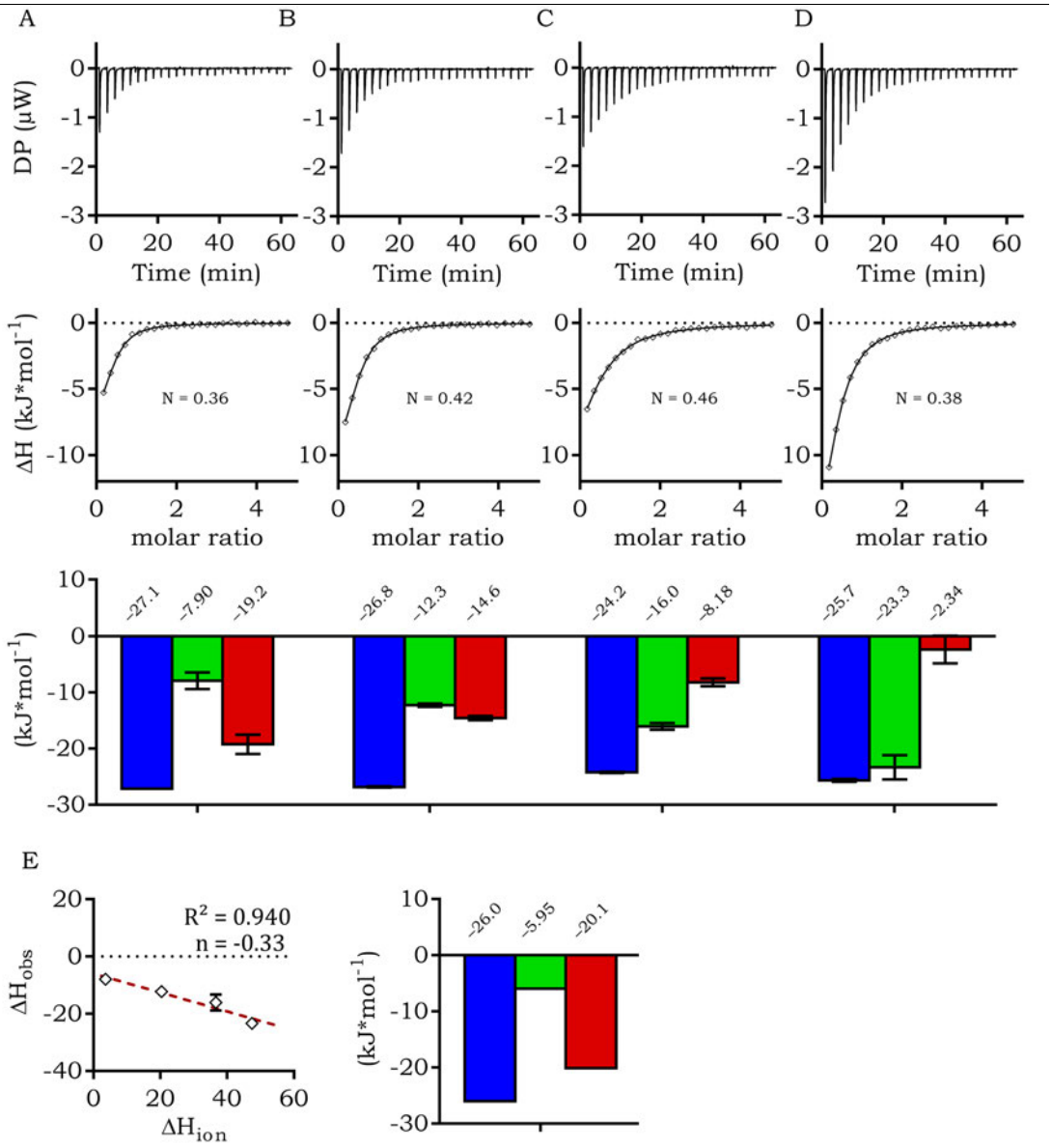


Figure S10. Buffer correction exemplarily performed for direct BA vs. trypsin titrations. **A–D** ITC experiments in phosphate (**A**) HEPES (**B**), imidazole (**C**), and TRIS (**D**) buffered solution. The thermograms (top) show the measured raw heat data. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line. The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **E** Signature plot of the BA-trypsin binding after buffer ionization correction.

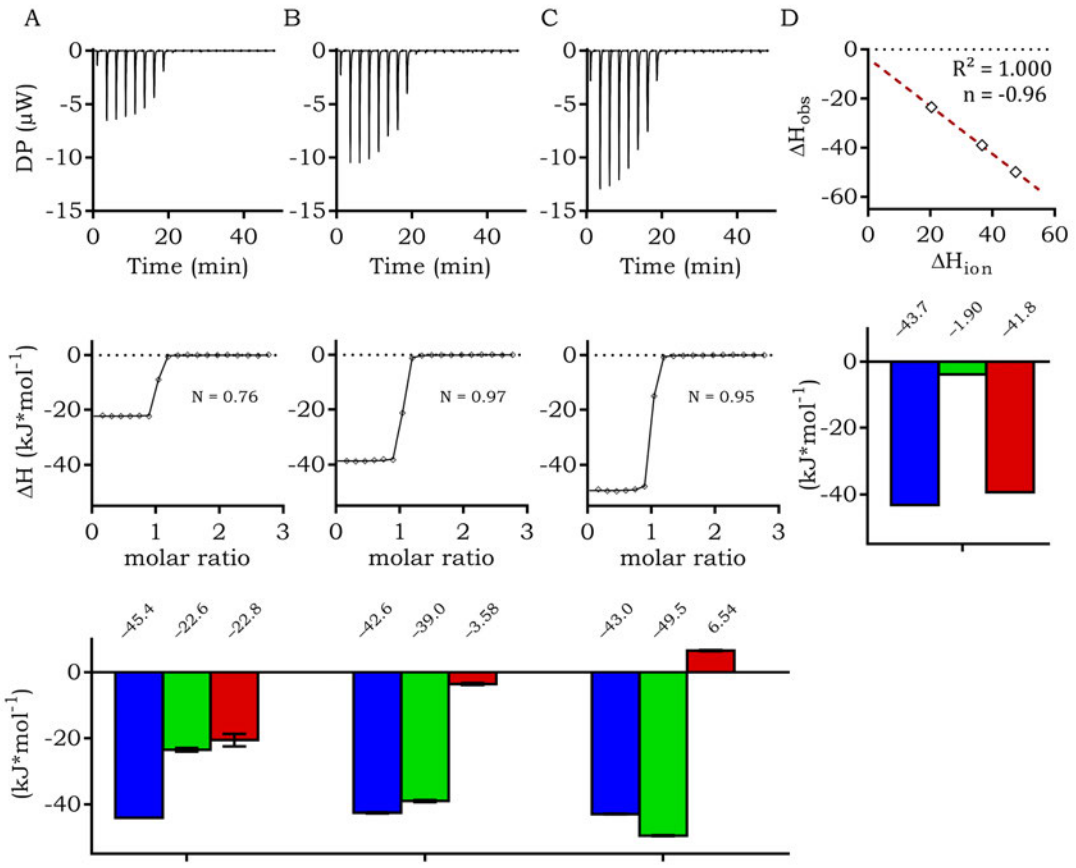


Figure S11. Buffer correction exemplarily performed for direct Ca^{2+} vs. EDTA titrations. **A–C** ITC experiments in HEPES (**A**), imidazole (**B**), and TRIS (**C**) buffered media. The thermograms (top) show the raw heat data. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line, calculated by the ITC software MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **D** Linear regression of ΔH_{obs} vs. ΔH_{ion} (top) and signature plot of the Ca^{2+}-EDTA binding after buffer ionization correction (mid).

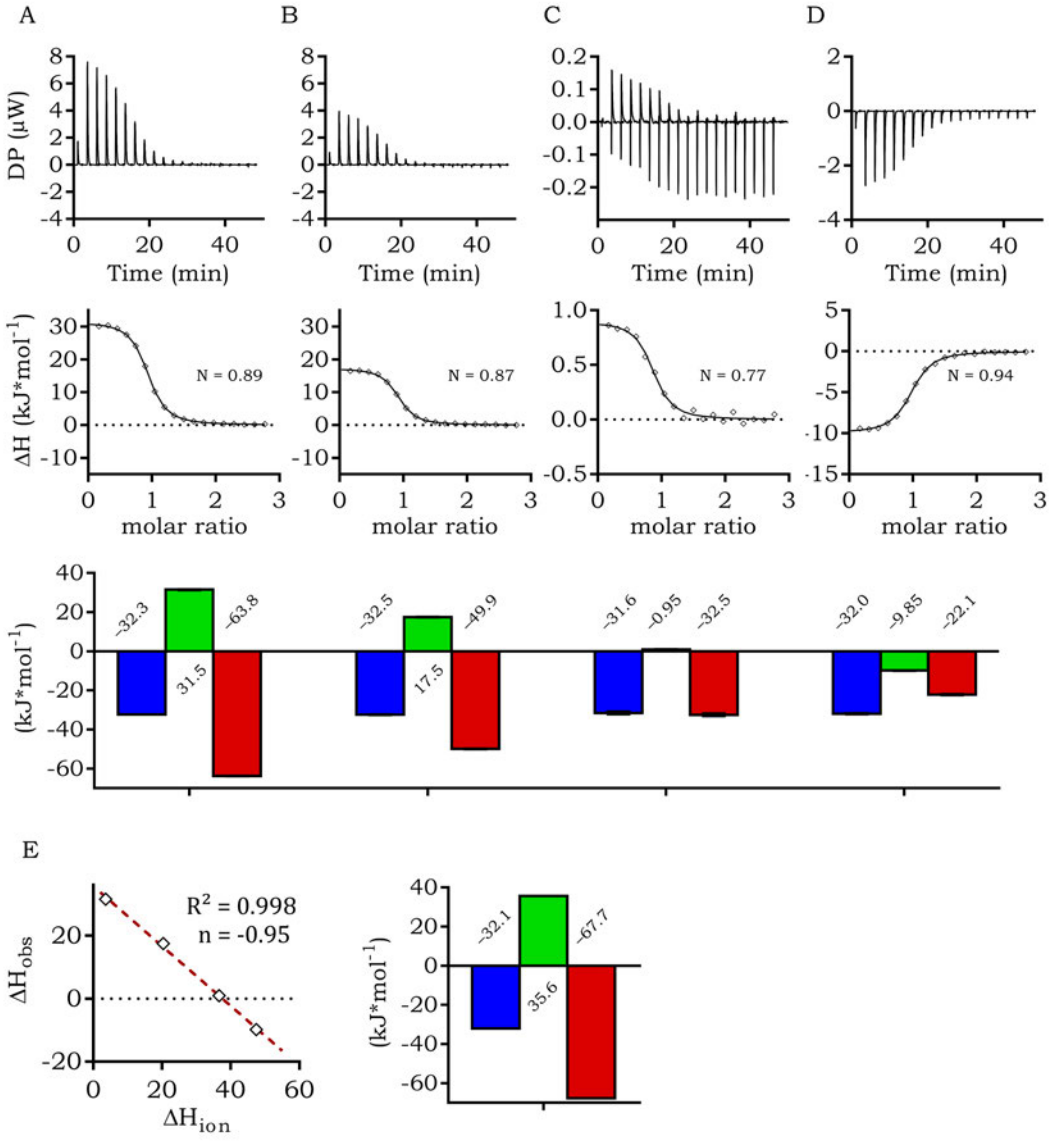


Figure S12. Buffer correction exemplarily performed for direct Mg^{2+} vs. EDTA. **A**–**D** ITC experiments in phosphate (**A**), HEPES (**B**), imidazole (**C**), and TRIS (**D**) buffered media. The thermograms (top) show the raw heat data. Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line, calculated by the ITC software MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG_{obs} (blue). **E** Linear regression of ΔH_{obs} vs. ΔH_{ion} (left) and signature plot of the Mg^{2+}-EDTA binding after buffer ionization correction (right).

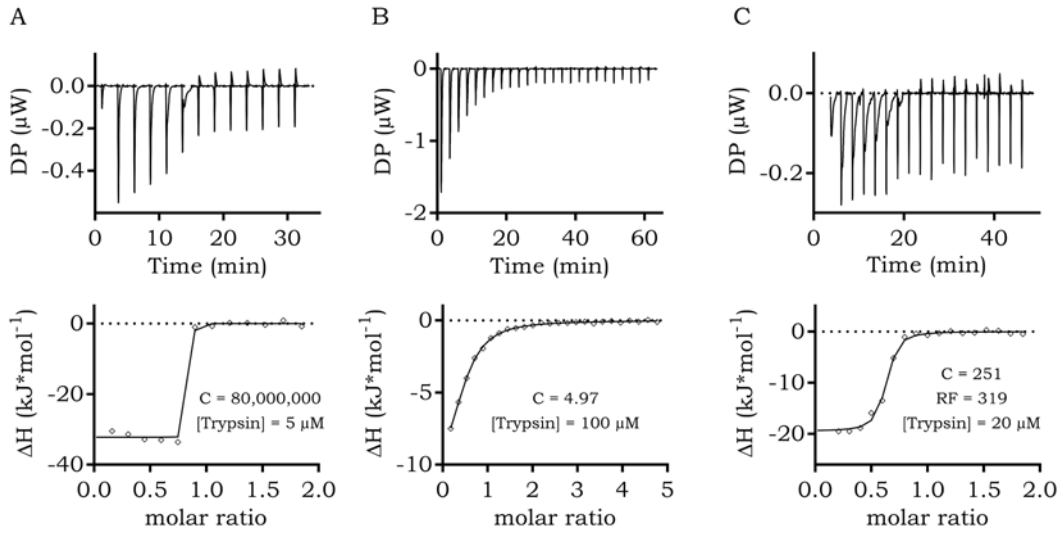


Figure S13. Exemplary curves for the displacement titrations of preincubated BA by aprotinin. **A** Direct titration of the aprotinin-trypsin binding. The C-value of ~80,000,000 is too high to obtain a reliable K_D. The C-value was calculated based on the literature value of 60 fM for the affinity of aprotinin to trypsin. **B** Direct titration of the BA-trypsin binding. The sample consumption is high with 100 µM trypsin in the reaction cell and 2.5 mM BA in the syringe. **C** Displacement titration of BA with aprotinin. 5 µM trypsin was preincubated in the reaction cell with 2500 µM BA. The RF of ~319 is sufficient to achieve a C-value in the desired range of ~250.

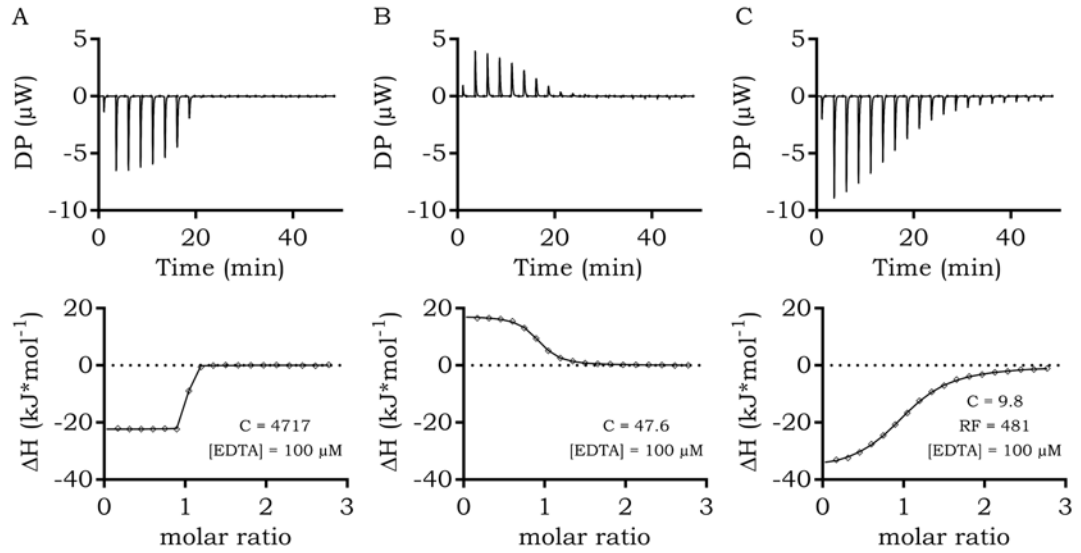


Figure S14. Exemplary curves for the improvement through displacement titrations. **A** Direct titration of the Ca^{2+}-EDTA binding. The C-value of ~4700 is too high to obtain a reliable K_D. **B** Direct titration of the Ca^{2+}-EDTA binding. The C-value is ~48. Thus the curve is well-shaped to obtain a reliable K_D. **C** Displacement titration of Mg^{2+} with Ca^{2+}. In the reaction cell, 100 µM EDTA is in the presence of 850 µM Mg^{2+}. The resulting C-value of ~10 is in the desired range. The RF is ~480.

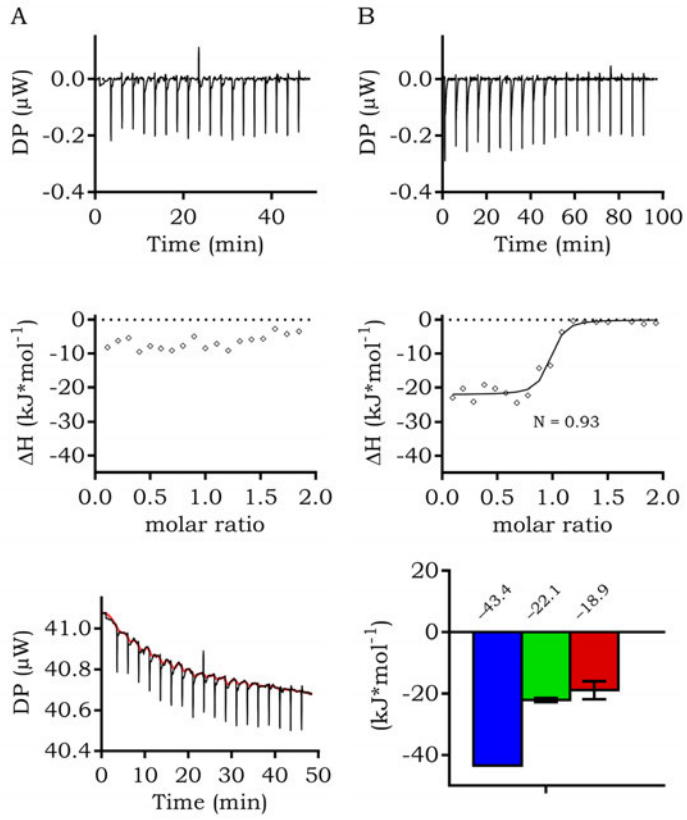


Figure S15. Exemplary curves for the signal improvement for direct leupeptin vs. trypsin titrations by adjusting the spacing time. The thermograms (top) show the raw heat data, and the red line indicates the baseline, calculated by the ITC software MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). Each point of the binding isotherm (mid) shows the integrated heat of the respective injection. The fit for calculating the thermodynamic parameters is shown as a black line. The signature plots (bottom) show the contributions of ΔH_{obs} (green) and $-T\Delta S_{obs}$ (red) to overall ΔG (blue). **A** Measurement at 25 °C and the default 150 s spacing time between each injection. **B** Increasing the spacing time to 300 s, the baseline could be calculated, and the binding isotherm could be fitted.

Table S1. Calculation of cost per experiment.

Binding partners							Cost per 20 experimental sets:
Procurement			Consumption				
Substance	Amount	Cost	Buffer correction	Heat capacity	Displacement titrations	Cost per Substance	
Trypsin	100 mg	34.44 €	11.2 mg	5.60 mg	3.20 mg	6.89 €	
Aprotinin	10 mg	32.70 €	2.67 mg	1.33 mg	1.33 mg	17.44 €	30.90 €
Leupeptin	100 mg	288.00 €	0.40 mg	0.20 mg	0.20 mg	2.30 €	
BA*HCl	5000 mg	23.54 €	1.20 mg	0.60 mg	0.83 mg	0.01 €	
Buffers							Cost per experiment
Procurement			Consumption				
Ingredient	Amount	Cost	for 1 L buffer			Sum (€)	
Na-H_2PO_4	1000 g	22.50 €	7.80 g			0.18 €	
HEPES	1000 g	163.20 €	11.9 g			1.94 €	
Imidazole	100 g	25.00 €	3.40 g			0.85 €	1.55 €
TRIS	1000 g	77.90 €	6.06 g			0.47 €	
NaCl	500 g	17.4 €	23.38 g			0.81 €	
In total, 20 experiments were performed as triplicates with control experiments.							

Table S2. Selected ligands with literature-known affinities for experimental design.

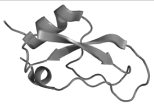
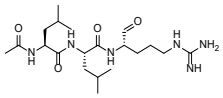
Ligand		Target	K_D	Method	Literature
Structure	Name				
	Aprotinin aa 1–58, RPDFCLEPPYTGPCKARIIRY FYNAKAGLCQTFVYGGCRA KRNNFKSAEDCMRTCGGA	Trypsin	60 fM	Displacement equilibrium measurements	Ref [1]
	Leupeptin	Trypsin	40.4 nM	Active site titration pH 7.2	Ref [2]
	BA	Trypsin	21.7 µM	ITC in TRIS pH 8.0	Ref [3]
Ca^{2+}	Calcium	EDTA	22.7 nM	ITC in HEPES pH 7.4	Ref [4]
Mg^{2+}	Magnesium	EDTA	1.67 µM	ITC in HEPES pH 7.4	Ref [4]

Table S3. Obtained thermodynamic values, including errors from direct titrations.

Ligand	buffer	T (°C)	N	K_D	ΔG_{obs} (kJ*mol^{-1})	ΔH_{obs} (kJ*mol^{-1})	$-T\Delta S_{obs}$ (kJ*mol^{-1})
Aprotinin	HEPES	15	0.68 ± 0.01	1.00 ± 6790 pM	≤ –45.1	–23.7 ± 1.2	≤ –21.4
	HEPES	25	0.72 ± 0.01	1.00 ± 2020 pM	≤ –49.6	–31.8 ± 0.7	≤ –17.8
	HEPES	35	0.67 ± 0.00	5.44 ± 1.97 nM	≤ –48.0	–51.4 ± 0.6	≤ 3.44
	Imidazole	25	0.75 ± 0.01	4.15 ± 4.13 nM	≤ –46.1	–37.0 ± 0.8	≤ –9.16
	TRIS	25	0.65 ± 0.00	3.77 ± 2.09 nM	≤ –46.9	–45.1 ± 0.6	≤ –1.94
Leupeptin	HEPES	15	0.83 ± 0.01	19.2 ± 13.6 nM	–43.2 ± 1.9	–8.4 ± 0.4	–34.8 ± 1.5
	HEPES	25	0.93 ± 0.01	28.8 ± 12.4 nM	–43.4 ± 1.7	–22.1 ± 0.7	–18.9 ± 2.9
	HEPES	30	0.70 ± 0.01	43.0 ± 10.8 nM	–42.9 ± 0.9	–39.0 ± 0.8	–3.9 ± 2.4
	Imidazole	25	0.80 ± 0.01	50.8 ± 14.7 nM	–41.7 ± 0.2	–22.4 ± 0.6	–17.8 ± 0.8
	TRIS	25	0.61 ± 0.01	15.9 ± 7.6 nM	–44.7 ± 0.9	–23.5 ± 1.2	–21.2 ± 3.6
BA	HEPES	15	0.30 ± 0.01	3.39 ± 0.87 µM	–29.0 ± 0.6	–6.6 ± 0.4	–22.5 ± 0.9
	HEPES	25	0.42 ± 0.01	20.1 ± 0.7 µM	–26.8 ± 0.1	–12.3 ± 0.3	–14.6 ± 0.4
	HEPES	35	0.43 ± 0.01	23.0 ± 0.8 µM	–27.4 ± 0.3	–22.7 ± 0.5	–4.7 ± 1.6
	Phosphate	25	0.36 ± 0.01	18.1 ± 1.4 µM	–27.1 ± 0.3	–7.9 ± 0.4	–19.2 ± 1.7
	Imidazole	25	0.46 ± 0.02	57.8 ± 3.6 µM	–24.2 ± 0.1	–16.0 ± 1.2	–8.2 ± 0.7
	TRIS	25	0.38 ± 0.01	32.2 ± 1.2 µM	–25.7 ± 0.3	–23.3 ± 0.8	–2.3 ± 2.5
Ca^{2+}	HEPES	15	0.76 ± 0.00	10.1 ± 2.6 nM	–44.2 ± 0.5	–23.9± 0.1	–20.3 ± 0.7
	HEPES	25	0.76 ± 0.00	21.2 ± 7.7 nM	–44.1 ± 1.4	–23.5 ± 0.3	–20.6 ± 1.9
	HEPES	35	0.78 ± 0.00	35.6 ± 5.7 nM	–44.0 ± 0.6	–21.2 ± 0.2	–22.9 ± 1.4
	Imidazole	25	0.97 ± 0.00	35.2 ± 3.4 nM	–42.6 ± 0.2	–39.0 ± 0.1	–3.6 ± 0.3
	TRIS	25	0.95 ± 0.00	30.2 ± 3.0 nM	–43.0 ± 0.1	–49.5 ± 0.1	6.5 ± 0.2
Mg^{2+}	HEPES	15	0.91 ± 0.00	870 ± 25.4 nM	–33.5 ± 0.2	15.9 ± 0.0	–49.4 ± 0.3
	HEPES	25	0.87 ± 0.00	2.10 ± 0.06 µM	–32.5 ± 0.1	17.5 ± 0.1	–49.9 ± 0.1
	HEPES	35	0.92 ± 0.00	592 ± 20.0 nM	–36.8 ± 0.0	18.1 ± 0.1	–54.9 ± 0.1
	Phosphate	25	0.89 ± 0.00	2.21 ± 0.04 µM	–32.3 ± 0.1	31.5 ± 0.1	–63.8 ± 0.1
	Imidazole	25	0.77 ± 0.02	3.08 ± 0.88 µM	–31.6 ± 0.7	1.0 ± 0.0	–32.5 ± 0.7
	TRIS	25	0.94 ± 0.00	2.54 ± 0.11 µM	–32.0 ± 0.2	–9.9 ± 0.1	–22.1 ± 0.3

Experiments were performed at least in triplicates.

Table S4. Obtained thermodynamic values, including errors from displacement titrations.

Syringe	Cell	buffer	T (°C)	N	K_{Dapp}	ΔG_{app} (kJ*mol^{-1})	ΔH_{app} (kJ*mol^{-1})	$-T\Delta S_{app}$ (kJ*mol^{-1})
Aprotinin	Trypsin + 2500 µM BA	HEPES	25	0.59 ± 0.01	79.7 ± 25.9 nM	−40.7 ± 0.9	−19.0 ± 0.8	−21.6 ± 1.3
Leupeptin	Trypsin + 250 µM BA	HEPES	25	0.75 ± 0.02	851 ± 203 nM	−34.8 ± 0.8	−22.0 ± 1.2	−13.4 ± 3.0
Ca^{2+}	EDTA + Mg^{2+}	HEPES	25	1.00 ± 0.00	10.2 ± 0.17 µM	−28.5 ± 0.0	−38.2 ± 0.2	9.66 ± 0.5

Experiments were performed at least in triplicates.

Table S5. ITC experiments at different temperatures to calculate the heat capacity.

Ligand	T (°C)	ΔG_{obs} (kJ*mol^{-1})	ΔH_{obs} (kJ*mol^{-1})	$-T\Delta S_{obs}$ (kJ*mol^{-1})	ΔC_p (kJ*mol^{-1}*K^{-1})	T_S (°C)	T_H (°C)
Ca^{2+}	15	−44.2	−23.9	−20.3	0.136	−139	192
	25	−44.1	−23.5	−20.6			
	35	−44.0	−21.2	−22.9			
Mg^{2+}	15	−33.5	15.9	−49.4	0.110	−161	-131
	25	−32.5	17.5	−49.9			
	35	−36.8	18.1	−54.9			

Experiments were performed at least in triplicates.

Table S7. Buffer ionization correction for EDTA binding ligands.

Ligand	buffer	$\Delta G°$ (kJ*mol^{-1})	ΔH_{obs} (kJ*mol^{-1})	$-T\Delta S_{obs}$ (kJ*mol^{-1})	$\Delta H°$ (kJ*mol^{-1})	$-T\Delta S°$ (kJ*mol^{-1})	n_{proton} (kJ*mol^{-1})
Ca^{2+}	HEPES	−43.2	−23.5	−20.6	−3.9	−39.8	−0.96
	Imidazole		−39.0	−3.6			
	TRIS		−49.5	6.5			
Mg^{2+}	Phosphate	−32.1	31.5	−63.8	35.6	−67.7	−0.95
	HEPES		17.5	−49.9			
	Imidazole		1.0	−32.5			
	TRIS		−9.9	−22.1			

Experiments were performed at least in triplicates.

Table S8. Obtained thermodynamic parameters from direct and displacement titrations.

Ligand	Titration protocol	K_D	ΔG (kJ*mol^{-1})	ΔH (kJ*mol^{-1})	−TΔS (kJ*mol^{-1})
Ca^{2+}	Direct	21.2 nM	−43.2	−23.5	−20.6
	Displacement of Mg^{2+}	25.1 nM	−43.4	−20.7	−22.7
Mg^{2+}	Direct	2.10 µM	−32.5	17.5	−49.9
	Displacement with Ca^{2+}	1.77 µM	−32.8	14.7	−47.6

Experiments were performed at least in triplicates.

EXPERIMENTAL SECTION

Reagents

If not stated differently, all reagents were purchased from Sigma Aldrich (St. Louis, MO, USA). Microbially-produced leupeptin was purchased from Sigma Aldrich in > 90% purity (HPLC). Trypsin was purchased as trypsin from bovine pancreas (Sigma Aldrich, Cat. No.: T8003).

Table S9. Buffers used for buffer ionization compensation.

Buffering	Buffer composition
$ITC^{EDTA}_{Phosphate}$	10 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 100 mM NaCl
ITC^{EDTA}_{HEPES}	10 mM HEPES pH 7.4, 100 mM NaCl
ITC^{EDTA}_{TRIS}	10 mM TRIS pH 7.4, 100 mM NaCl
$ITC^{EDTA}_{Imidazole}$	10 mM imidazole pH 7.4, 100 mM NaCl
$ITC^{Trypsin}_{Phosphate}$	50 mM NaH_2PO_4/Na_2HPO_4 pH 7.4, 100 mM NaCl
$ITC^{Trypsin}_{HEPES}$	50 mM HEPES pH 7.4, 100 mM NaCl
$ITC^{Trypsin}_{TRIS}$	50 mM TRIS pH 7.4, 100 mM NaCl
$ITC^{Trypsin}_{Imidazole}$	50 mM imidazole pH 7.4, 100 mM NaCl

Implementation of *ITCcalc*

The *ITCcalc* toolkit was developed with JavaScript as a client-site browser applet and optimized for use with Firefox. The toolkit can be accessed at https:\\itccalc.uni-mainz.de.

ITC experiments

All experiments were performed at least in technical triplicates on a MicroCal PEAQ-ITC automated system (Malvern Instruments, Worcestershire, UK) with a 200 µL Hastelloy cell and an injection syringe volume of 40 µL. $ITC^{Trypsin}_{HEPES}$ and ITC^{EDTA}_{HEPES} buffers (Table S9) were used as standard buffers except for experiments for the buffer ionization correction. For quick preparation of solutions, stock solutions in Millipore (MP) H_2O

of EDTA (10 mM), $CaCl_2$ (1 M), $MgCl_2$ (1 M), aprotinin (5 mM), leupeptin (10 mM), and BA (250 mM) were used. The 100 mM EDTA stock in MP H_2O was adjusted to a pH of 7.4 using NaOH before diluting the respective buffers (Table S10). Due to limited solubility, trypsin was directly dissolved to 100 µM in the respective $ITC^{Trypsin}$ buffer. Dilution of the aqueous stock solutions in the respective buffers resulted in a maximum of 2% (v/v) excess MP H_2O in the final solutions. Thus, to reduce dilution enthalpies, all final working solutions were exactly matched by adjusting to these 2% (v/v) additional MP H_2O.

For each experiment, a set of control experiments was performed to cover the dilution effects of analyte (analyte vs. buffer) or ligand (buffer vs. ligand) as well as interactions of buffer with the cell material or mismatched buffers (buffer vs. buffer). Since they did not show pronounced effects, control experiments were subtracted from (analyte vs. buffer and buffer vs. ligand) or summed up to (buffer vs. buffer) the raw data as a line. Additionally, the remaining offset was fitted as provided by the MicroCal PEAQ-ITC Analysis Software V1.21 (Malvern Panalytical, Worcestershire, UK). 19 injections à 2 µL were added to the reaction cell with an injection speed of 0.5 µL/s and a spacing time of 150 s.

Direct titrations

For direct titrations in HEPES buffer, a 10-fold molar excess of titrant was titrated to the analyte in the cell (Table S10). As default settings, experiments were performed at 25 °C with a stirring speed of 750 rpm and a reference power of 42 µW. 19 injections à 2 µL were added to the reaction cell with an injection speed of 0.5 µL/s and a spacing time of 150 s. After initial measurements for all direct titrations using this method, some adjustments were made.

The spacing time for titrations with leupeptin vs. trypsin was increased to 300 s (Figure S15). For the binding of BA to trypsin, the K_D is too low to give a sufficient slope at the inflection point (C-value too low) when measured with a 10-fold molar excess in the syringe. Therefore, a low-C titration was performed. Here, a 25-fold molar excess of BA was used to achieve an analyte complexation rate of >70%, which is necessary to interpret low-C titration curves.[5] To improve the resolution of the first injections near the equivalence point, the number of injections was increased to 25 à 1.5 µL for these low-C titrations.

Table S10. Concentrations used for direct titrations.

Cell	Syringe
100 µM EDTA	1 mM Ca^{2+}
100 µM EDTA	1 mM Mg^{2+}
5 µM trypsin	50 µM aprotinin
10 µM trypsin	100 µM leupeptin
100 µM trypsin	2.5 mM BA

Buffer Ionization compensation

Compensation for buffer ionization enthalpies was calculated using *ITCcalc*. Briefly, for each direct titration, ΔH_{obs} is plotted against the ionization enthalpy of the respective buffer. The slope of the regression line then reveals n_{proton}, and the intercept with the x-axis gives the corrected binding enthalpy (ΔH°). Four buffering agents with known ionization enthalpies were used (Table S6).[6,7]

Table S6. Buffer ionization enthalpies used for calibration curves.

Buffering agent	Buffer Ionization ΔH ($kJ*mol^{-1}$)
Phosphate	3.6 [a]
HEPES	20.4 [a]
Imidazole	36.6 [a]
TRIS	47.5 [a]

[a] Ref [6]

Heat capacity determination

Direct titrations in HEPES buffers were performed at differing temperatures (15 °C, 25 °C, and 30 °C) to obtain the changes in heat capacity (ΔC_P) upon ligand binding events. ΔH_{obs} from each titration set were plotted against T. Then, ΔC_P can be obtained as the slope from the regression of the ΔH vs. T plot.

Displacement titrations

Some direct titrations resulted in too-low or too-high C-values to give reliable K_D values. Displacement experiments were employed to circumvent these issues, where a high-affinity binder (e.g., Ca^{2+}, leupeptin, or aprotinin) was titrated to the target (e.g., EDTA or trypsin) in the presence of a relatively low-affinity binder (e.g., Mg^{2+} or BA; Table S11). All displacement experiments were performed in HEPES buffers. The resulting parameters for the unknown binder were calculated using *ITCcalc*.

Table S11. Used concentrations for displacement titrations.

Cell		Syringe
Target	Weak binder	Strong binder
100 µM EDTA	850 µM Mg^{2+}	1 mM Ca^{2+}
10 µM trypsin	250 µM BA	100 µM leupeptin
5 µM trypsin	250 µM BA	50 µM aprotinin

HAZARDS

No significant hazards emanate from the experiments presented in this paper. By adhering to good laboratory practice, contact with unprotected bare skin, inhalation of vapor, mist, or dust in the workplace atmosphere, or ingestion in any form should be avoided or taken care of according to the H and P statements of the globally harmonized system of classification and labeling of chemicals (GHS). Most solutions used in the described experiments are classified as nontoxic due to their low concentrations. Nevertheless, they should be handled cautiously and adequately disposed of according to local regulations. Safety goggles, a lab coat, and closed footwear should be worn throughout the experiment. The methanol used for cleaning is highly volatile, can form explosive air mixtures, and can be hazardous to humans. Therefore, storage containers should be tightly closed, and liquid handling should be performed in well-ventilated areas with no ignition sources. The automated ITC device is safe to use while the maintenance gate remains closed. Despite the slow movements of the automatically guided trays, they pose a risk of crushing. The nitrogen pressurized gas cylinder should be secured following laboratory guidelines and operated only by instructed personnel.

REFERENCES

(1) Vincent, J. P.; Lazdunski, M. Trypsin-Pancreatic Trypsin Inhibitor Association. Dynamics of the Interaction and Role of Disulfide Bridges. *Biochemistry* **1972**, *11* (16), 2967–2977. https://doi.org/10.1021/bi00766a007.

(2) Kuramochi, H.; Nakata, H.; Ishii, S. Mechanism of Association of a Specific Aldehyde Inhibitor, Leupeptin, with Bovine Trypsin. *J. Biochem.* **1979**, *86* (5), 1403–1410.

(3) Talhout, R.; Engberts, J. B. F. N. Thermodynamic Analysis of Binding of p -Substituted Benzamidines to Trypsin. *Eur. J. Biochem.* **2001**, *268* (6), 1554–1560. https://doi.org/10.1046/j.1432-1327.2001.01991.x.

(4) Henzl, M. T.; Larson, J. D.; Agah, S. Estimation of Parvalbumin Ca2+- and Mg2+-Binding Constants by Global Least-Squares Analysis of Isothermal Titration Calorimetry Data. *Anal. Biochem.* **2003**, *319* (2), 216–233. https://doi.org/10.1016/S0003-2697(03)00288-4.

(5) Turnbull, W. B.; Daranas, A. H. On the Value of c: Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Soc. Inf. Sci.* **2003**, *125* (48), 14859–14866. https://doi.org/10.1002/asi.4630310417.

(6) Goldberg, R. N.; Kishore, N.; Lennen, R. M. Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data* **2002**, *31* (2), 231–370. https://doi.org/10.1063/1.1416902.

(7) Fukada, H.; Takahashi, K. Enthalpy and Heat Capacity Changes for the Proton Dissociation of Various Buffer Components in 0.1 M Potassium Chloride. *Proteins Struct. Funct. Genet.* **1998**, *33* (2), 159–166. https://doi.org//10.1002/(SICI)1097-0134(19981101)33:2<159::AID-PROT2>3.0.CO;2-E.

4. Project 2: Chasing the Binding Conformation of NS2B/NS3 Protease Inhibitors

4.1. Elucidating the mode of action of allosteric Dengue- and Zika NS2B/NS3 Protease Inhibitors

4.1.1. Introduction

This chapter presents the results of experiments using biophysical methods that were either not suitable for tracing the binding mode and conformational switches, induced by allosteric inhibitors, or are not included in the publication and the manuscript of this project. Therefore, the two allosteric inhibitors **15** and **16**, and the two competitive reference inhibitors **17** and **4** were utilized as tool compounds (Figure 20).

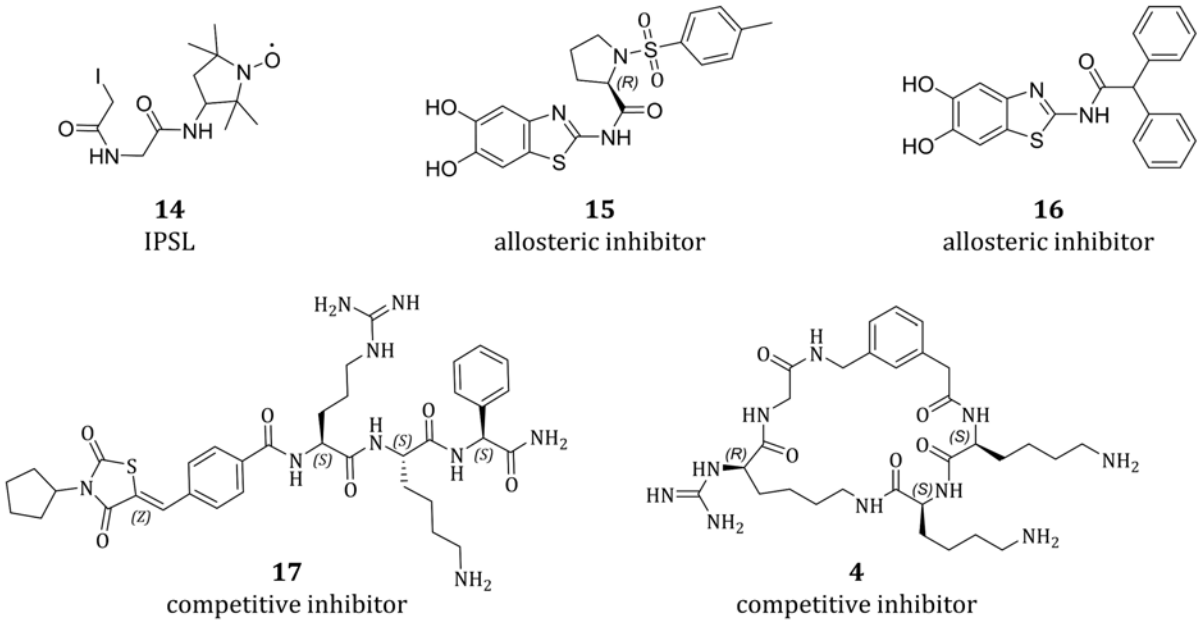


Figure 20. *Chemical substances used for biophysical methods to elucidate the allosteric binding mode.* Cpd **4** is used as a spin label for EPR experiments. Cpds **5** and **6** are allosteric inhibitors of NS2B/NS3 developed in the group of [redacted].[196,239] Cpds **7** and **8** are competitive reference cpd provided by [redacted],[240] and [redacted] ([redacted]).[57]

4.1.2. Results and Discussion

Electron paramagnetic resonance. Electron paramagnetic resonance (EPR) experiments were performed together with [redacted] of the Group of [redacted] ([redacted]). EPR spectroscopy is a magnetic resonance that observes the structure and dynamics of paramagnetic centers using microwaves. Such paramagnetic centers can be radicals or transition-metal ions. Site-specific labeling of the glycine-linked DENV2 NS2B/NS3 was achieved by introducing Cys residues into the amino acids sequence *via* site-directed mutagenesis. After enzyme expression and purification, proteins were spin-labeled with 3-(2-iodoacetamido)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy (IPSL, **14**). EPR experiments with pulsed double electron-electron resonance (DEER) analysis, allowing to determine distances of two spin labels, were conducted in cooperation with [redacted] from the group of [redacted].[241] Based on distance estimations with the crystal structures in the closed (PDB-ID: 3U1I)[54] and open conformation (PDB-ID: 2FOM)[34], the distance of Cys71* and Cys79* in NS2B to Cys118 and Cys158 in NS3 should differ significantly (Figure 21A). In the closed conformation, the *β* hairpin, formed by NS2B, brings Cys71* and Cys79* closer to the Cys residues in NS3. In the open conformation, the part of NS2B containing the Cys residues is further away from the Cys residues in NS3. Cys79* is not resolved in the crystal structures, indicating high flexibility. DEER-EPR spectra of NS2B/NS3 in the apo state showed several peaks, suggesting multiple conformations in solution. The prominent peak at a distance of 2–4 nm represents the closed conformation. Those at higher distances indicate the open conformation. In the presence of 100 μM of cpd **15**, the populations of the open conformation of all mutants slightly increased but did not adopt a defined distance. Also, the closed conformation seems to be still dominant. A less populated open conformation was revealed after adding 100 μM of the competitive inhibitor **17** (Figure 21).

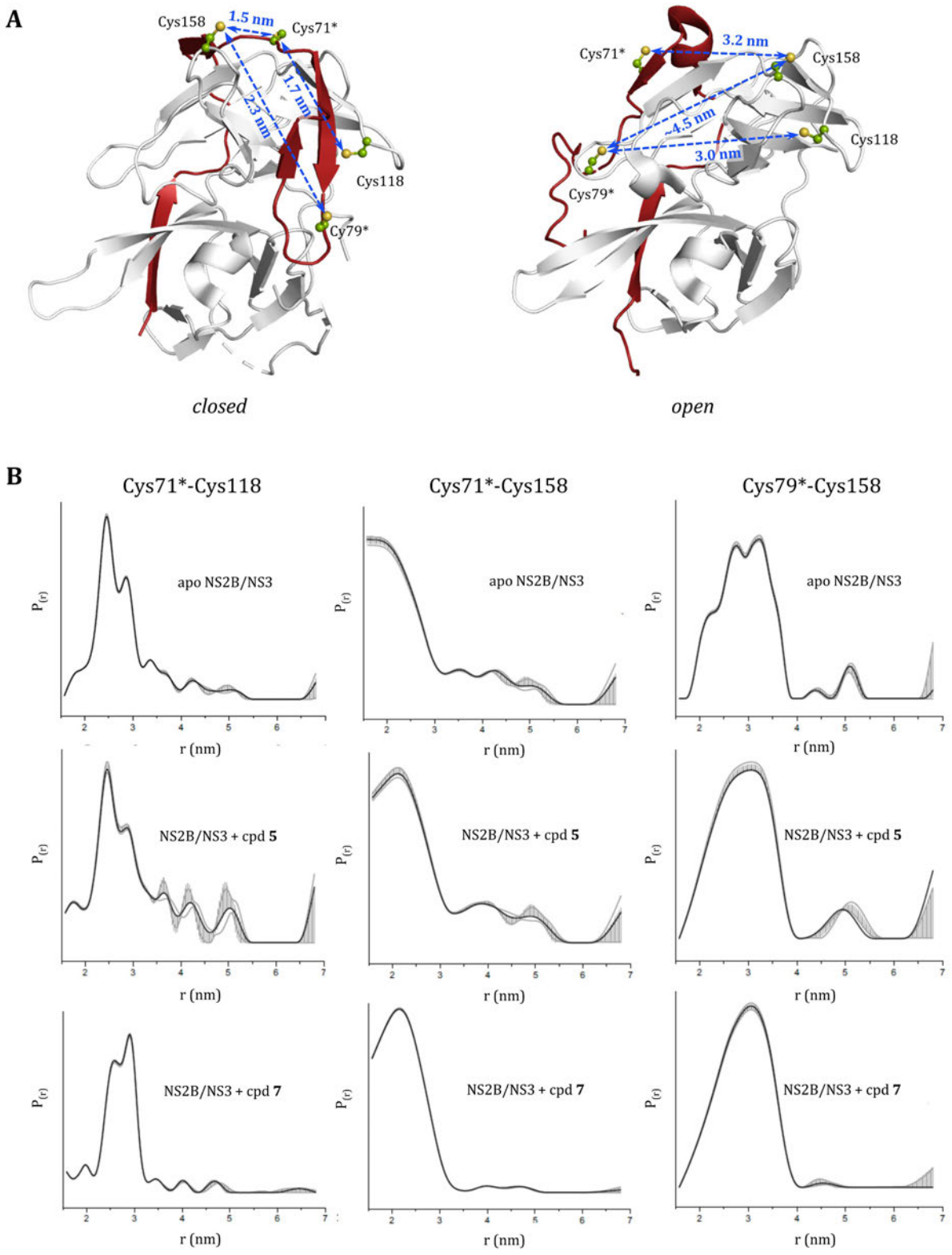


Figure 21. *DEER-EPR spectroscopy using DENV2 NS2B/NS3.* (**A**) Distance measurements of the Cys residues, introduced in DENV2 NS2B/NS3 for site-specific labeling with IPSL showing the changes upon a conformational transition from the closed (top) to the open conformation (bottom). (**B**) DEER-EPR spectra with apo NS2B/NS3 (top), NS2B/NS3 + 100 µM of cpd **5** (middle), and NS2B/NS3 + 100 µM of cpd **17** (bottom).

Microscale thermophoresis. The microscale thermophoresis (MST) technique exploits the migration of molecules along a temperature gradient. The target molecule needs to be labeled with a tracer fluorophore. A dilution series of a ligand is mixed with a fixed concentration of the target molecule and is transferred into capillaries (Figure 22). The capillaries are analyzed one by one using an infrared (IR)-laser to apply the temperature gradient and the emission light of the fluorophore to trace the migration of the target molecule. Ligand binding usually leads to changes in hydrodynamic radius, the hydration shell, or conformations that influence the thermophoretic behavior. Since this change is proportional to the bound fraction of the target molecule, the relative normalized fluorescence (F_{norm}) *vs.* ligand concentration can be sigmoidally fitted to obtain the K_d.

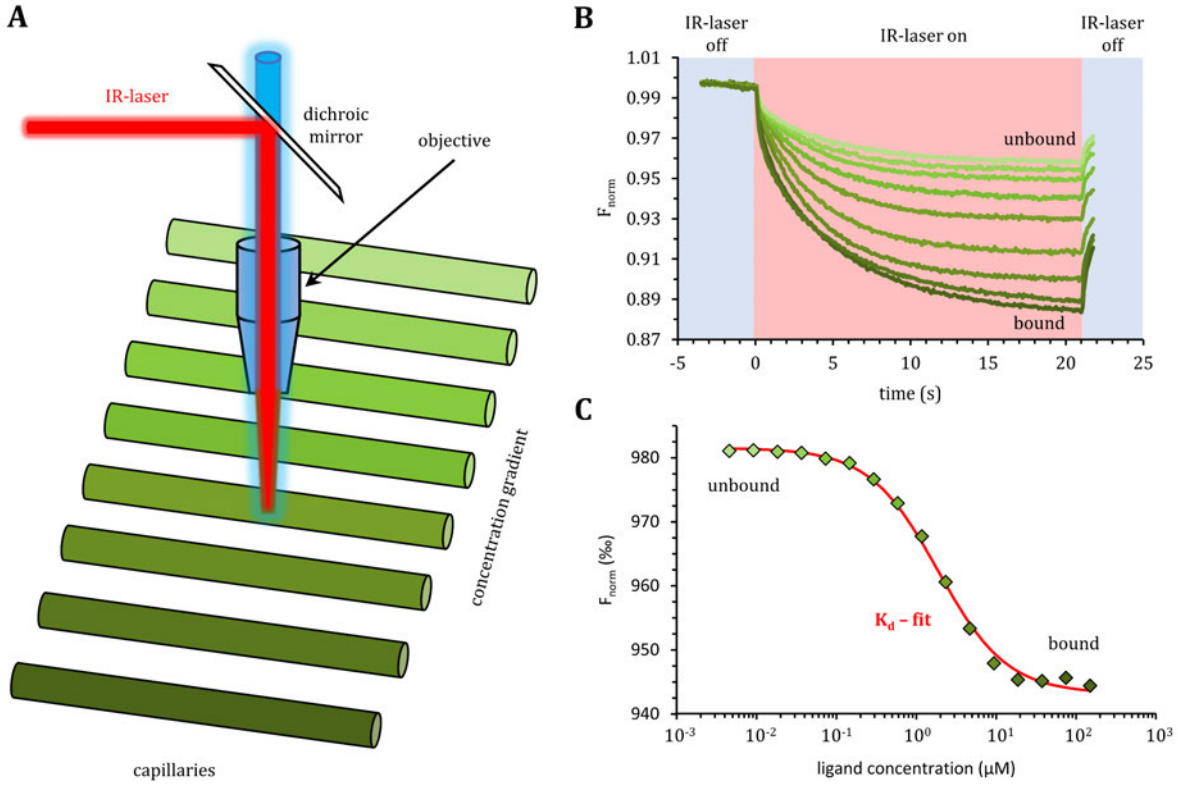


Figure 22. *Schematic MST-Setup and functionality.* (**A**) A tray filled with up to 16 capillaries containing a fluorophore-labeled target molecule at a fixed concentration and a dilution series of a ligand gets locally heated by an IR-laser. Simultaneously, the fluorophore in the heated spot gets excited so that the respective emission light can be used to trace the thermophoresis of each capillary. (**B**) The fluorescence traces can be divided into three parts. The laser is turned off in the first 5 s of the equilibrium phase (blue). Switching the laser on for 21 s induces a temperature gradient of 3–6 K (red). In this phase, two phenomena can be observed. During the heating process (in the first second), the fluorescence changes due to the changed temperature. Usually, the quantum yield drops at higher temperatures. After that, the thermophoresis can be monitored until an equilibrium is reached. After turning the laser off, the reverse temperature can be observed as the temperature decreases. (**C**) The relative normalized fluorescence (F_{norm}, the quotient of the fluorescence after thermophoresis, and the initial fluorescence) can be plotted against the ligand concentration to be sigmoidally fitted. This figure was created by ROBERT ALEXANDER ZIMMERMANN and modified with his permission.

In dilution series of cpd **15** against fluorescence labeled glycine-linked DENV2 NS2B/NS3 WT, a K_d of 31.5 ± 1.4 µM (Figure 23A) was yielded in good agreement with the values obtained from fluorometric enzyme activity assays (19.7 ± 1.5 µM).[239] Interestingly, higher concentrations of cpd **15** dramatically reduced the overall fluorescence of the labeled NS2B/NS3. A sigmoidal fit of the initial fluorescences yielded a K_d similar to that derived from thermophoresis (Figure 23B). This phenomenon can be interpreted in two ways: Either the ligand binds in the proximity of the fluorophore, thereby influencing its fluorescence by proximity sensing, or a conformational transition takes place, reducing the quantum yield of the fluorophore.

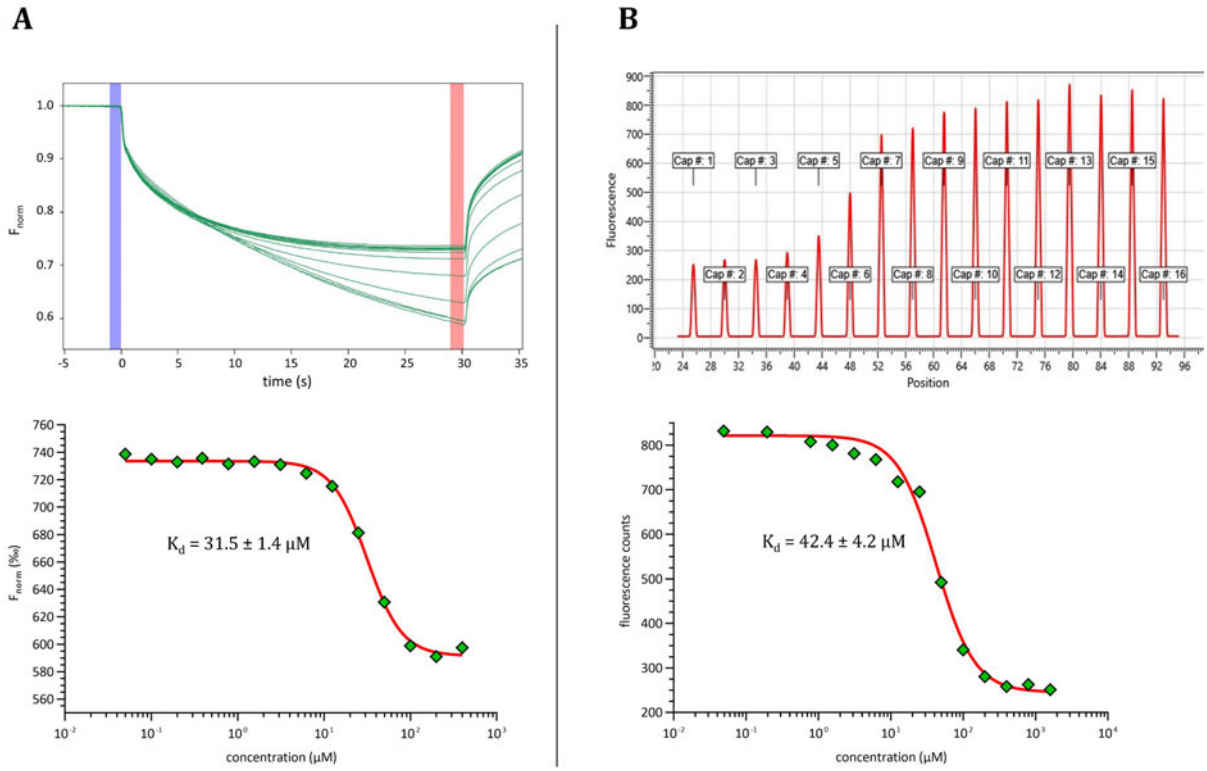


Figure 23. *K_d determination of allosteric binding using MST.* (**A**) The K_d of the allosteric inhibitor **15** was determined from the fluorescence traces, displaying the thermophoresis behavior of labeled DENV2 NS2B/NS3. (**B**) When fitting the maximal values of the capillary scan fluorescence, a K_d was obtained in good agreement with those from the thermophoretic behavior. This figure was created using the MO.Affinity Analysis v2.3 software (NanoTemper technologies, top) and GraphPad Prism[242] (bottom).

Nano surface plasmon resonance. In nano surface plasmon resonance (nanoSPR) experiments in cooperation with [redacted] from the group of [redacted]

DENV2 NS2B/NS3 was immobilized *via* its N-terminal hexahistidine (His_6) tag on gold-nanoparticles, modified with Ni-nitrilotriacetic acid (NTA). Focusing on single nanoparticles, the influence of the allosteric cpd **15** and the competitive cpd **17** should be investigated. However, there was a fast reduction of the bound protein, probably due to the shedding of the His_6 tag by the enzymatic activity of NS2B/NS3. This His_6 cleaving activity could be demonstrated for DENV2 NS2B/NS3 by incubating the protein at room temperature (rt.) and subsequent SDS-PAGE analysis and western blotting using the fluorescence generated by His Tag Horseradish Peroxidase-conjugated Antibodies (Bio-Techne, Minneapolis, MN, USA, Figure 24) for readout. Additionally, protein mass spectrometry of ZIKV NS2B/NS3 revealed slow autocleavage of the Hi_6 tag (*5.2 Insights into the autocleavage behavior of the ZIKV NS2B/NS3 protease*). An active site mutant of DENV or ZIKV NS2B/NS3 would probably overcome these issues.

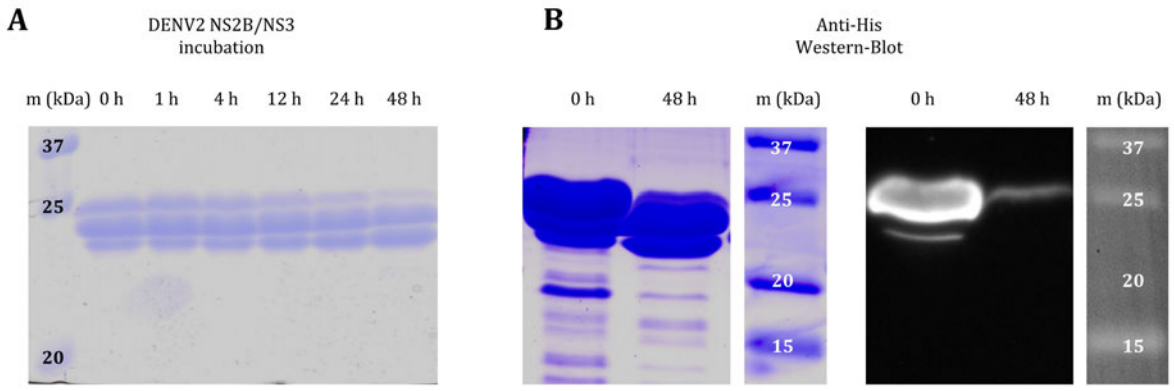


Figure 24. *Autocatalytic His_6 tag cleavage of DENV2 NS2B/NS3.* (**A**) Purification of DENV2 NS2B/NS3 with an N-terminal His_6 tag resulted in multiple bands in SDS-PAGE. Incubation and subsequent drawing of samples revealed the band's intensity with the highest molecular weight to decrease. Simultaneously, the density of the protein band with the lowest molecular weight increased. (**B**) Coomassie staining (left) and simultaneous Western Blotting followed by development with Anti-His Antibodies (right) of two identical performed SDS-PAGES proved the highest band to harbor a His_6, whereas the cleavage products only yielded low to no fluorescence.

Isothermal Titration Calorimetry. ITC was suitable for determining the thermodynamic binding profiles of competitive inhibitors of NS2B/NS3 (see *3.1 Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs*). In contrast, experiments with the allosteric inhibitors **15** and **16** did neither yield sufficient heat signals nor sigmoidally-shaped binding isotherms. Attempts to overcome this issue with different experimental temperatures (10 °C, 25 °C, and 35 °C), high concentrations of both enzyme and ligand, and varying injection volumes were unsuccessful. Even displacement titrations of pre-incubated NS2B/NS3 with cpd **16** and titration with the active site ligand **4** (Figure 20) did not show any effect on the binding behavior of the competitive ligand (data not shown).

SwitchSENSE. The DRX² instrument for SwitchSENSE measurements features a similar build as an SPR device. The target molecule (cargo) is immobilized on the gold surface of a chip, and different ligands (analytes) can be flushed through that chamber. The analyte must be labeled with a single-stranded (ss)DNA. The chip's gold surface is modified *via* thiol-gold coupling with a complementary ssDNA that harbors a fluorophore reporter. When flushing the chip with the target molecule, the labeled cargo is immobilized by DNA hybridization. When a positive charge is applied to the chip, the negatively charged double-stranded (ds)DNA is pulled down to the chip, causing quenching of the fluorescence readout. When the charge is reversed, the dsDNA is repelled from the surface, and the fluorescence intensity recovers (Figure 25). This speed of fluorescence recovery, called *dynamic response*, depends on the cargo's hydrodynamic radius and the properties of the analyte solution, e.g., viscosity, pH, and ionic strength. In standardized conditions, the viscosity, pH, ionic strength, and charge of the cargo can be neglected, allowing the detection of binding events of larger molecules or conformational changes of the cargo that occur upon ligand binding. Moreover, proximity sensing of the fluorophore can enable the detection of small-molecule interactions. Since the charge of the chip surface can be switched in the millisecond regiment, binding affinity and kinetics can be derived. Additionally, the hydrodynamic radius can be determined. In this work, a DRX² device (Dynamic Biosensors, Munich, Germany) was utilized. The conjugation of DENV2 and ZIKV NS2B/NS3 as cargo with the ssDNA was successfully performed by reaction of an *N*-hydroxysuccinimide (NHS) activated esters of the ssDNA with primary amines (lysine) of the protease, and subsequent purification *via* ion exchange chromatography (AEX) using the proFIRE (Dynamic Biosensors). With this, the single-labeled reaction product was separated from unreacted educts and multiple-labeled protein. Immobilization of the protein-DNA conjugate was successful. Although conformational changes upon binding of inhibitors **15** and **17** were expected, the sizing experiments did not show significant changes in the hydrodynamic radius. Hence, no binding affinity could be measured. An alternative attempt to utilize proximity sensing of six different dye reporters also failed to provide sufficient binding signals.

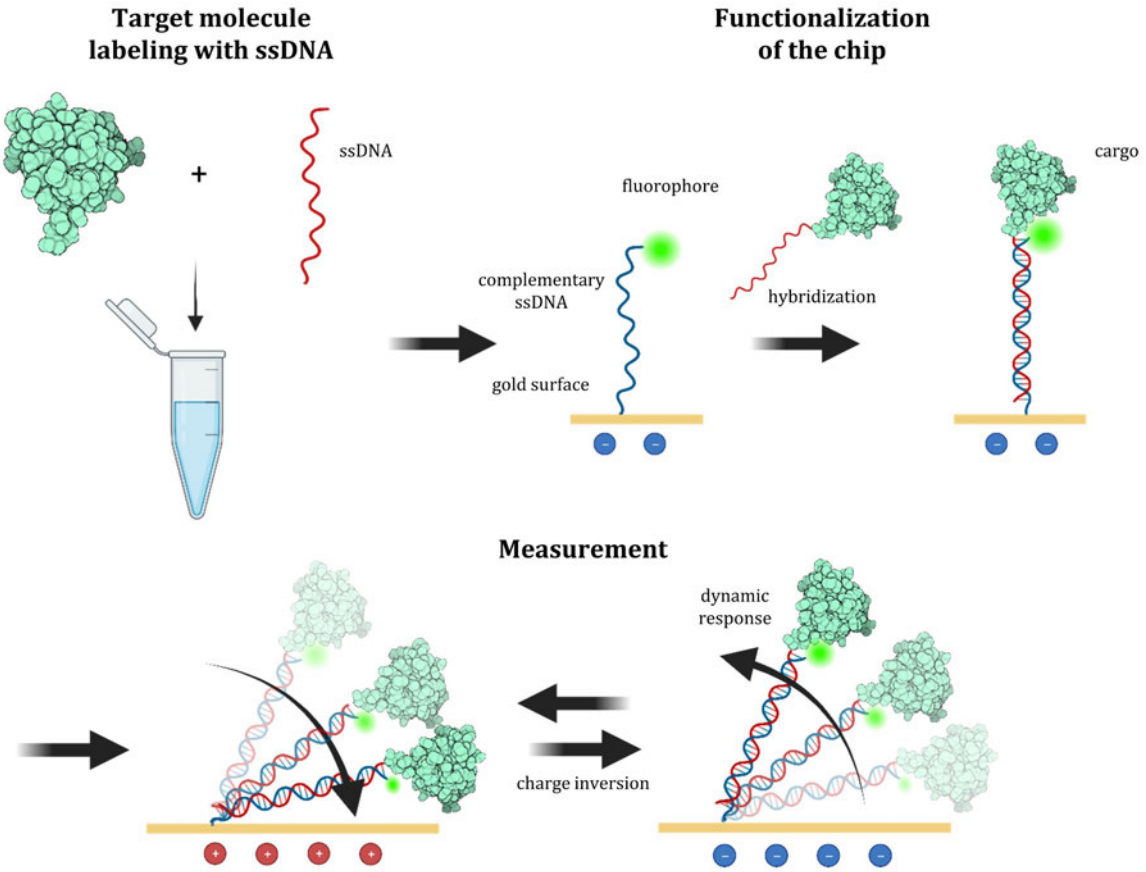


Figure 25. *Workflow and functionality of the SwitchSENSE technology.* A target molecule (DENV NS2B/NS3, PDB-ID: 3U1I, green)[54] was labeled with an ssDNA strand. In the DRX^2 device, a gold chip harboring a complementary ssDNA and a fluorophore reporter is functionalized by hybridization with the target molecule (cargo). If a positive charge is applied on the gold surface, the dsDNA is pulled down, causing quenching of the fluorophore. If this charge is inversed, the DNA with the cargo is pushed away, and the fluorescence regains. This hydrodynamic response is dependent on the hydrodynamic friction of the cargo. This figure was created using QuteMol[103] and BioRender (https://biorender.com/).

X-ray Crystallography. As stated earlier in this chapter, no undisputed crystal structure of NS2B/NS3 in complex with an allosteric inhibitor has been solved so far. Although excessive trials were undergone to crystallize the in-house glycine-linked DENV2 and ZIKV constructs, neither any published condition nor the RED Wings and SGC-I screening kits (The Structural Genomics Consortium, www.thesgc.org) yielded any crystal hits. For DENV2, this was probably the case due to two point mutations introduced into NS3 to enhance solubility.[195] Crystals could only be obtained using a new bivalent expressed ZIKV NS2B/NS3 construct (bZiPro, #86846, www.addgene.com) and using conditions reported for this very expression plasmid.[243] These crystals were obtained by [redacted] from the group of [redacted] in complex with cpds **5**–**7** (*3.1 Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs*).[244] These conditions for crystallization proved robust so that crystals of the apo state of bZiPro were also obtained in-house (Figure 26). Furthermore, crystals

of apo DENV4 full-length NS2B/NS3 (eDENV4) were successfully grown by exactly reproducing the respectively published conditions for this plasmid.[199] Although no co-crystals with the allosteric inhibitors **15** and **16** have been achieved, this work reached a milestone by obtaining apo crystals of bZiPro and eDENV4. These crystals can be readily used for fine-tuning, like macro- or microseeding procedures, to achieve larger crystals. Furthermore, soaking experiments with allosteric inhibitors can be pursued.

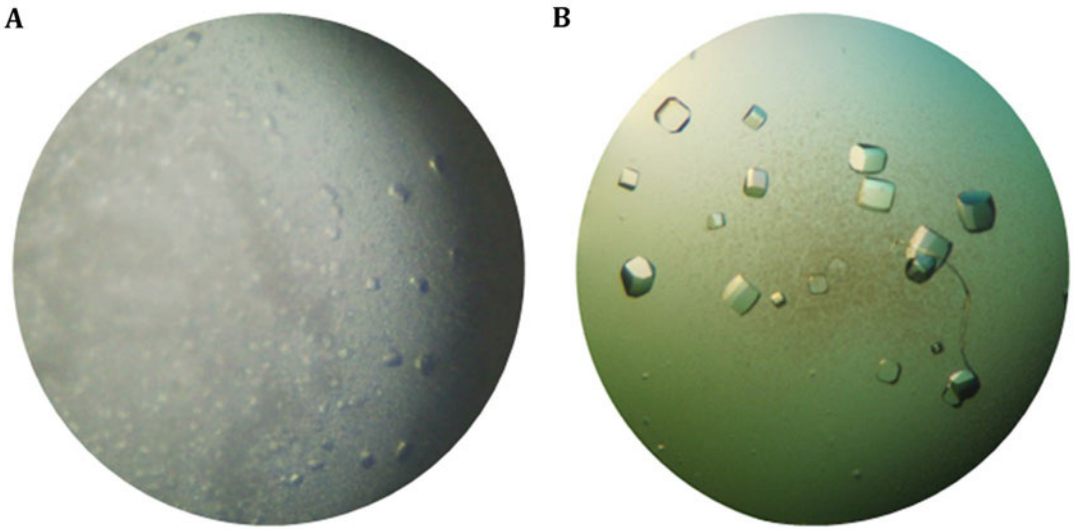


Figure 26. *Pictures of grown crystals of bZiPro.* (**A**) In the apo state, small cubic crystals were obtained that can be used for further optimization and seeding experiments. (**B**) Crystals of bZiPro grown by [redacted] of the group [redacted] under the same conditions in complex with a linear competitive cpd, used in *3.1 Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs*.

4.1.3. Experimental Section

Expression Constructs. For expression and mutagenesis of DENV2 NS2B/NS3, a pET-15b vector harboring the sequence of the DENV-2 NS2B/NS3 protease gene (GenBank ID: AY037116.1), covalently linked *via* a flexible G_4SG_4 linker and with the two-point mutations I30A and L31A in the $NS3_{pro}$ domain for better solubility, herein referred to as DENV2 NS2B/NS3 WT. As described earlier, this vector was kindly provided by [redacted].[196,239] The pET-Duet vector containing bZiPro was obtained from Addgene (#86846, www.addgene.com).[243]

Site-Directed Mutagenesis. The DENV2 NS2B/NS3 double Cys mutants were generated for successive reaction with IPSL as a label for EPR measurements performed based on DENV-2 NS2B/NS3 WT. The constructs were cloned together with [redacted] of the [redacted] group and [redacted] of the group of [redacted] using the Kapa HiFi PCR kit (Kapa Biosystems Inc. Woburn, MA, USA) and following primers. (i) S71*C: 5′–GTAAAAACCCGTGCGCCGTTCAAACCAAACCG–3′, and 5′–GAACGGCGCACGGGTTTTTACCCGGTTCC–3′; (ii) S79*C: 5′–GATCCTGTCAATTACGATCTGCGAAGATG–3′, and 5′–GACATGCTACCATCTTCGCAGATCGTAATTG–3′; (iii) T118C: 5′–GTTTAAGTGCAATACGGGTACCATTGGCGCGGTTAG–3′, and 5′–CGTATTGCACTTAAACAGGCCCGGTTTGGTTTGAAC–3′; (iv) S158C: 5′–GTGACCCGTTGTGGTGCCTACGTGTCCGCG–3′, and 5′–CGTAGGCACCACAACGGGTCACGACACCATTACC–3′

Recombinant Protein Expression and Purification. The DENV2 NS2B/NS3 WT, used for nanoSPR, MST, SwitchSENSE, autodigestion experiments, and the single and double Cys mutant constructs for EPR measurements, were expressed and purified exactly as described earlier.[239] The bivalent expressed bZiPro used in crystallization experiments was expressed and purified exactly as described previously.[244] bZiPro used for crystallization was performed exactly as described previously[244] but with some modifications.[243] An additional dialysis step after immobilized metal affinity chromatography (IMAC) for 16 h at rt. in 1 L crystallization buffer (20 mM Na-HEPES pH 7.5, 150 mM NaCl, 5 vol% glycerol, 2 mM DTT) and thrombin in a 25:1 (bZiPro: thrombin) concentration using Spectra/Por® dialysis membranes (molecular weight cut-off (MWCO): 12–14 kDa; Carl Roth, Karlsruhe, Germany). After dialysis, a reverse IMAC was performed, where the flowthrough was collected and concentrated by spin concentrators before resuming with size exclusion chromatography (SEC) in the same crystallization buffer. After SEC, bZiPro was concentrated to 40 mg/mL using Vivaspin 10 kDa MWCO spin concentrators (Sartorius, Göttingen, Germany).

Pulsed Electron Paramagnetic Resonance Spectroscopy. The DENV2 NS2B/NS3 single and double Cys mutants were spin-labeled by adding IPSL in a molar ratio of 10 to 1 (IPSL to Cys). The reaction was conducted for 4 h at rt. in SEC buffer (20 mM Tris·HCl, pH 8.0; 150 mM NaCl) under agitation and exclusion of light. Subsequent SEC under exclusion of light using a HiLoad 16/60 Superdex 200 PG column (GE Healthcare, Chicago, IL, USA) on an NGC chromatography system (Bio-Rad Laboratories, Hercules, CA, USA) was employed to remove unreacted IPSL. DEER-EPR experiments were performed as reported previously.[245] Briefly, pulsed EPR measurements were recorded at X band (9.4 GHz). Measurements were performed at 50 K, controlled by a continuous-flow helium cryostat (Oxford Instruments, Abington, UK) and a temperature controller (Oxford Instruments). Samples of 30–40 mL of 20 μM spin-labeled DENV2 NS2B/NS3 were loaded into EPR quartz tubes with a 3 mm outer diameter and shock frozen in liquid nitrogen. The four-pulse DEER sequence[246] was applied to utilize an Elexsys 580 spectrometer (Bruker, Billerica, MA, USA). Observer pulses of 16–32 ns and a pump pulse of 12 ns were used as previously established.[247] The pump pulse frequency was set to the maximum of the nitroxide EPR spectrum, the frequency separation to 65 MHz. The distance distributions were validated using the validation tool included in DeerAnalysis.[248] Poor fits (prune level of 1.15) were excluded. Interspin distances were predicted by the rotamer library approach included in the MMM software package.[249]

Microscale Thermophoresis. A Monolith NT.115 instrument (NanoTemper Technologies, Munich, Germany) was utilized to perform MST experiments. The purified DENV2 and ZIKV NS2B/NS3 proteases were labeled at lysine residues using the Protein Labeling Kit RED-NHS 2nd Generation (NanoTemper Technologies) according to the provided instructions. The labeled proteases were used in a final concentration of 5 nM, diluted in MST buffer (50 mM Tris·HCl, pH 9.0, 20 vol% glycerol, 1 mM CHAPS). A 1:1 dilution series of cpd **5** was prepared (1.60 mM, 800 μM, 400 μM, 200 μM, 100 μM, 50 μM, 25 μM, 12.5 μM, 6.25 μM, 3.13 μM, 1.56 μM, 781 nM, 391 nM, 195 nM, 97.7 nM, 48.9 nM), starting from a 20 mM DMSO stock. To assure constant experimental conditions, the first dilution was performed with MST buffer, followed by the 16 dilutions with MST buffer adjusted to 8% DMSO. Measurements were performed at 25 °C at medium MST-power using Monolith NT.115 Capillaries (NanoTemper Technologies). The obtained data were processed using the MO.Affinity Analysis software Version 2.3 (NanoTemper Technologies).

Western Blotting. Western blotting of DENV NS2B/NS3 to detect autocatalytic cleavage of the His_6 tag was performed exactly as described in the literature.[250]

Protein Crystallography. The conditions for crystallization of bZiPro were modeled after those published.[243] The purified bZiPro at 40 mg/mL was pre-incubated with the allosteric inhibitor 6 at 300 µM at rt for co-crystallization experiments. The Hangig-Drop vapor diffusion method was used on 24-well VDX hanging drop crystallization plates (Hampton Research, Aliso Viejo, CA, USA). Herefore, 1 µL protein solution with or without cpd **6** was mixed with 1 µL precipitation solution. In a matrix screening, the pH value of the 100 mM sodium acetate buffer was varied in 0.2 pH increments (pH 4.2–4.8), and the amount of the precipitant PEG-2000 was varied in 2% increments (24%-34%). The amount of ammonium sulfate was kept constant (0.2 M). Small, cube-shaped crystals grew after two weeks at rt. in several wells for apo and co-crystallized bZiPro. However, only small crystals have been obtained so far, unsuitable for X-ray diffraction.

4.2. Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases

4.2.1. Context, Project Summary, and Own Contribution

Flavonoids were the first non-competitive inhibitors of DENV NS2B/NS3, identified in the working group of [redacted], with myricetin (cpd **18**, Figure 27A) as the most promising inhibitor yielding an IC_{50}-value of 22 ± 2 µM.[251] Docking studies predicted the binding site to be the putative allosteric binding site of NS2B/NS3 adjacent to the substrate binding site, as proposed earlier.[252] The first synthetic allosteric inhibitors of flaviviral NS2B/NS3 were identified in a screening approach of an in-house library of ~250 cpds using a fluorescence-based enzyme activity assay with isolated DENV2 NS2B/NS3. This screening identified cpd **19** as a non-competitive inhibitor with an IC_{50} value of 98 ± 4 µM against DENV2 NS2B/NS3 (Figure 27A).[195] Successive SAR studies were performed by modifying the thiophene, the central amide, the nitro group, and the position of the thioether, revealing cpd **20** as a new lead cpd with an IC_{50} against DENV2 NS2B/NS3 of 4.2 ± 0.4 µM (Figure 27A) and high antiviral activity in DENV2 infected cells (DENV2 EC_{50} = 0.8 µM).[195] This study also confirmed antiviral activity in DENV-infected cells. The non-competitive mode of inhibition was determined by reduced v_{max} and constant K_M values in the presence of the inhibitors. Docking studies predicted the catechol to reside deeply buried in the cavity of the allosteric pocket, forming polar interactions with the sidechain of Asn152 and the backbone of Lys73 (Figure 27B). Besides, no further evidence for binding to the allosteric binding site of DENV2 was provided. In this publication, under the leadership of [redacted], a new series of allosteric inhibitors of DENV2 and ZIKV NS2B/NS3 that harbor a proline residue instead of the benzamide was developed. This approach improved the hydrophilicity and solubility of the cpds complying with Lipinski's rule of 5 regarding the logP[182] but lost some activity of cpd **15** against DENV2 NS2B/NS3 compared to cpd **20**. However, the inhibitory activity of cpd **15** against ZIKV NS2B/NS3 was even higher with an IC_{50} of 1.3 ± 0.1 µM, yielding a significant improvement in LLE[180] against ZIKV (cpd **20**: LLE = 0.24; cpd **15**: LLE = 3.07, Figure 27A).

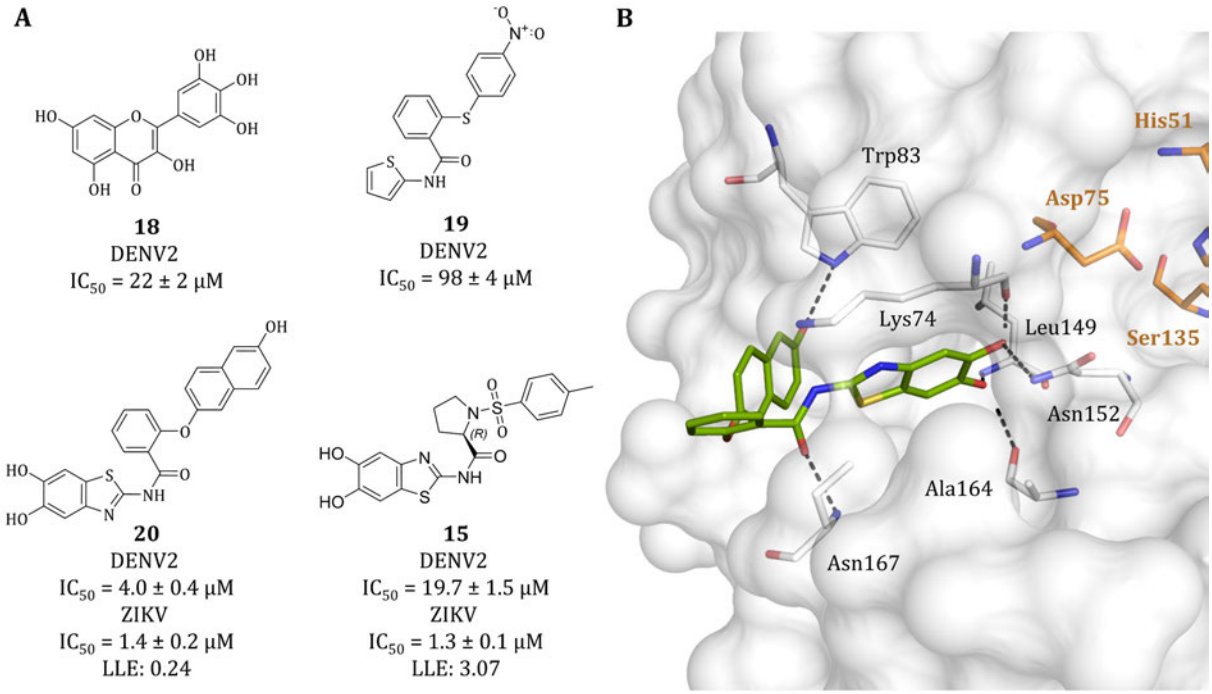


Figure 27. *Development of allosteric inhibitors in the working group of* [redacted]. (**A**) 2D structures, inhibitory activities, and ligand efficacy metrics of different stages of allosteric inhibitor development. (**B**) Predicted binding mode of cpd **20** (green sticks) in the open conformation of DENV2 (NS3: white surface. PDB-ID: 2FOM)[34]. For a clear view, only amino acids forming polar interactions (black dashed lines) with cpd **20** are shown as white sticks. The amino acids of the allosteric site are displayed as orange sticks. This figure was created using PyMOL.[44]

Using molecular docking studies, we assessed which conformation is more likely to be stabilized upon binding of allosteric inhibitors. Hence, the cpds of this study consisting of 47 binders and 16 nonbinders were docked against the allosteric site in open and closed conformations of DENV and ZIKV NS2B/NS3 (DENV open/closed PDB-ID: 2FOM[34]/3U1I[54]; ZIKV open/closed PDB-ID: 5GXJ[184]/5LC0[185]). Receiver operator characteristics (ROC) curves for this docking revealed slightly better discrimination of binders from nonbinders for 2FOM over 3U1I and 5GXJ over 5LC0, indicating binding to the open conformation of NS2B/NS3. These results suggest that open conformations represent a better suitable structural basis for the further development of allosteric inhibitors. Binding site verification was performed by introducing single Cys residues near the allosteric site of DENV2 NS2B/NS3, lacking native Cys residues. Thus, specific modification at the introduced Cys residues in the desired location was feasible (Figure 28A). Therefore, five single Cys mutants of DENV2 NS2B/NS3 of amino acids forming the allosteric site were generated (T118C, T120C, T122C, A164C, and A166C, Figure 28B) and subsequently modified by a covalent reaction with either *N*-ethylmaleimide (EMI) or *N*-benzylmaleimide (BMI). Enzyme activity studies with these modified enzymes revealed that T118C, T120C, and T122C retained most activity upon treatment with the maleimides, whereas A164C almost completely lost activity and A166C showed moderately reduced activity (Figure 28C). These results comply

with the respective positions of Ala164 and Ala166, which are deeper in the allosteric pocket than Thr118, Thr120, and Thr122. Furthermore, it suggests that this site is susceptible to allosteric inhibition by ligand binding. Fluorometric enzyme inhibition assays of cpd **15** against the three Thr to Cys mutants showed no considerable effect of these mutations but a significant reduction in inhibitory potency upon treatment of T120C and T122C with EMI and BMI, respectively (Figure 28D).

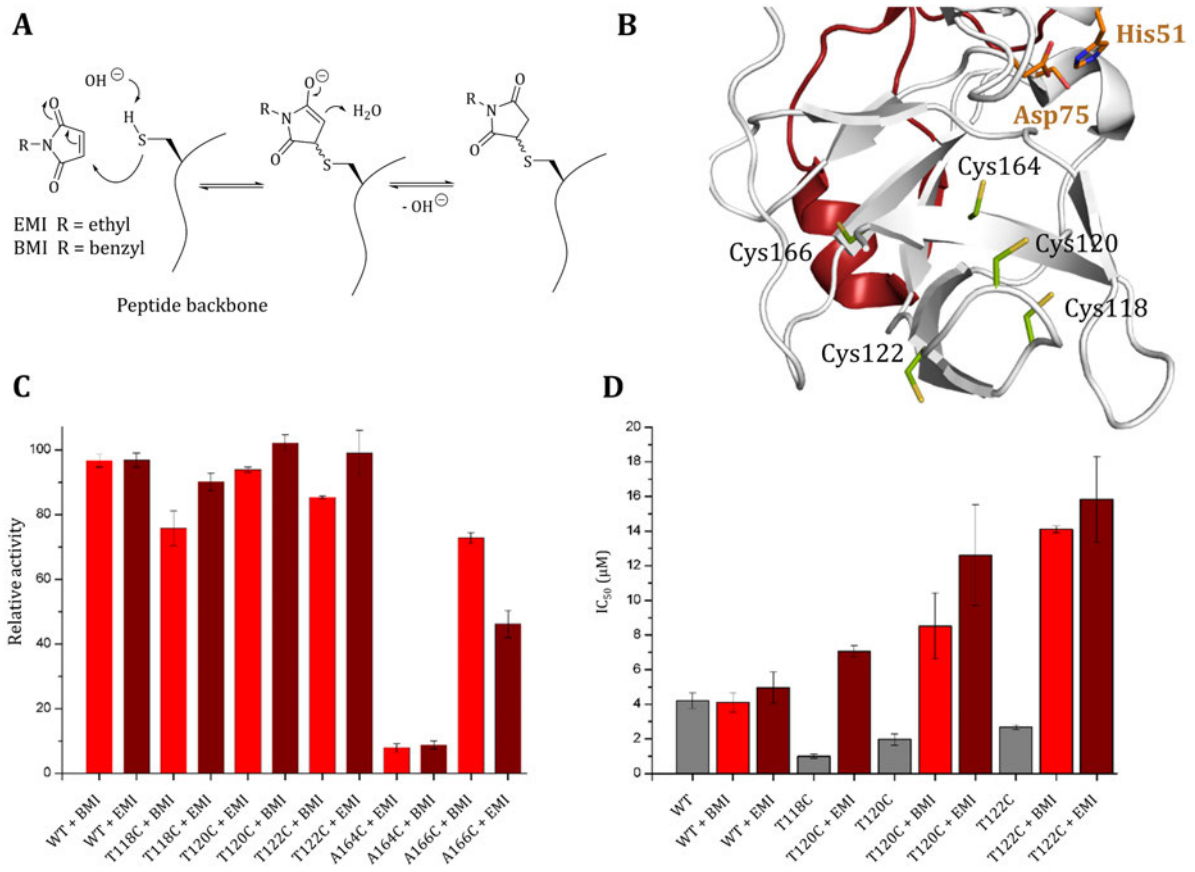


Figure 28. *Verification of the allosteric binding mode of DENV2 NS2B/NS3 by cysteine-maleimide coupling.* (**A**) The covalent reaction of maleimides with thiol groups in a Michael addition mechanism. (**B**) Position of Cys residues (green sticks), introduced into the sequence of $NS3_{pro}$ (white cartoon) by site-directed mutagenesis. $NS2B_{cf}$ (red cartoon) is in the open conformation (PDB-ID: 2FOM). The amino acids of the catalytic triad are displayed as orange sticks. This figure was created using PyMOL.[44] (**C**) Relative activities of the single Cys mutants unlabeled or reacted with either EMI (deep red) or BMI (light red). Gaussian progressed standard errors are indicated as black bars. (**D**) IC_{50} values of cpd **15** against the different constructs either unmodified (gray) or reacted with EMI or BMI.

(4) Millies, B.; Von Hammerstein, F.; Gellert, A.; **Hammerschmidt, S.**; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; Kersten, C.; Kiefer, W.; Bodem, J.; Hildt, E.; Windbergs, M.; Hellmich, U. A.; Schirmeister, T. 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

Own contribution: Enzyme expression and purification, site-directed mutagenesis of DENV2 NS2B/NS3 T122C, fluorometric enzyme activity assays, maleimide coupling, enzyme inhibition assays, writing of experimental parts & editing of the manuscript.

Contributions from other authors: Enzyme expression and purification, site-directed mutagenesis, fluorometric enzyme activity assays, enzyme inhibition assays, Inhibitor synthesis, docking studies, molecular dynamics studies, circular dichroism spectroscopy, cell permeability studies, metabolism studies, antiviral activity assays, writing of the original draft & editing of the manuscript.

This work has been published in *Journal of Medicinal Chemistry* (impact factor: 8.04).

The appended ‘Supporting Information’ represents an abridged version. The full version can be accessed online at doi: 10.1021/acs.jmedchem.9b01697.

4.2.2. Publication

The following publication quoted (within " and ") from page 190 to page 241 is exactly the same as the manuscript cited on page 189. "

Journal of Medicinal Chemistry

Article

Cite This: J. Med. Chem. 2019, 62, 11359−11382

pubs.acs.org/jmc

Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases

Benedikt Millies,[†] Franziska von Hammerstein,[†] Andrea Gellert,[†] Stefan Hammerschmidt,[†] Fabian Barthels,[†] Ulrike Göppel,[†] Melissa Immerheiser,[‡] Fabian Elgner,[§] Nathalie Jung,[#] Michael Basic,[§] Christian Kersten,[†] Werner Kiefer,[†] Jochen Bodem,[‡] Eberhard Hildt,[§] Maike Windbergs,[#] Ute A. Hellmich,*[,∥,⊥,∇] and Tanja Schirmeister*[,†,∇]

[†]Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz, Staudinger Weg 5, 55128 Mainz, Germany

[‡]Institut für Virologie und Immunbiologie, Julius-Maximilians-Universität Würzburg, Versbacher Straße 7, 97078 Würzburg, Germany

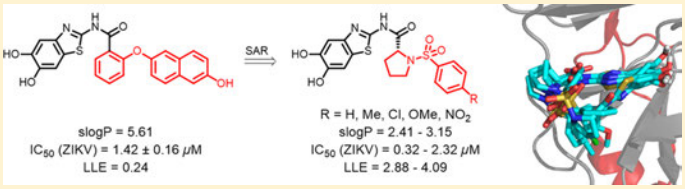
[§]Department of Virology, Paul-Ehrlich-Institut, 63225 Langen, Germany

[∥]Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz, Johann-Joachim-Becherweg 30, 55128 Mainz, Germany

[⊥]Center for Biomolecular Magnetic Resonance (BMRZ), Goethe-Universität, 60438 Frankfurt am Main, Germany

[#]Institute of Pharmaceutical Technology and Buchmann Institute for Molecular Life Sciences, Goethe University, Max-von-Laue-Straße 15, 60438 Frankfurt am Main, Germany

Supporting Information



ABSTRACT: The NS2B/NS3 serine proteases of the Zika and Dengue flaviviruses are attractive targets for the development of antiviral drugs. We report the synthesis and evaluation of a new, proline-based compound class that displays allosteric inhibition of both proteases. The structural features relevant for protease binding and inhibition were determined to establish them as new lead compounds for flaviviral inhibitors. Based on our structure−activity relationship studies, the molecules were further optimized, leading to inhibitors with submicromolar IC_{50} values and improved lipophilic ligand efficiency. The allosteric binding site in the proteases was probed using mutagenesis and covalent modification of the obtained cysteine mutants with maleimides, followed by computational elucidation of the possible binding modes. In infected cells, antiviral activity against Dengue virus serotype 2 using prodrugs of the inhibitors was observed. In summary, a novel inhibitor scaffold targeting an allosteric site shared between flaviviral NS2B/NS3 proteases is presented whose efficacy is demonstrated in vitro and in cellulo.

INTRODUCTION

In 2015, the Zika virus (ZIKV) became endemic in Brazil and infected between 400,000 and 1.6 million people. It is not only transmitted by the mosquito species *Aedes aegypti* and *Aedes albopticus*, typically associated with flaviviral distribution, but also from human to human.[1−4] In 2016, the World Health Organization declared ZIKV a public health emergency due to its appearance throughout South America and its correlation with a striking increase of neurological disorders, such as microcephaly in neonates.[5,6] Although the main symptoms of a ZIKV infection, fever, rash, and conjunctivitis are relatively mild and most infections remain asymptomatic, some of the (adult) patients develop severe neurological conditions such as the Guillain−Barré syndrome.[7−9]

ZIKV, the four different serotypes of the Dengue virus (DENV1−4), and West-Nile virus (WNV) belong to the family of Flaviviridae. The viral RNA encodes a polyprotein precursor that is processed by viral and cellular proteases. Here, the viral NS2B/NS3 serine protease plays an essential role. It cleaves the polyprotein precursor into the structural proteins that are components of the virion and the non-structural (NS) proteins involved in viral replication and maturation. Additionally, the NS2B/NS3 protease suppresses the human immune response by cleaving the stimulator of interferon genes (STING).[10,11] These functions render the

Received: October 14, 2019
Published: November 26, 2019

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

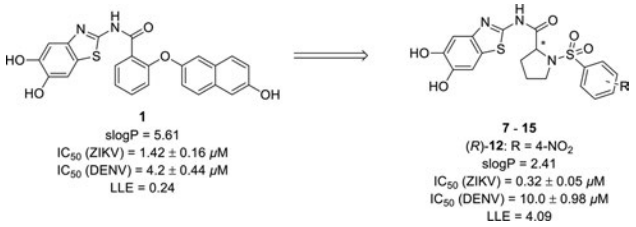


Figure 1. The inhibitor of the DENV NS2B/NS3-protease **1**[27] was modified to improve affinity, hydrophilicity, and lipophilic ligand efficiency (LLE).

flaviviral NS2B/NS3 protease a promising target for the development of antiflaviviral drugs.[12]

The protein NS3 consists of a helicase and a serine protease domain. However, NS3 requires the NS2B protein as a cofactor for full catalytic activity.[13,14] In structures of the NS2B/NS3 complex, two different conformations have been observed, the open and the catalytically active closed conformation, where NS2B is wrapped around NS3, forming a β-turn with its C-terminus, which contributes to the P2 and P3 binding sites of the protease.[13,15] In the X-ray structures of the open conformation, only the N-terminus of NS2B is resolved. The closed, active conformation is obtained upon binding of a substrate or a competitive inhibitor, while an equilibrium between the two conformations is observed in solution.[16−18]

Most of the reported competitive inhibitors for the ZIKV NS2B/NS3 protease or the related protease from, for example, WNV and DENV proteases,[19−21] share a peptidic or peptidomimetic structure. Some but not all of them additionally contain electrophilic warheads targeting the active site serine residue. However, because of the preference of the S1 and S2 binding pockets for cationic residues, effective competitive inhibitors display low membrane permeabilities.[22] As an alternative route to obtain antiflaviviral compounds, screening campaigns also identified noncompetitive inhibitors.[23−25] In the case of the DENV protease, noncompetitive inhibitors led to a shift of the protein conformation equilibrium to the open conformation. The putative allosteric binding site was proposed to be located close to Asn152. This residue was hence termed the "molecular switch" between the open and closed conformation.[15,26]

In previous studies,[27] we reported a new series of noncompetitive inhibitors for the DENV NS2B/NS3 protease (e.g. **1**, Figure 1). Since the use of these inhibitors as lead compounds for further drug development and as biophysical probes for structural and dynamic studies of flaviviral proteases was found to be hindered by their comparably high hydrophobicity, we aimed to optimize these inhibitors in terms of affinity, hydrophilicity, and lipophilic ligand efficiency (LLE). In this regard, the resulting new proline-based inhibitors indeed (Figure 1) showed improved properties.[27]

The antiflaviviral efficiency of the new inhibitors was demonstrated in cell-based assays. Based on the metabolism profiles of the compounds and the differences in inhibition potency between the methoxy-protected and the nonprotected compounds, we could determine the relevance of specific hydroxyl groups for allosteric inhibition capability.

The binding to the proposed allosteric site was verified using single cysteine point mutations and their reaction with maleimides, which block the allosteric site for the inhibitors and, intriguingly, also led to inhibition of the enzyme by itself, thus demonstrating cross-talk between the allosteric and the active site and paving the way for the future development of covalent inhibitors at this site.

The possible binding modes of the inhibitors were elucidated by molecular docking and molecular dynamic simulations.

RESULTS AND DISCUSSION

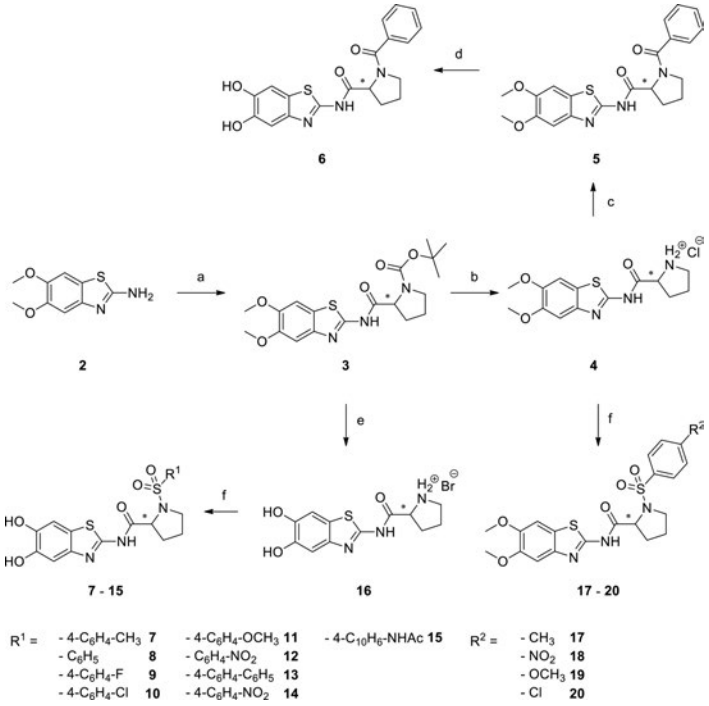
Considerations on New Lead Structures. Because of the high global incidence numbers and their severity, new drugs against flaviviral diseases are urgently needed. The flaviviral NS2B/NS3 protease essential for viral maturation seems to be a particularly promising inhibitor target. Due to the high sequential (>50% identity) and structural (1.11 Å RMSD) similarity between DENV and ZIKV NS2B/NS3 proteases (DENV2, PDB: 2FOM[28] and ZIKV, PDB: 5GXJ[15]), it should in principle be possible to use the same or structurally very similar inhibitors against both targets. Indeed, compound **1**, first developed by us against the DENV protease, was also active against the ZIKV protease and showed an even lower IC_{50} value (ZIKV: 1.41 ± 0.16 μM; DENV2: 4.2 ± 0.44 μM[27]). However, this molecule, based on a 5,6-dihydroxybenzo[*d*]thiazole-2-amine scaffold linked to a diaryl-ether moiety via an amide bond, suffers from poor water solubility and high lipophilicity. These issues significantly hamper not only the pharmacological use of this compound but also its application in cocrystallization or as a probe in, for example, NMR studies. Nonetheless, compound **1** is interesting because it displays a noncompetitive inhibition against both DENV2 and ZIKV proteases. It thus presumably interacts with the same, previously proposed, putative allosteric site. In this study, we therefore set out to systematically vary compound **1** for improved protease affinity and hydrophilicity. Since we previously identified the dihydroxy-substituted benzothiazole unit as an important structural feature, this fragment was kept constant.[27]

The ortho-substituted aromatic ring system in the center of **1** was exchanged against a proline residue to enhance solubility. Additionally, this opened the possibilities for further derivatization via facile amine coupling reactions to sulfonamides and carboxylic acid amides (Figure 1).

Chemistry. The benzothiazole building block 2-amino-5,6-dimethoxybenzo[*d*]thiazole **2** was synthesized as reported previously.[27] TBTU-mediated amide coupling with either (*S*)- or (*R*)-boc-protected proline resulted in the benzothiazole amides **3**, which were deprotected at the amine group with HCl in dioxane yielding the amines **4**. These were coupled

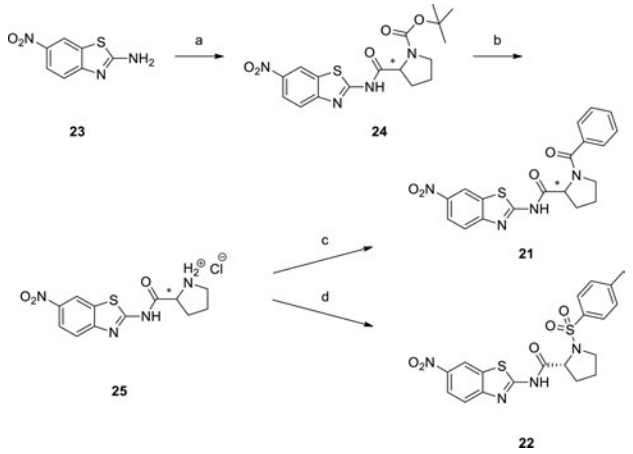
DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Scheme 1. Synthesis of Proline-Based Inhibitors 6, 7−15, and 17−20[a]



[a]Reagents and conditions: (a) boc-(*S*)- or -(*R*)-Pro, TBTU, HOBt, DIPEA, EtOAc, r.t., 97−99%; (b) HCl, dioxane, r.t. 91−94%; (c) benzoyl chloride, pyridine, r.t., acetonitrile, 90−96%; (d) BBr_3, DCM, −78 °C to r.t., 13−60%; (e) BBr_3, DCM, −78 °C to r.t., 86−97%; (f) R^1SO_2Cl, pyridine, acetonitrile, r.t. 39−99%.

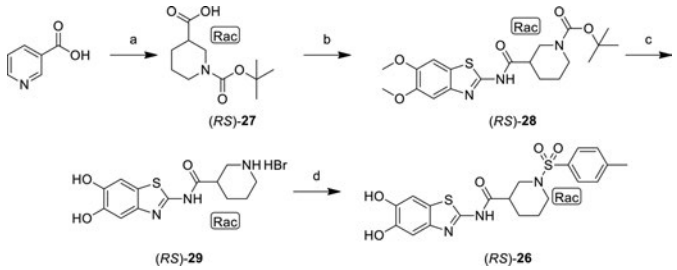
Scheme 2. Synthesis of Inhibitors with 7-Nitrobenzo[*d*]thiazole Units 21 and 22[a]



[a]Reagents and conditions: (a) boc-(*S*)- or -(*R*)-Pro, TBTU, HOBt, DIPEA, DMF, r.t., 59−93%; (b) HCl, dioxane, r.t., 99%; (c) benzoyl chloride, pyridine, acetonitrile, r.t., 68−79%; (d) TosCl, pyridine, acetonitrile, 88%.

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Scheme 3. Synthesis of Inhibitor (*RS*)-26[a]



[a]Reagents and conditions: (a) 1. PtO_2, H_2, HCl, r.t. 2. boc_2O, NaOH, dioxane, H_2O, r.t., 17%; (b) 2-amino-5,6-dimethoxybenzo[*d*]thiazole, TBTU, DIPEA, EtOAc, r.t., 84%; (c) BBr_3, DCM, −78 °C to r.t., 99%; (d) TosCl, pyridine, r.t., 27%.

with benzoyl chloride to obtain amides **5**, which contain two methoxy groups at the benzothiazole fragment. The inhibitors **6** with two free hydroxyl groups at the benzothiazole fragment were obtained by reaction with BBr_3 (Scheme 1).

The sulfonamide derivatives **7**−**15** were obtained by deprotection of either (*S*)-**3** or (*R*)-**3** with BBr_3 yielding concomitantly the free amino and the free hydroxyl groups (compounds **16**) and by subsequent reaction with various aromatic sulfonyl chlorides (Scheme 1). The different sulfonyl moieties were chosen to cover a broad range of compounds with polar and unpolar residues. The respective sulfonyl chlorides were either commercially available or prepared by aromatic substitution with chlorosulfonic acid.[29−31] The dimethoxy-substituted compounds with the sulfonamide structure (**17**−**20**) were synthesized by deprotection of the boc-protected proline fragment with HCl in dioxane and reaction with the respective aromatic sulfonyl chlorides (Scheme 1). To check if the amide coupling with boc-proline or one of the deprotection steps led to racemization, the enantiomeric purity, expressed as enantiomeric ratio (er), was determined exemplarily for (*R*/*S*)-**7** and (*R*/*S*)-**18** by HPLC using a chiral column. For (*R*)-**18** and (*S*)-**18**, no racemization was detected, while for (*R*)-**7** and (*S*)-**7**, in both cases, a ratio of 93:7 of the enantiomers was determined (Figure SI1), indicating that not the proline coupling step but the subsequent deprotection reaction may lead to partial racemization. The nitro-substituted derivatives (*R*)-**21**, (*S*)-**21**, and (*R*)-**22** were synthesized starting from the nitro-substituted benzothiazole **23** by a TBTU-mediated reaction with boc-protected proline to **24**, subsequent deprotection with HCl in dioxane yielding **25**, and final coupling with benzoyl chloride and tosyl chloride to yield the benzamide and the sulfonamide derivatives (Scheme 2).

(*RS*)-**26** was synthesized as a racemic mixture starting from nicotinic acid, which was reduced by H_2 with PtO_2 catalysis, followed by boc-protection to yield boc-protected nipecotic acid (*RS*)-**27**. After coupling with **2** to yield (*RS*)-**28** and deprotection with BBr_3 to yield (*RS*)-**29**, (*RS*)-**26** was obtained by conversion with tosyl chloride (Scheme 3).

The syntheses of the inhibitors (*R*)-**30**, (*S*)-**30**, **31**, and **32** were carried out similarly, starting from the commercially available amino acids or their boc-protected derivatives (Figure 2).

Derivative **33** was obtained by coupling of **2** with cyclohexanecarbonyl chloride to **34**, followed by deprotection of the methoxy groups with BBr_3 (Scheme 4).

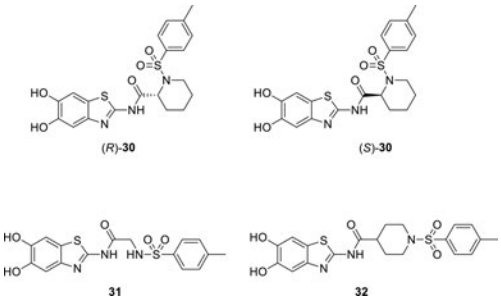


Figure 2. Overview of synthesized inhibitors derived from different amino acids.

Compound **35**, which is a hybrid molecule consisting of the 2-amino-5,6-dihydroxybenzo[*d*]thiazole moiety and the reported competitive inhibitor **36**,[19] was synthesized as shown in Scheme 5. First, the amino group of anthranilic acid **37** was boc-protected to yield **38**. TBTU-mediated amide coupling with 2-amino-5,6-dimethoxybenzo[*d*]thiazole resulted in **39**, which was deprotected at the amine and methoxy groups with BBr_3 generating amine **40**. The sulfonamide **35** was obtained by reaction with tosyl chloride.

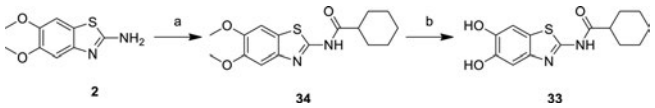
1 and its precursor **41** (Figure 3) were synthesized as reported previously.[27]

Structure−Activity Relationship. The newly synthesized inhibitor library consisting of 38 compounds was tested on the purified ZIKV and DENV2 NS2B/NS3 proteases in enzymatic assays employing the fluorogenic AMC-substituted peptide boc-GRR-7-amino-4-methylcoumarin (boc-GRR-AMC) (Table 1).

All proline-based inhibitors displayed weaker activity against the DENV protease compared to the reference compound **1**. This holds true for the nitro-substituted inhibitors **21** and **22**, as well as for the dihydroxy-substituted compounds. In contrast, with the exception of the aforementioned nitro-substituted compounds, similar or even better inhibition was observed against the ZIKV protease (IC_{50} values in the range of 0.32−3.17 μM), demonstrating a certain selectivity for this protease. Also, the introduction of the sulfonamide group ((*R*)-**8** and (*S*)-**8**) instead of the benzamide moiety ((*R*)-**6** and (*S*)-**6**) yielded similar inhibition of the ZIKV protease. Even if some of the (*R*)-configured enantiomers (e.g., (*R*)-**6** and (*R*)-**11**) show lower IC_{50} values compared to their (*S*)-

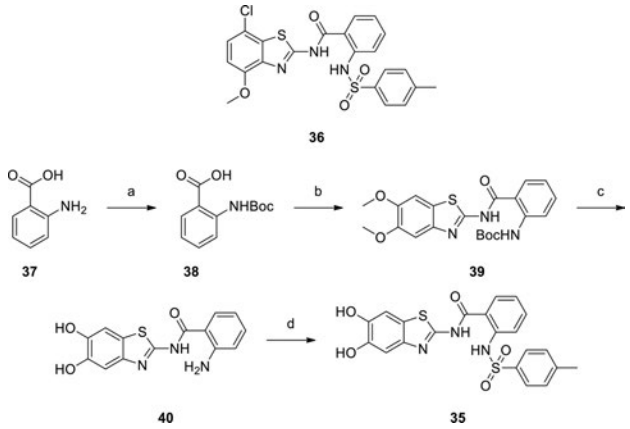
DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Scheme 4. Synthesis of 33[a]



[a]Reagents and conditions: (a) cyclohexanecarbonyl chloride, NEt_3, DCM, r.t., 71%; (b) BBr_3, DCM, −78 °C to r.t., 85%.

Scheme 5. Synthesis of 35, a Hybrid Molecule Consisting of the 5,6-Dihydroxybenzothiazole Moiety and the Reported Inhibitor 36[19a]



[a]Reagents and conditions: (a) boc_2O, dioxane, water, r.t., 69%; (b) 2-amino-5,6-dimethoxybenzo[*d*]thiazole, TBTU, HOBt, DIPEA, DMF, r.t., 97%; (c) BBr3, DCM, −78 °C to r.t., 86%; (d) TosCl, pyridine, r.t., 41%.

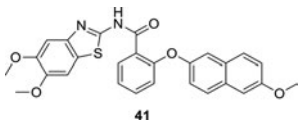


Figure 3. Structure of **41**.

enantiomers, the differences between the values are not significant in light of the obtained standard deviations. The reduced activities of the nitro-substituted compounds (*R*)-**21**, (*S*)-**21**, and **22** with a 7-nitrobenzo[*d*]thiazole moiety are consistent with our previous results.[27] Interestingly, a trend was observed regarding the hydrophobicity of the inhibitors: While the large, hydrophobic biphenyl or naphthyl groups of inhibitors (*R*)-**13**, (*S*)-**13**, (*R*)-**15**, and (*S*)-**15** led to IC_{50} values in the low micromolar range, the more hydrophilic ones, such as (*R*/*S*)-**10**, (*R*/*S*)-**11**, and (*R*/*S*)-**12**, showed submicromolar IC_{50} values.

The lowest IC_{50} value was found for the 4-nitrophenyl-substituted sulfonamide (*R*)-**12**. However, for further development, we did not select this nitro compound but the tosyl-substituted derivatives (*R*)-7 and (*S*)-7 as new lead structures due to several reasons: Aromatic nitro groups are considered problematic in drug design campaigns due to possible mutagenicity and toxicity.[32,33] Furthermore, nitro-substituted compounds may lead to artifacts in fluorometric assays due to quenching effects, which have to be accounted for by determination of correction factors.[34] The facile synthetic availability of the starting material in the case of the tosyl derivative in comparison to the similarly active *p*-chloro or *p*-methoxy analogues (*R*)-**10** and (*S*)-**10** and (*R*)-**11** and (*S*)-**11** and the better inhibition of the DENV protease by the tosyl compounds led to the choice to select (*R*)-7 and (*S*)-7 as new lead structures from which further compounds were derived (Table 2).

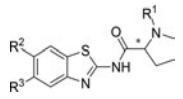
Exemplarily for different compound structures, noncompetitive inhibition of the ZIKV protease was confirmed by determining the IC_{50} values in the presence of different substrate concentrations for **1** as the first lead compound, (*R*)-7 as the new lead compound, (*R*)-**21** as one of the compounds with a nitro group, and the hybrid molecule **35** (Figure SI2). The inhibition was found to be independent of the substrate concentration, demonstrating a noncompetitive inhibition with respect to the substrate, in all cases.

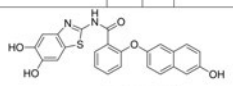
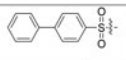
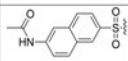
The K_D values of the tosyl derivatives' ((*R*)-7 and (*S*)-7) interaction with the ZIKV protease were also determined by microscale thermophoresis (MST) to be 0.49 μM ((*R*)-7) and 0.24 μM ((*S*)-7). IC_{50} values are a measure of the the inhibitory activities of the compounds, determined by a functional assay, which measures the residual hydrolysis of a substrate in the presence of an inhibitor. K_D values, on the other hand, are derived from the biophysical binding events between the inhibitor and enzyme alone, in the absence of a substrate. The slightly lower K_D values in comparison to the IC_{50} values thus show that the inhibitor binding does not lead to a decrease of the enzymatic activity to the same degree.

Since, as discussed above, the different substituents on the sulfonamide structure did not have a substantial impact on the affinity, in the next step, the sulfonamide group was kept

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Table 1. Inhibition of ZIKV and DENV2 Proteases[a]



Cpd ID	R^1	R^2	R^3	Configuration	ZIKV IC_{50} [μM] or %[a]	DENV2 IC_{50} [μM] or %[a]
1					1.41 ± 0.16	4.2 ± 0.44[27]
(*R*)-**6**		OH	OH	*R*	0.94 ± 0.22	33%
(*S*)-**6**		OH	OH	*S*	1.55 ± 0.18	28%
(*R*)-**21**		NO_2	H	*R*	21.9 ± 0.9	33.9 ± 0.6
(*S*)-**21**		NO_2	H	*S*	44%	41%
22		NO_2	H	*R*	31%	18%
(*R*)-**7**		OH	OH	*R*	1.32 ± 0.10	19.7 ± 1.5
(*S*)-**7**		OH	OH	*S*	2.32 ± 0.19	25.2 ± 3.0
(*R*)-**8**		OH	OH	*R*	1.86 ± 0.49	21%
(*S*)-**8**		OH	OH	*S*	1.43 ± 0.29	27%
9		OH	OH	*R*	1.03 ± 0.17	17%
(*R*)-**10**		OH	OH	*R*	1.40 ± 0.25	34%
(*S*)-**10**		OH	OH	*S*	0.93 ± 0.06	39%
(*R*)-**11**		OH	OH	*R*	0.86 ± 0.15	28%
(*S*)-**11**		OH	OH	*S*	1.32 ± 0.21	36%
(*R*)-**12**		OH	OH	*R*	0.32 ± 0.05	10.00 ± 0.98
(*S*)-**12**		OH	OH	*S*	0.93 ± 0.04	8.58 ± 1.37
(*R*)-**13**		OH	OH	*R*	2.47 ± 0.29	10.9 ± 2.3
(*S*)-**13**		OH	OH	*S*	3.17 ± 1.58	13.5 ± 1.6
(*R*)-**14**		OH	OH	*R*	1.11 ± 0.23	33%
(*S*)-**14**		OH	OH	*S*	1.78 ± 0.52	25%
(*R*)-**15**		OH	OH	*R*	1.35 ± 0.28	18%
(*S*)-**15**		OH	OH	*S*	2.95 ± 0.45	29%

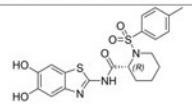
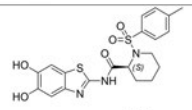
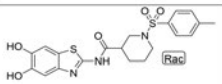
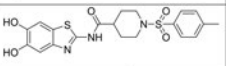
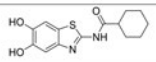
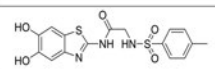
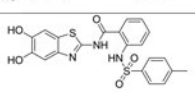
[a]IC_{50} values are indicated as means ± standard deviation from three independent measurements, each performed in duplicates. Percentage inhibition (%) was determined at 20 μM inhibitor concentration.

constant, while the proline moiety was exchanged against other amino acids.

Inhibitors (*R*)-**30** and (*S*)-**30**, containing a pipecolinic acid instead of the proline moiety, showed IC_{50} values of 0.58 and 0.51 μM on the ZIKV NS2B/NS3 protease (Table 2), respectively. Again, the configuration of the amino acid did not influence the inhibition.

Changing the position of the N atom of the cyclic amino acid and thus the conformation of the molecule yielded inhibitors (*RS*)-**26** and **32**. These compounds showed 10 fold higher IC_{50} values of 4.55 and 5.95 μM, respectively, compared to the pipecolinic acid derivatives (*R*)-**30** and (*S*)-**30**. On the other hand, the exchange of the ring system to the open, more flexible glycine-based inhibitor **31** led to an IC_{50} value of 0.81 μM. These results suggest that a defined inhibitor conformation, which is either fixed by the rigid pipecolinic acid or which can at least partially be adopted by the flexible

Table 2. Inhibition of ZIKV and DENV2 Proteases by Inhibitors Based on the Exchange of the Proline Structure[a]

Cpd ID	Structure	ZIKV IC_{50} [μM] or %[a]	DENV2 IC_{50} [μM] or %[a]
(*R*)-**30**		0.58 ± 0.05	29 ± 4
(*S*)-**30**		0.51 ± 0.44	13 ± 5
(*RS*)-**26**		4.55 ± 0.92	34%
32		5.95 ± 0.65	41%
33		36%	20%
31		0.81 ± 036	23± 4
35		6.43 ± 2.76	15.3 ± 2.3

[a]IC_{50} values are indicated as means ± standard deviation from three independent measurements, each performed in duplicates. Percentage inhibition (%) was determined at 20 μM inhibitor concentration.

glycine linker might be necessary for good inhibition potency. The relevance of the sulfonamide part is apparent since compound **33**, which lacks this moiety, is only weakly active.

In a recent study, Lee and co-workers reported the competitive inhibition of the ZIKV protease using nonpeptidic active site-directed inhibitors (**36**) (Scheme 5).[19] To evaluate which structural features are necessary for competitive or noncompetitive inhibition, we synthesized a hybrid compound (**35**), consisting of the 2-amino-5,6-dihydroxybenzo[*d*]thiazole moiety and the reported competitive inhibitor **36**. **35** displayed an IC_{50} value of 6.43 ± 2.76 μM at a substrate concentration of 100 μM. To test for competitive versus noncompetitive inhibition, the IC_{50} values were determined at different substrate concentrations (50, 75, 100, 150, and 200 μM) according to Dixon.[35] Since they were found to be independent of the substrate concentration (5.36 ± 1.40 μM), this strongly suggests a noncompetitive binding mode with respect to the substrate.

The different inhibition properties of **35** and **36** showing noncompetitive and competitive behaviors, respectively, demonstrated that the inhibition mode is mainly influenced by the substitution pattern of the benzothiazole moiety. This further highlights the importance of the hydroxy groups for noncompetitive binding.

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Similarly to previously reported compounds,[27] the dimethoxy-substituted derivatives **41** and **17**−**20** were shown to be inactive at 20 μM.

In summary, our results with the dimethoxy-substituted compounds, the hybrid compound **35** and the nitro-substituted compounds **21** and **22** highlight that the dihydroxy substitution pattern at the benzothiazole fragment is the primary determinant for accomplishing succesful noncompetitive inhibition.

To analyze the hydrophilicity of the new compounds, their *Slog*P values were calculated and compared to their retention times, measured with RP-HPLC, which showed a good correlation (see the Supporting Information part).[36] Using the *Slog*P and pIC_{50} values for inhibition of the ZIKV protease, lipophilic ligand efficiency (LLE) plots were prepared (Figure 4). These data confirmed the significant improvement of the new inhibitors compared to **1**, with several compounds exhibiting LLE values >3 or even >4 ((*R*)-**12**).[37]

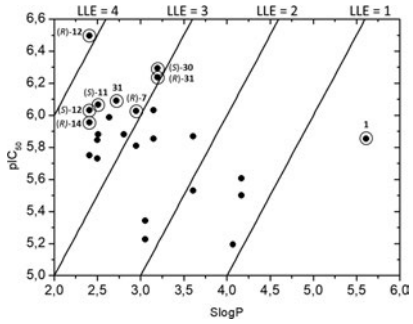


Figure 4. Lipophilic ligand efficiency (LLE) plots with the *Slog*P and pIC_{50} values for inhibition of ZIKV protease for all compounds with catechol and proline substructure **1**, **6**−**15**, **26**, **30**, **31**, **32**, and **35**.

Probing the Putative Allosteric Drug Binding Site by Site-Directed Mutagenesis. In previous docking studies with inhibitor **1**, the hydroxyl group of the naphthalene moiety was proposed to interact by hydrogen bonding with Thr120 of a putative allosteric pocket of the DENV protease. Similarly, the two hydroxyl groups of the benzothiazole moiety were proposed to interact with Asn152 and Lys73 in the same pocket.[27,38−40] To assess wether this inhibitor indeed binds into this putative allosteric pocket and to probe which residues of this binding pocket are relevant for interactions, several amino acids (T118, T120, T122, A164, and A166) were mutated individually to cysteine in the DENV protease to be able to modify the protease at single, specific sites with maleimides. The DENV protease was chosen because it does not contain any native cysteines. In contrast, the wild-type ZIKV protease inherently contains the Cys residues Cys80 and Cys143. A mutation of the ZIKV protease's cysteines to serines is principally also possible but affects the catalytic protease activity and thus possibly the overall structure.[13] Due to the overall high similarity of DENV and ZIKV NS2B/NS3, the DENV protease was thus deemed a suitable target for mutation and modification studies. Importantly, the inhibitor used for these studies with the maleimide-modified Cys mutants, namely, compound **1**, inhibits both proteases in the low micromolar range.

Thr120 was chosen for the mutation studies because of the interactions with the inhibitor proposed by docking studies. In some lower ranked poses, interactions with Thr118 were also proposed. Thr122 was chosen to evaluate if the inhibitor performs dynamic interactions with these closely located residues. Ala164 and Ala166 are located deeper inside the pocket and their Cys mutations and subsequent maleimide modifications were expected to block the binding pocket nearly completely. Subsequently, the newly introduced cysteine residues were irreversibly modified with the maleimide compounds *N*-benzylmaleimide (BMI) and *N*-ethylmaleimide (EMI). The enzyme activity of the modified DENV NS2B/NS3 cysteine mutants was determined by monitoring the

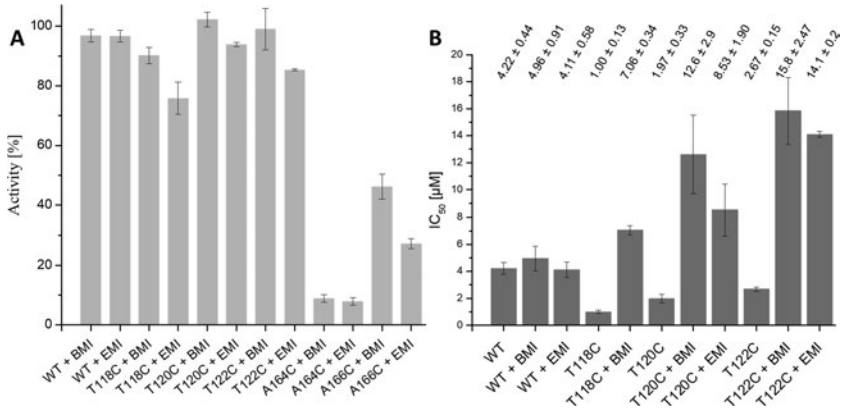


Figure 5. (A) Percentaged enzymatic activities of wild-type and single cysteine mutant DENV proteases modified with *N*-benzylmaleimide (BMI) or *N*-ethylmaleimide (EMI). Measurements were carried out with maleimide concentrations of 20 μM and their activity values were compared to the respective DMSO controls without maleimide (100%). (B) IC_{50} values of **1** on wild-type and mutant DENV proteases in the absence or presence of *N*-benzylmaleimide (BMI) or *N*-ethylmaleimide (EMI). Exact values are indicated above the respective bar. Activity and IC_{50} values are indicated as means ± standard deviation from three independent measurements, each performed in duplicates.

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

hydrolysis of the fluorogenic substrate boc-GRR-7-amino-4-methylcoumarin (boc-GRR-AMC) (Figure 5A).

For all mutants modified with maleimides except A164C and A166C, no dramatic changes in activity compared to the WT protein were observed. Intriguingly, the A164C mutant was almost completely inhibited by both maleimides even in the absence of additional inhibitory drugs, and the BMI-modified NS2B/NS3 A166C showed a reduction in activity to ~50% of the WT, while the EMI-modified NS2B/NS3 A166C showed one of ~70%. To demonstrate that the structural integrity of the DENV NS2B/NS3 A164C and A166C mutants was not compromised, circular dichroism (CD) spectra for both enzymes in the presence and absence of BMI or EMI were recorded. The resulting spectra (Figure SI5) showed that the respective protein's secondary structure is retained after maleimide binding. Overall, the decrease in activity was higher for EMI. While smaller differences between EMI- and BMI-modified enzymes were observed for T118C, T120C, and T122C, the highest difference was observed for A166C.

In the case of the wild-type enzyme, the IC_{50} of **1** is the same in the absence or presence of a maleimide as expected for an enzyme that cannot be modified by malemides due to the absence of native cysteine residues. Furthermore, the IC_{50} values of the lead structure **1** in the absence and presence of maleimide modifications were determined for those mutants, which still exhibited an activity >80% compared to WT in the presence of the maleimide (Figure 5B). While the IC_{50} values for inhibition by compound **1** of all constructs in the absence of the maleimide modification were in a similar range or even slightly lower compared to the WT enzyme, the inhibition of the maleimide-modified mutants was strongly impaired with 4- to 7-fold higher IC_{50} values. These increased IC_{50} values of **1** measured on the Cys constructs modified with maleimides strongly indicate that the interaction of the inhibitor **1** with the allosteric binding pocket is perturbed, thus providing a rational basis for docking studies and further development of inhibitors targeting this binding site. In the cases of the T120C and T122C mutants, where the modification experiment was done with both EMI and BMI, slightly higher IC_{50} values were observed with BMI. A possible explanation is the higher steric hindrance by the benzyl group of BMI compared to the smaller ethyl group of EMI.

A164 and A166 are in the center of the proposed binding pocket (Figure 6).[28] The strong inhibition of the respective cysteine mutant enzymes A164C and A166C and especially of the A164C mutant after their modification by maleimides shows that this part of the binding pocket is highly susceptible to allosteric inhibition.

Docking Studies. To date, no structure of a flaviviral NS2B/NS3 protease in complex with a noncompetitive, allosteric inhibitor is available. (For a discussion of the recently published complex structure 6MO0, see below.[41−43]) Therefore, molecular docking studies were performed to elucidate the potential binding modes of the inhibitors presented herein. Binding site identification using the druggability prediction tool LeadIT DoGSite Scorer proposed the same pocket as indicated by the above described experiments with the Cys mutants.[44,45] The allosteric site is in principle present in both the open and closed conformation of the protease.[38,40] To identify the most plausible conformation of the enzyme in the complex with the inhibitors and to identify the most plausible binding mode of the ligands within this complex, docking studies were performed on both

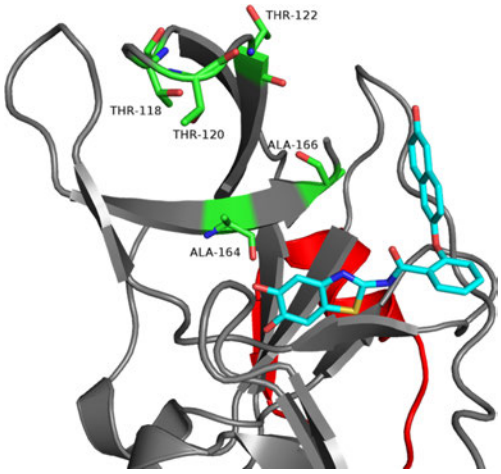


Figure 6. Position of the allosteric binding pocket on the DENV NS2B/NS3 protease with predicted binding modes of **1** in complex with the open conformation of DENV protease with residues that were mutated to cysteine highlighted in green. (Figure based on the open conformation, PDB: 2FOM).[28] NS2B is shown in red and NS3 is shown in gray. The atoms of compound **1** are colored according to the CPK color scheme, and the carbon atoms are shown in cyan. The amino acids of the allosteric site, which were mutated to cysteine residues, are indicated in the same way, while the carbon atoms are shown in green.

conformations of the DENV and the ZIKV proteases. For the structurally closely related inhibitors sharing the same 2-amino-5,6-dihydroxybenzo[*d*]thiazole core scaffold, the key interactions are expected to be conserved and their binding modes should be similar. However, for these compounds, the scoring function of the docking procedure should also be able to discriminate binders from nonbinders. Therefore, a visual binding mode inspection and receiver operator characteristic (ROC) analysis of the compounds, that were reported previously,[27] and herein were conducted for the closed and open conformation of the DENV and ZIKV NS2B/NS3 proteases.

The most stable predicted binding mode of ligands carrying the essential 2-amino-5,6-dihydroxybenzo[*d*]thiazole scaffold was found for the open conformation of the DENV protease (PDB code: 2FOM). In contrast, poses within the closed DENV (PDB code: 3U1I), as well as the open and closed ZIKV (PDB codes: 5LC0 and 5GXJ) conformations, were less stable (Figure SI6). The key interactions between the protein and the ligands are exemplarily described for (*R*)-7 and (*S*)-7 (Figure 7A).

Here, the catechol moiety of either compound (*R*)-7 or (*S*)-7 forms H-bonds to the backbone oxygen of Ala164 and Lys74. The hydroxyl group at the 5-position also acts as an H-bond acceptor for the Asn152 side chain. For compounds carrying a 5-nitro moiety instead of the catechol substructure, the nitro-oxygen atoms act as H-bond acceptors for the Asn152 side chain and Leu149 backbone (Figure SI7A). Residues Leu76, Gly148, and Ile165 surround the benzothiazole scaffold resulting in a hydrophobic contact area forming van der Waals interactions. The amide moiety acts as an H-

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

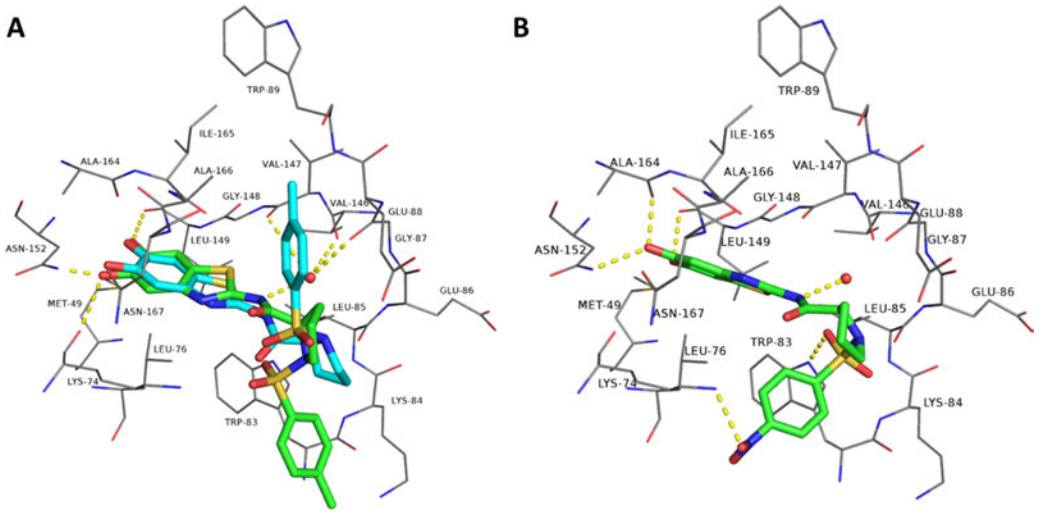


Figure 7. Predicted binding modes of (A) compounds (*R*)-7 (green carbon atoms) and (*S*)-7 (cyan carbon atoms) or (B) (*R*)-**12** in complex with the open conformation of DENV NS2B/NS3 protease (PDB code: 2FOM, gray carbon atoms) based on molecular docking. Electrostatic interactions illustrated as yellow dashes for (*R*)-7.

bond donor to form indirect interactions mediated through a water molecule (Water-247). This water molecule's presence is supported by electron density (found in the 2Fo-Fc map, PDB code: 2FOM) in the crystal structure, and it is likely to act as an H-bond donor to backbone oxygen atoms of Gly87, Val146, or Val147 and as an acceptor for a second water molecule (Water-238), which is likely to be displaced by the inhibitors (Figure SI7B). Depending on the respective proline configuration, the tosyl moiety is either oriented toward Trp83 ((*R*)-7) or Ala166 and Glu88 ((*S*)-7) to form hydrophobic interactions. Notably, the sulfonamide moiety does not form extensive H-bonds, and only for a few ligand interactions with Lys74, Trp84, or Asn167 side chains were predicted (Figure 7). These results indicate that the type of linker (amide or sulfonamide), as well as the configuration of the proline substructure, is of minor importance as long as hydrophobic moieties can address one of these mainly hydrophobic subpockets. This also explains the overall similar binding affinity of inhibitors containing a hydrophobic moiety attached to their proline linker (Table 1). For (*R*)-**12**, an additional H-bond to the Lys74 side chain was found, which might be the reason for the slightly increased affinity of this compound. Further, the predicted binding modes agree with the observations for the mutated proteases, where the introduced cysteine residues covalently react with EMI or BMI. C164-EMI/BMI and C166-BMI seem to occupy the allosteric site reducing the enzymatic activity, while C120/122-BMI blocks or at least alters the binding site for the inhibitors reducing their IC_{50} values (Figures 5 and 6).

Besides a stable binding mode of the core scaffold, also the discrimination between binders and nonbinders was most evident for the docking against the DENV NS2B/NS3 open conformation (Figure 8 and Table SI3).

In ROC studies, high values up to a maximum of 1 mean a strong discrimination between binders and nonbinders and low numbers indicate a reduced discrimination. Using this protein

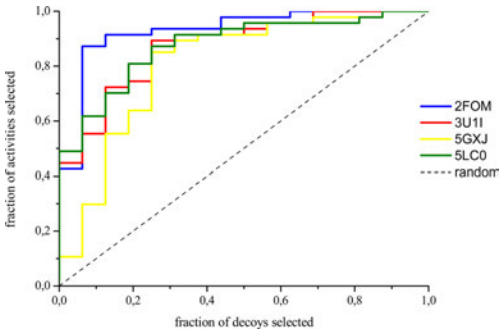


Figure 8. Receiver operator characteristic (ROC) curves for docking against DENV NS2B/NS3 open (PDB code: 2FOM, blue line, ROC area under curve (AUC) 0.93) and closed (PDB code: 3U1I, orange line, ROC AUC 0.87) and ZIKV NS2B/NS3 open (PDB code: 5GXJ, yellow line, ROC AUC 0.81) and closed conformation (PDB code: 5LC0, green line, ROC AUC 0.88). The dataset consisted of 47 binders and 16 nonbinders (defined as <10% inhibition at 20 μM concentration in the enzyme inhibition assay). Docking was performed with LeadIT-2.3.2.[46,47]

as a receptor, an ROC AUC of 0.93 was obtained, whereas it was 0.87, 0.81, and 0.88 for the DENV protease closed, ZIKV protease open, and closed conformations, respectively. Notably, the most negative docking scores were calculated for the binders in complex with the DENV protease open conformation as well (Table SI3). Considering the docking scores, the ROC values, and the stability of the binding modes, it seems very likely that the inhibitors prefer to bind to the open conformation of the protein.

However, there seem to be slight differences between the proteins that may explain the slightly higher inhibitor affinities

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

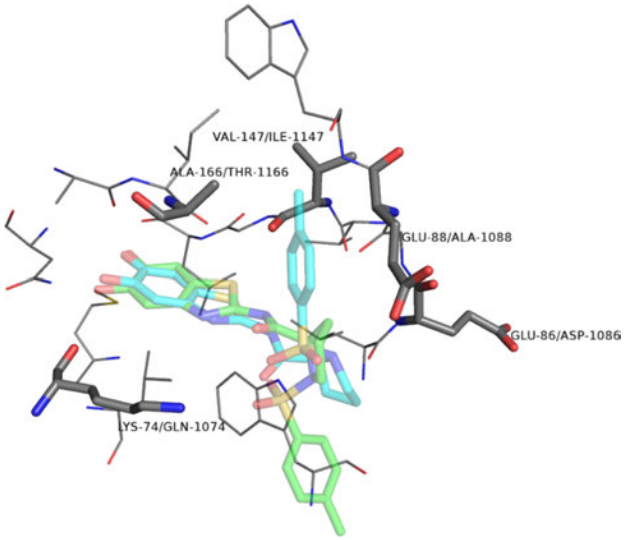


Figure 9. Different residues between DENV and ZIKV protease within the proposed binding site shown as sticks and labeled as DENV/ZIKV using DENV NS2B/NS3 open conformation (PDB code: 2FOM) as structural template. Lead compounds (*R*)-7 (green carbon atoms) and (*S*)-7 (cyan carbon atoms) are illustrated as transparent sticks for the binding site definition based on molecular docking. Point of view as in Figure 7.

for the ZIKV protease and the lower ROC AUC (0.81) for ZIKV NS2B/NS3 open in the docking studies. The crystallographically determined open conformations of the proteases differ to some extent (Cα RMSD = 0.78 Å, Figure SI11A, PDB-IDs: 2FOM for DENV and 5GXJ for ZIKV), while the closed conformations are highly similar (Cα RMSD = 0.61 Å, Figure SI11B, PDB-IDs: 3U1I for DENV and 5LC0 for ZIKV). This especially affects the proposed allosteric binding site (Figure SI11C). The C-terminal residues subsequent to the Asn152, which have been described as a switch between the active and inactive state,[15] of DENV NS2B/NS3 form a loop toward the active site, while in ZIKV, these residues subsequent to the corresponding Asn1152 are oriented directly toward residues 1088−1092 (Figure SI11B). As the crystal structure of the ZIKV NS2B/NS3 consists of a non-physiological homodimer in the asymmetric unit, the second monomer may have affected the open conformation in this structure due to crystal contacts. Indeed, a structure analogous to the one found in DENV NS2B/NS3 is prevented by ZIKV residues 1029−1032 from one monomer by clashing with the residues subsequent to Asn1152 of the second monomer (Figure SI11C). We therefore hypothesize that for ZIKV NS2B/NS3 proteases in solution, a conformation highly similar to that observed for DENV protease should be present and that this is the conformation targeted by our inhibitors. Assuming that a DENV-analogue open conformation of NS2B/NS3 also exists for the ZIKV protease, we compared the ZIKV protease allosteric site sequence with the open DENV protease structure (PDB code: 2FOM). Between the two enzymes, only five residues differ within the proposed binding site (Figure 9).

Those residues, which orient their side chains away from the pocket (E86/D1086, V147/I1147, A166/T1166; DENV protease/ZIKV protease), should have only minor impact on selectivity. The displacement of $E88_{DENV}$ against $A1088_{ZIKV}$, however, reduces the polarity of this subpocket and further enhances the space for hydrophobic substituents such as the prolinyl-tosyl moiety. The exchange of $K74_{DENV}$ against $Q1074_{ZIKV}$ might also have a minor impact on the affinity since hydrogen bonds with the ligand are predicted to be formed to the backbone (Figure 7), while side chain interactions were rarely observed in the docking. Therefore, it might not have an impact on selectivity.

To further investigate the proposed interaction mode in the open conformation, molecular dynamics (MD) simulations were carried out as they can yield additional information to identify the native binding mode from a variety of docking poses.[48] Ten nanosecond MDs of the DENV NS2B/NS3 protease open conformation (PDB code: 2FOM) were conducted in the presence and absence of lead structure compound (*R*)-7 as a high affinity binder displaying an extensive contact network with the protein (Figure 7A). Within the simulation time course, the protease showed a high stability with a backbone RMSD of 1.26 Å for the apo and 1.29 Å for the compound (*R*)-7-bound structure (Figure SI10A). The high stability of the ligand's binding mode (nonhydrogen atom RMSD of 1.46 Å) throughout the MD is a further hint for the feasibility of the predicted binding mode, as native poses usually remain stable during simulations (Figure SI10B,C).[49]

Intriguingly, both, the docking studies and MD simulations, predict polar interactions of the inhibitors with Asn152. This residue may present a formidable position to be addressed by inhibitors for flaviviral proteases, because it is not only completely conserved across NS3 proteins but is also essential for their function. Others have shown that mutagenesis of Asn152 led to inactive protein,[26] and, similarly, in our hand, the N152Q mutant of the DENV protease showed no more activity (data not shown). Another potential site to be addressed by inhibitors is the S1 loop (residues 152−167),

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

which has been proposed, at least for the ZIKV protease, to mediate the switch between open and closed conformations.[15] Binding of molecules in this region is predicted to influence the equilibrium between the active and inactive conformation, leading to an inhibition of the protease by stabilizing the open, inactive conformation.[26] Due to the high structural similarities between flaviviral proteases, one could assume that the inhibitors' binding modes should be similar for the DENV and ZIKV proteases (as well as other flaviviral proteases).

A recently published X-ray structure of DENV NS2B/NS3 in an open conformation raised the intriguing possibility of a high-resolution structure of an open, allosterically inhibited flaviviral protease.[41] However, it is currently unclear whether the observed electron density in this structure, which is in a similar location as the residues identified by us to be relevant for allosteric inhibitor binding, can be attributed to a small molecule inhibitor or parts of the protein.[42,43] Nonetheless, this study confirms the importance of this allosteric site and highlights the importance for further studies in this direction.

Antiviral Activity. To analyze the possible effects of the compounds on DENV2 replication in cells, the cytotoxicity against Vero cells was determined by MTS tests (at varying concentrations of compounds (*R*/*S*)-7, (*R*/*S*)-**12**, and (*R*/*S*)-**17**). While most of the compounds had no measurable effects, (*R*)-7 and (*R*)-**12** showed a significant reduction in substrate turnover at concentrations ≥30 μM. Therefore, (*R*)-7 and (*R*)-**12** were used at concentrations of 3 μM, while all other compounds were used at concentrations of 10 μM.

To analyze the suppression of viral replication, Vero cells were preincubated for 1 h with the respective compounds and subsequently infected with DENV2 (Figure 10). Cellular supernatants were collected after 4 days and viral RNAs were isolated. The relative viral genome copy numbers were determined with an RNA standard curve. The compounds (*S*)-**12**, (*R*)-**17**, and (*S*)-**17** decreased viral genome copy numbers at a concentration of 10 μM about one order of magnitude, while (*R*)-**12** showed also a decrease at 3 μM.

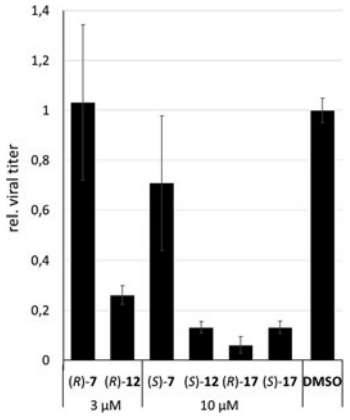


Figure 10. Compounds (*R*)-**12**, (*S*)-**12**, (*R*)-**17**, and (*S*)-**17** suppress DENV2 replication. Vero cells were preincubated for 1 h with the compounds at 3 or 10 μM and subsequently infected with DENV2. Viral RNAs were isolated from the supernatant, relatively quantified by real-time quantitative polymerase chain reaction (RTqPCR) and normalized on the DMSO control.

In preparation for testing on ZIKV replication, compounds **1**, (*R*)-7, (*S*)-7, (*R*)-**17**, (*S*)-**17**, and **41** were also tested on A549 cells and no effects on cell viability were observed up to concentrations of 5 μM. Therefore, in the subsequent assays, concentrations of 5 μM were used for all compounds. After treatment of A549 cells with the compounds and ZIKV, for **41**, a slight reduction of the viral RNA could be seen (Figure SI15), as well as a slight reduction of plaque forming units in Vero cells (Figure SI16). Furthermore, **41** (in contrast to **1**) leads to a redistribution of the ZIKV Env protein from the ER to the cytoplasm (Figure SI17). No significant reduction of viral RNA could be observed with **1**, (*R*)-7, (*S*)-7, (*R*)-**17**, and (*S*)-**17** (Figure SI18). For better comparability, the experiments were repeated for **1**, (*R*/*S*)-7, (*R*/*S*)-**12**, (*R*/*S*)-17, and **41** with Vero cells, with the same procedure as described for DENV2. Also, no significant inhibition of the virus replication was observed (Figure 11).

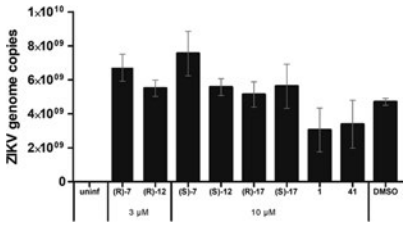


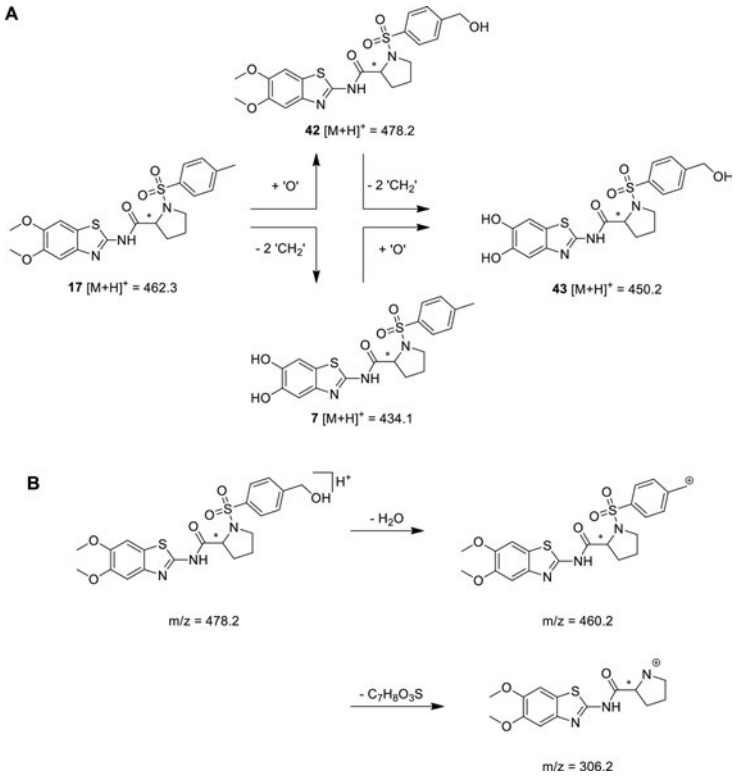
Figure 11. Effect of the treatment of ZIKV-infected Vero cells with **1**, (*R*/*S*)-7, (*R*/*S*)-**12**, (*R*/*S*)-**17**, and **41**. A concentration of 3 μM was used for (*R*)-7 and (*R*)-**12** and 10 μM for all other compounds. Vero cells were incubated with the compounds for 1 h and subsequently infected. Viral RNAs were isolated, quantified by RTqPCR, and normalized to the untreated control.

(*R*/*S*)-7 and (*R*/*S*)-**12** exhibited inhibition of both proteases, but only antiviral activity on DENV2 could be found. The reason for this unexpected observation may be found in the different organization of the virus inside of host cells: Both viruses have been shown to form replication factories containing RNA and NS3 with the difference that the ZIKV replication factories are additionally surrounded by intermediate filaments in cage-like structures.[50] This could lead to a stronger protection of the ZIKV protease from inhibitors compared to DENV. **41** showed a slight effect in the experiments (Figures SI13–SI15), which may be explainable by its lipophilicity, which is the highest amongst all compounds. This could lead to a better permeability of **41** through the protection mechanism employed by the ZIKV inside the cell. The cellular mechanisms of these effects need to be elucidated in future experiments.

For (*R*)-**12** and (*S*)-**12**, a decrease in DENV2 replication in the cell-based assay agrees with our enzyme assay data. However, when we compared the ability of (*R*)-7, (*S*)-7, (*R*)-**17**, and (*S*)-**17** to inhibit the isolated DENV2 NS2B/NS3 protease in the fluorometric enzyme assay, only (*R*)-7 and (*S*)-7 were able to efficiently inhibit the enzyme. While compounds (*R*)-7 and (*S*)-7 contain hydroxyl groups on the benzothiazole group, (*R*)-**17** and (*S*)-**17** carry methoxy groups instead. The same holds true for compound **1** and its methoxy-protected precursor **41** concerning inhibition of the isolated ZIKV protease compared to the antiviral activity against ZIKV in the cell-based assays (Figures SI15–SI17). It therefore seemed

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359–11382

Scheme 6. Metabolism of 17 in Rat Liver Microsomes[a]



[a](A) The metabolism pathways of **17** by consecutive demethylation and hydroxylation in different orders. (B) Two observed fragmentation reactions of **42**, indicating the metabolic hydroxylation at an aliphatic carbon.

very likely that the differences in the cell-based and in vitro assays could be traced back to the cellular metabolism and that compounds (*R*)-**17** and (*S*)-**17** or **41** may represent prodrugs for the respective hydroxy compounds.

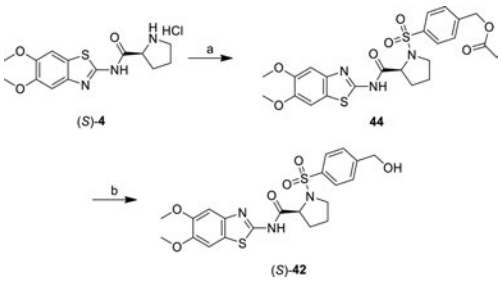
Metabolism Studies. To test whether the compounds indeed constitute prodrugs, metabolism studies using rat liver microsomes were performed exemplarily with (*R*)-**17** and (*S*)-**17** and the metabolites were analyzed by LC−MS/MS. Both enantiomers were readily transformed into the same metabolites. The most unpolar metabolites showed an m/z of 448.2 with a loss of m/z = 14 compared to the mass of the $[M + H]^+$ ions (m/z = 462.2) of (*R*)-**17** and (*S*)-**17**, pointing to a single demethylation. Also, the peaks for the doubly demethylated metabolites (*R*)-7 and (*S*)-7 were observed. The main metabolite shows an increase in m/z (478.2) compared to the $[M + H]^+$ peaks of (*R*)-**17** and (*S*)-**17**. This difference of m/z = 16 can be explained by the insertion of an oxygen atom. Additional masses were detected for m/z = 464.2 and 450.2, which could be explained by single and double demethylation, respectively, at the benzothiazole units of this oxygenated species. Isolation, followed by fragmentation of the m/z = 478.2 signal and comparison to the fragmentation of (*R*)-**17** and (*S*)-**17**, showed that the oxygen atom is added to the tosyl group, leading to **42** and **43** (Scheme 6).

The fragment of m/z = 460.1, caused by dehydration of the metabolite with m/z = 478.2, showed that the hydroxyl group is added to an aliphatic moiety, namely, the methyl group of the tosyl unit, which is a well-known biotransformation reaction.[51] To show that it is really **42** that is produced when (*R*)-**17** or (*S*)-**17** is metabolized, this compound was synthesized by reaction of (*S*)-**4** with 4-bromomethylpenylsulfonylchloride and sodium acetate to **44**, followed by the deprotection of the hydroxy group to yield (*S*)-**42** (Scheme 7). The synthesized compound (*S*)-**42** showed the same retention time and fragmentation pattern as the metabolite identified in the metabolism study proving their identity.

Overall, these metabolism studies show that the demethylation of (*R*)-**17** and (*S*)-**17** to (*R*)-7 and (*S*)-7, respectively, but also the hydroxylation at the tosyl group to yield (*R*)-**42** and (*S*)-**42** is catalyzed by cellular enzymes. These enzymes are known to be also present in the cells used for the antiviral activity studies.[52,53] In the cell-based antiviral activity assays, the methyl-protected hydroxy groups at the benzothiazole unit of (*R*)-**17** and (*S*)-**17** obviously increased the membrane permeability, and within cells, the active compounds with free hydroxy groups at the benzothiazole unit are released by oxidation. Protection of the hydroxyl groups of compounds 7 and its derivatives thus yields potent prodrugs for the

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Scheme 7. Synthesis of the Metabolite (*S*)-42[a]



[a]Reactions and conditions: (a) 4-$BrCH_2PhSO_2Cl$, NaOAc, DMF, 50 °C, 74%; (b) LiOH, THF, H_2O, r.t., 85%.

application in cellular systems. To investigate this hypothesis, cell permeability studies were performed.

Cell Permeability. Cell permeability studies were conducted exemplarily with (*R*)-7 and (*R*)-**17** involving a donor–acceptor setup. A layer of human epithelial cells (Caco-2), which forms a functional cellular barrier separating the donor and acceptor compartment, was treated with 10 μM of each compound in three separate experiments. After 3 h, the solutions above (donor) and below the cellular layer (acceptor) were analyzed by LC–MS together with cell-free controls (Figure SI19). While (*R*)-7 could not be detected in the solution below the layer, (*R*)-**17** was able to penetrate the cellular barrier with 91% efficiency compared to the control. This showed clearly that (*R*)-**17** is able to penetrate the cell layer much better than (*R*)-7, which is in accordance with the higher *S* log *P* value.

CONCLUSIONS

We identified new allosteric inhibitors for the ZIKV and DENV NS2B/NS3 protease. Starting from compound **1**, we obtained the enantiomers (*R*)-7 and (*S*)-7 as the new lead structures with improved LLE values (**1**: LLE = 0.24; **7**: LLE = 3.07) by the exchange of the aromatic ether moiety to the 1-tosylprolinyl fragment. Various structural modifications of **7** showed that neither the configuration of the proline residue nor the substitution pattern of the (sulfone)amide moiety do have a major influence on the IC_{50} values. The most potent inhibitor within this series turned out to be the sulfonamide (*R*)-**12** (IC_{50} = 0.32 μM, LLE = 4.09). The exchange of the proline to other amino acids led to compounds (*R*)-**30** and (*S*)-**30** (IC_{50} = 0.51/0.58 μM; LLE = 3.04/3.09). With these compounds ((*R*)-**12**, (*R*)-**30**, and (*S*)-**30**), we identified some of the most potent nonpeptidic, allosteric ZIKV protease inhibitors known.

The introduction of cysteine residues to the suspected allosteric pocket of the DENV2 protease, which were modified by maleimides, led to weaker binding of **1**, indicating binding of the inhibitor to this pocket. Addressability of this allosteric pocket for enzyme inhibition was further highlighted by the loss of activity upon modification of specific cysteine residues.

Furthermore, the binding modes were analyzed by docking studies. A more stable predicted binding mode of the 5,6-dihydroxybenzo[*d*]thiazole core scaffold and a better discrimination between structurally closely related binders and nonbinders for the open conformation of the flaviviral protease compared to the closed one were observed.

A selection of inhibitors and their methoxy derivatives was also tested in cell-based assays and those showing efficient inhibition of DENV2 replication in cells were identified. The results of the antiviral testing were further supported by metabolism studies, where the hypothesis that methoxy groups can act as prodrugs for phenols could be confirmed. With the oxidation of the *p*-methyl group, an additional metabolism pathway of these structures was discovered.

The compounds showed slightly better inhibition of the ZIKV protease compared to the DENV protease. In the cell-based assays, however, the contrary was observed. This points to differences in the structural organization of the viruses inside of host cells.[50] Thus, in future studies, a detailed look at the influence of ZIKV on the cellular biochemistry is needed, for example, on metabolism or uptake.

Overall, our data show that the new inhibitors are promising starting points for further optimization toward higher affinities against flaviviral proteases, which a conserved allosteric binding pocket in these enzymes constitutes a very promising site that can be addressed with our inhibitors and, due to their improved hydrophilicity, that these inhibitors can be used in further studies to assess the molecular details of flaviviral protease inhibition mechanisms and inhibitor binding details.

EXPERIMENTAL SECTION

Protein Preparation. *NS2B/NS3 Constructs and Cloning.* The DENV-2 NS2B/NS3 protease gene (GenBank ID: AY037116.1), with two point mutations in the NS3 region (I30A and L31A) cloned into a pET15b vector, was received from the working group of Professor Diederich (University of Marburg, Germany).[54] The construct contains an *N*-terminal hexa-histidine tag with a thrombin protease cleavage site. Point mutations were introduced through Quikchange mutagenesis using the Kapa HiFi PCR kit (KapaBiosystems) using the following primers: (i) T118C: 5′gtttaagtgcaatacgggtac-cattggcgcggttag3′ and 5′cgtattgcacttaaacaggcccggtttggtttgaac3′, (ii) T120C: 5′ccaattgcggtaccattggcgcggttagcctgg3′ and 5′ggtaccg-caattggtcttaaacaggcccggtttgg3′, (iii) T122C: 5′caatacgggttg-cattggcgcggttagcctgg3′ and 5′cgccaatgcaacccgtattggtcttaaacaggccc3′, (iv) N152Q: 5′gtatggtcagggtgtcgtgacccgtagtggtg3′ and 5′gacaccctgac-catacaggccaaccactttaccc3′, (v) A164C: 5′cgtgtcctgcattgccaacacg-gaaaagtccattgaag3′ and 5′gcaatgcaggacacgtaggcaccactacgggtc3′, and (vi) A166C: 5′cgatttgcaacacggaaaagtccattgaagataacccgg3′ and 5′gtgttgcaaatcgcggacacgtaggcaccac3′. The ZIKV NS2B/NS3 protease construct was designed according to the genome of French Polynesia ZIKV (GenBank ID: KJ776791.1).[55] The gene, cloned into a pET11a vector, was commercially obtained by GenScript. The construct contains a point mutation (R65A) and an N-terminal hexa-histidine tag with a tobacco etch virus (TEV) protease cleavage site.[14]

Protein Expression and Purification. Both NS2B/NS3 flaviviral proteases and their mutants were expressed in *Escherichia coli* (*E. coli*) strain BL21-Gold (DE3) (Agilent Technologies) according to previously published protocols.[14] In short, *E. coli* cells were grown in LB medium with the presence of ampicillin at 37 °C to an OD600 nm value of ~0.5 and induced with 1 mM isopropyl-D-thiogalactoside. Cells were then allowed to grow at 20 °C for 12–14 h. Harvested cells were shock frozen in liquid nitrogen and stored at −20 °C until further use. For purification, cells were resuspended in lysis buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 0.1 v/v Triton X-100, RNase, DNase, lysozyme, 1 mM PMSF, 1 mM DTT, 1 mM benzamidine) and lysed by sonication (Sonopuls, Bandelin). The lysates were cleared by centrifugation and the protein was purified by gravity flow chromatography with Ni-NTA Agarose (Qiagen). After extensive washing (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole), the protein was eluted (250 mM imidazole). The His-tags were removed by the corresponding protease (Thrombin or TEV)

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359–11382

during a dialysis step (20 mM Tris-HCl pH 8, 150 mM NaCl). After reverse IMAC, a gel filtration step (HiLoad 16/60 Superdex 75 or 200 column, GE Healthcare) was carried out (gel filtration buffer: 50 mM Tris-HCl pH 8, 150 mM NaCl). Throughout all steps, protein concentration was measured via absorbance at 280 nm and the sample purity was assessed through SDS-PAGE. Protein stock solutions in the gelfiltration buffer were flash frozen with liquid nitrogen and stored at −80 °C.

Molecular Modeling. All structures used for molecular docking studies were freely available in the Protein Data Bank (PDB).[56] The following PDB accession codes of the NS2B/NS3 protease were used for docking: 2FOM (DENV2 open conformation), 3U1I (DENV3 closed conformation), 5GXJ (ZIKV open conformation), and 5LC0 (ZIKV closed conformation). The DENV3 closed conformation (3U1I) was used as a surrogate for the DENV2 closed conformation. Similar inhibitory results were reported previously for both serotypes[27] and only minor differences within the proposed binding site were found in the crystal structures, where dissimilar amino acids orient their side chains away from the binding pocket. Prior docking, missing residue side chains were corrected using MOE2015.1001.[57] Additionally for 2FOM, residue Thr168 was added because it was not resolved in the crystal structure but close to the proposed allosteric site. Further receptor preparation was performed within the LeadIT-2.3.2 worksuite[46,47] using the automated binding site detection routine based on DoGSiteScorer[44,45] and residue protonation of ProToss. Water molecules forming three hydrogen bonds to the receptor were selected to be displaceable (water molecules 214, 230, and 247 in 2FOM, water molecule 231 in 3U1I, and no water molecules in 5GXJ and 5LC0), while water molecules with two or fewer interactions were removed. All ligands were energy minimized using the MMFF94 force field[58] within MOE2015.1001 prior docking. To secure the planarity of the amino-benzothiazole S−C−N−C torsion angle, an additional parameter was included in the torsion library of FlexX to only allow dihedrals of 0 ± 20° (Figure SI12). The relevance of this torsion parameter was identified by QM calculations (relaxed dihedral scan) of the corresponding dihedral angle for compound **1** using Gaussian 09[59] with B3LYP-D3[60−62] where an energy difference of 6.0 kcal/mol between the favored planar and the perpendicular conformation was found. The docking protocol was performed under default parameters within LeadIT-2.3.2 using the hybrid approach (enthalpy/entropy) for ligand placement.

MD simulations were performed with the open conformation of DENV NS2B/NS3 starting from PDB structure 2FOM (residues 43−76 of NS2B and residues 18−168 of NS3, where Thr168 was added as described above) without ligand and in complex with compound (*R*)-**7** generated by docking. Crystallographic water molecules were kept as part of the structure and only removed in the NS2B/NS3-(*R*)-**7** complex, when within 2 Å of the ligand. The protein structures were subsequently built within tleap of the AmberTools17.[63,64] All Asp and Glu residues were in their deprotonated form, while Arg and Lys were protonated. His45 (NS2B) and His60 (NS3) were protonated in the ε position, His47 (NS3) was δ-protonated, and His51 (NS3) was protonated two times. Ligand (*R*)-**7** parameters were derived from the Generalized Amber Force Field (GAFF)[65] using AM1-BCC[66] atomic partial charges within an antechamber.[63,64] The initial structures were relaxed over 200 time steps with a sander. Counter ions (Cl^-) were added and a TIP3P[67] waterbox exceeding the structure by 10.0 Å was built with tleap. Subsequent equilibration and production simulations were carried out using GPU nodes of the high-performance computing cluster Mogon of the University of Mainz using NAMD-2.11[68] and the Amber99SB force field.[69] During equilibration, harmonic constraints were applied to all nonwater atoms of the system, which was heated from 100 to 300 K over 500 ps, while constraints were gradually released in an NVT ensemble followed by further 500 ps without constraints. Ten nanoseconds of production runs were performed with an NPT ensemble using periodic boundary conditions and a van der Waals cutoff at 14 Å. Time steps of 2 fs were used in combination with rigid bond lengths and trajectories that were written every picosecond. Trajectories were analyzed with VMD-1.9.2[70] using snapshots extracted every 10 ps resulting in 1000 frames for 10 ns MDs.

Fluorometric Assays. For determination of the inhibitory activity of the compounds, assays based on the fluorogenic substrate boc-GRR-AMC were used for both ZIKV and DENV2 proteases.[71] The substrate and inhibitors were prepared as stock solutions in DMSO. The assay buffer consisted of 1 mM CHAPS, 50 mM Tris, and 20% glycerol at pH 9. The measurements were done in flat-bottom 96-well microtiter plates from Greiner Bio-One with a Tecan Infinite F2000 PRO fluorimeter or a Tecan Spark 10M in three independent experiments, each performed in duplicates. In each well, a total volume of 200 μL was used, consisting of 180 μL of buffer, 5 μL of the enzyme solution, 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. For the first screening, the inhibitors were used at a concentration of 20 μM, and for selected compounds, the IC_{50} values were determined with dilution series between 0.01 and 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 calculation of the IC_{50} values was done with GraFit from Erithacus Software Limited 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 an inhibitor concentration of [I] = 0 μM, Y_{min} as the minimum value, obtained at high inhibitor concentrations, and s as the Hill coefficient.[34]

Testing for noncompetitive behavior was done by measuring the IC_{50} values at five different substrate concentrations.[35]

Microscale Thermophoresis. To determine the dissociation constant (K_D), reflecting the interaction between the inhibitors and the ZIKV protease, MST was used. The purified protease was labeled with the Monolith NTTM Protein Labeling Kit RED-NHS, containing NT-647-NHS as a fluorescent dye for primary amine groups of lysine residues. The protein labeling was executed by the labeling protocol provided by the supplier of the kit. Ten microliter of the 150 nM labeled protein stock solution was combined with 16 different concentrations (ranging from 500 μM to 60 nM) of the inhibitors (10 μL). The compounds were diluted with a buffer containing 1 mM CHAPS, 50 mM Tris, and 20% glycerol at pH 9. To ensure equal DMSO concentrations, DMSO was added to a final concentration of 5% (v/v). The measurement was carried out on a NanoTemper Monolith NT.115 instrument (MST power of 10%, LED power of 50%, 30 s laser-on time, and 5 s laser-off time) with standard capillaries.

Antiviral Activity and Cytotoxicity. *Cell Toxicity Test on Vero Cells.* To exclude toxic side effects of the compounds, the cell survival and metabolism were measured by a cell proliferation assay (Promega, Germany). Vero cells were incubated with decreasing amounts of compounds solubilized in DMSO or with DMSO alone as a control. The assays were performed in triplicates according to the manufacturer's instructions. After 4 days, 10 μL of the MTS substrate was added, cells were further incubated for 60 min, and the OD490 was measured. Substance concentrations inhibiting the MTS substrate conversion were excluded from further analyses.

Cell Toxicity Test on A549 Cells. The cell viability was determined by the PrestoBlue Cell viability reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Cycloheximide (CHX) was used in a concentration of 70 μM as a positive control.

Cell Culture in Vero Cells. To analyze the influence of the compounds on DENV (serotype 2, as described in Wu et al.[27] and verified by nucleotide sequencing) and ZIKV (PF13/251013-18) replication, cells were seeded in 48-well plates. After 24 h, compounds were added, and 1 h after that, the cells were infected with DENV2 or ZIKV. Cell culture supernatants were harvested 3 days post-infection (p.i.) and centrifuged at 2000 rpm to remove detached cells.

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Cell Culture in A549 Cells. Human A549 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 4.5 g/L glucose, 10% fetal calf serum (FCS), 2 mM L-glutamine, 0.1 U/mL penicillin, and 100 μg/mL streptomycin. Six hours after seeding, the cells were infected with ZIKV strain 976 U or ZIKV PF13/251013-18 from FP with a multiplicity of infection (MOI) of 0.1. The substances **1**, **41**, and DMSO were diluted in the inoculum. At 16, 40, and 64 h p.i., the cells were washed once with warm PBS and covered with fresh DMEM including the substances to ensure their presence during the whole time course. The cells were washed with PBS and harvested at 24, 48, or 72 h p.i. in peqGOLD TriFast (peqlab Biotechnologie GmbH, Erlangen, Germany) for RNA isolation. The supernatant was harvested for the determination of extracellular viral RNA and particles.

RNA Isolation and Quantification DENV. Viral nucleic acids were isolated from 200 μL cell culture supernatants using the Roche HP Viral Nucleic Acid kit according to the manufacturer's manual. Relative viral genome copy numbers were determined using 5 μL of the eluted RNA and the RTqPCR using the LightMix Modular Dengue kit (TIB MOLBIOL, Germany) in combination with the LightCycler RNA Process Control kit (Roche). All reactions were performed in triplicates on Roche LightCycler96 or Roche LightCycler480 II qPCR machines. The relative RNA concentration was determined using an RNA standard curve and the respective cycler software. The quality of the RTqPCRs was ensured by following the MIQE guidelines.[72] In brief: First, the recursion coefficient of the relative standard curve r^2 had to be >0.95. Second, triplicate assays with a standard deviation of >0.5 were excluded from the analyses. Resulting genome copy numbers were expressed as genome copies per milliliter of the cell culture supernatants.

RNA-Isolation ZIKV. The intracellular RNA was isolated using peqGOLD Trifast (Peqlab Biotechnologie GmbH) according to the manufacturer's instructions.

The cell culture supernatant was centrifuged at 1000*g* for 5 min prior to isolation of the extracellular RNA with the QIAamp Viral RNA Mini kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

RT-qPCR ZIKV. The extracellular ZIKV RNA was quantified using the Zika Virus detection kit (TIB Molbiol, Berlin, Germany) and the LightCycler Multiplex RNA Virus Mastermix (Roche, Basel, Switzerland) according to the manufacturer's protocol in a LightCycler480 (Roche). The intracellular RNA was transcribed into cDNA by reverse transcription using the RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA). The cDNA was quantified in a LightCycler480 (Roche) using the SYBR Green Mastermix (Thermo Fisher Scientific, Waltham, MA, USA) and the following primers: ZIKV-fwd (AGATCCCGGCTGAAACACTG), ZIKV_rev (TTGCAAGGTCCATCTGTCCC), hRPL27_fw (AAAGCTGTCATCGTGAAGAAC), and hRPL27_rv (GCTGCTACTTTGCGGGGGTAG).

The amount of ZIKV RNA was then normalized to the amount of RPL27 transcripts using the ddcp-method.

Metabolism Studies. Rat liver microsomes were purchased by Sigma-Aldrich and characterized for cytochrome P450, cytochrome B5, and the activity of CYP1A, CYP3A, CYP2C, and cytochrome c reductase by the company. For the assay, the NADPH generating system was first generated by incubation at 37 °C of potassium phosphate buffer (790 μL, 100 mM, pH 7.4), $MgCl_2$ (50 μL, 80 mM), glucose-6-phosphate (50 μL, 100 mM), NADP disodium salt (50 μL, 20 mM), and glucose-6-dehydrogenase (50 μL, 100 IU/mL) for 10 min. After addition of the microsomes (50 μL, 20 mg/mL) and further 10 min of incubation at 37 °C, the inhibitors (2 μL, 5.21 mM in acetonitrile) were added. Aliquots of 50 μL were taken at 0, 15, 30, 45, and 60 min and added to 100 μL of ice-cold acetonitrile to stop the reaction. After centrifugation, the supernatant was analyzed by LC–MS/MS (Agilent Poroshell 120 EC-C_{18} 150 × 2.10 mm 4 μm column; mobile phase: 40% acetonitrile, 50% H_2O, 10% of a 0.1% solution of formic acid in water). Ion chromatograms were obtained using electronic filters for the ions of interest. Control incubations were performed with potassium phosphate buffer instead of microsomes.

Cell Permeability Studies. Permeability of (*R*)-**7** and (*R*)-**17** across a cellular barrier was assessed in transport studies using the human colon cell line C2BBe1, a clone of the Caco-2 cell line (passage 48, ATCC/LGC Standards, Wesel, Germany). The cells were grown on semipermeable Transwell inserts with a pore diameter of 0.4 μm and an area of 1.12 cm^2 (Corning Life Science, Corning, USA) in DMEM cell culture medium (Gibco, Thermo Fisher Scientific, Waltham, USA) supplemented with 10% fetal calf serum (Sigma-Aldrich, St. Louis, USA) and 1% nonessential amino acids (Gibco, Thermo Fisher Scientific, Waltham, USA). The donor compartment of the Transwell contained 500 μL of medium, and the acceptor compartment contained 1500 μL of medium and was exchanged every 2 days. To ensure biological barrier functionality of the cell layer, measurements of the transepithelial electrical resistance (TEER) were carried out using chopstick electrodes (Millicell ERS-2, Merck Millipore, Burlington, USA). Cells were grown for 20 days until they reached a TEER value of approximately 300 Ω·cm^2. (*R*)-**7** and (*R*)-**17** were dissolved in DMSO and applied in a final concentration of 10 μM in cell culture medium (containing 1% DMSO v/v) in the donor compartment of the Transwell setup and incubated for 3 h. After the incubation time, the cell culture media from the donor and acceptor compartment were collected and TEER measurements were repeated to ensure that the cellular barrier was not damaged by the compounds. The samples were frozen in liquid nitrogen and the solvent was removed under reduced pressure, and methanol (donor sample: 200 μL, acceptor sample: 600 μL) was added and mixed. After centrifugation, the supernatant was analyzed by LC–MS as described for the metabolism studies.

Chemistry. All chemicals were purchased from Sigma-Aldrich, TCI, Alfa Aesar, Acros, or Carbolution and used without further purification, if not stated otherwise. Solvents were purified by distillation and dichloromethane was desiccated by refluxing with phosphorous pentoxide for 1 h, followed by distillation. For reactions susceptible to moisture or air, an inert atmosphere of argon was used. Column chromatography was performed with silica gel (0.030–0.063 or 0.015–0.040 mm) obtained from Merck. Reaction progress was monitored by thin-layer chromatography using Machery-Nagel Alugram Xtra Sil G/UV_{254} silica gel 60 plates for detection at 254 nm. Specific rotations $[\alpha]_D^{24}$ were measured on a P3000 polarimeter from Krüss and are reported in $cm^3\ g^{-1}\ dm^{-1}$.

^1H and ^{13}C NMR spectra were measured on a Bruker Fourier 300 (300 MHz) spectrometer and ^{19}F NMR spectra were measured on a Bruker Avance III HD 300 (300 MHz) spectrometer at ambient temperature. COSY, HSQC, and HMBC experiments were used for assignments. Chemical shifts are reported in parts per millions (ppm, δ units). Trimethylsilane for ^1H and ^{13}C NMR spectra and trichlorfluorsilane for ^{19}F spectra were used as references. Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, dd = doublet of doublet, ddd = doublet of doublet of doublet, t = triplet, m = multiplet), coupling constants, and number of protons. Melting points were determined on an SMP10 from Stuart in open capillaries.

The purities of the inhibitors were higher than 95% in all cases. These were determined on an 1100 series HPLC system from Agilent with an Agilent Poroshell 120 EC-C_{18} 150 × 2.10 mm 4 μm column. The mobile phase was 45% acetonitrile, 45% H_2O, and 10% of a 0.1% solution of formic acid in water. The detection wavelength was 254 nm. The molecular mass was detected by an Agilent 1100 series LC/MSD trap with electron spray ionization (ESI) in the positive mode.

Enantiomeric ratios were determined on a Chiralcel IF3 3 μm 4.6 × 250 mm column with a Waters 2695 Separations module, a Column Thermostat Jetstream 2, and a Waters 996 photodiode array detector at 35 °C. Ethanol and hexane were used as the mobile phase in the annotated ratios. The enantiomeric ratio was calculated from the signal areas at 254 nm.

General Procedures. A: Boc-Protection of Amino Groups. Amino groups of amino acids were protected according to reported procedures.[73]

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359–11382

B: Amide Bond Formation with TBTU. The appropriate amine (1.0 equiv), carboxylic acid (1.0 equiv), and *O*-(benzotriazole-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluroniumtetrafluoroborate (TBTU) (1.1 equiv) were dissolved in DMF (20 mL) while cooling with an ice-water bath under argon atmosphere. Ethyldiisopropylamine (4.0 equiv) was added, the ice bath was removed after 30 minutes, and the reaction mixture was stirred at room temperature for 2 days. Water was added and the mixture was extracted with ethyl acetate three times. The organic layer was separated and washed three times each with 2 M HCl, saturated $NaHCO_3$ solution, and saturated NaCl solution. The organic layer was dried with Na_2SO_4 and the solvent was removed by reduced pressure to yield the product.

C: Amide Formation via Acyl Chlorides. A round-bottom flask was charged with the amine hydrochloride (1 equiv), pyridine (2.1 equiv), and acetonitrile as the solvent. The acyl chloride (1.1 equiv) was added dropwise while cooling with an ice-water bath. The mixture was stirred for 18 h, followed by evaporation of the solvent. The crude product was purified by column chromatography (ethyl acetate/cyclohexane = 1:2).

D: Coupling of Sulfonyl Chlorides with Amines. A round-bottom flask was charged with the appropriate amine (1.0 equiv), sulfonyl chloride (1.0 equiv), and acetonitrile (15 mL). Pyridine (2.1 equiv) was added, and after stirring for 14 h, the solution was concentrated, and purification by column chromatography (ethyl acetate/cyclohexane = 1:1 + 0.1% TFA) yielded the respective products.

E: Deprotection with BBr_3. The appropriate compound (1 equiv) was dissolved in dichloromethane (10 mL) and a 1 M solution of BBr_3 in dichloromethane (8 equiv) was added slowly at −79 °C. After 30 min, the cooling bath was removed and the solution was stirred at room temperature overnight. Methanol (5 mL) was added under cooling conditions with an ice bath and it was stirred for 30 min. In the case of amines, the hydrobromide was obtained by the addition of methanol (1 mL) and filtration. In other cases, the solvent was evaporated and the residue was purified by column chromatography (ethyl acetate/cyclohexane 1:1).

F: Deprotection of Boc-Groups by HCl. A round-bottom flask was charged with the starting material (1 equiv) and HCl (4 M in 1,4-dioxane, 20 equiv) was added. The mixture was stirred for 15 min and pentane was added to complete the crystallization. Collection by filtration afforded the products.

*2-Amino-5,6-dimethoxybenzo[d]thiazole (**2**).* 3,4-Dimethoxyaniline (1.0 equiv, 5.00 g, 32.7 mmol) and KSCN (2.2 equiv, 6.98 g, 71.9 mmol) were suspended in acetic acid (150 mL). Bromine (1.0 equiv, 1.64 mL, 32.7 mmol) was added slowly under cooling with an ice-water bath to keep the temperature below 20 °C. The mixture was stirred for 16 h at room temperature and filtrated. The residue was refluxed for 20 min in water, which was acidified with concentrated HCl (1 mL) and filtrated. The procedure of refluxing the residue with water followed by filtration was repeated. The combined filtrates were alkalized to pH 9 with KOH and filtrated to obtain 2-amino-5,6-dimethoxybenzo[*d*]thiazole (5.975 g, 28.42 mmol, 87% yield, lit.: 66%[74]) as an off-white solid. ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.28 (s, 1H), 7.19 (s, 2H), 6.98 (s, 1H), 3.75 (s, 3H), 3.72 ppm (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 166.26, 148.65, 147.16, 145.09, 121.81, 105.30, 102.84, 56.67, 56.14 ppm.

*(R)-tert-Butyl-2-((5,6-dimethoxybenzo[d]thiazol-2-yl)-carbamoyl)pyrrolidine-1-carboxylate ((R)-**3**).* General procedure A; colorless solid; yield 86% (703 mg, 1.72 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.31 (s, 1H), 7.22 (s, 1H), 4.74−4.29 (m, 1H), 3.94 (s, 3H), 3.92 (s, 3H), 3.74−3.29 (m, 2H), 2.59−2.36 (m, 1H), 2.09−1.81 (m, 3H), 1.70−1.31 ppm (m, 9H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 171.3, 156.6, 149.4, 147.6, 142.6, 142.5, 123.8, 103.5, 102.8, 81.5, 60.5, 56.5, 56.2, 47.4, 28.5, 28.5, 28.5 (3 × C), 27.8, 24.3 ppm. $[\alpha]_D^{24}$ + 100 (*c* 10 mg cm^{-3} in chloroform).

*(S)-tert-Butyl-2-((5,6-dimethoxybenzo[d]thiazol-2-yl)-carbamoyl)pyrrolidine-1-carboxylate ((S)-**3**).* General procedure A; colorless solid; yield: 99% (808 mg, 1.98 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.31 (s, 1H), 7.22 (s, 1H), 4.66−4.37 (m, 1H), 3.94 (s, 3H), 3.93 (s, 3H), 3.66−3.29 (m, 2H), 2.60−2.35 (m, 1H), 2.08−1.86 (m, 3H), 1.60−1.30 ppm (m, 9H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 170.3, 156.6, 149.4, 149.4, 147.6, 142.5, 123.8, 103.5, 102.8, 81.5, 60.5, 56.5, 56.2, 47.4, 28.5 (3 × C), 27.7, 24.7 ppm. $[\alpha]_D^{24}$ − 86 (*c* 10 mg cm^{-3} in chloroform).

*(R)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrochloride ((R)-**4**).* General procedure F; colorless solid; yield 94% (329 mg, 0.96 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.58 (s, 1H), 7.33 (s, 1H), 4.60−4.45 (m, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.39−3.20 (m, 2H), 2.46−2.34 (m, 1H), 2.08−1.84 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 167.7, 155.8, 149.1, 147.3, 142.2, 123.0, 103.8, 103.6, 59.1, 56.0, 55.8, 45.7, 29.5, 23.6 ppm. $[\alpha]_D^{24}$ + 43 (*c* 10 mg cm^{-3} in DMSO).

*(S)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrochloride ((S)-**4**).* General procedure F; colorless solid; yield 91% (131 mg, 0.38 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.59 (s, 1H), 7.33 (s, 1H), 4.62−4.47 (s, 1H), 3.82 (s, 3H), 3.80 (s, 3H), 3.35−3.22 (s, 2H), 2.47−2.34 (m, 1H), 2.10−1.86 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 167.7, 155.7, 149.1, 147.3, 142.2, 123.0, 103.7, 103.6, 59.1, 56.0, 55.8, 45.7, 29.5, 23.6 ppm. $[\alpha]_D^{24}$ − 40 (*c* 10 mg cm^{-3} in DMSO).

*(R)-1-Benzoyl-N-(5,6-dimethoxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**5**).* General procedure C; colorless solid; yield 90% (114 mg, 0.26 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.79−7.70 (m, 2H), 7.52−7.36 (m, 3H), 7.22 (s, 1H), 7.20 (s, 1H), 4.95−4.70 (m, 1H), 4.19−4.06 (m, 1H), 3.98 (s, 3H), 3.97 (s, 3H), 3.76−3.56 (m, 1H), 2.60−2.43 (m, 1H), 2.28−1.91 ppm (m, 3H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 171.8, 161.9, 152.9, 151.1, 149.2, 134.5, 131.1, 128.3, 128.3, 128.0, 128.0, 118.7, 115.8, 103.0, 98.8, 62.2, 56.7, 56.6, 50.6, 30.0, 26.0 ppm. $[\alpha]_D^{24}$ + 61 (*c* 10 mg cm^{-3} in $CHCl_3$).

*(S)-1-Benzoyl-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-pyrrolidine-2-carboxamide ((S)-**5**).* General procedure C; colorless solid; yield 96% (230 mg, 0.56 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.62−7.56 (m, 2H), 7.54 (s, 1H), 7.52−7.42 (m, 2H), 7.30 (s, 1H), 4.77−4.68 (m, 1H), 3.82 (s, 3H), 3.81 (s, 3H), 3.71−3.58 (m, 1H), 3.58−3.46 (m, 1H), 2.38−2.44 (m, 1H), 2.04−1.82 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 171.1, 164.7, 156.5, 149.0, 147.0, 142.6, 136.2, 130.3, 128.3, 128.3, 127.3, 127.3, 123.0, 103.7, 103.6, 60.1, 56.0, 55.8, 50.0, 29.7, 25.3 ppm. $[\alpha]_D^{24}$ − 55 (*c* 10 mg cm^{-3} in $CHCl_3$).

*(R)-1-Benzoyl-N-(5,6-dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**6**).* General procedure E; yellow solid, yield 13% (20 mg, 0.05 mmol); mp 248 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.66−7.53 (m, 2H), 7.54−7.44 (m, 3H), 7.29 (s, 1H,), 7.17 (s, 1H), 4.76−4.63 (m, 1H), 3.81−3.50 (m, 2H), 2.39−2.19 (m, 1H), 2.05−1.72 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO, δ): 170.56, 168.17, 155.20, 145.28, 143.62, 135.88, 132.3, 129.93, 127.95, 127.95, 127.02, 127.02, 121.63, 106.26, 105.86, 59.87, 49.72, 29.37, 24.91 ppm. MS (ESI) *m*/*z*: $[M + H]^+$ calcd for $C_{19}H_{17}N_3O_4S$, 384.1; found, 384.1. $[\alpha]_D^{24}$ + 103 (*c* 10 mg cm^{-3} in methanol).

*(S)-1-Benzoyl-N-(5,6-dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((S)-**6**).* General procedure E; yellow solid, yield 60% (43 mg, 0.11 mmol); mp 251 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.61−7.54 (m, 2H), 7.52−7.42 (m, 3H), 7.30 (s, 1H), 7.16 (s, 1H), 4.76−7.63 (m, 1H), 3.86−3.53 (m, 2H), 2.42−2.11 (m, 1H), 2.07−1.62 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.83, 168.45, 155.47, 145.56, 143.89, 136.29, 132.5, 130.24, 128.29, 128.29, 127.30, 127.30, 121.91, 106.54, 106.39, 60.15, 50.02, 29.64, 25.20 ppm. MS (ESI) *m*/*z*: $[M + H]^+$ calcd for $C_{19}H_{17}N_3O_4S$, 384.1; found, 384.1. $[\alpha]_D^{24}$ − 86 (*c* 10 mg cm^{-3} in methanol).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide ((R)-**7**).* General procedure D; colorless solid; yield 71% (86 mg, 0.20 mmol); mp 205−209 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.79 (d, *J* = 8.2 Hz, 2H), 7.40 (d, *J* = 8.2 Hz, 2H), 7.21 (s, 1H), 7.20 (s, 1H), 4.43−4.34 (m, 1H), 3.69−3.54 (m, 1H), 3.38−3.24 (m, 1H), 2.41 (s, 3H), 2.12−1.87 (m, 3H), 1.77−1.58 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 172.5, 158.1, 147.4, 145.9, 141.5, 135.0, 131.1 (2 × C), 128.9, (2 × C), 123.5, 117.9, 107.0, 106.7, 63.1, 50.6, 32.0, 25.7, 21.5 ppm. MS (ESI) *m*/*z*: $[M + H]^+$ calcd for $C_{19}H_{19}N_3O_5S_2$, 434.1; found, 434.1. $[\alpha]_D^{24}$ + 113 (*c* 10 mg cm^{-3} in methanol). er = 93:7 (hexane/EtOH 70:30).

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide ((S)-**7**).* General procedure D; colorless solid; yield 33% (52 mg, 0.12 mmol); mp 210 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.78 (d, J = 8.2 Hz, 2H), 7.46 (d, J = 8.1 Hz, 2H), 7.24 (s, 1H), 7.11 (s, 1H), 4.42−4.28 (m, 1H), 3.57−3.38 (m, 1H), 3.26−3.02 (m, 1H), 2.41 (s, 3H), 1.97−1.79 (m, 3H), 1.56−1.44 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.5, 155.2, 145.6, 144.0, 143.7, 141.8, 133.9, 129.9 (2 × C), 127.4 (2 × C), 122.0, 106.5, 106.4, 60.8, 49.1, 31.0, 24.5, 21.0 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{19}H_{19}N_3O_5S_2$, 434.1; found, 434.1. $[\alpha]_D^{24}$ − 126 (c 10 mg cm^{-3} in methanol). er = 93:7 (hexane/EtOH 70:30).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-(phenylsulfonyl)-pyrrolidine-2-carboxamide ((R)-**8**).* General procedure D; colorless solid; yield 66% (78 mg, 0.19 mmol); mp 170−173 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.97−7.89 (m, 2H), 7.78−7.55 (m, 3H), 7.20 (s, 1H), 7.20 (s, 1H), 4.49−4.35 (m, 1H), 3.71−3.51 (m, 1H), 3.41−3.23 (m, 1H), 2.17−1.83 (m, 3H), 1.76−1.53 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 172.4, 157.7, 147.2, 145.8, 142.4, 138.1, 134.6, 130.5 (2 × C), 128.9 (2 × C), 123.8, 107.1, 106.9, 63.1, 50.6, 32.1, 25.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{17}N_3O_5S_2$, 420.1; found, 420.1. $[\alpha]_D^{24}$ + 122 (c 10 mg cm^{-3} in acetone).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-(phenylsulfonyl)-pyrrolidine-2-carboxamide ((S)-**8**).* General procedure D; colorless solid; yield 64% (76 mg, 0.18 mmol); mp 170−173 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.94−7.84 (m, 2H), 7.71−7.52 (m, 3H), 7.21 (s, 1H), 7.19 (s, 1H), 4.40 (m, 1H), 3.66−3.52 (m, 1H), 3.30 (m, 1H), 2.22−1.82 (m, 3H), 1.75−1.53 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 172.4, 157.9, 147.2, 145.8, 141.9, 137.9, 134.5, 130.5 (2 × C), 128.8 (2 × C), 123.6, 107.0, 106.9, 63.1, 50.6, 32.0, 25.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{17}N_3O_5S_2$, 420.1; found, 420.1. $[\alpha]_D^{24}$ − 141 (c 10 mg cm^{-3} in acetone).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((4-fluorophenyl)-sulfonyl)pyrrolidine-2-carboxamide (**9**).* General procedure D; green solid; yield: 67% (81 mg, 0.19 mmol); mp 135−137 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 8.05−7.92 (m, 2H), 7.43−7.30 (m, 2H), 7.23−7.13 (m, 2H), 4.48−4.35 (m, 1H), 3.67−3.54 (m, 1H), 3.41−3.24 (m, 1H), 2.13−1.90 (m, 3H), 1.82−1.65 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 172.3, 147.0, 145.5, 143.1, 134.6, 131.9 (2 × C), 131.8, 122.6, 117.7 (2 × C), 117.4, 107.4, 106.8, 63.0, 50.5, 32.2, 20.1 ppm. ^{19}F NMR (282 MHz, methanol-d_4, δ): −76.25 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{16}FN_3O_5S_2$, 438.1; found, 438.1. $[\alpha]_D^{24}$ + 48 (c 5 mg cm^{-3} in methanol).

*(R)-1-((4-Chlorophenyl)sulfonyl)-N-(5,6-dihydroxybenzo[d]-thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**10**).* General procedure D; colorless solid; yield: 47% (60 mg, 0.13 mmol); mp 190−192 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.89 (d, J = 8.6 Hz, 2H), 7.62 (d, J = 8.6 Hz, 2H), 7.20 (s, 1H), 7.19 (s, 1H), 4.56−4.33 (m, 1H), 3.68−3.49 (m, 1H), 3.43−3.26 (m, 1H), 2.12−1.87 (m, 3H), 1.84−1.65 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 172.2, 157.6, 147.1, 145.7, 142.7, 140.8, 137.1, 130.7 (2 × C), 130.5 (2 × C), 124.0, 107.2, 106.9, 63.0, 50.5, 32.2, 25.7 ppm. MS (ESI) m/z: $[M + H]^+$ calculated for $C_{18}H_{16}ClN_3O_5S_2$, 454.0; found, 454.0. $[\alpha]_D^{24}$ + 128 (c 5 mg cm^{-3} in acetone).

*(S)-1-((4-Chlorophenyl)sulfonyl)-N-(5,6-dihydroxybenzo[d]-thiazol-2-yl)pyrrolidine-2-carboxamide ((S)-**10**).* General procedure D; colorless solid; yield: 71% (90 mg, 0.20 mmol); mp 190−192 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.90 (d, J = 8.5 Hz, 2H), 7.72 (d, J = 8.5 Hz, 2H), 7.24 (s, 1H), 7.12 (s, 1H), 4.46−4.23 (m, 1H), 3.61−3.38 (m, 1H), 3.35−3.18 (m, 1H), 2.02−1.83 (m, 3H), 1.70−1.45 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.7, 155.6, 146.1, 144.5, 142.2, 138.7, 136.2, 130.0 (2 × C), 129.7 (2 × C), 122.4, 107.0, 106.9, 61.3, 49.6, 31.5, 24.9 ppm. MS (ESI) m/z: $[M + H]^+$ calculated for $C_{18}H_{16}ClN_3O_5S_2$, 454.0; found, 454.1. $[\alpha]_D^{24}$ − 111 (c 10 mg cm^{-3} in acetone).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((4-methoxyphenyl)sulfonyl)pyrrolidine-2-carboxamide ((R)-**11**).* General procedure D; gray solid; yield 25% (32 mg, 0.07 mmol); mp 117−120 °C; ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.83 (d, J = 8.7 Hz, 2H), 7.21 (s, 1H), 7.19 (s, 1H), 7.07 (d, J = 8.7 Hz, 2H), 4.42−4.28 (m, 1H), 3.83 (s, 3H), 3.65−3.53 (m, 1H), 3.43−3.20 (m, 1H), 2.18−1.82 (m, 3H), 1.78−1.57 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 171.1, 163.7, 146.0, 144.5, 140.0, 129.7 (2 × C), 127.8, 122.0, 114.2 (2 × C), 105.6, 105.3, 61.7, 54.8, 49.2, 30.6, 24.4, 19.5 ppm. MS (ESI) m/z: $[M + H]^+$ calculated for $C_{19}H_{19}N_3O_6S_2$, 450.1; found, 450.1. $[\alpha]_D^{24}$ + 142 (c 10 mg cm^{-3} in acetone).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((4-methoxyphenyl)sulfonyl)pyrrolidine-2-carboxamide ((S)-**11**).* General procedure D; gray solid; yield 79% (99 mg, 0.22 mmol); mp 117−120 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.83 (d, J = 8.0 Hz, 2H), 7.24 (s, 1H), 7.20−7.05 (m, 3H), 4.41−4.24 (m, 1H), 3.85 (s, 3H), 3.59−3.35 (m, 1H), 3.23−3.02 (m, 1H), 1.94−1.79 (m, 3H), 1.61−1.41 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.4, 162.7, 155.1, 145.5, 143.9, 141.7, 129.5 (2 × C), 128.2, 121.8, 114.5 (2 × C), 106.4, 106.3, 60.7, 55.6, 49.0, 30.9, 24.4 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{19}H_{19}N_3O_6S_2$, 450.1; found, 450.1. $[\alpha]_D^{24}$ − 136 (c 2.5 mg cm^{-3} in acetone).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((4-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((R)-**12**).* General procedure D; colorless solid; yield 41% (53 mg, 0.11 mmol); mp 222−224 °C;^{1}H NMR (300 MHz, methanol-d_4, δ): 8.38 (d, J = 8.4 Hz, 2H), 8.11 (d, J = 8.4 Hz, 2H), 7.37−7.10 (m, J = 3.9 Hz, 2H), 4.65−4.34 (m, 1H), 3.80−3.57 (m, 1H), 3.53−3.34 (m, 1H), 2.28−1.94 (m, 3H), 1.91−1.60 ppm (m, 1H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 170.6, 150.4, 145.7, 144.3, 142.9, 139.1, 136.8, 128.7 (2 × C), 124.1 (2 × C), 122.6, 105.8, 105.5, 61.6, 49.1, 30.9, 24.4 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{16}N_4O_7S_2$, 465.1; found, 465.0. $[\alpha]_D^{24}$ + 230 (c 1 mg cm^{-3} in DMF).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((4-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((S)-**12**).* General procedure D; colorless solid; yield 73% (122 mg, 0.26 mmol); mp 228 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.44 (d, J = 8.7 Hz, 2H), 8.13 (d, J = 8.7 Hz, 2H), 7.23 (s, 1H), 7.12 (s, 1H), 4.50−4.34 (m, 1H), 3.63−3.45 (m, 1H), 3.39−3.24 (m, 1H), 2.07−1.84 (m, 3H), 1.74−1.54 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.1, 150.1, 145.8, 144.2, 142.5, 128.9 (2 × C), 124.7 (2 × C), 122.0, 119.0, 115.0, 106.6, 106.4, 61.0, 49.2, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{16}N_4O_7S_2$, 465.1; found, 465.0. $[\alpha]_D^{24}$ − 210 (c 1 mg cm^{-3} in DMF).

*(R)-1-([1,1′-Biphenyl]-4-ylsulfonyl)-N-(5,6-dihydroxybenzo[d]-thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**13**).* General procedure D; colorless solid; yield 79% (110 mg, 0.22 mmol); mp 89−91 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.01−7.87 (m, 3H), 7.80−7.70 (m, 2H), 7.60−7.39 (m, 3H), 7.24 (s, 1H), 7.13 (s, 1H), 4.48−4.37 (m, 1H), 3.61−3.45 (m, 1H), 3.32−3.20 (m, 1H), 2.06−1.85 (m, 3H), 1.67−1.51 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.4, 155.2, 145.6, 144.7, 144.0, 141.8, 138.3, 135.6, 129.2 (2 × C), 128.7, 128.1 (2 × C), 127.6 (2 × C), 127.2 (2 × C), 122.0, 106.5, 106.4, 60.8, 49.1, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{24}H_{21}N_3O_5S_2$, 496.1; found, 496.1. $[\alpha]_D^{24}$ + 160 (c 10 mg cm^{-3} in acetone).

*(S)-1-([1,1′-Biphenyl]-4-ylsulfonyl)-N-(5,6-dihydroxybenzo[d]-thiazol-2-yl)pyrrolidine-2-carboxamide ((S)-**13**).* General procedure D; colorless, yield 89% (125 mg, 0.25 mmol); mp 89−91 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.04−7.88 (m, 3H), 7.81−7.71 (m, 2H), 7.58−7.40 (m, 3H), 7.24 (s, 1H), 7.13 (s, 1H), 4.47−4.36 (m, 1H), 3.64−3.44 (m, 1H), 3.36−3.17 (m, 1H), 2.08−1.84 (m, 3H), 1.65−1.52 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 172.0, 155.2, 145.6, 144.7, 144.0, 141.8, 138.3, 135.6, 129.2 (2 × C), 128.7, 128.0 (2 × C), 127.6 (2 × C), 127.2 (2 × C), 122.0, 106.5, 106.4, 60.8, 49.1, 31.0, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{24}H_{21}N_3O_5S_2$, 496.1; found, 496.1. $[\alpha]_D^{24}$ − 152 (c 10 mg cm^{-3} in acetone).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((3-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((R)-**14**).* General procedure D; colorless solid, yield 38% (49 mg, 0.1 mmol); mp 210−212 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.59−8.45 (m, 2H), 8.30 (d, J = 7.1 Hz, 1H), 8.01−7.87 (m, 1H), 7.23 (s, 1H), 7.12 (s, 1H), 4.64−4.34 (m, 1H), 3.36−3.20 (m, 2H), 2.20−1.79 (m, 3H), 1.74−1.56 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.1, 156.9, 155.0,

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

148.1, 145.7, 144.1, 138.7, 133.2, 131.5, 127.9, 127.8, 122.0, 106.6, 106.4, 60.9, 49.2, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{18}H_{16}N_4O_7S_2$, 465.1; found, 465.1. $[\alpha]_D^{24}$ + 102 (c 5 mg cm^{-3} in acetone).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-1-((3-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((S)-**14**).* General procedure D; colorless solid, yield 70% (91 mg, 0.20 mmol); mp 210−212 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.58−8.48 (m, 2H), 8.36−8.26 (m, 1H), 8.02−7.88 (m, 1H), 7.23 (s, 1H), 7.12 (s, 1H), 4.51−4.35 (m, 1H), 3.63−3.48 (m, 1H), 3.37−3.23 (m, 1H), 2.11−1.80 (m, 3H), 1.78−1.53 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 171.4, 156.3, 149.5, 149.5, 146.9, 145.3, 140.1, 134.6, 132.9, 129.2, 123.5, 123.4, 107.7, 107.7, 62.2, 50.6, 32.5, 25.9 ppm. MS (ESI) m/z: $[M + H]^+$ calculated for $C_{18}H_{16}N_4O_7S_2$, 465.1; found, 465.1. $[\alpha]_D^{24}$ − 96 (c 10 mg cm^{-3} in acetone).

*(R)-1-((6-Acetamidonaphthalen-2-yl)sulfonyl)-N-(5,6-dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**15**).* General procedure D; black solid; yield 93% (137 mg, 0.26 mmol); mp 97−99 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.56−8.28 (m, 2H), 8.13−7.89 (m, 2H), 7.84−7.61 (m, 2H), 7.20 (s, 1H), 7.11 (s, 1H), 4.50−4.34 (m, 1H), 3.63−3.38 (m, 1H), 3.35−3.00 (m, 1H), 2.09 (s, 3H), 1.95−1.69 (m, 3H), 1.59−1.33 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.6, 169.2, 158.9, 145.7, 144.1, 142.7, 141.8, 139.7, 135.5, 132.2, 130.1, 128.8, 128.4, 128.2, 123.2, 121.4, 114.4, 106.6, 106.3, 60.9, 50.0, 31.1, 24.5, 24.1 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{24}H_{22}N_4O_6S_2$, 527.1; found, 527.1. $[\alpha]_D^{24}$ + 90 (c 1 mg cm^{-3} in DMF).

*(S)-1-((6-Acetamidonaphthalen-2-yl)sulfonyl)-N-(5,6-dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((S)-**15**).* General procedure D; black solid; yield 29% (41 mg, 0.08 mmol); mp 102−104 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.52−8.38 (m, 2H), 8.14−7.94 (m, 2H), 7.88−7.66 (m, 2H), 7.24 (s, 1H), 7.14 (s, 1H), 4.54−4.37 (m, 1H), 3.61−3.46 (m, 1H), 3.35−3.15 (m, 1H), 2.14 (s, 3H), 1.98−1.74 (m, 3H), 1.59−1.36 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO, δ): 170.5, 169.1, 159.0, 145.7, 144.0, 141.8, 141.4, 139.7, 135.4, 132.1, 130.1, 128.8, 128.4, 128.1, 123.1, 121.9, 114.4, 106.5, 106.4, 60.8, 49.2, 31.0, 24.5, 24.2 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{24}H_{22}N_4O_6S_2$, 527.1; found, 527.2. $[\alpha]_D^{24}$ − 80 (c 1 mg cm^{-3} in DMF).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrobromide ((R)-**16**).* General procedure E; off-white solid; yield 95% (419 mg, 1.18 mmol); ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.32 (s, 1H), 7.26 (s, 1H), 4.84−4.63 (m, 1H), 3.60−3.39 (m, 2H), 2.78−2.51 (m, 1H), 2.37−2.05 ppm (m, 3H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 169.3, 160.4, 148.7, 147.2, 135.4, 120.8, 107.7, 104.3, 61.8, 47.7, 30.7, 24.9 ppm. $[\alpha]_D^{24}$ + 27 (c 10 mg cm^{-3} in methanol).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrobromide ((S)-**16**).* General procedure E; off-white solid; yield 97% (428 mg, 1.18 mmol); ^{1}H NMR (300 MHz, methanol-d_4, δ): 7.33 (s, 1H), 7.27 (s, 1H), 4.51−4.43 (m, 1H), 3.60−3.42 (m, 1H), 2.75−2.53 (m, 2H), 2.36−2.06 ppm (m, 3H). ^{13}C NMR (75 MHz, methanol-d_4, δ): 169.3, 160.4, 148.7, 147.2, 135.4, 120.8, 107.7, 104.3, 61.8, 47.7, 30.7, 24.9 ppm. $[\alpha]_D^{24}$ − 24 (c 10 mg cm^{-3} in methanol).

*(R)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide ((R)-**17**).* General procedure D; colorless solid; yield 91% (122 mg, 0.26 mmol); mp 171−173 °C; ^{1}H NMR (300 MHz, chloroform-d, δ): 7.81 (d, J = 8.2 Hz, 2H), 7.35 (d, J = 8.0 Hz, 2H), 7.28 (s, 1H), 7.23 (s, 1H), 4.46−4.30 (m, 1H), 3.99 (s, 3H), 3.98 (s, 3H) 3.86−3.73 (m, 1H), 3.34−3.28 (m, 1H), 2.42 (s, 3H), 2.16−2.08 (m, 3H), 1.79−1.65 ppm (m, 1H). ^{13}C NMR (75 MHz, chloroform-d, δ): 172.4, 161.4, 151.4, 149.5, 144.6, 133.4, 132.0, 130.1 (2 × C), 128.0 (2 × C), 118.9, 103.0, 99.2, 61.9, 56.8, 56.7, 49.5, 31.6, 25.1, 21.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{21}H_{23}N_3O_5S_2$, 462.1; found, 462.1. $[\alpha]_D^{24}$ + 152 (c 10 mg cm^{-3} in acetonitrile).

*(S)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide ((S)-**17**).* General procedure D; colorless solid; yield 80% (132 mg, 0.30 mg); mp 172−176 °C; ^{1}H NMR (300 MHz, chloroform-d, δ): 7.82 (d, J = 8.2 Hz, 2H), 7.34 (d, J = 8.0 Hz, 2H), 7.26 (s, 1H), 7.23 (s, 1H), 4.42−4.26 (m, 1H), 3.99 (s, 3H), 3.97 (s, 3H), 3.86−3.67 (m, 1H), 3.40−3.18 (m, 1H), 2.42 (s, 3H), 2.20−1.93 (m, 3H), 1.84−1.56 ppm (m, 1H). ^{13}C NMR (75 MHz, chloroform-d, δ): 172.0, 160.9, 151.1, 149.2, 144.4, 133.5, 133.3, 130.1 (2 × C), 128.0 (2 × C), 119.5, 103.0, 99.6, 61.9, 56.8, 56.6, 49.6, 31.5, 25.1, 21.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{21}H_{23}N_3O_5S_2$, 462.1; found, 462.1. $[\alpha]_D^{24}$ − 144 (c 10 mg cm^{-3} in acetonitrile).

*(R)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-((4-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((R)-**18**).* General procedure D; colorless solid; yield 92% (132 mg, 0.27 mmol); mp 183−186 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.45 (d, J = 8.8 Hz, 2H), 8.14 (d, J = 8.8 Hz, 2H), 7.57 (s, 1H), 7.31 (s, 1H), 4.55−4.37 (m, 1H), 3.84 (s, 3H), 3.82 (s, 3H), 3.62−3.50 (m, 1H), 3.41−3.23 (m, 1H), 2.14−1.86 (m, 3H), 1.74−1.55 ppm (m, 1H). ^{13}C NMR (75 MHz, chloroform-d, δ): 170.3, 158.1, 156.0, 150.1, 149.0, 147.1, 142.4, 128.9 (2 × C), 124.7 (2 × C), 123.0, 103.7, 103.5, 60.9, 56.0, 55.7, 49.2, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{20}N_4O_7S_2$, 493.1; found, 493.1. $[\alpha]_D^{24}$ + 178 (c 5 mg cm^{-3} in acetonitrile). er = 100:0 (hexane/EtOH 50:50).

*(S)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-((4-nitrophenyl)-sulfonyl)pyrrolidine-2-carboxamide ((S)-**18**).* General procedure D; colorless solid; yield 94% (134 mg, 0.27 mmol); mp 196−199 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ):a 8.44 (d, J = 8.8 Hz, 2H), 8.14 (d, J = 8.8 Hz, 2H), 7.57 (s, 1H), 7.31 (s, 1H), 4.55−4.39 (m, 1H), 3.84 (s, 3H), 3.81 (s, 3H), 3.65−3.46 (m, 1H), 3.44−3.26 (m, 1H), 2.10−1.84 (m, 3H), 1.75−1.54 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.3, 158.1, 156.0, 150.1, 149.0, 147.1, 142.4, 128.9 (2 × C), 124.7 (2 × C), 123.0, 103.7, 103.5, 60.8, 56.0, 55.8, 49.2, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{20}N_4O_7S_2$, 493.1; found, 493.1. $[\alpha]_D^{24}$ − 158 (c 5 mg cm^{-3} in acetonitrile). er = 100:0 (hexane/EtOH 50:50).

*(R)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-((4-methoxyphenyl)sulfonyl)pyrrolidine-2-carboxamide ((R)-**19**).* General procedure D; colorless solid; yield 99% (138 mg, 0.29 mmol); mp 186−188 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.84 (d, J = 8.4 Hz, 2H), 7.58 (s, 1H), 7.31 (s, 1H), 7.16 (d, J = 8.4 Hz, 2H), 4.48−4.32 (m, 1H), 3.86 (s, 3H), 3.84 (s, 3H), 3.82 (s, 3H), 3.59−3.37 (m, 1H), 3.32−3.07 (m, 1H), 3.01−1.78 (m, 3H), 1.54−1.43 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.7, 162.8, 148.9, 147.1, 140.3, 129.6 (2 × C), 128.3, 123.0, 121.4, 114.6 (2 × C), 103.7, 103.5, 60.7, 56.0, 55.9, 55.7, 48.8, 31.0, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{20}N_4O_7S_2$ 478.1; found, 478.1. $[\alpha]_D^{24}$ + 188 (c 10 mg cm^{-3} in DMSO).

*(S)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-((4-methoxyphenyl)sulfonyl)pyrrolidine-2-carboxamide ((S)-**19**).* General procedure D; colorless solid; yield 97% (94 mg, 0.20 mmol); mp 186−188 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.84 (d, J = 8.8 Hz, 1H), 7.57 (s, 1H), 7.31 (s, 1H), 7.16 (d, J = 8.9 Hz, 1H), 4.51−4.27 (m, 0H), 3.86 (s, 2H), 3.84 (s, 2H), 3.82 (s, 2H), 3.56−3.40 (m, 1H), 3.23−3.06 (m, 1H), 1.98−1.80 (m, 2H), 1.65−1.45 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.7, 162.8, 149.0, 147.1, 146.0, 142.5, 129.6 (2 × C), 128.4, 123.0, 114.6 (2 × C), 103.7, 103.5, 60.7, 55.9, 55.7, 55.7, 49.1, 31.0, 24.4 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{20}N_4O_7S_2$, 478.1; found, 478.1. $[\alpha]_D^{24}$ − 182 (c 10 mg cm^{-3} in DMSO).

*(S)-1-((4-Chlorophenyl)sulfonyl)-N-(5,6-dimethoxybenzo[d]-thiazol-2-yl)pyrrolidine-2-carboxamide (**20**).* General procedure D; colorless solid; yield 95% (133 mg, 0.28 mmol); mp 180−182 °C; ^{1}H NMR (300 MHz, chloroform-d, δ): 7.90 (d, J = 8.5 Hz, 2H), 7.54 (d, J = 8.5 Hz, 2H), 7.29 (s, 1H), 7.23 (s, 1H), 4.49−4.26 (m, 1H), 4.00 (s, 3H), 3.99 (s, 3H), 3.88−3.70 (m, 1H), 3.39−3.17 (m, 1H), 2.25−2.05 (m, 3H), 1.83−1.70 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO, δ): 170.5, 156.1, 149.0, 147.1, 142.5, 138.3, 135.7, 129.6 (2 × C), 129.3 (2 × C), 123.0, 103.7, 103.5, 60.8, 56.0, 55.8, 49.1, 31.1, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{20}N_4O_7S_2$, 482.1; found, 482.1. $[\alpha]_D^{24}$ − 171 (c 10 mg cm^{-3} in DMSO).

*(R)-1-Benzoyl-N-(6-nitrobenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((R)-**21**).* General procedure C; yellow solid, yield 68% (68 mg, 0.17 mmol); mp 238 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ):

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

9.05 (d, J = 2.4 Hz, 1H), 8.28 (dd, J = 9.0, 2.4 Hz, 1H), 7.90 (d, J = 9.0 Hz, 1H), 7.67−7.54 (m, 2H), 7.54−7.43 (m, 3H), 4.86−4.69 (m, 1H), 3.78−3.61 (m, 1H), 3.61−3.49 (m, 1H), 2.44−2.23 (m, 1H), 2.05−1.79 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 172.3, 168.5, 163.6, 153.5, 143.1, 135.9, 132.3, 130.3, 128.3 (2 × C), 127.3, 127.3, 121.8, 120.6, 119.1, 60.3, 50.0, 29.5, 25.3 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{19}H_{16}N_4O_4S$, 397.1; found, 397.1. $[\alpha]_D^{24}$ + 122 (c 5 mg cm^{-3} in DMF).

*(S)-1-Benzoyl-N-(6-nitrobenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide ((S)-**21**).* General procedure C; yellow solid, yield 79% (94 mg, 0.23 mmol); mp 245 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 9.06 (d, J = 2.4 Hz, 1H), 8.33−8.25 (m, 1H), 7.91 (d, J = 9.0 Hz, 1H), 7.60 (dd, J = 7.4, 2.1 Hz, 2H), 7.52−7.45 (m, 3H), 4.86−4.69 (m, 1H), 3.78−3.61 (m, 1H), 3.61−3.49 (m, 1H), 2.44−2.23 (m, 1H), 2.06−1.82 (m, 3H) ppm. ^{13}C NMR (75 MHz, DMSO-d_6, δ): 172.2, 168.4, 163.5, 153.4, 143.0, 135.9, 132.2, 130.3, 128.2 (2 × C), 127.3 (2 × C), 121.8, 120.6, 119.1, 60.3, 49.9, 29.4, 25.2 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{19}H_{16}N_4O_4S$, 397.1; found, 397.1. $[\alpha]_D^{24}$ − 106 (c 5 mg cm^{-3} in MeOH).

*(R)-N-(6-Nitrobenzo[d]thiazol-2-yl)-1-tosylpyrrolidine-2-carboxamide (**22**).* General procedure C; yellow solid, yield 88% (110 mg, 0.25 mmol); mp 175−177 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 9.07 (d, J = 2.4 Hz, 1H), 8.29 (dd, J = 8.9, 2.4 Hz, 1H), 7.92 (d, J = 8.9 Hz, 1H), 7.80 (d, J = 8.0 Hz, 2H), 7.46 (d, J = 8.0 Hz, 2H), 4.47−4.34 (m, 1H), 3.63−3.37 (m, 1H), 3.27−3.06 (m, 1H), 2.41 (s, 3H), 2.06−1.79 (m, 3H), 1.66−1.48 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 172.0, 163.3, 153.3, 143.8, 143.1, 133.7, 132.3, 129.9 (2 x C), 127.4 (2 x C), 121.8, 120.7, 119.1, 60.9, 49.1, 31.0, 24.5, 21.0 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{19}H_{18}N_4O_5S_2$, 447.1; found, 447.1. $[\alpha]_D^{24}$ + 197 (c 10 mg cm^{-3} in acetone).

*(R)-tert-Butyl-2-((6-nitrobenzo[d]thiazol-2-yl)carbamoyl)-pyrrolidine-1-carboxylate ((R)-**24**).* General procedure B; colorless solid; yield 93% (850 mg, 2.17 mmol); ^{1}H NMR (300 MHz, chloroform-d, δ): 8.73 (d, J = 2.3 Hz, 1H), 8.30 (dd, J = 9.0, 2.3 Hz, 1H), 7.84 (d, J = 9.0 Hz, 1H), 4.74−4.45 (m, 1H), 3.68−3.19 (m, 2H), 2.70−2.49 (m, 1H), 2.10−1.89 (m, 3H), 1.52 ppm (s, 9H). ^{13}C NMR (75 MHz, chloroform-d, δ): 173.3, 162.1, 157.0, 153.4, 143.9, 132.7, 122.1, 121.2, 118.1, 82.1, 60.2, 47.6, 28.5, 27.0 (3 × C), 24.7 ppm. $[\alpha]_D^{24}$ + 124 (c 10 mg cm^{-3} in chloroform).

*(S)-tert-Butyl-2-((6-nitrobenzo[d]thiazol-2-yl)carbamoyl)-pyrrolidine-1-carboxylate ((S)-**24**).* General procedure B; yellow solid; yield 59% (1.26 g, 3.21 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 9.07 (d, J = 2.4 Hz, 1H), 8.28 (dd, J = 9.0, 2.4 Hz, 1H), 7.90 (d, J = 9.0 Hz, 1H), 4.52−4.37 (m, 1H), 3.55−3.43 (m, 1H), 3.41−3.33 (m, 1H), 2.36−2.16 (m, 1H), 2.02−1.73 (m, 3H), 1.23 ppm (s, 9H). ^{13}C NMR (75 MHz, DMSO, δ): 173.4, 163.4, 153.4, 152.7, 143.1, 132.2, 121.8, 120.6, 119.1, 78.9, 59.6, 46.5, 30.7, 27.8 (3 × C), 23.5 ppm. $[\alpha]_D^{24}$ − 112 (c 10 mg cm^{-3} in chloroform).

*(R)-N-(6-Nitrobenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrochloride ((R)-**25**).* General procedure F; colorless solid, yield 99% (704 mg, 2.14 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 9.08 (d, J = 2.4 Hz, 1H), 8.28 (dd, J = 9.0, 2.4 Hz, 1H), 7.93 (d, J = 9.0 Hz, 1H), 4.69−4.54 (m, 1H), 3.40−3.21 (m, 2H), 2.47−2.35 (m, 1H), 2.14−1.85 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 169.4, 163.3, 153.5, 143.7, 132.7, 122.3, 121.4, 119.7, 59.7, 46.2, 29.8, 23.9 ppm. $[\alpha]_D^{24}$ + 82 (c 10 mg cm^{-3} in DMSO).

*(S)-N-(6-Nitrobenzo[d]thiazol-2-yl)pyrrolidine-2-carboxamide Hydrochloride ((S)-**25**).* General procedure F; yellow solid, yield 99% (705 mg, 2.14 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 9.08 (d, J = 2.4 Hz, 1H), 8.28 (dd, J = 9.0, 2.4 Hz, 1H), 7.93 (d, J = 9.0 Hz, 1H), 4.67−4.53 (m, 1H), 3.43−3.16 (m, 2H), 2.47−2.38 (m, 1H), 2.22−1.78 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO, δ): 169.0, 162.9, 153.1, 143.3, 132.3, 121.9, 121.0, 119.2, 59.3, 45.7, 29.3, 23.5 ppm. $[\alpha]_D^{24}$ − 80 (c 10 mg cm^{-3} in DMSO).

*(RS)-N-(5,6-Dihydroxybenzo[d]-thiazol-2-yl)-1-tosylpiperidine-3-carboxamide ((RS)-**26**).* General procedure D; colorless solid; yield 27% (32 mg, 0.07 mmol); mp 137−139 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.65 (d, J = 7.9 Hz, 2H), 7.46 (d, J = 7.9 Hz, 3H), 7.20 (d, J = 1.6 Hz, 1H), 7.09 (d, J = 1.6 Hz, 1H), 3.84−3.69 (m, 1H), 3.63−3.50 (m, 1H), 2.87−2.71 (m, 1H), 2.46−2.35 (m, 4H), 2.29−2.14 (m, 1H), 1.94−1.72 (m, 2H), 1.58−1.31 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 171.3, 155.2, 145.6, 143.9, 143.7, 141.8, 132.3, 129.9 (2 × C), 127.5 (2 × C), 121.9, 106.5, 106.3, 47.7, 46.0, 41.4, 26.2, 23.5, 21.0 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{21}N_3O_5S_2$, 448.1; found, 448.2.

*(RS)-1-(tert-Butoxycarbonyl)piperidine-3-carboxylic Acid ((RS)-**27**).* Nicotinic acid (1.0 equiv, 5.00 g, 40.0 mmol) was dissolved in 2 M HCl (40 mL), PtO_2 (40 mg) was added, and the mixture was stirred under a hydrogen atmosphere of 4 bar for 4 days, followed by filtration over celite. The filtrate was alkalized with NaOH and boc_2O (1.0 equiv, 8.90 g, 40.0 mmol) and dioxane (40 mL) was added subsequently. After stirring for 16 h, the organic solvent was evaporated under reduced pressure and crystallization was induced by addition of 2 M HCl. 1-(*tert*-Butoxycarbonyl)piperidine-3-carboxylic acid was obtained by filtration as a colorless solvent with a yield of 17% (1.56 g, 6.80 mmol).

^{1}H NMR (300 MHz, DMSO-d_6, δ): 3.83 (m, 2H), 2.93−2.73 (m, 2H), 2.47−2.31 (m, 1H), 1.86−1.69 (m, 2H), 1.47−1.28 ppm (m, 11H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 175.6, 153.9, 78.6, 42.8, 42.5, 40.4, 28.1, 27.7 ppm (3 × C).

*(RS)-tert-Butyl-3-((5,6-dimethoxybenzo[d]thiazol-2-Yl)-carbamoyl)piperidine-1-carboxylate ((RS)-**28**).* General procedure B; colorless solid; yield 84% (620 mg, 1.47 mmol); ^{1}H NMR (300 MHz, chloroform-d, δ): 7.25 (s, 1H), 7.21 (s, 1H), 4.22−4.04 (m, 2H), 3.94 (s, 3H), 3.92 (s, 3H), 3.36−3.20 (m, 1H), 2.76−2.54 (m, 1H), 1.98−1.79 (m, 1H), 1.79−1.63 (m, 2H), 1.54−1.35 ppm (m, 11H). ^{13}C NMR (75 MHz, chloroform-d, δ): 171.7, 155.7, 154.8, 150.0, 148.3, 138.0, 122.1, 102.9, 102.9, 80.5, 77.4, 56.5, 56.4, 44.7, 43.0, 28.5 (3 × C), 27.7, 24.1 ppm.

*(RS)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)piperidin-3-carboxamide Hydrobromide ((RS)-**29**).* General procedure E; colorless solid; 99% (266 mg, 0.71 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.23 (s, 1H), 7.11 (s, 1H), 3.48−3.31 (m, 1H), 3.26−2.80 (m, 4H), 2.14−2.00 (m, 1H), 1.87−1.51 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO, δ): 171.1, 155.2, 145.7, 144.0, 141.6, 121.8, 106.5, 106.5, 43.6, 43.0, 38.9, 25.9, 21.1 ppm.

*(R)-N-(5,6-Dihydroxybenzo[d]-thiazol-2-yl)-1-tosylpiperidine-2-carboxamide ((R)-**30**).* General procedure D; colorless solid; yield 41% (50 mg, 0.11 mmol); mp 214−216 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.59 (d, J = 8.0 Hz, 2H), 7.28 (d, J = 8.0 Hz, 2H), 7.19 (s, 1H), 7.10 (s, 1H), 4.83−4.71 (m, 1H), 3.76−3.47 (m, 2H), 2.30 (s, 3H), 1.72−1.56 (m, 2H), 1.58−1.44 (m, 1H), 1.38−1.21 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 169.4, 155.0, 145.6, 144.0, 143.2, 141.7, 136.3, 129.6 (2 × C), 126.7 (2 × C), 122.0, 106.5, 106.3, 54.0, 42.6, 27.9, 24.1, 21.0, 18.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{21}N_3O_5S_2$, 448.1; found, 448.1. $[\alpha]_D^{24}$ + 31 (c 10 mg cm^{-3} in methanol).

*(S)-N-(5,6-Dihydroxybenzo[d]-thiazol-2-yl)-1-tosylpiperidine-2-carboxamide ((S)-**30**).* General procedure D; colorless solid; yield 36% (43 mg, 0.10 mmol); mp 214−216 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.59 (d, J = 7.9 Hz, 2H), 7.28 (d, J = 7.9 Hz, 2H), 7.20 (s, 1H), 7.10 (s, 1H), 4.81−4.70 (m, 1H), 3.75−3.49 (m, 2H), 1.73−1.58 (m, 2H), 1.59−1.46 (m, 1H), 1.41−1.22 ppm (m, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 169.5, 155.0, 145.6, 144.0, 143.2, 141.7, 136.3, 129.7 (2 x C), 126.8 (2 x C), 122.0, 106.5, 106.3, 54.0, 42.7, 27.9, 24.1, 21.0, 18.7 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{21}N_3O_5S_2$, 448.1; found, 448.1. $[\alpha]_D^{24}$ − 40 (c 10 mg cm^{-3} in methanol).

*N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-2-tosylacetamid (**31**).* General procedure D; gray solid; yield 12% (15 mg, 0.04 mmol); mp 183−185 °°C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.69 (d, J = 7.9 Hz, 2H), 7.36 (d, J = 7.9 Hz, 2H), 7.21 (s, 1H), 7.08 (s, 1H), 3.79 (d, J = 5.9 Hz, 2H), 2.32 ppm (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 167.0, 155.0, 145.6, 143.9, 142.8, 141.7, 137.5, 129.5 (2 × C), 126.6 (2 × C), 121.9, 106.5, 106.4, 45.0, 20.9 ppm. MS (ESI) m/z: $[M + H]^+$ calculated for $C_{16}H_{15}N_3O_5S_2$, 394.1; found, 394.1.

*N-(5,6-Dihydroxybenzo[d]-thiazol-2-yl)-1-tosylpiperidine-4-carboxamide (**32**).* General procedure D; colorless solid; yield 43% (51 mg, 0.11 mmol); mp 227−230 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.63 (d, J = 7.9 Hz, 2H), 7.45 (d, J = 7.9 Hz, 2H), 7.19 (s, 1H),

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

7.07 (s, 1H), 3.79−3.48 (m, 2H), 2.41 (s, 3H), 2.33−2.19 (m, 2H), 2.04−1.82 (m, 3H), 1.75−1.53 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 173.6, 156.3, 146.4, 144.7, 144.5, 142.7, 133.3, 130.8 (2 × C), 128.4 (2 × C), 122.7, 118.3, 107.3, 107.2, 46.2 (2 × C), 28.2 (2 × C), 21.9 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{20}H_{21}N_3O_5S_2$, 448.1; found, 448.2.

*N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-cyclohexanecarboxamide (**33**).* General procedure E; colorless solid; yield 85% (61 mg, 0.21 mmol); mp 232−234 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.20 (s, 1H), 7.08 (s, 1H), 1.90−1.54 (m, 6H), 1.50−1.11 ppm (m, 5H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 174.9, 156.2, 145.9, 144.2, 141.9, 122.1, 106.8, 106.7, 43.9, 29.2 (2 × C), 25.7, 25.5 ppm (2 × C). MS (ESI) m/z: $[M + H]^+$ calcd for $C_{14}H_{16}N_2O_3S_1$, 293.1; found, 293.1.

*N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-cyclohexanecarboxamide (**34**).* General procedure C; colorless solid; yield 71% (324 mg, 1.01 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.28−7.24 (m, 2H), 3.96 (s, 3H), 3.93 (s, 3H), 2.30 (tt, J = 11.6, 7.0 Hz, 1H), 1.91−1.79 (m, 2H), 1.79−1.66 (m, 2H), 1.66−1.43 (m, 4H), 1.20−0.97 ppm (m, 2H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 174.5, 158.3, 149.5, 147.7, 141.9, 123.7, 103.0, 103.0, 56.5, 56.2, 45.2, 29.3 (C × 2), 25.5 (C × 2), 25.5 ppm.

*N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)-2-(4-methylphenylsulfonamido)benzamide (**35**).* General procedure D; yellow solid; yield 41% (75 mg, 0.16 mmol); mp 204−206 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.07−7.91 (m, 1H), 7.62 (d, J = 8.2 Hz, 2H), 7.53−7.37 (m, 2H), 7.35−7.14 (m, 4H), 7.13−7.05 (m, 1H), 2.25 ppm (s, 3H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 175.8, 158.0, 146.6, 145.6, 144.7, 144.3, 144.2, 138.2, 136.4, 133.3, 130.7, 130.2 (2 × C), 127.3 (2 × C), 124.3, 123.8, 121.1, 107.6, 107.5, 21.4 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{21}H_{17}N_3O_5S_2$, 456.1; found, 456.1.

*2-((tert-Butoxycarbonyl)amino)benzoic Acid (**38**).* Based on a reported procedure, 2-aminobenzoic acid (1 equiv, 1.000 g, 7.29 mmol) and di-*tert*-butyl dicarbonate (2 equiv, 3.183 g, 14.58 mmol) were dissolved in a 2:1 mixture of 1,4-dioxane and H_2O (50 mL).[75] Triethylamine (2 equiv, 2.02 mL, 14.58 mmol) was added and the mixture was stirred for 14 h. Dioxane was evaporated under reduced pressure and the aqueous layer was acidified to pH 2 with concentrated HCl and extracted three times with ethyl acetate (3 × 15 mL). The organic layer was separated and the ethyl acetate was removed in vacuo. The residue was dissolved in diethyl ether (20 mL) and crystallization was induced by the addition of *n*-pentane (15 mL). After filtration, 2-((*tert*-butoxycarbonyl)amino)benzoic acid was obtained as colorless crystals with a yield of 69% (1.195 g, 5.037 mmol). ^{1}H NMR (300 MHz, DMSO-d_6, δ): 13.59 (s, 1H), 10.52 (s, 1H), 8.28 (dd, J = 8.6, 1.1 Hz, 1H,), 7.96 (dd, J = 8.0, 1.7 Hz, 1H), 7.56 (ddd, J = 8.6, 7.3, 1.7 Hz, 1H), 7.07 (ddd, J = 8.0, 7.3, 1.1 Hz, 1H), 1.48 ppm (s, 9H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.1, 152.5, 142.0, 134.7, 131.7, 121.9, 118.5, 115.6, 80.6, 28.4 ppm (3 × C).

*tert-Butyl-(2-((5,6-dimethoxybenzo[d]thiazol-2-yl)carbamoyl)-phenyl)carbamate (**39**).* General procedure B; colorless solid, yield 97% (593 mg, 1.38 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 10.04 (s, 1H), 8.41 (d, J = 8.5 Hz, 1H), 7.78 (d, J = 7.9 Hz, 1H), 7.54−7.40 (m, 1H), 7.20 (s, 1H), 6.89−6.79 (m, 1H), 6.60 (s, 1H), 4.94 (s, 3H), 3.64 (s, 3H), 1.55 ppm (s, 9H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 167.1, 158.2, 152.9, 149.3, 147.7, 141.4, 141.1, 134.0, 128.4, 123.1, 122.0, 120.4, 118.4, 102.6, 102.4, 80.9, 56.3, 55.9, 28.4 ppm (3 × C).

*2-Amino-N-(5,6-dihydroxybenzo[d]thiazol-2-yl)benzamide (**40**).* General procedure E; yellow solid; yield 86% (211 mg, 0.70 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 8.05 (dd, J = 7.9, 1.2 Hz, 1H), 7.55 (ddd, J = 8.6, 7.3, 1.5 Hz, 1H), 7.36−7.20 (m, 3H), 7.14 ppm (s, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 159.7, 150.0, 147.7, 146.6, 139.0, 136.0, 133.0, 126.9, 124.8, 122.8, 119.7, 118.2, 116.1, 115.9 ppm.

*(S)-N-(5,6-Dimethoxybenzo[d]thiazol-2-yl)-1-((4-(hydroxymethyl)phenyl)sulfonyl)pyrrolidine-2-carboxamide ((S)-**42**).* (*S*)-**44** (90 mg, 0.17 mmol, 1.0 equiv) was dissolved in THF (12 mL) and LiOH (41 mg, 1.73 mmol, 10.0 equiv) and water (3 mL) was added. The solution was stirred for 16 h, followed by the addition of water (10 mL) and extraction of the aqueous phase with ethyl acetate (3 × 10 mL). The organic phase was washed with saturated NaCl solution (10 mL) and the solvent was evaporated under reduced pressure. The residue was purified by column chromatography (ethyl acetate/cyclohexane = 1:1) to yield (*S*)-**42** (70 mg, 0.15 mmol, yield 85%) as a yellow solid.

Mp 64−66 °C; ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.86 (d, J = 8.0 Hz, 2H), 7.63−7.53 (m, 3H, *H*-8), 7.31 (s, 1H), 5.47 (t, J = 5.7 Hz, 1H), 4.62 (d, J = 5.7 Hz, 2H), 4.46−4.33 (m, 1H), 3.84 (s, 3H), 3.82 (s, 3H), 3.57−3.43 (m, 1H), 3.26−3.13 (m, 1H), 1.96−1.81 (m, 3H), 1.59−1.42 ppm (m, 1H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 170.7, 156.1, 149.0, 148.4, 147.1, 142.5, 134.9, 127.3 (2 x C), 126.9 (2 × C), 123.0, 103.7, 103.5, 62.2, 60.8, 56.0, 55.8, 49.2, 31.0, 24.5 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{21}H_{23}N_3O_6S_2$, 478.1; found, 478.1. $[\alpha]_D^{24}$ − 129 (c 10 mg cm^{-3} in ACN).

*(S)-4-((2-((5,6-Dimethoxybenzo[d]thiazol-2-yl)carbamoyl)-pyrrolidine-1-yl)sulfonyl)benzyl Acetate ((S)-**44**).* (*S*)-**4** (200 mg, 0.58 mmol, 1.0 equiv) and sodium acetate (191 mg, 2.33 mmol, 4.0 equiv) were dissolved in DMF (15 mL) and 4-bromomethylphenylsulfonyl chloride (157 mg, 0.58 mmol, 1.0 equiv) was added. After stirring at 50 °C for 8 h, ethyl acetate (20 mL) and water (20 mL) were added and the aqueous phase was extracted with ethyl acetate (3 × 20 mL). The combined organic phases were washed with 2 M HCl, saturated $NaHCO_3$ solution, and saturated NaCl solution three times each (20 mL). The organic solvent was removed under reduced pressure to yield (*S*)-**44** (224 mg, 0.43 mmol, yield 74%) as a colorless solid.

Mp 165−168 °C; ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.87 (d, J = 8.3 Hz, 2H), 7.53 (d, J = 8.3 Hz, 2H), 7.31 (s, 1H), 7.22 (s, 1H), 5.17 (s, 2H), 4.38−4.26 (m, 1H), 3.95−3.92 (s, 6H), 3.70−3.54 (m, 1H), 3.30−3.16 (m, 1H), 2.39−2.25 (m, 1H), 2.13 (s, 3H), 1.95−1.64 ppm (m, 3H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 170.6, 169.4, 155.8, 149.5, 147.8, 142.6, 142.4, 135.2, 128.6 (2 x C), 128.3 (2 × C), 123.8, 103.6, 102.7, 64.9, 62.4, 56.5, 56.2, 50.1, 30.1, 24.6, 20.9 ppm. MS (ESI) m/z: $[M + H]^+$ calcd for $C_{23}H_{25}N_3O_7S_2$, 520.1; found, 520.1. $[\alpha]_D^{24}$ − 182 (c 10 mg cm^{-3} in chloroform).

*(R)-1-(tert-Butoxycarbonyl)pipecolinic Acid ((R)-**45**).* General procedure A; colorless solid, 89% (635 mg, 2.77 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 4.74−4.44 (m, 1H), 3.93−3.66 (m, 1H), 3.04−2.69 (m, 1H), 2.16−2.02 (m, 1H), 1.73−1.48 (m, 3H), 1.37 (s, 9H), 1.20−1.02 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 173.0, 155.0, 79.0, 54.4, 41.6, 28.0 (3 × C), 26.4, 24.4, 20.4 ppm. $[\alpha]_D^{24}$ + 44 (c 10 mg cm^{-3} in methanol).

*(S)-1-(tert-Butoxycarbonyl)pipecolinic Acid ((S)-**45**).* General procedure A; colorless solid, 89% (667 mg, 2.91 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 4.67−4.45 (m, 1H), 3.87−3.71 (m, 1H), 3.05−2.66 (m, 1H), 2.13−2.01 (m, 1H), 1.72−1.50 (m, 3H), 1.37 (s, 9H), 1.22−1.02 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 173.0, 155.0, 79.0, 54.3, 41.6, 28.0 (3 × C), 26.4, 24.4, 20.4 ppm. $[\alpha]_D^{24}$ − 37 (c 10 mg cm^{-3} in methanol).

*1-(tert-Butoxycarbonyl)piperidine-4-carboxylic Acid (**46**).* General procedure A; colorless solid, 94% (1.67 mg, 7.74 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 3.94−3.73 (m, 2H), 2.96−2.71 (m, 2H), 2.45−2.34 (m, 1H), 1.87−1.69 (m, 2H), 1.51−1.28 ppm (m, 11H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 175.7, 153.9, 78.6 (2 × C), 42.9, 39.4, 28.1 (3 × C), 27.8 ppm (2 × C).

*(R)-tert-Butyl-2-((5,6-dimethoxybenzo[d]thiazol-2-yl)-carbamoyl)piperidine-1-carboxylate ((R)-**47**).* General procedure A; colorless solid; 89% (655 mg, 1.55 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.30 (s, 1H), 7.26 (s, 1H), 4.24−4.04 (m, 1H), 3.98 (s, 3H), 3.97 (s, 3H), 2.97−2.78 (m, 1H), 2.42−2.32 (m, 1H), 1.78−1.59 (m, 4H), 1.59−1.42 ppm (m, 11H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 170.0, 156.5, 155.8, 149.6, 147.8, 141.9, 123.5, 103.2, 102.8, 81.6, 60.5, 56.5, 56.3, 42.8, 28.5 (3 × C), 25.2, 24.8, 20.6 ppm. $[\alpha]_D^{24}$ + 73 (c 6.9 mg cm^{-3} in methanol).

*(S)-tert-Butyl-2-((5,6-dimethoxybenzo[d]thiazol-2-yl)-carbamoyl)piperidine-1-carboxylate ((S)-**47**).* General procedure A; colorless solid; 98% (715 mg, 1.70 mmol); ^{1}H NMR (300 MHz,

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

chloroform-*d*, δ): 7.27 (s, 1H), 7.22 (s, 1H), 4.21−4.01 (m, 1H), 3.95 (s, 3H), 3.94 (s, 3H), 2.99−2.83 (m, 1H), 2.41−2.28 (m, 1H), 1.78−1.57 (m, 4H), 1.57−1.39 ppm (m, 11H). ^{13}C NMR (75 MHz, $CDCl_3$, δ): 170.2, 157.0, 155.8, 149.8, 148.0, 140.6, 122.9, 102.8, 102.8, 81.5, 60.5, 56.5, 56.3, 42.8, 28.5 (3 × C), 25.4, 24.7, 20.6 ppm. $[\alpha]_D^{24}$ − 40 (*c* 6.7 mg cm^{-3} in methanol).

tert-Butyl-4-((5,6-dimethoxybenzo[d]thiazol-2-yl)carbamoyl)-piperidine-1-carboxylate (**48**). General procedure A; colorless solid; yield 50% (366 mg, 0.87 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.24 (s, 1H), 7.21 (s, 1H), 4.20−3.99 (m, 2H), 3.95 (s, 3H), 3.91 (s, 3H), 2.76−2.46 (m, 3H), 1.90−1.59 (m, 4H), 1.44 ppm (s, 9H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 173.0, 158.5, 154.6, 150.0, 148.2, 139.3, 122.5, 103.0, 102.1, 80.0, 77.4, 56.6, 56.4, 43.1 (2 × C), 28.5 (3 × C), 28.2 ppm (3 × C).

tert-Butyl-(2-((5,6-dimethoxybenzo[d]thiazol-2-yl)amino)-2-oxoethyl)carbamate (**49**). General procedure A; off-white solid; yield 66% (551 mg, 1.50 mmol); ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.25 (s, 1H), 7.20 (s, 1H), 4.17 (d, *J* = 5.6 Hz, 2H), 3.94 (s, 3H), 3.91 (s, 3H), 1.45 ppm (s, 9H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 168.7, 157.4, 156.3, 149.7, 147.9, 140.3, 122.9, 102.8, 102.7, 80.9, 56.5, 56.3, 44.7, 28.4 ppm (3 × C).

*(R)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)piperidin-2-carboxamide Hydrobromide ((R)-***50**). General procedure E; off-white solid; yield 55% (220 mg, 0.59 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.27 (s, 1H), 7.13 (s, 1H), 4.19−3.96 (m, 1H), 3.39−3.21 (m, 1H), 3.10−2.87 (m, 1H), 2.32−2.12 (m, 1H), 1.89−1.39 ppm (m, 5H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 174.3, 155.5, 145.9, 144.3, 140.9, 121.7, 106.6, 106.3, 57.3, 43.5, 26.7, 21.6, 21.2 ppm. $[\alpha]_D^{24}$ + 5 (*c* 10 mg cm^{-3} in methanol).

*(S)-N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)piperidin-2-carboxamide Hydrobromide ((S)-***50**). General procedure E; off-white solid; yield 62% (248 mg, 0.66 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.27 (s, 1H), 7.13 (s, 1H), 4.14−3.99 (m, 1H), 3.35−3.24 (m, 1H), 3.06−2.90 (m, 1H), 2.30−2.19 (m, 1H), 1.87−1.45 ppm (m, 5H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 176.5, 154.8, 145.9, 144.3, 140.8, 123.2, 107.3, 106.6, 57.2, 43.5, 26.7, 21.6, 21.2 ppm. $[\alpha]_D^{24}$ − 7 (*c* 10 mg cm^{-3} in methanol).

N-(5,6-Dihydroxybenzo[d]thiazol-2-yl)piperidin-4-carboxamide Hydrobromide (**51**). General procedure E; colorless solid; yield 74% (197 mg, 0.53 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.22 (s, 1H), 7.09 (s, 1H), 3.42−3.26 (m, 2H), 3.05−2.74 (m, 3H), 2.08−1.94 (m, 2H), 1.92−1.71 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO-d_6, δ): 172.3, 155.4, 145.6, 143.9, 141.6, 121.8, 106.5, 106.4, 42.4 (2 × C), 38.8, 24.7 ppm (2 × C).

2-Amino-N-(5,6-dimethoxybenzo[d]thiazol-2-yl)acetamide Hydrobromide (**52**). General procedure E; colorless solid; yield 96% (418 mg, 1.31 mmol); ^{1}H NMR (300 MHz, DMSO-d_6, δ): 7.26 (s, 1H), 7.13 (s, 1H), 4.00−3.82 ppm (m, 2H). ^{13}C NMR (75 MHz, DMSO, δ): 165.6, 154.8, 145.8, 144.2, 141.5, 121.8, 106.6, 106.5, 40.9 ppm.

Benzenesulfonyl Chloride (**53**). According to a reported procedure,[29] a round-bottom flask was charged with chlorosulfonic acid (3.0 equiv, 5.0 mL, 75 mmol). Benzene (1.0 equiv, 2.2 mL, 25 mmol) was added while cooling with an ice-water bath and the mixture was stirred at room temperature until gas development stopped. After addition of crushed ice, the mixture was extracted with chloroform (3 × 10 mL). The organic phase was washed with water (10 mL), saturated $NaHCO_3$ solution (10 mL), dried with Na_2SO_4, and evaporated. The residue was distilled at 107 °C (5 mbar) to afford benzene sulfonyl chloride as a colorless liquid (1.52 g, 8.6 mmol, 34% yield, lit: 62%[29]). ^{1}H NMR (300 MHz, chloroform-*d*, δ): 8.13−7.95 (m, 2H), 7.81−7.70 (m, 1H), 7.70−7.56 ppm (m, 2H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 143.8, 134.8, 129.2 (2 × C), 126.4 ppm (2 × C).

4-Methoxybenzene-1-sulfonyl Chloride (**54**). According to a reported procedure,[30] anisole (1 equiv, 4.1 mL, 37.5 mmol) was dissolved in chloroform (20 mL) and chlorosulfonic acid (2 equiv, 5.0 mL, 75 mmol) was added at −10 °C. After warming up to room temperature, the mixture was stirred until the expiration of the gas development. Crushed ice was added, the phases were separated, and the aqueous phase was extracted with chloroform (3 × 10 mL). The organic phases were combined and washed with water (10 mL) and saturated $NaHCO_3$ solution (10 mL), dried with Na_2SO_4, and evaporated. The residue was purified by column chromatography (ethyl acetate/cyclohexane = 1:2) to afford 4-methoxybenzene-1-sulfonyl chloride as colorless crystals (3.4 g, 16 mmol, yield: 66%, lit.: 72%[30]). ^{1}H NMR (300 MHz, chloroform-*d*, δ): 7.97 (d, *J* = 9.1, 2H), 7.05 (d, *J* = 9.1, 2H), 3.92 ppm (s, 3H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 165.0, 136.2, 129.7 (2 × C), 114.9 (2 × C), 56.1 ppm.

3-Nitrobenzene-1-sulfonyl Chloride (**55**). According to a reported procedure,[31] chlorosulfonic acid (3 equiv, 5.0 mL, 75 mmol) and nitrobenzene (1 equiv, 2.6 mL, 25 mmol) were heated to 110 °C until the development of gas was slowing down. Then, the temperature was increased to 130 °C until expiration of the gas development. After cooling down, crushed ice was added and the reaction mixture was extracted with chloroform (3 × 10 mL). The organic phase was washed with water (10 mL) and saturated $NaHCO_3$ solution, dried with Na_2SO_4, and evaporated. The residue was purified by column chromatography (ethyl acetate/cyclohexane = 5: 1) to yield 3-nitrobenzene-1-sulfonyl chloride as colorless crystals (500 mg, 2.2 mmol, yield: 9%, lit.: 56%[31]). ^{1}H NMR (300 MHz, chloroform-*d*, δ): 8.89 (m, 1H), 8.65−8.56 (m, 1H), 8.43−8.32 (m, 1H), 7.96−7.84 ppm (m, 1H). ^{13}C NMR (75 MHz, chloroform-*d*, δ): 148.5, 145.7, 132.4, 131.5, 129.7, 122.6 ppm.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.9b01697.

Additional data of the enantiomeric ratios, inhibition studies, LLE calculations, site-directed mutagenesis, docking and molecular dynamics studies, test results for ZIKV, and cell permeability (PDF)

HPLC traces of key target compounds (PDF)

Molecular formula strings (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: U.Hellmich@uni-mainz.de (U.A.H.).

*E-mail: schirmei@uni-mainz.de (T.S.).

ORCID

Benedikt Millies: 0000-0002-9125-7308

Ute A. Hellmich: 0000-0001-7162-285X

Author Contributions

[V]U.A.H. and T.S. share senior authorship. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

U.A.H. acknowledges support by the Carl Zeiss Foundation and the Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe University Frankfurt, funded by the State of Hesse. E.H. obtained a grant by the LOEWE Center DRUID (Novel Drug Targets against Povertry-Related and Neglected Tropical Diseases).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Didier Musso from the Institute Louis in Papeete, Malardé, Tahiti, for the ZIKV PF13/251013 strain and the European Virus Archive for the ZIKV 976 Uganda strain. The contribution of Dr. Thien Anh Le and Prof. Bernd

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Engels from the Institute of Physical and Theoretical Chemistry, Würzburg, concerning the QM calculations is gratefully acknowledged. We thank Leander Gerke from the group of Prof. Till Opatz, JGU Mainz for assisting the enantiomeric ratios' determination, Collin Zimmer for assisting the metabolism studies, and Jonas Busold for assisting the syntheses.

■ ABBREVIATIONS

AMC, 7-amino-4-methylcoumarin; BMI, *N*-benzylmaleimide; CHX, cycloheximide; CTCF, corrected total cell fluorescence; DENV, Dengue virus; DMEM, Dulbecco's Modified Eagle Medium; er, enantiomeric ratio; EMI, *N*-ethylmaleimide; FP, French Polynesia; LLE, lipophilic ligand efficiency; MOI, multiplicity of infection; MST, microscale thermophoresis; MTS, (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium); QM, quantum mechanics; ROC, receiver operator characteristics; RTqPCR, real time quantitative polymerase chain reaction; STING, stimulator of interferon genes; U, Uganda; WNV, West Nile virus; ZIKV, Zika virus

■ REFERENCES

(1) Campos, G. S.; Bandeira, A. C.; Sardi, S. I. Zika Virus Outbreak, Bahia, Brazil. *Emerging Infect. Dis.* **2015**, *21*, 1885−1886.
(2) Musso, D.; Gubler, D. J. Zika Virus. *Clin. Microbiol. Rev.* **2016**, *29*, 487−524.
(3) Hennessey, M.; Fischer, M.; Staples, J. E. Zika Virus Spreads to New Areas - Region of the Americas, May 2015-January 2016. *MMWR Morb. Mortal. Wkly. Rep.* **2016**, *65*, 55−58.
(4) Hazin, A. N.; Poretti, A.; Di Cavalcanti Souza Cruz, D.; Tenorio, M.; van der Linden, A.; Pena, L. J.; Brito, C.; Gil, L. H. V.; de Barros Miranda-Filho, D.; Marques, E. T. D. A.; Turchi Martelli, C. M.; Alves, J. G. B.; Huisman, T. A. Computed Tomographic Findings in Microcephaly Associated with Zika Virus. *N. Engl. J. Med.* **2016**, *374*, 2193−2195.
(5) Chan, J. F. W.; Choi, G. K. Y.; Yip, C. C. Y.; Cheng, V. C. C.; Yuen, K.-Y. Zika Fever and Congenital Zika Syndrome: An Unexpected Emerging Arboviral Disease. *J. Infect.* **2016**, *72*, 507−524.
(6) Mlakar, J.; Korva, M.; Tul, N.; Popović, M.; Poljšak-Prijatelj, M.; Mraz, J.; Kolenc, M.; Resman Rus, K.; Vesnaver Vipotnik, T.; Fabjan Vodušek, V.; Vizjak, A.; Pižem, J.; Petrovec, M.; Avšič Šupanc, T. Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* **2016**, *374*, 951−958.
(7) Duffy, M. R.; Chen, T.-H.; Hancock, W. T.; Powers, A. M.; Kool, J. L.; Lanciotti, R. S.; Pretrick, M.; Marfel, M.; Holzbauer, S.; Dubray, C.; Guillaumot, L.; Griggs, A.; Bel, M.; Lambert, A. J.; Laven, J.; Kosoy, O.; Panella, A.; Biggerstaff, B. J.; Fischer, M.; Hayes, E. B. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. *N. Engl. J. Med.* **2009**, *360*, 2536−2543.
(8) Aubry, M.; Teissier, A.; Huart, M.; Merceron, S.; Vanhomwegen, J.; Roche, C.; Vial, A.-L.; Teururai, S.; Sicard, S.; Paulous, S.; Desprès, P.; Manuguerra, J.-C.; Mallet, H.-P.; Musso, D.; Deparis, X.; Cao-Lormeau, V.-M. Zika Virus Seroprevalence, French Polynesia, 2014-2015. *Emerging Infect. Dis.* **2017**, *23*, 669−672.
(9) Cao-Lormeau, V.-M.; Blake, A.; Mons, S.; Lastère, S.; Roche, C.; Vanhomwegen, J.; Dub, T.; Baudouin, L.; Teissier, A.; Larre, P.; Vial, A.-L.; Decam, C.; Choumet, V.; Halstead, S. K.; Willison, H. J.; Musset, L.; Manuguerra, J.-C.; Desprès, P.; Fournier, E.; Mallet, H.-P.; Musso, D.; Fontanet, A.; Neil, J.; Ghawché, F. Guillain-Barré Syndrome Outbreak Associated with Zika Virus Infection in French Polynesia: A Case-Control Study. *Lancet* **2016**, *387*, 1531−1539.
(10) Chambers, T. J.; Grakoui, A.; Rice, C. M. Processing of the Yellow Fever Virus Nonstructural Polyprotein: A Catalytically Active NS3 Proteinase Domain and NS2B are Required for Cleavages at Dibasic Sites. *J. Virol.* **1991**, *65*, 6042−6050.
(11) Aguirre, S.; Maestre, A. M.; Pagni, S.; Patel, J. R.; Savage, T.; Gutman, D.; Maringer, K.; Bernal-Rubio, D.; Shabman, R. S.; Simon, V.; Rodriguez-Madoz, J. R.; Mulder, L. C. F.; Barber, G. N.; Fernandez-Sesma, A. DENV Inhibits Type I IFN Production in Infected Cells by Cleaving Human STING. *PLoS Pathog.* **2012**, *8*, No. e1002934.
(12) Luo, D.; Vasudevan, S. G.; Lescar, J. The Flavivirus NS2B-NS3 Protease-Helicase as a Target for Antiviral Drug Development. *Antiviral Res.* **2015**, *118*, 148−158.
(13) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. *Science* **2016**, *353*, 503−505.
(14) von Hammerstein, F.; Lauth, L. M.; Hammerschmidt, S.; Wagner, A.; Schirmeister, T.; Hellmich, U. A. *Cis*-Autocatalytic Cleavage of Glycine-linked Zika Virus NS2B-NS3 Protease Constructs. *FEBS Lett.* **2019**, 2204−2213.
(15) Chen, X.; Yang, K.; Wu, C.; Chen, C.; Hu, C.; Buzovetsky, O.; Wang, Z.; Ji, X.; Xiong, Y.; Yang, H. Mechanisms of Activation and Inhibition of Zika Virus NS2B-NS3 Protease. *Cell Res.* **2016**, *26*, 1260−1263.
(16) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D. Structure of the NS2B-NS3 Protease from Zika Virus after Self-Cleavage. *Nat. Commun.* **2016**, *7*, 13410.
(17) Mahawaththa, M. C.; Pearce, B. J.; Szabo, M.; Graham, B.; Klein, C. D.; Nitsche, C.; Otting, G. Solution Conformations of a Linked Construct of the Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *142*, 141−147.
(18) Hill, M. E.; Yildiz, M.; Hardy, J. A. Cysteine Disulfide Traps Reveal Distinct Conformational Ensembles in Dengue Virus NS2B-NS3 Protease. *Biochemistry* **2018**, *58*, 776−787.
(19) Lee, H.; Ren, J.; Nocadello, S.; Rice, A. J.; Ojeda, I.; Light, S.; Minasov, G.; Vargas, J.; Nagarathnam, D.; Anderson, W. F.; Johnson, M. E. Identification of Novel Small Molecule Inhibitors Against NS2B/NS3 Serine Protease from Zika Virus. *Antiviral Res.* **2017**, *139*, 49−58.
(20) Nitsche, C.; Zhang, L.; Weigel, L. F.; Schilz, J.; Graf, D.; Bartenschlager, R.; Hilgenfeld, R.; Klein, C. D. Peptide-Boronic Acid Inhibitors of Flaviviral Proteases: Medicinal Chemistry and Structural Biology. *J. Med. Chem.* **2017**, *60*, 511−516.
(21) Kuno, G.; Chang, G. J.; Tsuchiya, K. R.; Karabatsos, N.; Cropp, C. B. Phylogeny of the Genus Flavivirus. *J. Virol.* **1998**, *72*, 73−83.
(22) Behnam, M. A. M.; Graf, D.; Bartenschlager, R.; Zlotos, D. P.; Klein, C. D. Discovery of Nanomolar Dengue and West Nile Virus Protease Inhibitors Containing a 4-Benzyloxyphenylglycine Residue. *J. Med. Chem.* **2015**, *58*, 9354−9370.
(23) Li, Z.; Brecher, M.; Deng, Y.-Q.; Zhang, J.; Sakamuru, S.; Liu, B.; Huang, R.; Koetzner, C. A.; Allen, C. A.; Jones, S. A.; Chen, H.; Zhang, N.-N.; Tian, M.; Gao, F.; Lin, Q.; Banavali, N.; Zhou, J.; Boles, N.; Xia, M.; Kramer, L. D.; Qin, C.-F.; Li, H. Existing Drugs as Broad-Spectrum and Potent Inhibitors for Zika Virus by Targeting NS2B-NS3 Interaction. *Cell Res.* **2017**, *27*, 1046−1064.
(24) Li, Z.; Sakamuru, S.; Huang, R.; Brecher, M.; Koetzner, C. A.; Zhang, J.; Chen, H.; Qin, C.-F.; Zhang, Q.-Y.; Zhou, J.; Kramer, L. D.; Xia, M.; Li, H. Erythrosin B is a Potent and Broad-Spectrum Orthosteric Inhibitor of the Flavivirus NS2B-NS3 Protease. *Antiviral Res.* **2018**, *150*, 217−225.
(25) Nitsche, C.; Passioura, T.; Varava, P.; Mahawaththa, M. C.; Leuthold, M. M.; Klein, C. D.; Suga, H.; Otting, G. De Novo Discovery of Nonstandard Macrocyclic Peptides as Noncompetitive Inhibitors of the Zika Virus NS2B-NS3 Protease. *ACS Med. Chem. Lett.* **2019**, *10*, 168−174.
(26) 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.
(27) Wu, H.; Bock, S.; Snitko, M.; Berger, T.; Weidner, T.; Holloway, S.; Kanitz, M.; Diederich, W. E.; Steuber, H.; Walter, C.;

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Hofmann, D.; Weißbrich, B.; Spannaus, R.; Acosta, E. G.; Bartenschlager, R.; Engels, B.; Schirmeister, T.; Bodem, J. Novel Dengue Virus NS2B/NS3 Protease Inhibitors. *Antimicrob. Agents Chemother.* **2015**, *59*, 1100−1109.
(28) 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.
(29) Ullmann, F. Notiz über die Darstellung von Benzolsulfochlorid. *Ber. Dtsch. Chem. Ges.* **1909**, *42*, 2057−2058.
(30) Morgan, M. S.; Cretcher, L. H. A Kinetic Study of Alkylation by Ethyl Arylsulfonates. *J. Am. Chem. Soc.* **1948**, *70*, 375−378.
(31) Mehrotra, S.; Angamuthu, R. Janus Head Type Lone Pair−π−Lone Pair and S···F···S Interactions in Retaining Hexafluorobenzene. *CrystEngComm* **2016**, *18*, 4438−4444.
(32) Brenk, R.; Schipani, A.; James, D.; Krasowski, A.; Gilbert, I. H.; Frearson, J.; Wyatt, P. G. Lessons Learnt from Assembling Screening Libraries for Drug Discovery for Neglected Diseases. *ChemMedChem* **2008**, *3*, 435−444.
(33) Nepali, K.; Lee, H.-Y.; Liou, J.-P. Nitro-Group-Containing Drugs. *J. Med. Chem.* **2019**, *62*, 2851−2893.
(34) 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*, 368−382.
(35) Dixon, M. The Determination of Enzyme Inhibitor Constants. *Biochem. J.* **1953**, *55*, 170−171.
(36) Wildman, S. A.; Crippen, G. M. Prediction of Physicochemical Parameters by Atomic Contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 868−873.
(37) Leeson, P. D.; Springthorpe, B. The Influence of Drug-Like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* **2007**, *6*, 881.
(38) 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.
(39) Mukhametov, A.; Newhouse, E. I.; Aziz, N. A.; Saito, J. A.; Alam, M. Allosteric Pocket of the Dengue Virus (Serotype 2) NS2B/NS3 Protease: In Silico Ligand Screening and Molecular Dynamics Studies of Inhibition. *J. Mol. Graphics Modell.* **2014**, *52*, 103−113.
(40) 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.* **2011**, *86*, 438−446.
(41) Yao, Y.; Huo, T.; Lin, Y.-L.; Nie, S.; Wu, F.; Hua, Y.; Wu, J.; Kneubehl, A. R.; Vogt, M. B.; Rico-Hesse, R.; Song, Y. Discovery, X-ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease. *J. Am. Chem. Soc.* **2019**, *141*, 6832−6836.
(42) Grinter, R. *Questionable Ligand Density: 6MO0, 6MO1, 6MO2*. https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47072.html (accessed Aug 20, 2019).
(43) Green, T. J. *Questionable Ligand Density: 6MO0, 6MO1, 6MO2*. https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47086.html (accessed Aug 22, 2019).
(44) Volkamer, A.; Kuhn, D.; Grombacher, T.; Rippmann, F.; Rarey, M. Combining Global and Local Measures for Structure-Based Druggability Predictions. *J. Chem. Inf. Model.* **2012**, *52*, 360−372.
(45) Volkamer, A.; Griewel, A.; Grombacher, T.; Rarey, M. Analyzing the Topology of Active Sites: on the Prediction of Pockets and Subpockets. *J. Chem. Inf. Model.* **2010**, *50*, 2041−2052.
(46) BioSolveIT. *LeadIT, Version 2.3.2*; BioSolveIT GmbH: Sankt Augustin, Germany, 2017.
(47) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.
(48) Liu, K.; Watanabe, E.; Kokubo, H. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations. *J. Comput.-Aided Mol. Des.* **2017**, *31*, 201−211.
(49) Cavalli, A.; Bottegoni, G.; Raco, C.; de Vivo, M.; Recanatini, M. A Computational Study of the Binding of Propidium to the Peripheral Anionic Site of Human Acetylcholinesterase. *J. Med. Chem.* **2004**, *47*, 3991−3999.
(50) Cortese, M.; Goellner, S.; Acosta, E. G.; Neufeldt, C. J.; Oleksiuk, O.; Lampe, M.; Haselmann, U.; Funaya, C.; Schieber, N.; Ronchi, P.; Schorb, M.; Pruunsild, P.; Schwab, Y.; Chatel-Chaix, L.; Ruggieri, A.; Bartenschlager, R. Ultrastructural Characterization of Zika Virus Replication Factories. *Cell Rep.* **2017**, *18*, 2113−2123.
(51) Rieutord, A.; Stupans, I.; Shenfield, G. M.; Gross, A. S. Gliclazide Hydroxylation by Rat Liver Microsomes. *Xenobiotica* **2008**, *25*, 1345−1354.
(52) Hukkanen, J.; Lassila, A.; Päivärinta, K.; Valanne, S.; Sarpo, S.; Hakkola, J.; Pelkonen, O.; Raunio, H. Induction and Regulation of Xenobiotic-Metabolizing Cytochrome P450s in the Human A549 Lung Adenocarcinoma Cell Line. *Am. J. Respir. Cell Mol. Biol.* **2000**, *22*, 360−366.
(53) Ye, Y.; Zhu, J.; Ai, Q.; Wang, C.; Liao, M.; Fan, H. Quantitative Proteomics Reveals Changes in Vero Cells in Response to Porcine Epidemic Diarrhea Virus. *J. Proteome Res.* **2019**, *18*, 1623−1633.
(54) 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.
(55) Baronti, C.; Piorkowski, G.; Charrel, R. N.; 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*, e00500-14−e00500-14.
(56) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(57) Chemical Computing Group ULC. *Molecular Operating Environment (MOE) 2019.1001*; 1010 Sheerbroke St. West, Suite #910, Montreal, QC, Canada H3A 2R7, 2019.
(58) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490−519.
(59) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09 Revision D.01*; Gaussian. Inc.: Wallingford, CT, 2009.
(60) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(61) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(62) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(63) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.;

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *Amber*; University of California, San Francisco: San Francisco, CA, 2018.

(64) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(65) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(66) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(68) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(69) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(70) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

(71) Steuer, C.; Heinonen, K. H.; Kattner, L.; Klein, C. D. Optimization of Assay Conditions for Dengue Virus Protease: Effect of Various Polyols and Nonionic Detergents. *J. Biomol. Screen.* **2009**, *14*, 1102−1108.

(72) Bustin, S. A.; Benes, V.; Garson, J. A.; Hellemans, J.; Huggett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffl, M. W.; Shipley, G. L.; Vandesompele, J.; Wittwer, C. T. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **2009**, *55*, 611−622.

(73) Heller, B.; Sundermann, B.; Buschmann, H.; Drexler, H.-J.; You, J.; Holzgrabe, U.; Heller, E.; Oehme, G. Photocatalyzed [2 + 2 + 2]-Cycloaddition of Nitriles with Acetylene: An Effective Method for the Synthesis of 2-Pyridines under Mild Conditions. *J. Org. Chem.* **2002**, *67*, 4414−4422.

(74) 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*, 209−216.

(75) Guo, H.; Wang, Z.; Ding, K. PEG-Polymer-Supported Liquid-Phase Combinatorial Synthesis of Structurally Diverse 2,3-Dihydro-4-pyridones. *Synthesis (Stuttg)* **2005**, 1061−1068.

DOI: 10.1021/acs.jmedchem.9b01697
J. Med. Chem. 2019, 62, 11359−11382

Proline-based allosteric inhibitors of Zika and Dengue virus NS2B/NS3 proteases

Benedikt Millies,[†] Franziska von Hammerstein,[†] Andrea Gellert,[†] Stefan Hammerschmidt,[†] Fabian Barthels,[†] Ulrike Göppel,[†] Melissa Immerheiser,[⊥] Fabian Elgner,[‖] Nathalie Jung[⊣], Michael Basic,[‖] Christian Kersten,[†] Werner Kiefer,[†] Jochen Bodem,[⊥] Eberhard Hildt,[‖] Maike Windbergs[⊣], Ute A. Hellmich,[‡,§,#] Tanja Schirmeister[†,#]

[†]Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz, Staudinger Weg 5, 55128 Mainz, Germany

[⊥]Institut für Virology und Immunbiologie, Julius-Maximilians-Universität Würzburg, Versbacher Str. 7, 97078 Würzburg, Germany

[‖]Department of Virology, Paul-Ehrlich-Institut, 63225 Langen, Germany

[‡]Institute of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz, Johann-Joachim-Becherweg 30, 55128 Mainz, Germany

[§]Center for Biomolecular Magnetic Resonance (BMRZ), Goethe-Universität, 60438 Frankfurt am Main, Germany

[⊣] Institute of Pharmaceutical Technology and Buchmann Institute for Molecular Life Sciences, Goethe University, Max-von-Laue-Straße 15, 60438 Frankfurt am Main, Germany

Table of Contents

1 DETERMINATION OF ENANTIOMERIC RATIOS

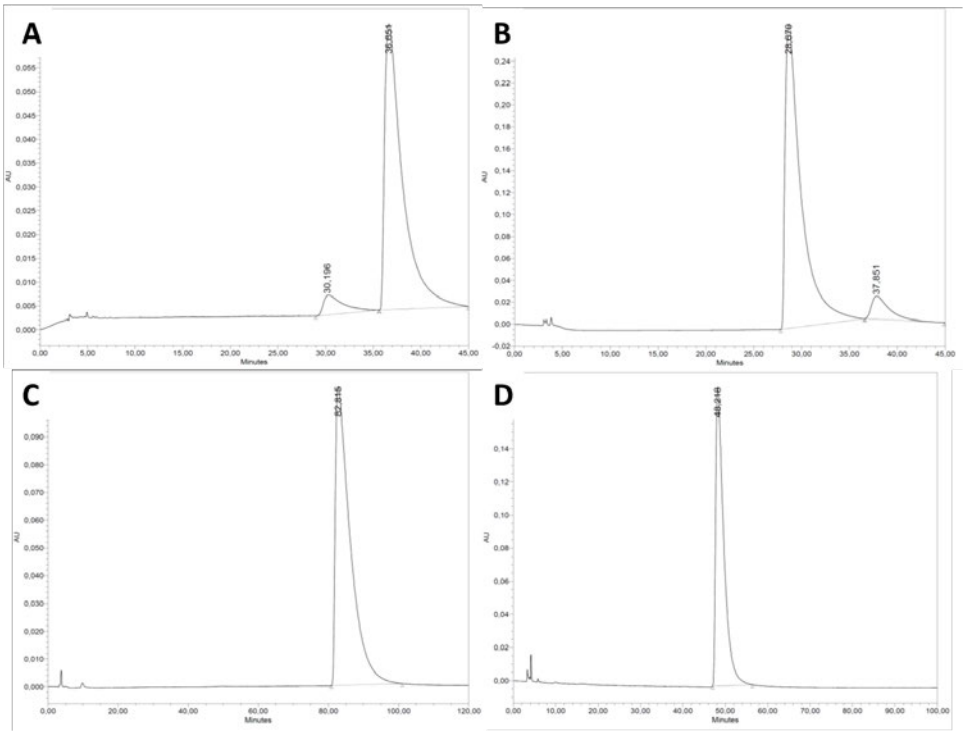


Figure SI1. Chromatograms for the determination of enantiomeric ratios of (*R*)-**7** (A), (*S*)-**7** (B), (*R*)-**18** (C) and (*S*)-**18** (D) by HPLC at 254 nm.

2 IC_{50} DETERMINATIONS AT DIFFERENT SUBSTRATE CONCENTRATIONS

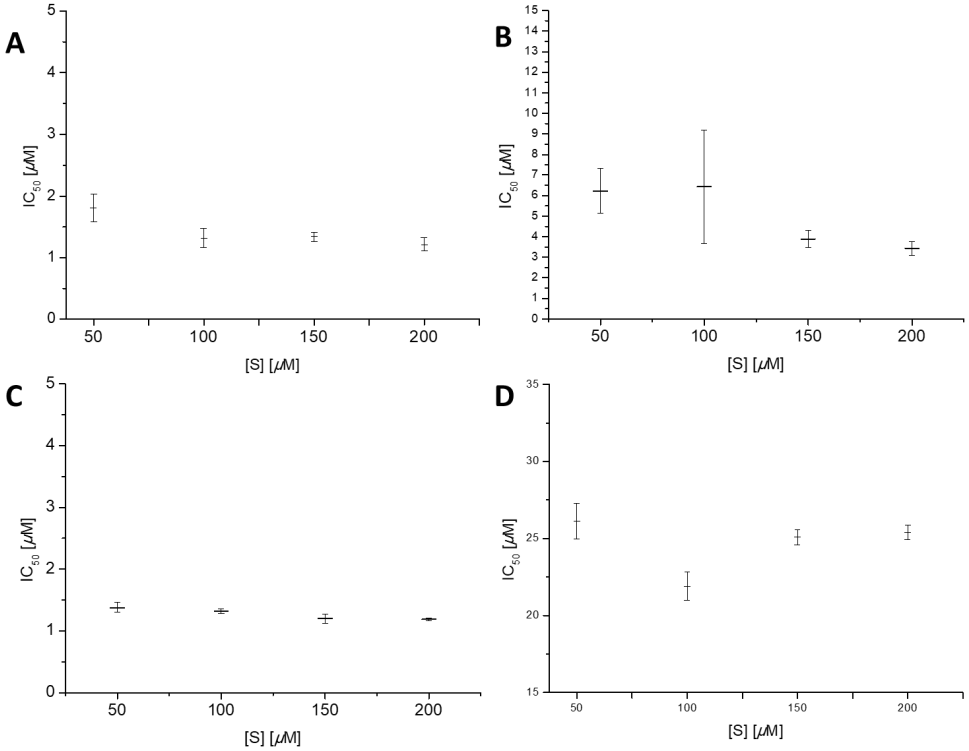


Figure SI 2. IC_{50} values of inhibitors **1** (**A**), **35** (**B**), (*R*)-**7** (**C**) and (*R*)-**21** (**D**) measured on the ZIKV NS2B/NS3 protease at different substrate concentrations. IC_{50} values are indicated as means ± standard deviation from three independent measurements each, performed in duplicates.

3 HYDROPHILICITY DETERMINATION OF SYNTHESIZED COMPOUNDS

SlogP values were calculated with Moe, version 2015.1001, with the SlogP descriptor.[1] Retention times were collected on the LC-MS system described in the experimental section. The lipophilic ligand efficiencies LLE were calculated using the following equation:

$$LLE = pK_i - clogP$$

Table SI1. SlogP values, retention times and lipophilicity ligand efficiencies (LLE) of the synthesized compounds.

Compound	SlogP	t_R	LLE
1	5.61	1.98	0.24
6a	2.95	0.78	3.08
6b	2.95	0.78	2.86
7a	2.81	1.26	3.07
7b	2.81	1.30	3.07
8a	2.50	1.03	3.23
8b	2.50	1.01	3.35
9	2.64	1.16	3.35
10a	3.15	1.54	2.70
10b	3.15	1.57	2.88
11a	2.51	1.11	3.56
11b	2.51	1.09	3.37
12a	2.41	1.16	4.09
12b	2.41	1.18	3.62
13a	4.17	2.84	1.44
13b	4.17	2.93	1.33
14a	2.41	1.12	3.55
14b	2.41	1.11	3.34
15a	3.61	0.82	2.26
15b	3.61	0.83	1.92
26	3.06	1.38	2.29
30a	3.20	1.79	3.04
30b	3.20	1.82	3.09

31	2.72	0.80	3.37
32	3.06	1.33	2.17
35	4.07	2.15	1.12

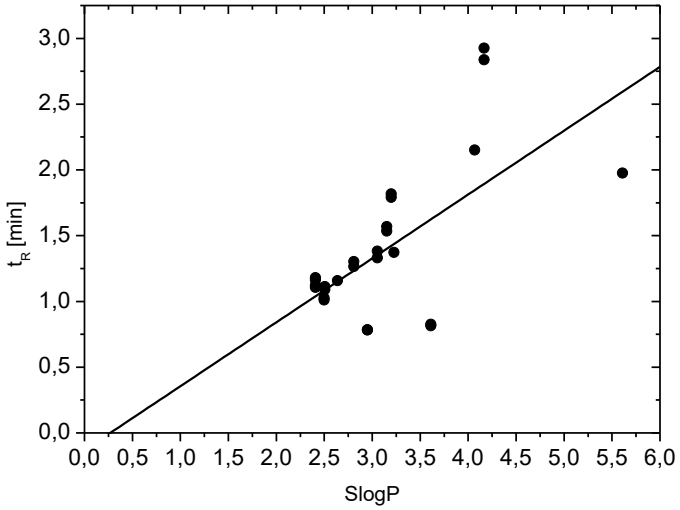


Figure SI3. Correlation of calculated SlogP values with retention times from RP-HPLC runs. Shown is the regression line.

4 CD SPECTROSCOPY STUDIES OF CYSTEINE MUTANTS

MATERIALS AND METHODS Samples of maleimide (MI) -coupled proteases were prepared by diluting 20 mM stock solutions of BMI and EMI with H_2O to a concentration of 200 μM. To 81 μM stock solutions of DENV2 A164C and A166C mutants, a twofold excess of either BMI or EMI was added and the mixture was incubated for coupling over 10 minutes at room temperature and afterwards diluted with H_2O and gel filtration buffer to 5 μM protein concentration, Tris HCl 1.25 mM and NaCl 10 mM.

All CD spectra were measured at room temperature using a JASCO J.815 CD-spectrometer. Control references were carried out with identically treated buffer of the same composition.

The far-UV measurements were conducted 3 times each in the 260 – 190 nm spectral range using a scanning speed of 50 nm min^{-1}, a 5 nm spectral bandwidth, 0.5 s data integration time, a 1 nm data pitch and an accumulation cycle of 5. Spectra were processed with the Jasco Spectra Analysis software (Tokyo, Japan) and analyzed using the Beta Structure Selection method (BeStSel).

RESULTS AND DISCUSSIONS The voltage occurring while performing CD measurements at low wavelengths is a limiting factor of obtaining accurate spectra.[2] Therefore, the salt buffer and DMSO concentrations should be kept constant and as low as tolerated by the measured protein.

Diluting the maleimide (MI) to 200 μM with DMSO prior to reaction with the protease and subsequent CD measurements resulted in a voltage too high to get reliable data, thus the DMSO concentration in the final solutions was lowered. To analyze whether a dilution of the MI in H_2O and successive coupling leads to comparable yields of MI-modified proteases, we performed an enzyme activity assay as described in the main manuscript, with unmodified DENV2 NS2B/NS3 proteases in comparison with the proteases incubated 10 min at room temperature in presence of

10 μM MI diluted with DMSO and H_2O, respectively. In both cases, the same residual activities were seen, indicating no influence of the solvent on the modification reaction (**Figure SI4**).

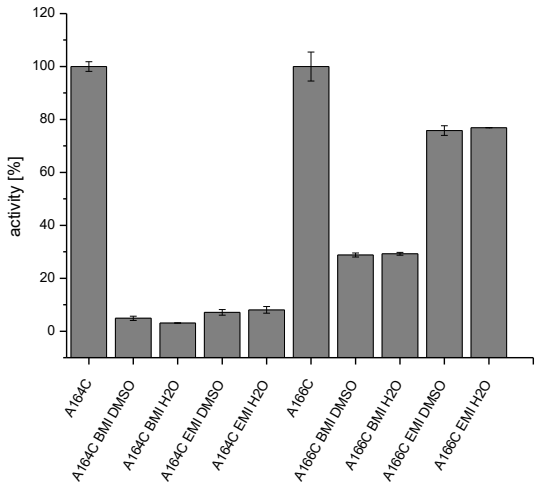


Figure SI4: Enzyme activity assay of the DENV2 A164C and the A166C mutants modified with either 10 μM of the water- or DMSO-diluted MIs. Error bars show the standard deviations of the duplicate measurements.

CD spectra measured for the uncoupled and MI-coupled proteases are shown in **Figure SI5A** for A164C and **Figure SI5B** for A166C, respectively. No shift in secondary structures caused by modification with maleimide could be observed, pointing at an inactivation of the proteases just because of the MI-modification, while retaining the overall folding.

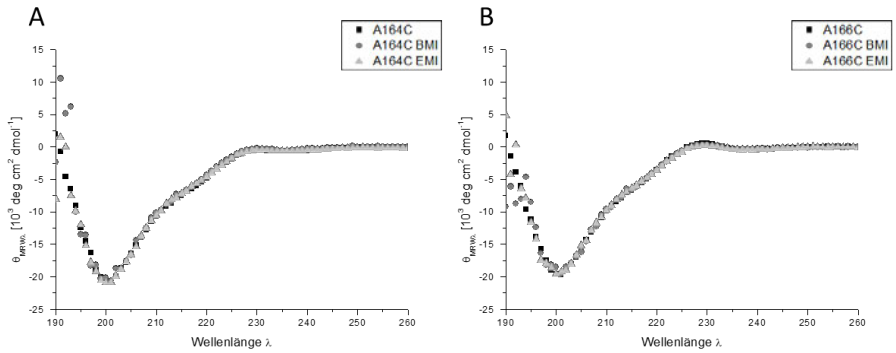


Figure SI5: Overlay of CD spectra as a comparison of maleimide-modified and unmodified DENV2 NS2B/NS3 protease. **A** represents the spectra of the A164C mutant, **B** shows the spectra of A166C, respectively.

As a result of the lack of α-helical structures of the DENV2 NS2B/NS3 protease (

Figure SI11), at a first sight, the spectra look like those of random coiled structures. Therefore, further analyzes using the Beta Structure Selection method (BeStSel) online tool particularly suitable for *β*-rich structures were performed (**Table SI2**).[3]

Table SI2: Estimated secondary structures [%] of the unmodified and MI-modified DENV2 NS2B/NS3 protease cysteine mutants using the BeStSel online tool for analyzing the spectral data ranging from 200 to 250 nm.

Estimated secondary structure content (%)	A164C	A164C BMI	A164C EMI	A166C	A166C BMI	A166C EMI
Helix1 (regular)	0.40	0.00	0.70	0.00	0.00	0.00
Helix2 (distorted)	0.00	0.00	0.00	0.00	0.00	0.00
Anti1 (left-twisted)	0.00	0.00	0.00	0.00	0.00	0.00

Anti2 (relaxed)	2.40	0.90	2.40	3.40	3.90	3.80
Anti3 (right-twisted)	23.00	23.20	22.80	22.60	20.80	22.40
Parallel	0.00	0.00	0.00	0.00	0.00	0.00
Turn	13.20	13.90	13.60	13.30	13.40	13.80
Others	61.00	62.00	60.50	60.60	61.90	60.00
Helix	0.40	0.00	0.70	0.00	0.00	0.00
Antiparallel	25.40	24.20	25.30	26.10	24.60	26.20
Parallel	0.00	0.00	0.00	0.00	0.00	0.00
Turn	13.20	13.90	13.60	13.30	13.40	13.80
Others	61.00	62.00	60.50	60.60	61.90	60.00
Spectral deviation						
RMSD	0.0606	0.0928	0.0650	0.0907	0.1005	0.0852
NRMSD	0.0097	0.0148	0.0104	0.0147	0.0167	0.0142

5 DOCKING STUDIES

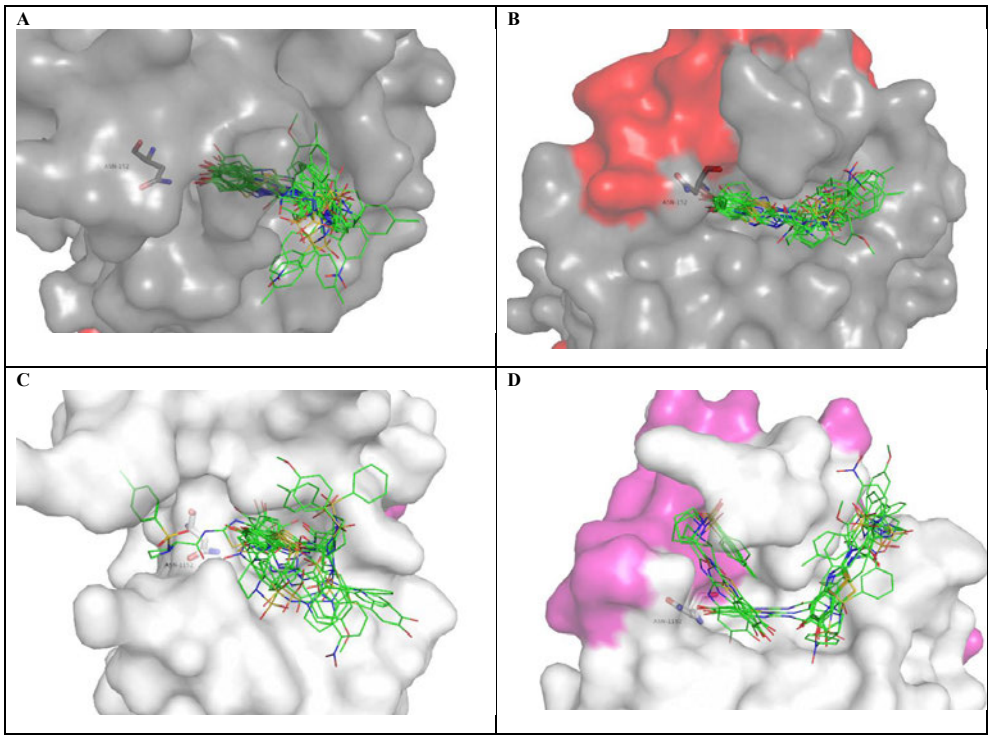


Figure SI6. Superposition of predicted binding modes of compounds **6a**/**b**, **7a**/**b**, **8a**/**b**, **10a**, **11a**/**b**, **12a**/**b**, **14a**/**b**, **35** (as lines with green carbon atoms) in complex with A) DENV protease open conformation (PDB code 2FOM). B) DENV protease closed conformation (PDB code 3U1I). C) ZIKV protease open conformation (PDB code 5GXJ) and D) ZIKV protease closed conformation (PDB code 5LC0). Asn152 (DENV protease) and Asn1152 (ZIKV protease), respectively, are represented as sticks for orientation. DENV NS3 is illustrated as grey surface, NS2B as red surface; ZIKV NS3 as white surface, NS2B as magenta surface.

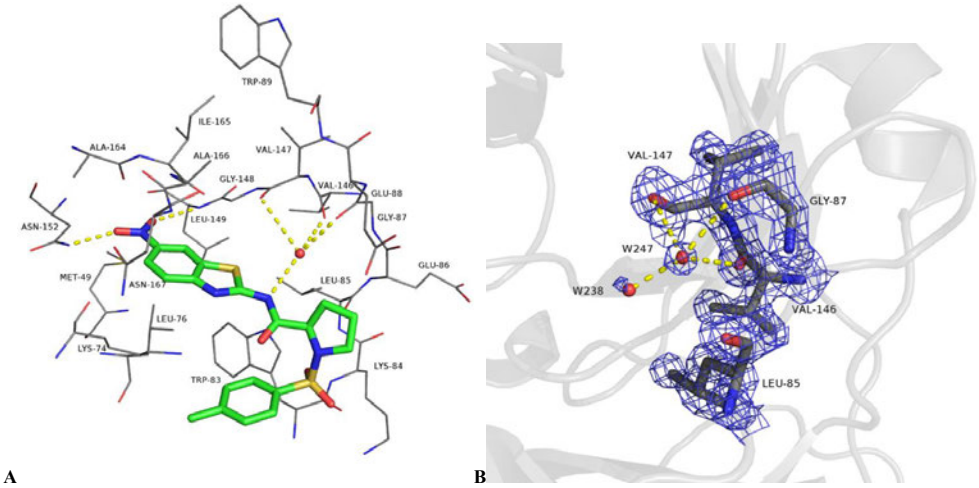


Figure SI7. A) Predicted binding mode of the 6-nitro-benzothiazole compound **22** (green carbon atoms) in complex with the open conformation of DENV protease (grey carbon atoms. PDB code 2FOM). Electrostatic interactions are illustrated as yellow dashes. B) 2Fo-Fc electron density map (PDB code 2FOM) of water molecule W247 and surrounding residues are contoured at 1.5 σ. Possible H-bonds of W247 are illustrated as yellow dashes. W238 with lower electron density is supposed to be less stable bound and displaced by the amide NH-moiety of the ligand.

Table SI3: Rank and FlexX-scores of all 62 compounds within the binder and non-binder discrimination test set for ROC analysis. Compounds **Wu 1** – **SI19** were taken from Wu et al. 2015.[4] Docking was performed against DENV protease *open* (PDB code 2FOM), *closed* (PDB code 3U1I) and ZIKV protease *open* (PDB code 5GXJ) and *closed* (PDB code 5LC0) conformation using LeadIT-2.3.2.

Compound	Binder / Non-binder	Receptor							
		2FOM		3U1I		5GXJ		5LC0	
		Rank	Score	Rank	Score	Rank	Score	Rank	Score
6b	binder	16	-28.08	18	-22.69	8	-21.47	7	-21.81
7b	binder	13	-28.28	29	-20.48	23	-19.9	25	-18.84
12b	binder	20	-27.55	4	-24.25	40	-18.43	18	-19.55
6a	binder	6	-30.04	2	-25.21	18	-20.51	10	-20.89
21a	binder	4	-32.05	26	-20.99	11	-21.11	12	-20.75
35	binder	8	-28.93	1	-25.44	9	-21.36	2	-24.32
7a	binder	28	-26.81	20	-22.01	47	-18.13	37	-17.56
12a	binder	9	-28.39	9	-23.1	21	-20.17	41	-16.78
17b	non-binder	59	-18.32	51	-16.28	48	-18.05	58	-13.69
17a	non-binder	62	-17.81	61	-12.71	50	-17.93	52	-14.57
9	binder	30	-26.71	10	-23.04	42	-18.24	26	-18.62
18b	non-binder	43	-23.84	28	-20.62	6	-22.7	42	-16.4
18a	non-binder	21	-27.54	39	-18.99	29	-19.44	31	-17.99
10a	binder	31	-26.44	12	-22.93	43	-18.24	21	-19.3
15a	binder	15	-28.1	8	-23.42	2	-24.44	5	-22.28
11a	binder	37	-25.37	31	-20.18	37	-18.9	44	-16.22
10b	binder	34	-26.26	16	-22.82	38	-18.74	38	-17.5
15b	binder	26	-26.98	27	-20.96	3	-24.31	6	-22.18
8a	binder	10	-28.38	23	-21.41	10	-21.2	14	-20.3
8b	binder	42	-24.34	19	-22.16	32	-19.1	17	-19.75
20	non-binder	60	-18.14	53	-15.98	51	-17.92	49	-14.99
19b	non-binder	61	-17.84	50	-16.29	58	-16.64	62	-12.73
19a	non-binder	58	-19.65	62	-12.44	52	-17.78	63	-12.68
14b	binder	1	-34.89	5	-24.09	5	-22.8	23	-19.23
14a	binder	12	-28.36	6	-23.78	15	-20.9	3	-22.86

13a	binder	24	-27.09	13	-22.88	41	-18.37	9	-21.29
11b	binder	38	-25.36	30	-20.36	19	-20.38	34	-17.78
13b	binder	11	-28.37	3	-24.72	17	-20.69	32	-17.98
22	binder	17	-27.95	25	-21.11	49	-18.01	16	-19.85
30a	binder	35	-25.65	36	-19.27	28	-19.44	29	-18.22
30b	binder	41	-24.35	42	-18.2	24	-19.77	22	-19.29
31	binder	52	-22.14	11	-22.95	54	-17.32	45	-15.8
21b	binder	32	-26.36	38	-19.14	33	-19.07	1	-24.91
26 (R config.)	binder	27	-26.91	14	-22.84	30	-19.4	35	-17.58
32	binder	48	-22.89	32	-20.16	20	-20.34	4	-22.54
26 (S config.)	binder	19	-27.93	7	-23.6	13	-21.05	30	-18.03
33	non-binder	33	-26.36	46	-17.16	14	-20.99	15	-20.28
Wu 1	binder	36	-25.47	40	-18.98	35	-19.02	33	-17.83
Wu 2	binder	2	-32.47	52	-16.11	7	-22.56	40	-17.12
Wu 3	binder	14	-28.24	57	-14.82	53	-17.71	48	-15.41
Wu 4	binder	3	-32.07	45	-17.51	31	-19.20	19	-19.49
Wu 5	binder	18	-27.93	17	-22.71	27	-19.50	8	-21.3
Wu 6 (**1**)	binder	25	-27.05	15	-22.83	1	-24.49	20	-19.47
Wu 7	non-binder	55	-20.7	59	-14.34	59	-16.29	56	-14.19
Wu 8	binder	39	-24.62	35	-19.5	12	-21.06	28	-18.31
Wu SI1	non-binder	49	-22.78	60	-13.98	55	-17.12	50	-14.63
Wu SI2	binder	40	-24.58	44	-17.58	46	-18.14	51	-14.61
Wu SI3	non-binder	47	-23.14	55	-15.49	34	-19.06	46	-15.75
Wu SI4	binder	57	-19.98	54	-15.56	44	-18.22	53	-14.49
Wu SI5	binder	45	-23.68	58	-14.73	61	-15.98	59	-13.49
Wu SI6	binder	29	-26.71	24	-21.39	4	-23.68	11	-20.8

Wu SI7	non-binder	46	-23.54	22	-21.6	16	-20.82	24	-19.01
Wu SI8	non-binder	56	-20.68	37	-19.21	56	-17.08	36	-17.56
Wu SI9	binder	7	-29.23	21	-21.88	26	-19.68	27	-18.47
Wu SI10	non-binder	54	-20.71	47	-17.09	62	-15.04	54	-14.33
Wu SI11	non-binder	50	-22.56	49	-16.32	60	-16.09	55	-14.27
Wu SI12	binder	5	-30.25	41	-18.4	22	-19.95	39	-17.22
Wu SI13	binder	23	-27.41	48	-16.85	57	-16.86	61	-12.83
Wu SI14	non-binder	63	-16.06	63	-10.6	63	-13.56	57	-13.86
Wu SI15	non-binder	51	-22.41	56	-14.85	45	-18.18	60	-13.48
Wu SI16	binder	44	-23.82	33	-20.1	36	-18.99	47	-15.67
Wu SI17	binder	53	-21.46	34	-19.58	25	-19.73	43	-16.38
Wu SI18	binder	22	-27.5	43	-17.63	39	-18.52	13	-20.35

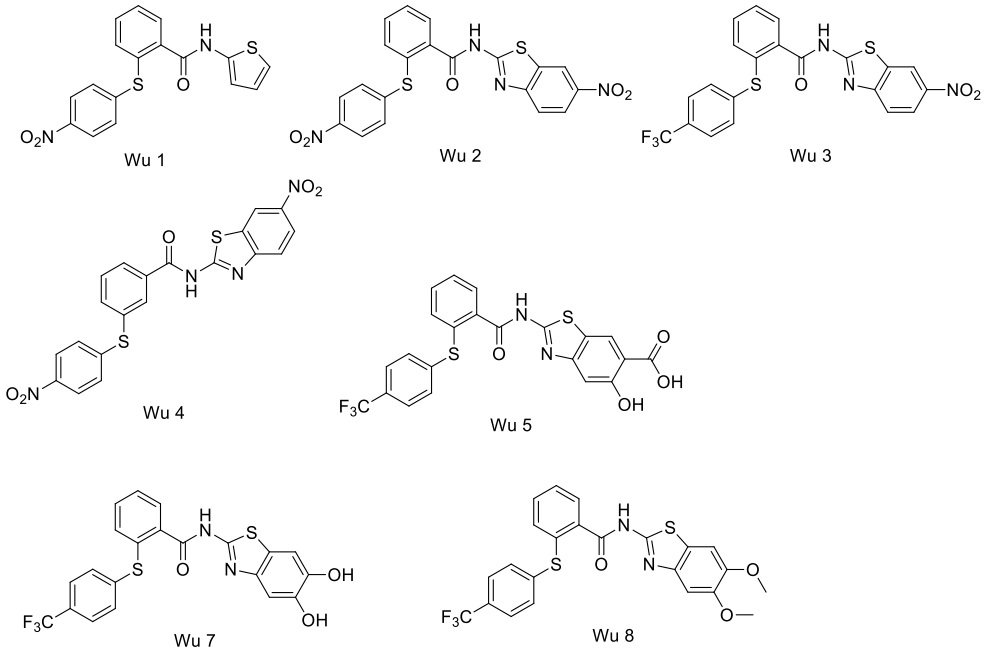


Figure SI8. Structures of Wu 1 – Wu 8.[4]

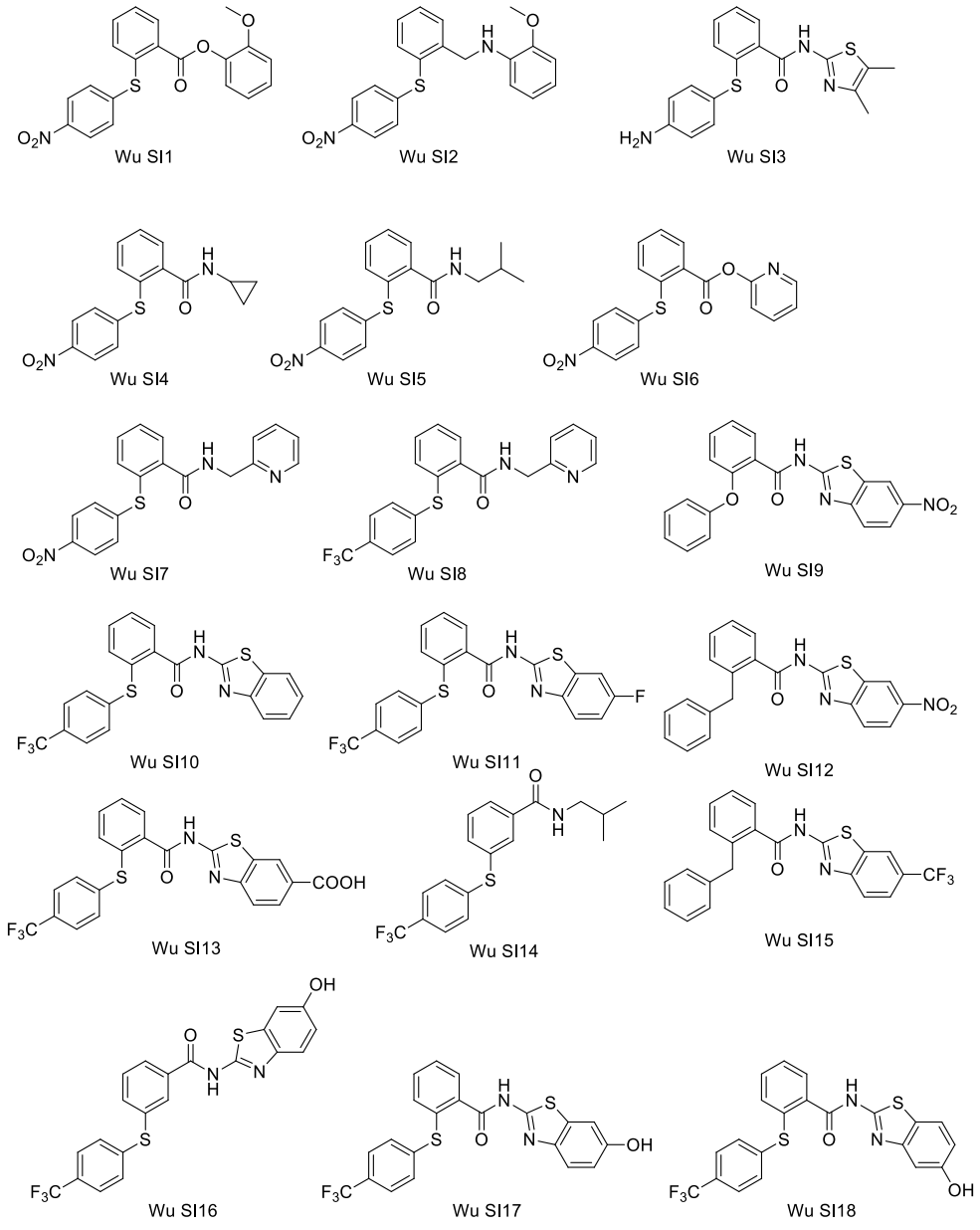


Figure SI 9. Structures of Wu SI1 – SI18.[4]

6 MD

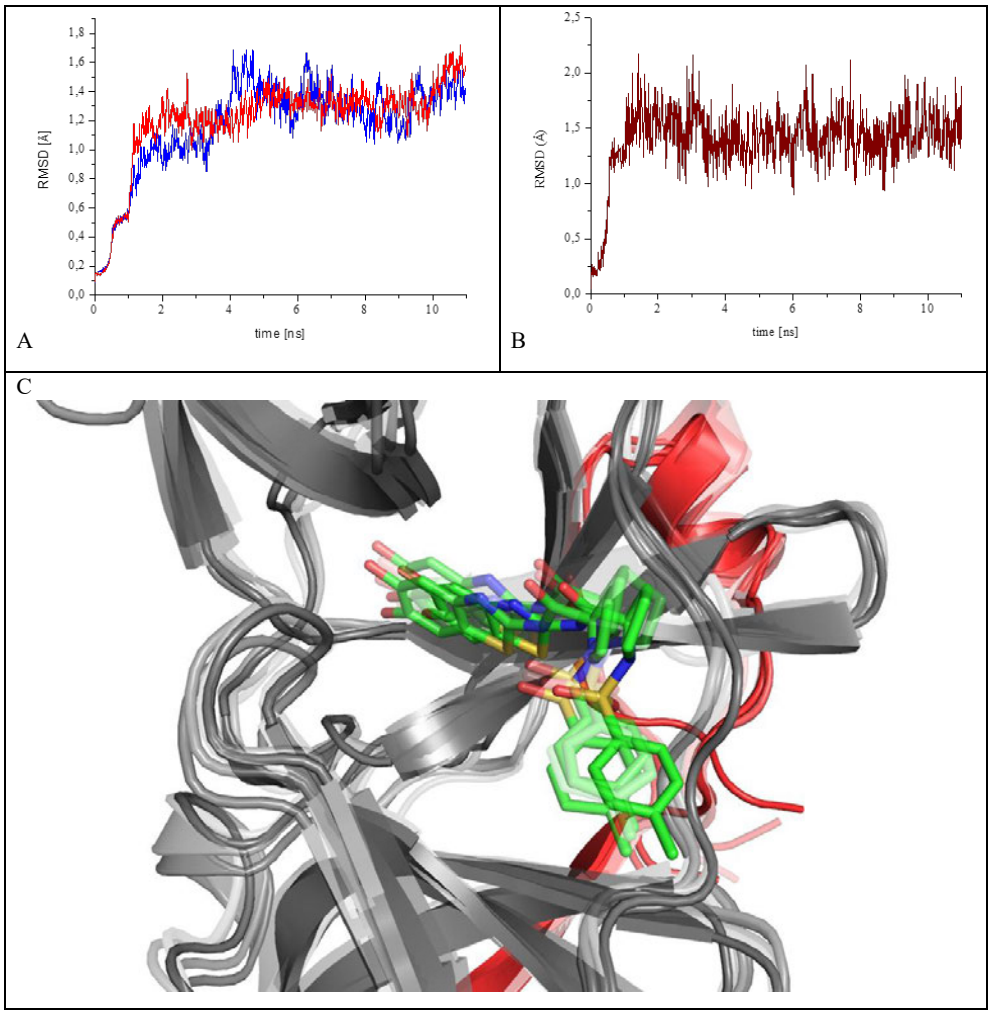


Figure SI10. A) Backbone RMSD-plot of DENV NS3/NS2B apo structure (blue line) and in complex with compound **7a** (red line) over 1 ns equilibration and 10 ns production run. B) **7a** non-hydrogen-atom RMSD-plot in complex with DENV NS2B/NS3 over 1 ns equilibration and 10 ns production run (average RMSD of 1.46 Å compared to starting structure generated by docking).

C) Superposition of trajectory snapshots after 2, 4, 6, 8 and 10 ns (with decreasing transparency) of the production run of DENGV NS2B/NS3-**7a** complex.

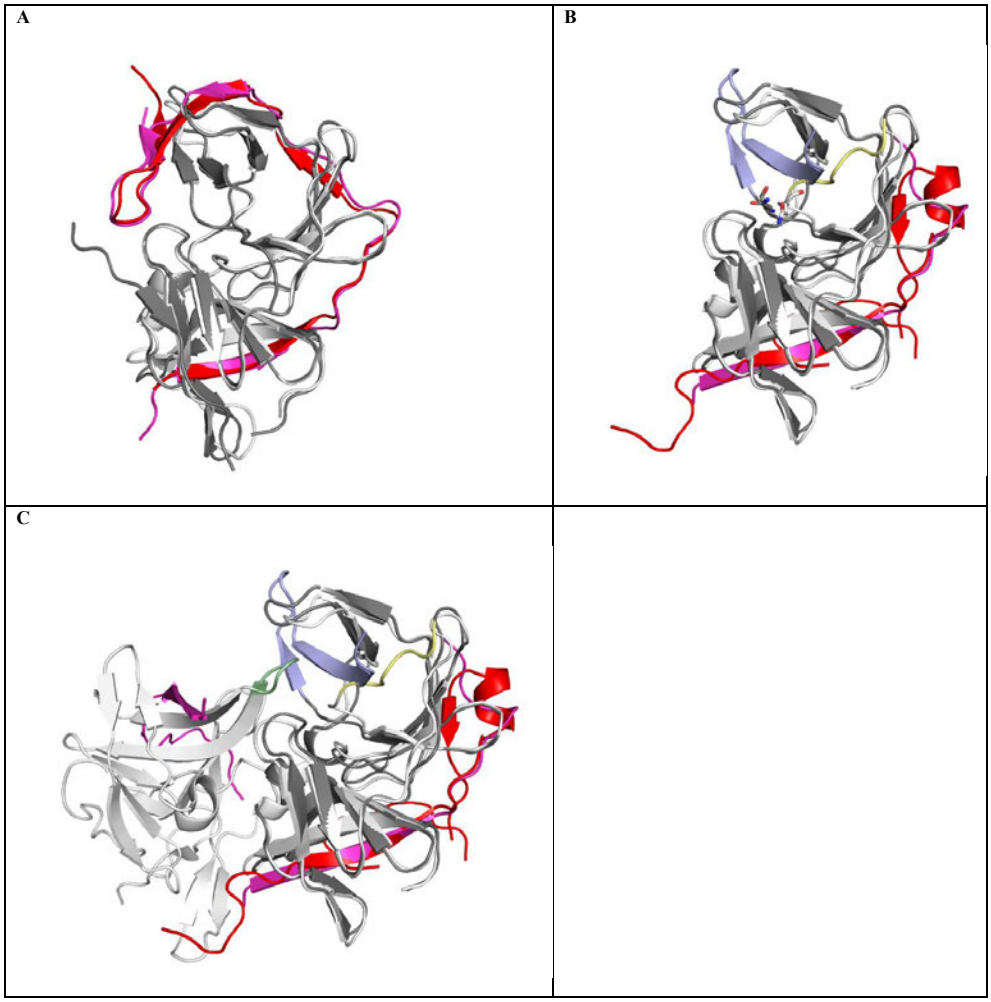
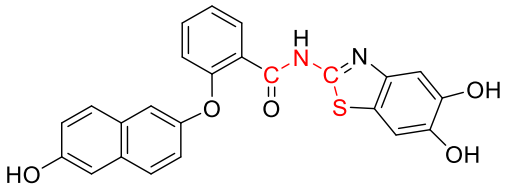
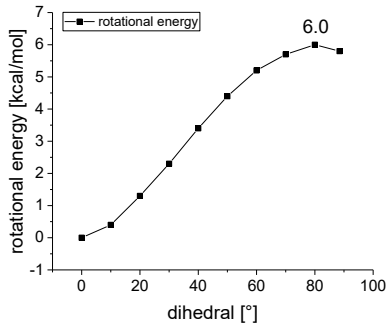


Figure SI11. A) Superposition of DENV (PDB code 3U1I) and ZIKV (PDB code 5LC0) NS2B/NS3 protease in the closed conformation (C_α RMSD = 0.61 Å). DENV NS3 in grey. NS2B in red; ZIKV NS3 in white and NS2B in magenta; the color scheme also accounts for B and C. B) Superposition of DENV and ZIKV protease in open conformation (PDB codes 2FOM and 5GXJ, respectively; C_α RMSD = 0.78 Å). Asn(1)152 is shown as sticks; subsequent residues differing in orientation are shown in light blue for the DENV protease and light yellow for the ZIKV protease. C) Additionally to B, the second monomer of the asymmetric unit from the ZIKV NS2B/NS3 is illustrated. Residues 1029-1032 preventing an open conformation analogue to DENV NS2B/NS3 by clashing are shown in light green.

A



B



C

```
@subgraph **aminothiazole planarity**
atom 1 S.3
atom 2 C.2
atom 3 N.AM
atom 4 C.2
atom 5 N.2
atom 6 H
bond 1 2 1
bond 2 3 1
bond 3 4 1
bond 2 5 2
bond 3 6 1
data
tangle/te 0 0
tangle/te 20 1
tangle/te 160 1
```

	tangle/te 180 0 period 360 symmetry 180 ***end***

Figure SI12. A) Structure of compound **1** with atoms for relaxed dihedral scan marked in red. B) Relaxed dihedral scan on B3LYP-D3/TZVP/PCM (solvent = water) level for red marked atoms. C) Torsion parameters to allow only dihedral angels 20° out of plane for marked atoms within FlexX docking.

7 ANTIVIRAL ACTIVITY

PLAQUE ASSAY ZIKV The titer of cell culture supernatants was quantified by plaque assays. Vero cells were seeded in a 6-well plate in a concentration of 3×10^5 cells/well and infected 6 h later with cell culture supernatant in a serial dilution in DMEM. The cells were washed once with PBS and covered with 37 °C pre-warmed DMEM containing 0.4% agarose 2 h p.i. After 15 min at room temperature, the plates were incubated for 96 h at 37 °C. To visualize the plaques, the agarose was removed gently, the cells were fixed for 20 min at room temperature with 4% formaldehyde in PBS and then stained for 15 min with 0.1% crystal violet in 20% ethanol. After the cells were washed once with water the titer (PFU/mL) was determined by counting the plaques in the well with the respective dilution.

IMMUNOFLUORESCENCE MICROSCOPY ZIKV The experiments were performed in tissue culture plates harbouring cover slides. The A549 cells were then fixed with 4% formaldehyde in PBS for 20 min and then permeabilized with 0.5% Triton X-100 in PBS for 15 min at room temperature. After blocking with 1% BSA in PBS for 15 min at room temperature, the cells were stained with the antibody Anti-Flavivirus Group Antigen Antibody, clone D1-4G2-4-15 (Merck millipore, Darmstadt, Germany) in the dilution 1:200 and afterwards with the secondary antibody anti-mouse-AlexaFluor488 (Thermo Fisher Scientific) and DAPI for 1 h at room temperature in a humid chamber. After both antibody incubations, the coverslides were washed three times with PBS at room temperature. Finally, the cells were mounted with Mowiol on microscope slides. Immunofluorescence staining was analyzed using a confocal laser scanning microscope (CLSM 510 Meta) with Zen 2009 Software (both from Carl Zeiss, Oberkochen, Germany). The same software was used to calculate the corrected total cell fluorescence (CTCF).

STATISTICAL ANALYSIS ZIKV Results are presented as means ± standard errors of the means (SEMs) from at least three independent experiments. The significance of the results was analyzed by unpaired two-tailed Student's *t*-test, using GraphPad Prism, version 6.07 for Windows (GraphPad Software, San Diego, CA, USA). In all figures the statistical significance is compared to the control group. Statistical significance is represented in figures as follows:

ns = not significant = $p > 0.05$; * = $p \leq 0.05$; ** = $p \leq 0.01$; *** = $p \leq 0.001$; **** = $p \leq 0.0001$.

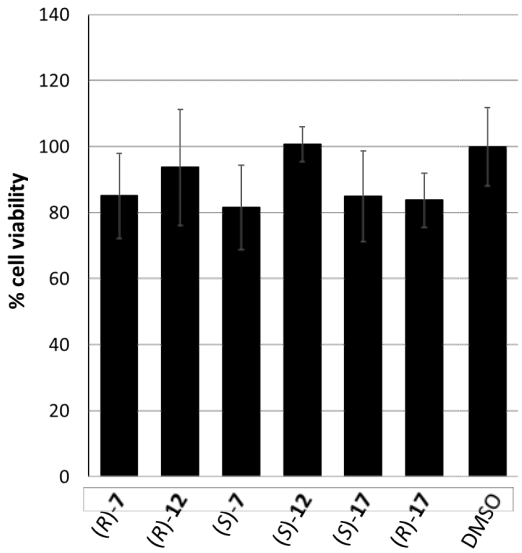


Figure SI 13. Cell toxicity tests of (*R*)-**7**, (*S*)-**7**, (*R*)-**12**, (*S*)-**12**, (*R*)-**17** and (*S*)-**17** on Vero cells. Cell cytotoxicity of the compounds at 10 μM after 4 days of incubation was determined by a cell proliferation assay with DMSO as a control.

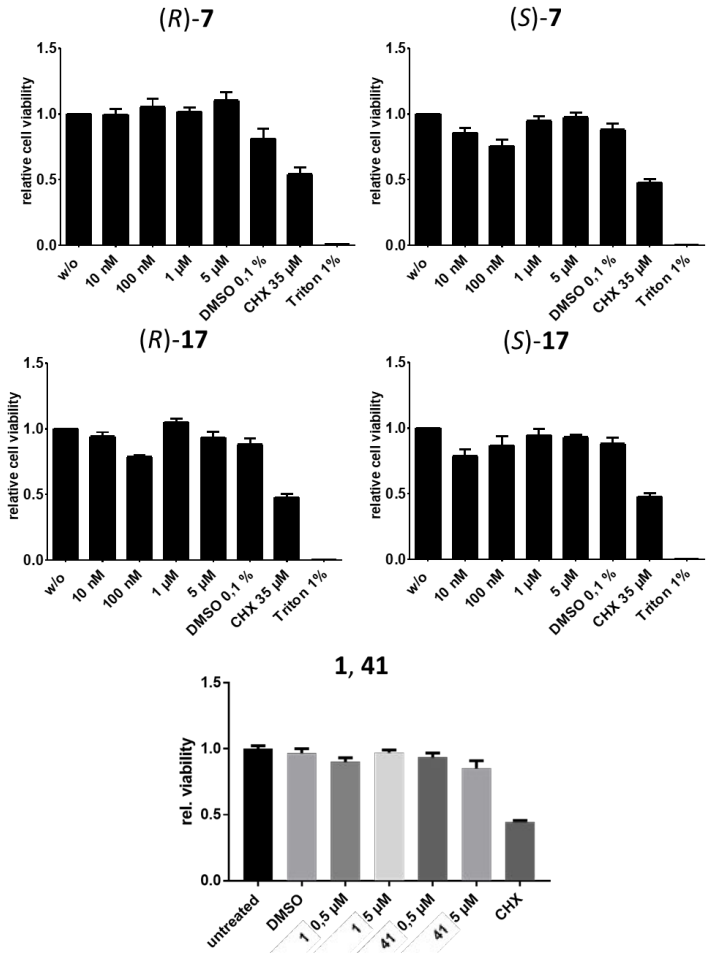


Figure SI 14. Cell toxicity tests of (*R*)-**7**, (*S*)-**7**, (*R*)-**17**, (*S*)-**17**, **1** and **41** on A549 cells. Cell cytotoxicity of the compounds at the indicated concentrations was determined by the PrestoBlue Cell viability reagent after 24 h. Cycloheximide (CHX) was used as a positive control at a concentration of 70 μM.

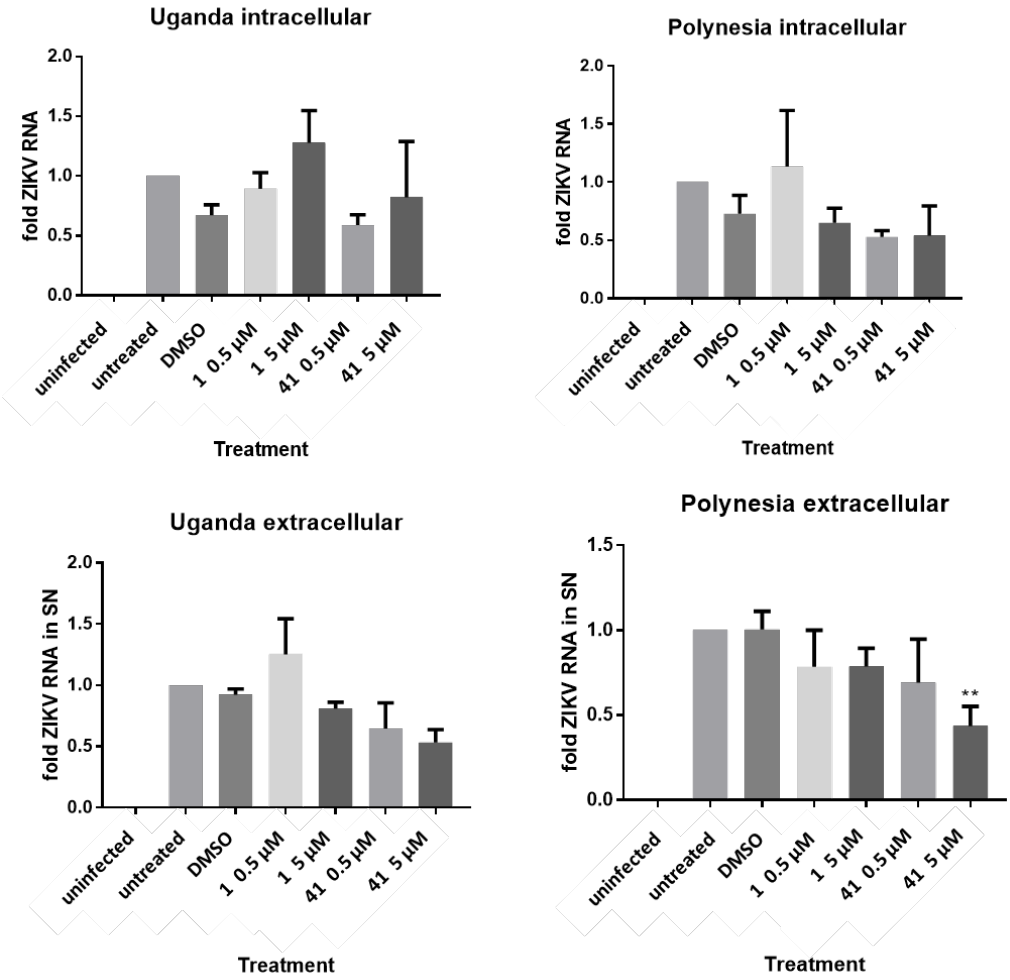


Figure SI15. Effect of the treatment of ZIKV Uganda or French Polynesia Strain infected cells with **1** and **41**. A549 cells were infected with ZIKV virus and the compounds were added at different concentrations. Viral RNAs were isolated, quantified by RTqPCR and normalized to the untreated control.

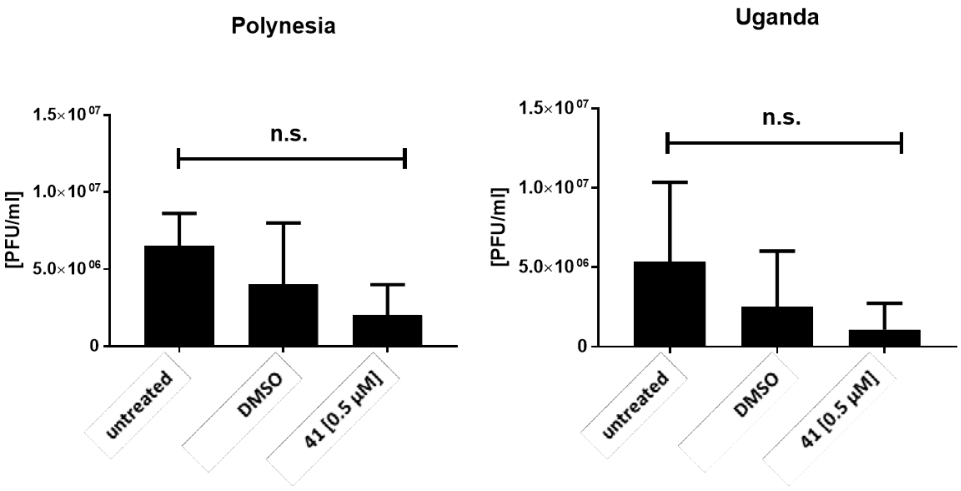


Figure SI16. Effect of **41** on the inhibition of the French Polynesia or Uganda ZIKV strain. Vero cells were seeded, infected with ZIKV cell culture supernatant and covered with agarose. After incubating for 96 h, the cells were fixed with formaldehyde, stained with crystal violet and the titer as plaque forming units (PFU) was determined by counting.

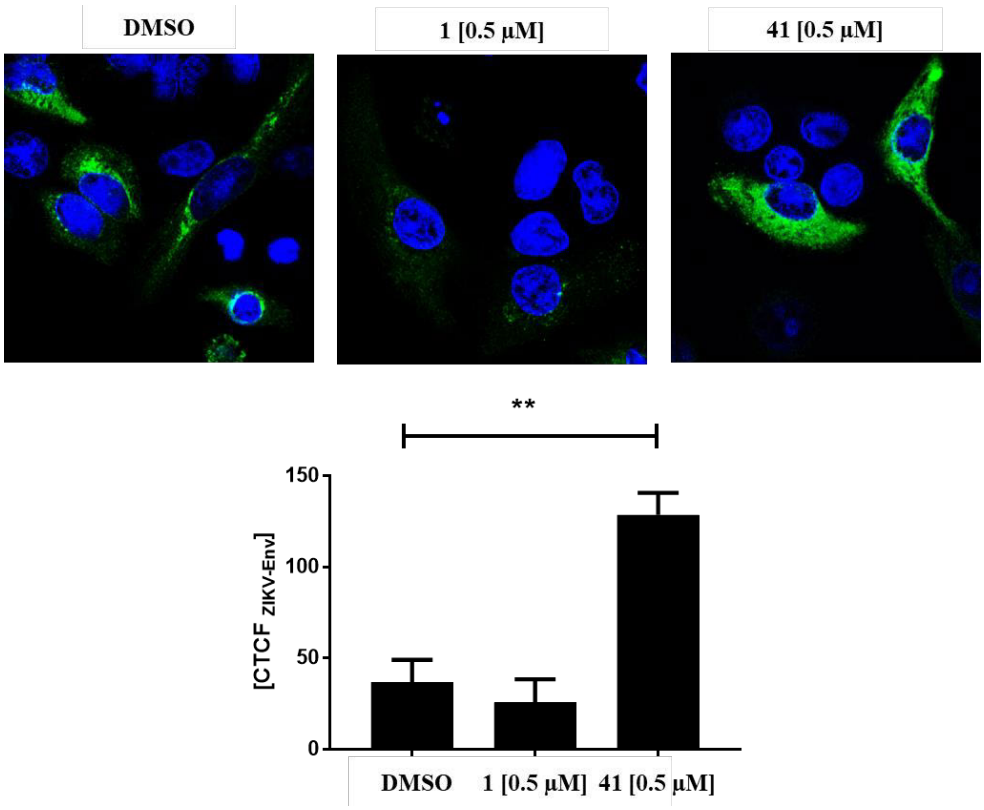


Figure SI17. **41** leads to a redistribution of ZIKV Env protein (green) from the ER to the cytoplasm. A549 Cells were infected with ZIKV and treated with **1** and **41**. After fixing and permeabilizing, the cells were stained with an anti-flavivirus antigen antibody and a secondary antibody, inspection by confocal scanning microscopy was done and the corrected total cell fluorescence (CTCF) was determined. DNA was stained with 4´,6-diamidin-2-phenylindol (blue).

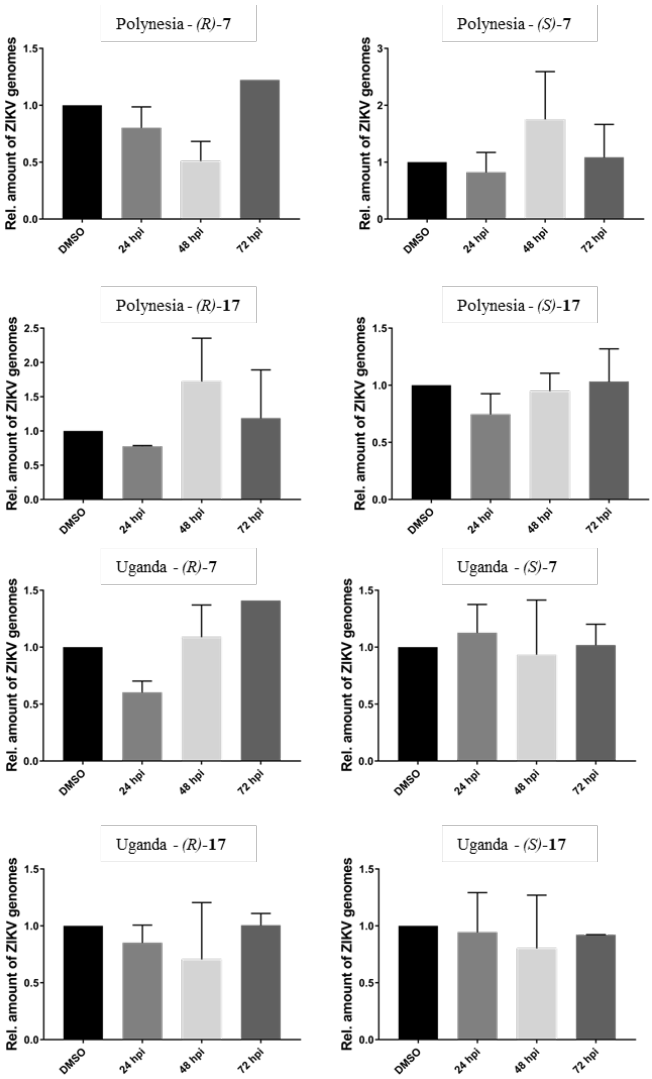


Figure SI18. Effect of the treatment of ZIKV Uganda or French Polynesia Strain infected cells with (*R*)-**7**, (*S*)-**7**, (*R*)-**17**, and (*S*)-**17** at a concentration of 5 μM. A549 cells were infected with ZIKV virus and the compounds were added at 5 μM. Viral RNAs were isolated, quantified by RTqPCR and normalized to the untreated control.

8 CELL PERMEABILITY

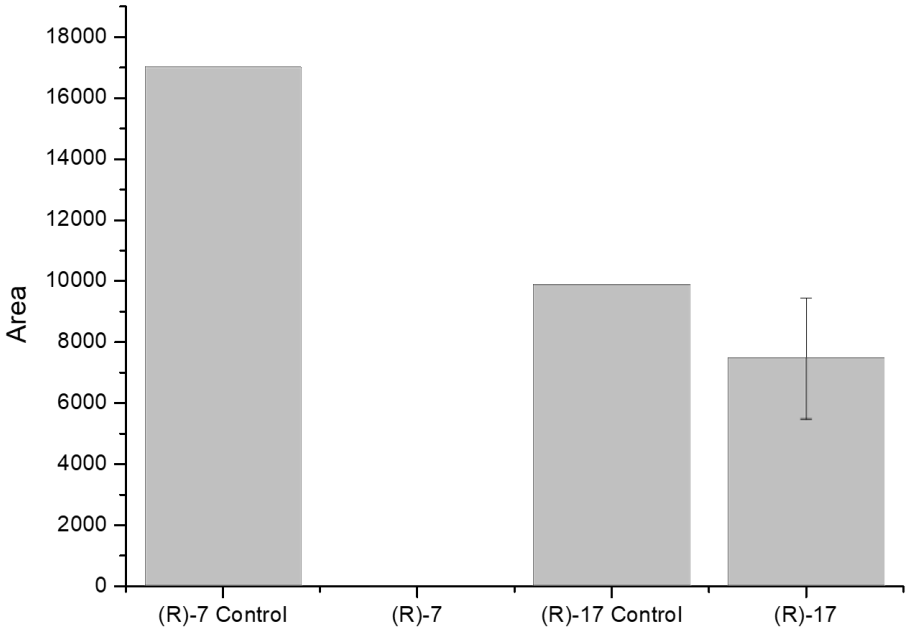


Figure SI19. MS areas from cell-permeability studies. A layer of Caco-2 cells was incubated with 10 μM (*R*)-**7** and (*R*)-**17** for 3 hours and the donor and acceptor solutions were analyzed by LC-MS. Shown are the areas of the MS signals of the acceptor solution.

“

4.3. The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR

4.3.1. Context, Project Summary, and Own Contribution

In a follow-up project to Millies *et al.*[239], [redacted] from the [redacted] group further optimized the allosteric inhibitors of NS2B/NS3. Since phenolic groups, particularly catechol structures as present in cpds **15** and **20**, are prone to oxidative instabilities,[253] more stable cpds were yielded by inversion of the central amide, leading to cpd **21**.[196] Docking experiments with cpd **15** and its *S*-configurated analog with similar inhibitory activities, revealed two poses of the allosteric inhibitors in which the catechol stays in the same position but the molecule is rotated by 180° (Figure 29A).[239] Based on these binding modes, we developed the idea to simultaneously address both possible interactions by designing Y-shaped, branched inhibitors. Another objective was reducing the inhibitors' molecular weight by truncating the central proline to increase the ligand efficacy, thereby providing a better basis for further inhibitor development. Both approaches produced the potent allosteric inhibitors **16** and **22** (Figure 29B).[196] Cpd **22** has an improved LE of 0.40 kcal·(mol·HA)$^{-1}$ compared to cpd **15** with an LE of 0.28 kcal·(mol·HA)$^{-1}$.

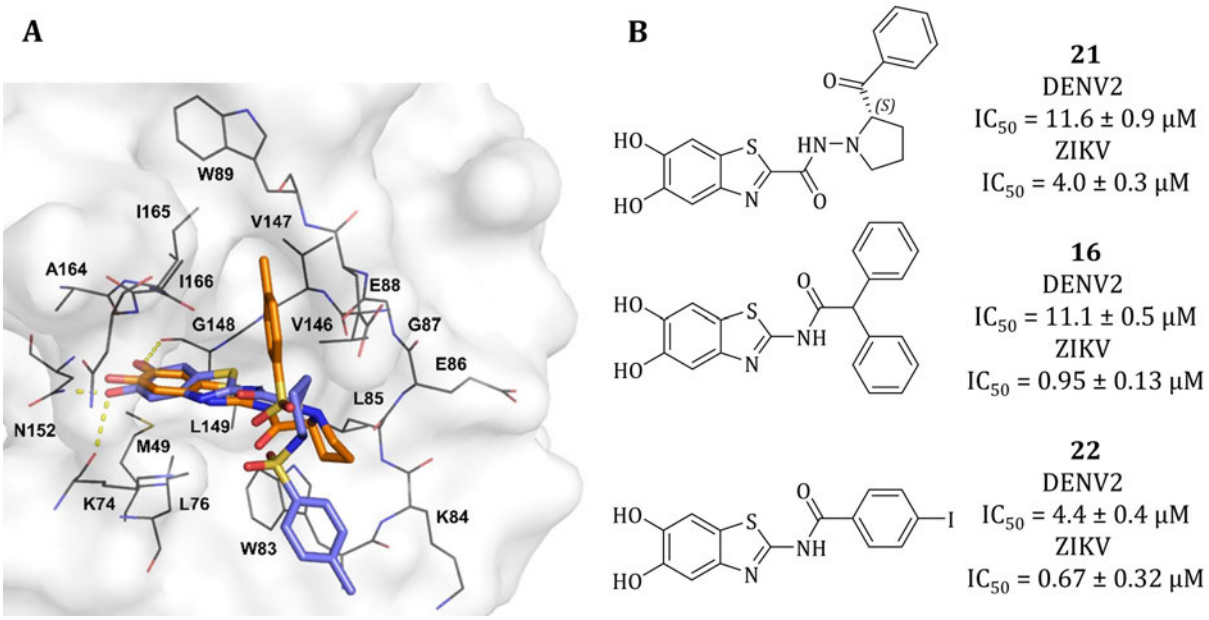


Figure 29. (**A**) *Predicted binding modes of cpd **15** (blue sticks) and its enantiomer (orange sticks) in the DENV2 NS2B/NS3 allosteric binding site* (PDB-ID: 2FOM).[34,239] The catechol resides deeply in the allosteric binding pocket, forming conserved interactions (yellow dashed lines) with Lys74, Gly148, and Asn152 for both the (*R*)- and (*S*)- enantiomers. The aromatic *para*-tosyl moieties interact with distinct hydrophobic subsites. Reprinted with permission from J Bioorg. Med. Chem. 47 (2021) 116392. Copyright © 2021 Elsevier Ltd. All rights reserved.[196] (**B**) 2D structures and IC_{50} values of the new lead cpds, yielded by Maus *et al.*[196]

In parallel, [redacted] of the group of [redacted] worked with [redacted] of the [redacted] group on single-molecule (sm) Förster resonance energy transfer (FRET) experiments using DENV2 NS2B/NS3. Again, it was exploited that the DENV2 NS2B/NS3 does not contain any Cys in its sequence (compare *4.2 Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases*). Two Cys residues were introduced at positions whose distances change significantly between both conformations (closed: 18 Å; open: 50 Å) and were subsequently labeled with a FRET pair of ATTO dyes by maleimide coupling (Figure 30A, B).

The resulting energy transfer efficiencies (E_{ET}) can be used as a "molecular ruler" to determine the distances of the dyes and conclude the conformation of NS2B/NS3 adopted in solution.[255,256] By observing two point clouds in the FRET efficiencies, it could be visualized that both conformations of DENV2 NS2B/NS3 exist in equilibrium. With the addition of the competitive cpd **17**,[240] this distribution was shifted to higher E_{ET}, indicating stabilization of closed conformation (Figure 30C). In a succeeding project, [redacted] performed smFRET experiments on DENV2 NS2B/NS3 to elucidate the effects of allosteric inhibitors on the conformational equilibrium.[257] To provide direct evidence for the open conformation to be stabilized upon binding of allosteric inhibitors, cpd **16** was employed in this study. Adding the allosteric cpd **16** did not influence the conformational equilibrium (Figure 31A). This finding supports the hypothesis that the main population of labeled apo NS2B/NS3 with low energy transfer efficiencies represents the open conformation (Figure 31, left).

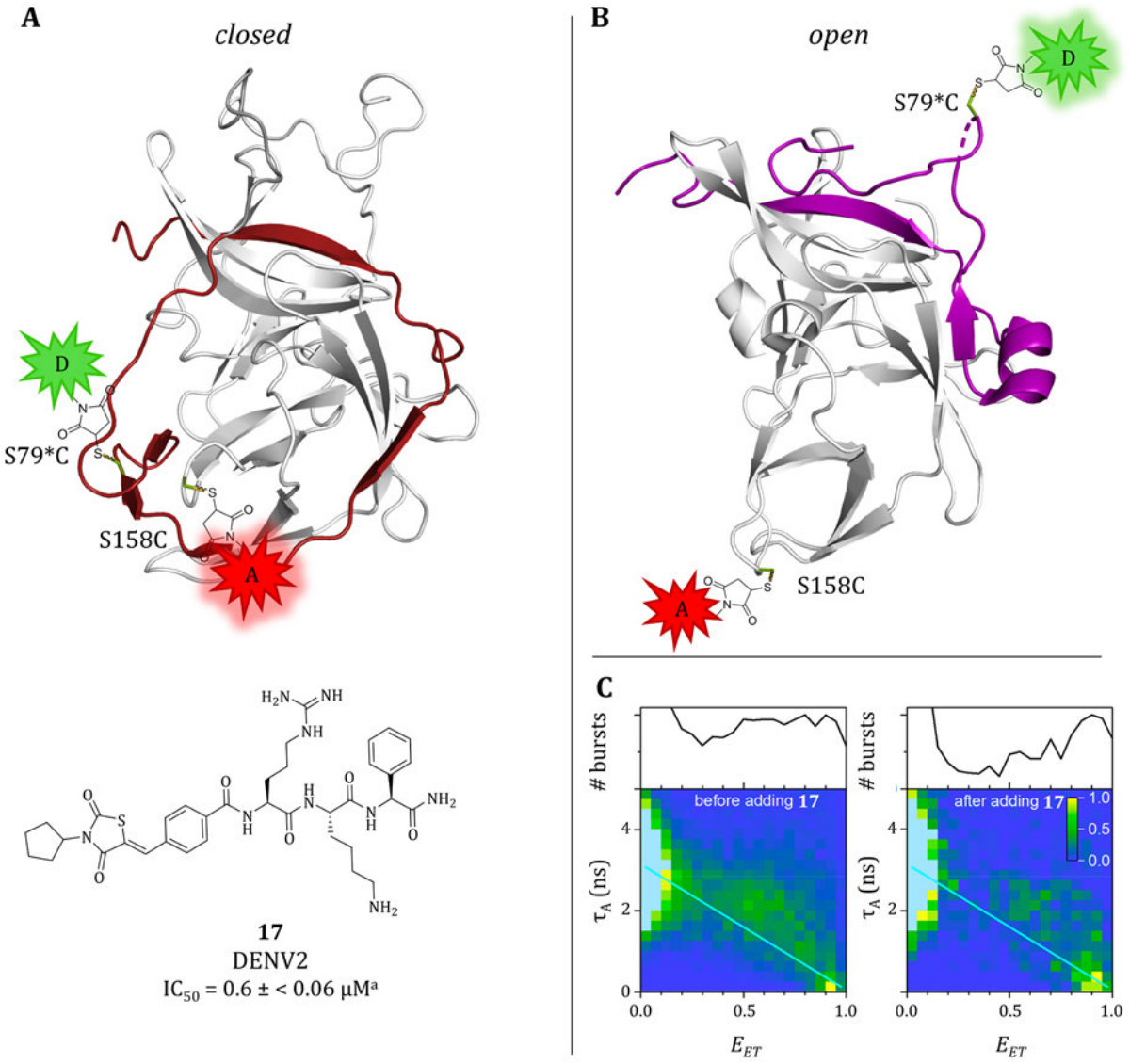


Figure 30. *smFRET using a double Ser to Cys mutant of DENV2 NS2B/NS3.* (**A**) In the closed conformation of NS2B (red)/NS3 (white), which is induced by the competitive inhibitor **17**, the FRET pair consisting of a donor (green) and an acceptor dye (red) get close together so that after excitation of the donor, the fluorescence of the acceptor can be recorded (PDB-ID: 2M9P).[254] (**B**) In the open conformation (NS2B: violet, NS3: white, PDB-ID: 2FOM)[34], the distance of the dyes increases, hence the energy transfer efficiency (E_{ET}) is reduced so that only the acceptor reveals high fluorescence. (**C**) Top: A plot of the normalized occurrences of the individual bursts, representing individual molecules within the intensity time trace of the FRET pair labeled DENV2 NS2B/NS3 in a 2D histogram (bottom), separated according to acceptor lifetime τ_A and E_{ET}. A shift towards higher E_{ET} upon adding cpd **7** can be observed. Reprinted (modified) with permission from The Journal of Physical Chemistry B 2021 125 (25), 6837-6846.[51] Copyright © 2021 American Chemical Society.

These results can be explained by previous reports suggesting that both the presence of a glycine-rich linker between DENV2 NS2B and NS3 and high pH values, as used in our experiments, favor the open conformation.[53,199,258] To overcome this issue, a competition setup was employed, were the conformational equilibrium was shifted towards the closed conformation by adding a fixed concentration of the competitive cpd **17** (Figure 31B). Then, a dilution series of the allosteric inhibitor **16** was added to shift the equilibrium back to the open conformation at intermediate energy transfer efficiencies (Figure 31C). This experiment not only provided the first qualitative

proof that allosteric inhibitors shift DENV2 NS2B/NS3 to the open conformation but also allowed us to determine the K_d of cpd **6** based on the proportion of the open conformation.[257]

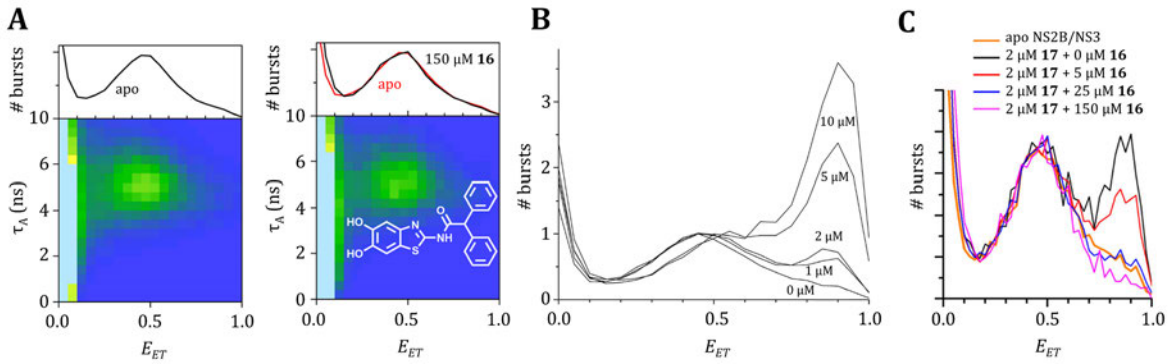


Figure 31. *A competitive smFRET assay to determine the effects of allosteric inhibitors.* (**A**) Top: E_{ET} frequency distributions, a plot of the normalized occurrences of the individual bursts, representing individual molecules within the intensity time trace of the FRET pair labeled DENV2 NS2B/NS3 in a 2D histogram (bottom), separated according to acceptor lifetime τ_A and FRET efficiency E_{ET}. No direct effect upon adding cpd **16** can be observed. Reprinted (modified) with permission from Protein Science. 2023;32:e4526.[257] Copyright © 2022 The Authors. Protein Science published by Wiley Periodicals LLC on behalf of The Protein Society. (**B**) E_{ET} frequency distributions of FRET pair labeled DENV2 NS2B/NS3 with different concentrations of the cpd **17**. Reprinted (modified) with permission from Protein Science. 2023;32:e4526.[257] Copyright © 2022 The Authors. Protein Science published by Wiley Periodicals LLC on behalf of The Protein Society. (**C**) E_{ET} frequency distributions of FRET pair labeled DENV2 NS2B/NS3 with fixed concentrations of the competitive cpd **17** and varying concentrations of the allosteric cpd **16**. Reprinted (modified) with permission from Protein Science. 2023;32:e4526.[257] Copyright © 2022 The Authors. Protein Science published by Wiley Periodicals LLC on behalf of The Protein Society.

However, since I only contributed moderately to the three publications described above (References: Maus *et al.*: 'SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases'[196], Götz *et al.*: 'Conformational Dynamics of the Dengue Virus Protease Revealed by Fluorescence Correlation and Single-Molecule FRET Studies'[51] and Maus *et al.*: 'A competition smFRET assay to study ligand-induced conformational changes of the dengue virus protease'[257]) they are listed in *2.2 Research Articles Beyond this Doctoral Thesis* and are not included in this thesis.

The following publication, which is part of this thesis, was dedicated to transferring the knowledge gained by smFRET studies to results of other biophysical techniques that allow the observation of conformational differences between the open and the closed state of NS2B/NS3. Herein, we used ZIKV NS2B/NS3 as a model system, and the macrocyclic competitive cpd **4**, and the allosteric cpd **16** as tool compounds (Figure 32A, B, C).[57,196]

To demonstrate that smFRET is not only possible for proteins lacking native Cys residues, ZIKV NS2B/NS3 was mutated by site-directed mutagenesis to only harbor two Cys in the desired positions. Therefore, a quintuple-mutant ZIKV NS2B/NS3 (^{5}ZiPro, #198432, www.addgene.com) was generated, where the three native Cys residues Cys80, Cys143, and Cys178 were mutated to Ser residues. Contrarily, Ser84* and Ser160 were mutated to Cys residues since they are located near the respective positions of Cys79* and Cys158 in DENV2 NS2B/NS3 (Figure 32A, B).[257] For ^{19}F-NMR experiments of ZIKV NS2B/NS3, fluor-labeling was performed by expressing the protease in a defined media containing 7F-Tryptophane to be incorporated into the protein. Six single Trp to Phe mutants were generated (ZIKV NS2B/NS3 W61*F, W5F, W50F, W69F, and W89F — #198435–198440, www.addgene.com) to assign the six observed signals in ^{19}F-NMR to the respective Trp residue. After labeling, the vanished peak of the NMR spectra of the individual mutants could be assigned to the respective position of the Trp to Phe mutation (Figure 32D).

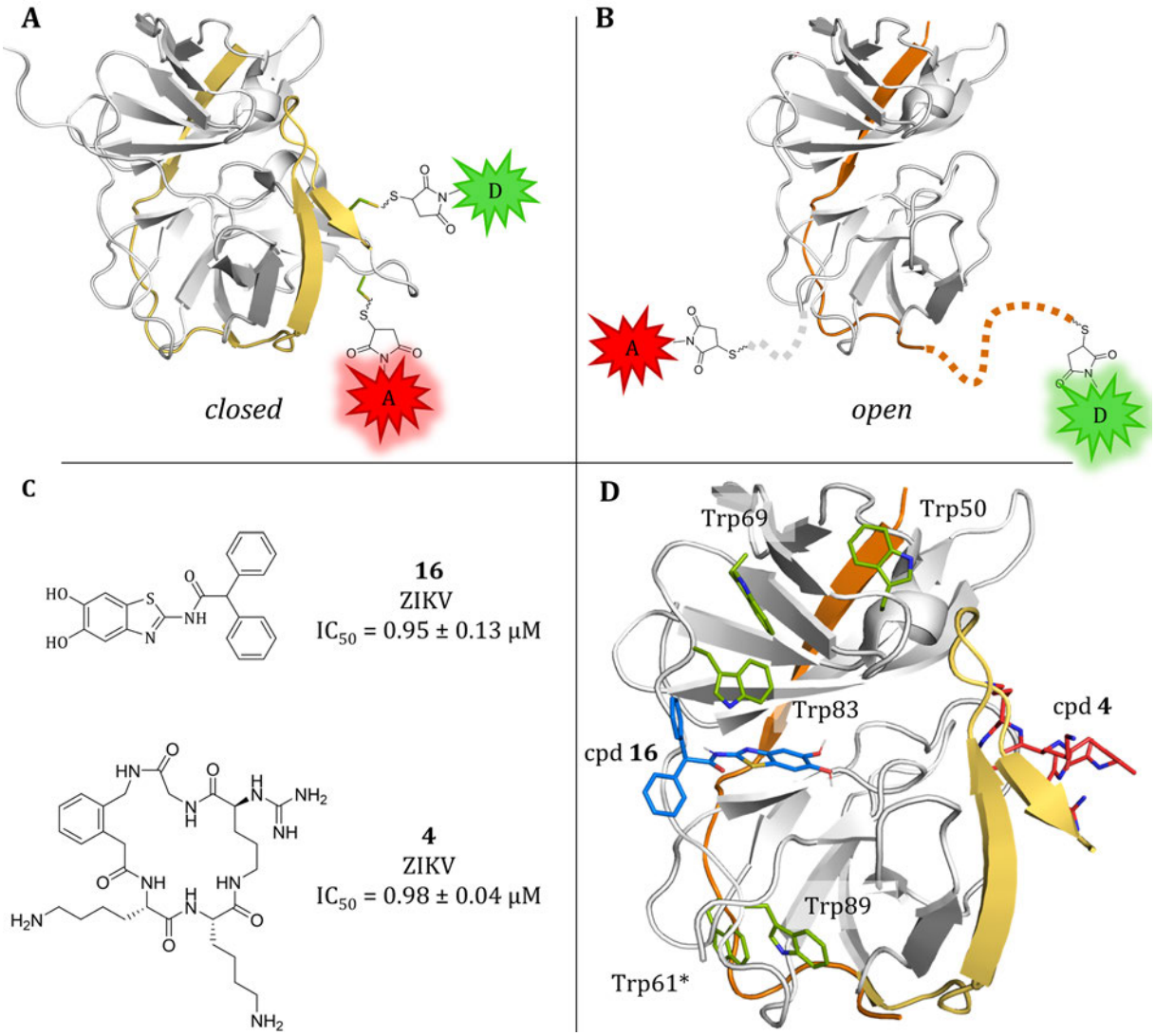


Figure 32. *Modifications of ZIKV NS2B/NS3 for smFRET and ^{19}F-NMR.* (**A**) In the closed conformation of NS2B (yellow)/NS3 (white), which is induced by the competitive cpd **4**, the FRET pair consisting of a donor (green) and an acceptor dye (red) get close together so that with excitation of the donor, the fluorescence of the acceptor can be recorded (PDB-ID open: 5GXJ).[184] (**B**) In the open conformation (PDB-ID: 6Y3B)[57], the distance of the dyes increases. Hence the E_{ET} is reduced so that only the acceptor exhibits high fluorescence. (**C**) 2D-Structures of the allosteric cpd **6** and the competitive cpd **4**. (**D**) Structure of the ZIKV protease (PDB-ID open: 5GXJ[184] and closed: 6Y3B[57]) with the position of the native Trp residues highlighted in green. The docking pose of cpd **16** (blue) and cpd **4** (red) demonstrate different Trp residues in proximity. For a clear view, only $NS3_{pro}$ of the open conformation is displayed. Trp5 in the N-terminus of $NS3_{pro}$ is not resolved in these structures and thus is not shown.

Using ^{5}ZiPro labeled with a FRET pair, the smFRET experiments of the previous publication using DENV2 could be reproduced,[257] providing further evidence that the observed effects of different conformations being stabilized by the same competitive and allosteric inhibitors is a general aspect of flaviviral NS2B/NS3 proteases. This knowledge was then utilized to explain other observations of the behavior of NS2B/NS3 proteases like the diverging melting points in differential scanning fluorimetry (DSF) and nanoDSF or the peaks in ^{19}F-NMR spectra of ZIKV NS2B/NS3, shifting in other patterns upon addition of the two inhibitor classes. These findings can also be used to rationalize some of the results of other biophysical methods described above (*4.1 Elucidating the mode of action of allosteric Dengue- and Zika NS2B/NS3 Protease Inhibitors*).

(5) [redacted] **Hammerschmidt, S. J.**[#]; [redacted] The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR. *Manuscript submitted to Eur. J. Med. Chem.*

Own contribution: Enzyme expression and purification, site-directed mutagenesis, ^{19}F-Labeling of the proteases, ^{19}F-NMR experiments, writing parts of the original draft & editing of the manuscript.

Contributions from other authors: smFRET labeling and experiments, DSF, nanoDSF, CD-spectroscopy, fluorometric enzyme activity assays, ^{19}F-NMR experiments, writing of the original draft & editing of the manuscript.

This work has been submitted to *European Journal of Medicinal Chemistry* (impact factor 7.09).

4.3.2. Publication

The following manuscript, from page 249 to page 275, contains unpublished data.

The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics, explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR.

[redacted] Stefan J. Hammerschmidt [a#], [redacted]

[#] authors contributed equally

* corresponding author

[redacted]

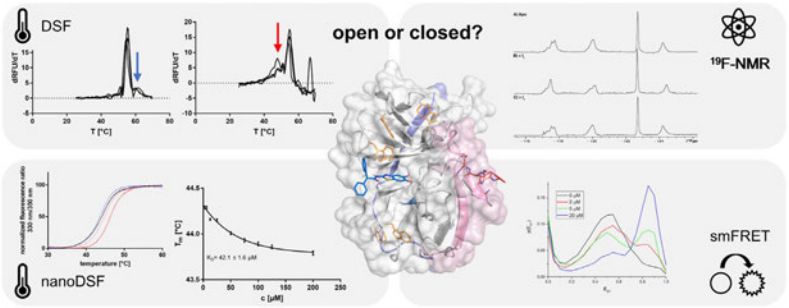
[a] Institute of Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg-University, Mainz, Germany

[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.

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.



HIGHLIGHTS

- Allosteric and competitive ligands favor different conformations of the ZIKV protease.
- smFRET allows the assignment of stabilized conformations (allosteric – open; competitive – closed).
- Effects in more accessible assays such as DSF, nanoDSF and ^{19}F-NMR can be explained based on the conformational changes of the protease.

INTRODUCTION

Zika and dengue viruses cause two of the most wide-spread 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 Zika 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 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 the 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. As for the DENV protease, we found that a competitive inhibitor[41] (**I$_c$**, Figure 1) stabilized the closed conformation, whereas an allosteric inhibitor[42] (**I$_a$**, Figure 1) stabilized the open conformation.

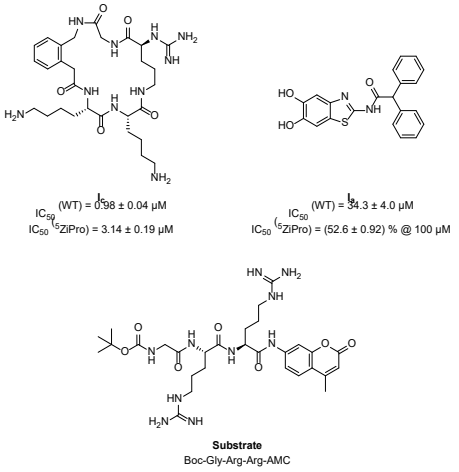


Figure 1. Structures of the competitive inhibitor **I$_c$**,[41] and the allosteric inhibitor **I$_a$**[42], as well as the fluorogenic-substrate for enzyme activity assays used in this study.

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.

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, S84C, 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 (Figure 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 (Figure 1) was used to detect the fluorescence increase upon proteolytic release of AMC by an active protease. In the presence of a reducing agent (5 mM DTT), the k_{cat} values for the WT and the ^{5}ZiPro were similar (k_{cat} = 0.11 s^{-1} and k_{cat} = 0.13 s^{-1}, Table 1). Likewise, the K_M values determined for ^{5}ZiPro (K_M = 417 ± 16 µM) and the WT enzyme (K_M = 361 ± 27 µM) under the same buffer conditions are very similar (Table 1). To ascertain the structural integrity of ^{5}ZiPro, circular dichroism spectroscopy was used (Figure 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 (Figure 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 activity and 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 (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 (Figure 4). Relatively low labeling efficiencies were found, with 14% for ATTO 488 and 20% for ATTO 643. Nonetheless, these proved sufficient for smFRET experiments (see below).

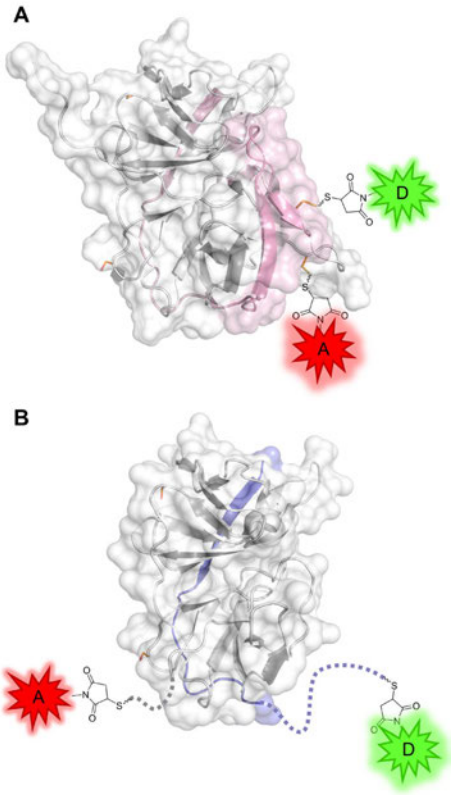


Figure 2. Closed (**a**) and open (**b**) conformation of the ZIKV protease quintuple mutant and relative smFRET label positions. $NS3^{pro}$ is shown in white, and NS2B of the closed conformation is colored light red or light blue for the open conformation. Mutated positions are shown as orange sticks. 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. A dashed line indicates missing amino acids.

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_{DTT}) 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}]	$k_{cat} \cdot K_M^{-1}$ [$M^{-1} \cdot s^{-1}$]
WT_{lit}[43]	558 ± 62	$2.6 \cdot 10^{-2} \pm 1.4 \cdot 10^{-3}$	47.2 ± 5.8
WT_{DTT}	361 ± 27	$0.13 \pm 5.2 \cdot 10^{-3}$	356 ± 30
WT_{TCEP}	641 ± 20	$2.4 \cdot 10^{-2} \pm 1.1 \cdot 10^{-3}$	37.5 ± 2.1
$^{5}ZiPro_{DTT}$	417 ± 16	$0.11 \pm 2.5 \cdot 10^{-3}$	272 ± 12
$^{5}ZiPro_{TCEP}$	794 ± 115	$1.4 \cdot 10^{-4} \pm 1.4 \cdot 10^{-5}$	0.18 ± 0.03
$^{5}ZiPro^{ATTO\ 488/ATTO\ 643}{}_{DTT}$	397 ± 16	$6.3 \cdot 10^{-2} \pm 1.4 \cdot 10^{-3}$	158 ± 7
^{19}F-Trp $eZiPro_{TCEP}$	885 ± 56	$5.3 \cdot 10^{-2} \pm 3.2 \cdot 10^{-3}$	59.6 ± 5.2
$eZiPro_{TCEP}$	1202 ± 111	$2.0 \cdot 10^{-2} \pm 1.6 \cdot 10^{-3}$	16.6 ± 2.1

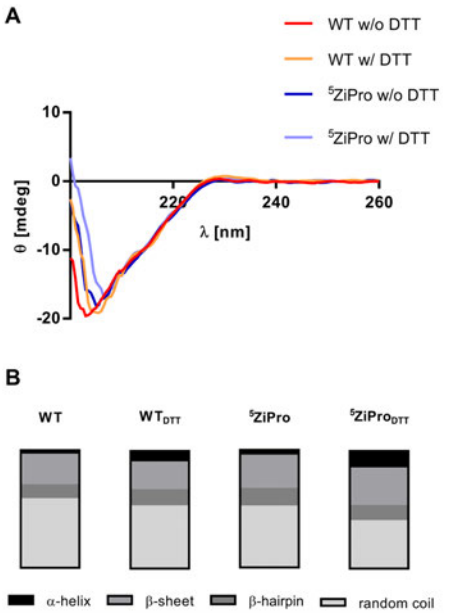


Figure 3. CD spectra of the ZIKV NS2B-NS3 wildtype protease (WT) and its quintuple mutant (^{5}ZiPro) with and without the addition of a reducing agent (5 mM DTT) at a protein concentration of 5 µM. (**a**) recorded CD spectra. (**b**) CD spectroscopy based secondary structure prediction of the WT and ^{5}ZiPro with and without reducing agent (5 mM DTT) determined with Bestsel.[44]

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 Figure 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 (Figure 6). After adding the competitive ligand, the cross-correlation function G_{AD} (gray in Figure 6) approached zero faster than G_{AD} without the inhibitor (black in Figure 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.[30] Additionally, counts in the acceptor channel mainly originate from the 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 (Figure 7 a). The maximum of the corresponding frequency distribution of E_{ET} was around $E_{ET} \approx 0.5$. Adding 5 µM of $\mathbf{I_c}$ led to a second point cloud at higher E_{ET} (Figure 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 (Figure 8).

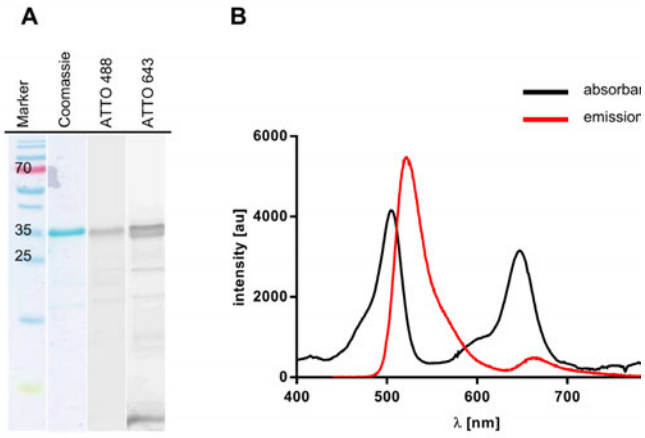


Figure 4. Successful fluorophore labeling of ^{5}ZiPro. (**a**) SDS-PAGE gel of ATTO 488/ATTO 643 FRET pair labeled ZIKV protease. Marker PageRuler™ Prestained Protein Ladder (in kDa), Coomassie stain, ATTO 488 Laser scan with excitation at 500 nm, ATTO 643 Laser scan with excitation at 630 nm. (**b**) Absorption (black) and emission (red) spectra. The latter was measured at 500 nm excitation

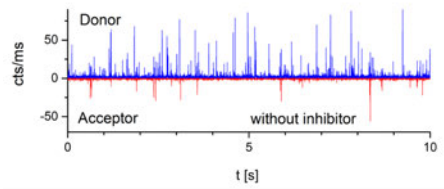


Figure 5. Section of the fluorescence intensity time trace of ATTO 488/ATTO 643 FRET pair labeled ZIKV protease. The sample was excited with excitation pulses at 502 nm, and the binning time was 1 ms. The donor channel is shown in blue; the acceptor channel is shown in red.

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} (Figure 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,[30] 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[30] 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 (Figure 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 $\mathbf{I_a}$.

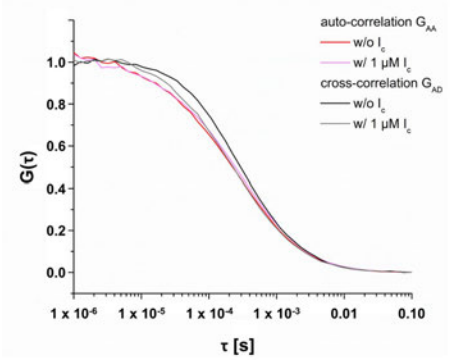


Figure 6. Auto- (G_{AA}) and cross-correlations (G_{AD}) of the ATTO 488 and ATTO 643 labeled ^{5}ZiPro before and after adding 1 µM of $\mathbf{I_c}$.

Thermal Shift Assay. The thermal shift assay, or DSF, is commonly used for detecting protein-ligand interactions.[45,46] 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)[47] or based on intrinsic tryptophan and tyrosine fluorescence (nanoDSF).[48] Ligands that interact with the protein usually influence the unfolding transition and shift the midpoint of the unfolding transition (i.e., the melting temperature T_m).[47] The simplicity of the method as well as the low protein consumption, make it an interesting method for the study of protein-ligand interactions.[45,47]

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 $\mathbf{I_a}$ (Figures 10 and 11).

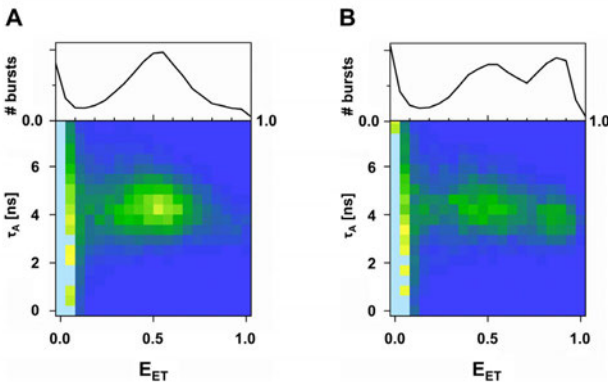


Figure 7. A plot of the normalized occurrences of individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 643 FRET pair labeled ^{5}ZiPro plotted in a 2D histogram, depending on acceptor lifetime τ_A and FRET efficiency E_{ET}. (**a**) Without competitive inhibitor and (**b**) with competitive inhibitor ($\mathbf{I_c}$ = 5 µM). The respective maxima were normalized to one 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.

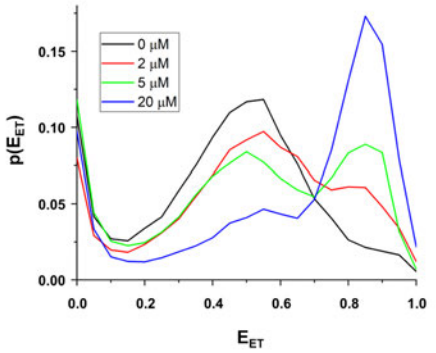


Figure 8. E_{ET} frequency distributions of ATTO 488/ATTO 643 FRET pair labeled ZIKV protease with different concentrations of the competitive inhibitor.

The DSF approach with SYPRO Orange yielded a T_m of 54.29 ± 0.50 °C for the ZIKV protease in the apo state (Figure 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 (Figure 10, blue arrow), the presence of $\mathbf{I_c}$ gave rise to a peak-shoulder at lower temperatures (Figure 10, red arrow). Higher concentrations of the competitive inhibitor (100 µM, Figure 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 nm and plotted against the temperature.[49] 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 (Figure 11 a). The presence of $\mathbf{I_a}$ led to a lowering of the melting temperature to 43.35 ± 0.03 °C (Figure 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 (Figure 11 a and b, light red line), which is further increased to 46.40 ± 0.01 °C by higher concentrations (100 µM, Figure 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).

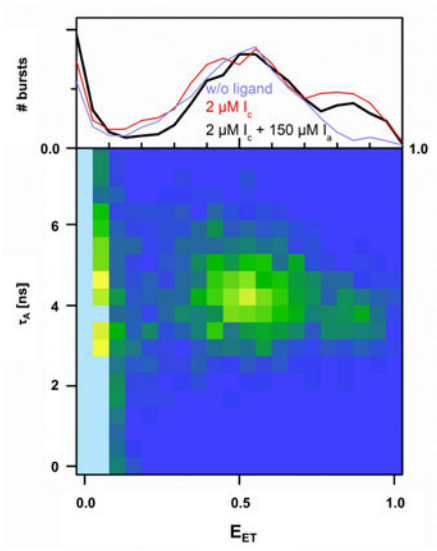


Figure 9. A plot of the normalized occurrences of the individual bursts (individual molecules) within the intensity time trace of the ATTO 488/ATTO 643 FRET pair labeled ^{5}ZiPro in a 2D histogram, separated according to acceptor lifetime τ_A and FRET efficiency E_{ET}. Blue: without ligand, red: with 2 µM $\mathbf{I_c}$, black: with 2 µM $\mathbf{I_c}$ and 150 µM $\mathbf{I_a}$. 2D histogram: with $\mathbf{I_c}$ and $\mathbf{I_a}$.

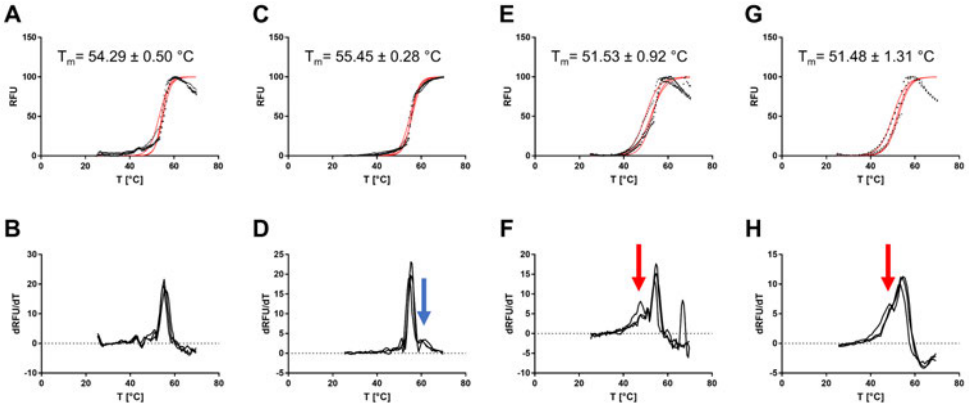


Figure 10. Melting curves of ZIKV WT determined by DSF. Three independent measurements (black) and their corresponding sigmoidal fit curves (red) are shown. (**a, b**) measured fluorescence in the temperature range from 20 °C to 70 °C for the ZIKV protease without ligand and the corresponding 1st derivative (**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 $\mathbf{I_c}$.

^{19}F-NMR. Fluorine has become a valuable reporter to probe interactions between biomacromolecules and to investigate protein structural dynamics by NMR spectroscopy.[50–52] A fluorine reporter can be incorporated post-translationally through labeling cysteine residues[53,54] or the co-translational insertion of fluorinated amino acids such as fluoro-tryptophan.[54,55] The ZIKV protease contains six tryptophan residues, five within the core NS3 protease domain and one within the NS2B^{cf} (Figure 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., introducing an amber stop codon recognized by a mutant aminoacyl-tRNA synthetase that site-specifically incorporates 7F-Trp for NMR studies.[56] Here, we aimed to simplify this approach by using uniform fluoro-tryptophan labeling, thus circumventing mutagenesis, dedicated expression strains, or using specific tRNA synthetases. Instead, uniform incorporation of the labeled amino acid can be achieved by including it in the bacterial growth medium.[54] 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[57], we incorporated 5F-Trp into ZIKV NS2B-NS3 thus complementing the available data with 7F-Trp (Figure 11 a). Activity and binding affinity for the substrate (Figure 1) of the labeled protease were similar to that of the wild type (WT) under the same buffer conditions (Table 1).

In the ^{19}F-NMR spectrum of the labeled protease, six peaks of differing linewidth were observed (Figure 12 a), displaying an overall broader dispersion than previously seen for the 7F-Trp labeled samples.[56] Using single tryptophan point mutations, a subset of residues could be tentatively assigned (Figure S3). As expected for a highly flexible residue, W5 at the N-terminus of NS3 gives rise to the narrowest peak (Figure 12). This agrees with the fact that this residue could not be resolved in available crystal structures of NS2B-NS3 (Figure 12 b). Of note, ZIKV NS2B-NS3 W83F and W89F could be expressed and purified but were 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 elimination.

While the addition of the inhibitors caused only subtle effects on the spectra, adding $\mathbf{I_c}$ 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 (Figure 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.[56] 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.

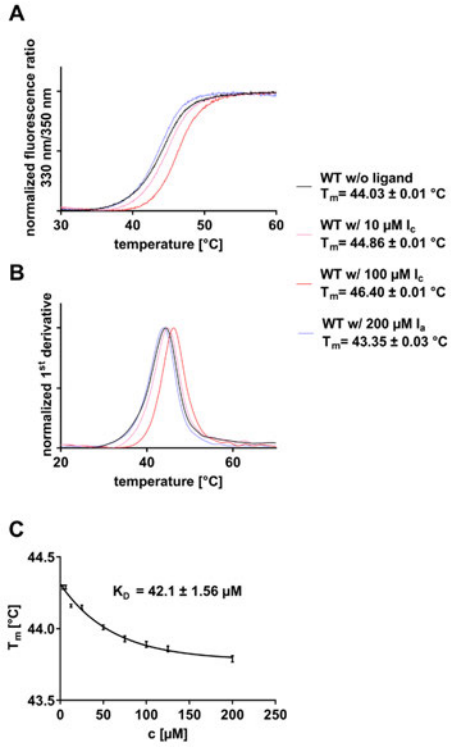


Figure 11. Thermal shift assay performed by nanoDSF. Shown is the mean of three independent measurements. (**a**) Melting curves of the WT determined by nanoDSF with and without competitive and allosteric inhibitors. Black: without ligand, pink: 10 µM of $\mathbf{I_c}$, red 100 µM of $\mathbf{I_c}$, 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 ± 1.6 µM.

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 could confirm its functional and structural integrity. Importantly, this shows that the ZIKV protease is relatively tolerant towards mutagenesis, a vital 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 for painting 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 relatively low, presumably due to the formation of disulfide bridges and/or reduced solvent accessibility of the respective labeling sites, smFRET experiments could be carried out with a sufficiently increased protein concentration. The faster decrease in the cross-correlation function G_{AD} after adding $\mathbf{I_c}$ 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.[58] Without a ligand, a single distinct point cloud at intermediate E_{ET} was seen in the 2D histogram (Figure 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,[28,30], 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 structurally closely related proteins, this nonetheless must be shown unambiguously for each protein. Despite its high information content, sample preparation for smFRET and the subsequent data analysis is challenging. Furthermore, not all laboratories have easy access to instrumental equipment. Thus, we applied other, more accessible methodologies to observe whether the allosteric and competitive inhibitors have conformational consequences. The results from the smFRET study were used to guide the respective interpretation.

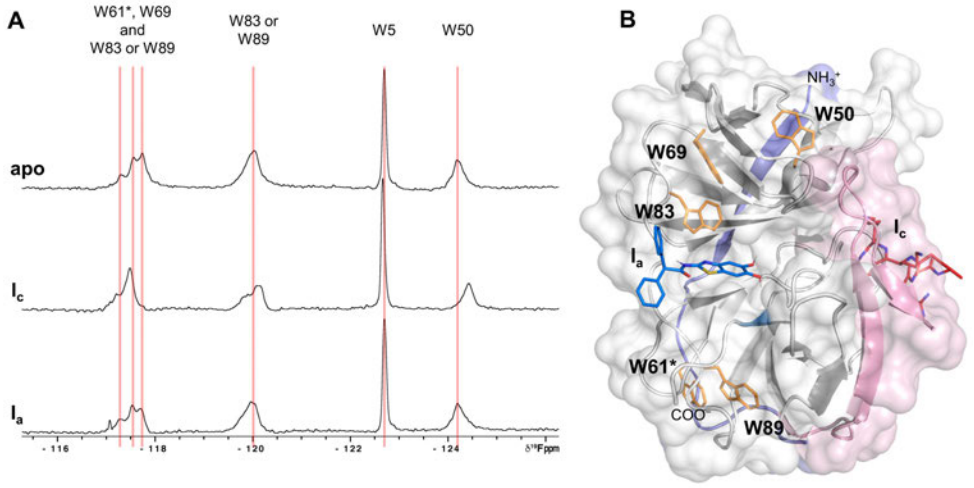


Figure 12. Sensing of inhibitor binding by 5F-Trp and their respective positions. (**a**) ^{19}F-NMR spectra of the 5F-tryptophan labeled ZIKV protease with and without ligands. From up to down: spectrum of the apo protease; the spectrum of the protease with 100 µM of the competitive inhibitor; the spectrum of the protease with 100 µM of the allosteric inhibitor. 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 the allosteric inhibitor is shown in blue. The competitive inhibitor is shown in red. NS2B is shown in light red (closed conformation) and light blue (open conformation). Only NS3 of the open conformation (PDB: 5GXJ) is shown in white for a clear view. Residue W5 in the NS3 N-terminus was not resolved in these structures and is thus not shown.

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. Notably, SYPRO Orange used for DSF may associate with hydrophobic regions as present in the allosteric binding site of NS2B-NS3 (Figure S1), thereby possibly affecting the protein's conformational equilibrium.[13,28,45,59] Other disadvantages of SYPRO Orange are its susceptibility to non-protein-induced fluorescence, such as EDTA aggregates, plastic surfaces, lipids, glycerol, and detergents.[45,60,61] In these cases, nanoDSF, 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 (Figure S2).[41] All other tryptophan residues are not or only marginally affected by the open/closed transition (Figure 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 the 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 experiments.

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 potent inhibitors.

MATERIAL and METHODS

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}: amino acids 1423–1467 and NS3^{pro}: amino acids 1503–1688)[40] as used by us before[62] 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 (amino acids 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 65–75 °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.[63] Here, the following primers were used: (vi) 5′–GAAAAAACCGGCAAACGCAGCGGCGCGCTGTGGGAC GTGCCG–3′ and 5′–GCGTTTGCCGGTTTTTTCACGCATCGGCGGACCATCG TCTTCAACC–3′. For the signal assignment of ^{19}F-NMR experiments, the single W/F mutants of gZiPro as well as eZiPro were generated using the primers: (vii) W61#F 5′–GACATTACCTTTGAAAAGGATGCGGAAGTTACC–3′ and 5′–CTTTTCAAAGGTAATGTCGCCCGCACGC–3′, (viii) 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′.

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.[42,64] 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*-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% 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 5ml column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany) on an ÄKTA start FPLC device (Cytiva). After a washing step with 5 column volumes 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 ^{5}ZiPro (with additional 1 mM DTT). Fractions containing NS2B/NS3 were concentrated, flash-frozen in liquid nitrogen, and stored at –80 °C. SDS-PAGE confirmed the purity and identity of NS2B/NS3 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 amino acids but containing only 10% of non-fluorinated tryptophane and 90% of the labeled amino acid (4-fluoro-tryptophane, 5-fluoro-tryptophane, 6-fluoro-tryptophane).[65] 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).

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.

CD Spectroscopy

Circular dichroism measurements were performed on a JASCO-815 spectrometer from JASCO (Tokyo, Japan) with an enzyme concentration of 5–10 µM in buffer (10 mM HEPES pH 7.4). 2.5 mM DTT was used as a reducing agent. The scan speed was 50 nm/min, and the bandwidth and scan intervals were set to 5 nm and 1 nm, respectively. For each sample, 12 measurements (at room temperature) were recorded, 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).[44]

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.[28,30] 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.

smFRET Experiments

The smFRET experiments were performed as described previously[28,30] 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 50ct, 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.

Analysis of smFRET Data

The data were analyzed as described before.[28,30] 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.

DSF Experiments

The DSF experiments were performed in quadruplicate on the openDSF instrument.[66] 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, and 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.

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.[67] The melting temperature was calculated as the inflection point of the resulting sigmoidal curve.

^{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 3 600 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.

Acknowledgments

We thank [redacted] for providing us with the competitive inhibitor.

We thank [redacted] group for providing the nanoDSF. In particular, we thank [redacted] for the instruction on device usage.

We thank [redacted] for the generation of the eZiPro mutant.

[redacted] 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 [redacted] and [redacted] (Johannes Gutenberg-University, Mainz) for their experimental support.

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 ^{5}ZiPro; 4. smFRET Experiments; 5. Analysis of smFRET Data; 6. Fluorometric Enzyme Assay; 7. K_M Curves

(1) Harapan, H.; Michie, A.; Mudatsir, M.; et al. Epidemiology of Dengue Hemorrhagic Fever in Indonesia: Analysis of Five Decades Data from the National Disease Surveillance. *BMC Res. Notes* **2019**, *12* (1), 4–9. https://doi.org/10.1186/s13104-019-4379-9.

(2) Sharma, V.; Sharma, M.; Dhull, D.; et al. Zika Virus: An Emerging Challenge to Public Health Worldwide. *Can. J. Microbiol.* **2020**, *66* (2), 87–98. https://doi.org/10.1139/cjm-2019-0331.

(3) WHO. Prevention and Control of Dengue and Dengue and Haemorrhagic Fever. *New Delhi World Heal. Organ. South East Asia Reg. Off.* **2011**.

(4) Behnam, M. A. M.; Nitsche, C.; Vechi, S. 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* (9), 1037–1042. https://doi.org/10.1021/ml500245v.

(5) Nitsche, C. Proteases from Dengue, West Nile and Zika Viruses as Drug Targets. *Biophys. Rev.* **2019**, *11* (2), 157–165. https://doi.org/10.1007/s12551-019-00508-3.

(6) Yildiz, M.; Ghosh, S.; Bell, J. A.; et al. Allosteric Inhibition of the NS2B-NS3 Protease from Dengue Virus. *ACS Chem. Biol.* **2013**, *8* (12), 2744–2752. https://doi.org/10.1021/cb400612h.

(7) Rothan, H. A.; Han, H. C.; Ramasamy, T. S.; et al. Inhibition of Dengue NS2B-NS3 Protease and Viral Replication in Vero Cells by Recombinant Retrocyclin-1. *BMC Infect. Dis.* **2012**, *12* (314), 1471–2334. https://doi.org/10.1186/1471-2334-12-314.

(8) Amberg, S. M.; Nestorowicz, A.; McCourt, D. W.; et al. NS2B-3 Proteinase-Mediated Processing in the Yellow Fever Virus Structural Region: In Vitro and in Vivo Studies. *J. Virol.* **1994**, *68* (6), 3794–3802.

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

(10) 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.

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

(12) Shannon, A. E.; Chappell, K. J.; Stoermer, M. J.; et al. Simultaneous Uncoupled Expression and Purification of the Dengue Virus NS3 Protease and NS2B Cofactor Domain. *Protein Expr. Purif.* **2016**, *119*, 124–129. https://doi.org/10.1016/j.pep.2015.11.022.

(13) 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.

(14) Noble, C. G.; Seh, C. C.; Chao, A. T.; et al. 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.

(15) Chen, W. N.; Loscha, K. V.; Nitsche, C.; et al. The Dengue Virus NS2B-NS3 Protease Retains the Closed Conformation in the Complex with BPTI. *FEBS Lett.* **2014**, *588* (14), 2206–2211. https://doi.org/10.1016/j.febslet.2014.05.018.

(16) Erbel, P.; Schiering, N.; Arcy, A. D.; 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.

(17) Yang, H., Chen, X., Ji, X., Xiong, Y., Yang, K. 5GXJ Zika Virus NS2B-NS3 Protease. 2017. https://doi.org/10.2210/pdb5GXJ/pdb.

(18) Lei, J.; Hansen, G.; Nitsche, C.; et al. Crystal Structure of Zika Virus NS2B-NS3 Protease in CompleLEI, Jian e Colab. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. Science, v. 353, n. 6298, p. 503–505, Jul 2016.x with a Boronate Inhibitor. *Science* **2016**, *353* (6298), 503–505. https://doi.org/10.1126/science.aag2419.

(19) 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.

(20) Lei, J.; Hansen, G.; Nitsche, C.; et al. 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.

(21) Roy, A.; Lim, L.; Srivastava, S.; et al. Solution Conformations of Zika NS2B-NS3pro and Its Inhibition by Natural Products from Edible Plants. *PLoS One* **2017**, *12* (7), 1–22. https://doi.org/10.1371/journal.pone.0180632.

(22) Evenäs, J.; Thulin, E.; Malmendal, A.; et al. 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), 3448–3457. https://doi.org/10.1021/bi9628275.

(23) Wand, A. J. Dynamic Activation of Protein Function: A View Emerging from NMR Spectroscopy. *Nat. Struct. Biol.* **2001**, *8* (11), 926–931. https://doi.org/10.1038/nsb1101-926.

(24) Kauk, M.; Hoffmann, C. Intramolecular and Intermolecular FRET Sensors for GPCRs – Monitoring Conformational Changes and Beyond. *Trends Pharmacol. Sci.* **2018**, *39* (2), 123–135. https://doi.org/10.1016/j.tips.2017.10.011.

(25) Zhu, L.; Yang, J.; Li, H.; et al. Conformational Change Study of Dengue Virus NS2B-NS3 Protease Using 19F NMR Spectroscopy. *Biochem. Biophys. Res. Commun.* **2015**, *461* (4), 677–680. https://doi.org/10.1016/j.bbrc.2015.04.090.

(26) Ha, T.; Ting, A. Y.; Liang, J.; et al. Single-Molecule Fluorescence Spectroscopy of Enzyme Conformational Dynamics and Cleavage Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (3), 893–898. https://doi.org/10.1073/pnas.96.3.893.

(27) Santoso, Y.; Joyce, C. M.; Potapova, O.; et al. Conformational Transitions in DNA Polymerase I Revealed by Single-Molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (2), 715–720. https://doi.org/10.1073/pnas.0910909107.

(28) 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.

(29) Ruer, M.; Krainer, G.; Gröger, P.; et al. ATPase and Protease Domain Movements in the Bacterial AAA+ Protease FtsH Are Driven by Thermal Fluctuations. *J. Mol. Biol.* **2018**, *430* (22), 4592–4602. https://doi.org/10.1016/j.jmb.2018.07.023.

(30) 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.* **2023**, *32* (1), e4526.

https://doi.org/10.1002/pro.4526.

(31) Li, Q.; Kang, C. Insights into Structures and Dynamics of Flavivirus Proteases from NMR Studies. *Int. J. Mol. Sci.* **2020**, *21* (7). https://doi.org/10.3390/ijms21072527.

(32) Li, Y.; Zhang, Z.; Phoo, W. W.; et al. Structural Dynamics of Zika Virus NS2B-NS3 Protease Binding to Dipeptide Inhibitors. *Structure* **2017**, *25* (8), 1242-1250.e3. https://doi.org/10.1016/j.str.2017.06.006.

(33) Li, Y.; Zhang, Z.; Phoo, W. W.; et al. Structural Insights into the Inhibition of Zika Virus NS2B-NS3 Protease by a Small-Molecule Inhibitor. *Structure* **2018**, *26* (4), 555-564.e3. https://doi.org/10.1016/j.str.2018.02.005.

(34) Mahawaththa, M. C.; Pearce, B. J. G.; Szabo, M.; et al. Solution Conformations of a Linked Construct of the Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *142*, 141–147. https://doi.org/10.1016/j.antiviral.2017.03.011.

(35) Li, Y.; Phoo, W. W.; Loh, Y. R.; et al. Structural Characterization of the Linked NS2B-NS3 Protease of Zika Virus. *FEBS Lett.* **2017**, *591* (15), 2338–2347. https://doi.org/10.1002/1873-3468.12741.

(36) Kang, C. B.; Keller, T. H.; Luo, D. Zika Virus Protease: An Antiviral Drug Target. *Trends Microbiol.* **2017**, *25* (10), 797–808. https://doi.org/10.1016/j.tim.2017.07.001.

(37) Kim, Y. M.; Gayen, S.; Kang, C. B.; et al. NMR Analysis of a Novel Enzymatically Active Unlinked Dengue NS2B-NS3 Protease Complex. *J. Biol. Chem.* **2013**, *288* (18), 12891–12900. https://doi.org/10.1074/jbc.M112.442723.

(38) Bräuchle, C.; Lamb, D. C.; Michaelis, J. *Single Particle Tracking and Single Molecule Energy Transfer*; Wiley VCH: Weinheim, 2009. https://doi.org/10.1002/9783527628360.

(39) Foster, T. Zwischenmolekulare Energiewanderung Und Fluoreszenz. *Ann. Phys.* **1948**, *437*, 55–75.

(40) Baronti, C.; Piorkowski, G.; Charrel, R. N.; et al. Complete Coding Sequence of Zika Virus from a French Polynesia Outbreak in 2013. *Genome Announc.* **2014**, *2* (3), e00500-14. https://doi.org/10.1128/genomeA.00500-14.

(41) Braun, N. J.; Quek, J. P.; Huber, S.; et al. Structure-Based Macrocyclization of Substrate Analogue NS2B-NS3 Protease Inhibitors of Zika, West Nile and Dengue Viruses. *ChemMedChem* **2020**, *15* (15), 1439–1452. https://doi.org/10.1002/cmdc.202000237.

(42) Maus, H.; Barthels, F.; Hammerschmidt, S. J.; et al. SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B/NS3 Proteases. *Bioorganic Med. Chem.* **2021**, *47*. https://doi.org/10.1016/j.bmc.2021.116392.

(43) 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.

(44) Micsonai, A.; Wien, F.; Bulyáki, É.; et al. BeStSel: A Web Server for Accurate Protein Secondary Structure Prediction and Fold Recognition from the Circular Dichroism Spectra. *Nucleic Acids Res.* **2018**, *46* (W1), W315–W322. https://doi.org/10.1093/nar/gky497.

(45) Gao, K.; Oerlemans, R.; Groves, M. R. Theory and Applications of Differential Scanning Fluorimetry in Early-Stage Drug Discovery. *Biophys. Rev.* **2020**, *12* (1), 85–104. https://doi.org/10.1007/s12551-020-00619-2.

(46) Vivoli, M.; Novak, H. R.; Littlechild, J. A.; et al. Determination of Protein-Ligand Interactions Using Differential Scanning Fluorimetry. *J. Vis. Exp.* **2014**, No. 91, 1–13. https://doi.org/10.3791/51809.

(47) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat. Protoc.* **2007**, *2* (9), 2212–2221. https://doi.org/10.1038/nprot.2007.321.

(48) Ahmad, Y.; Shamsi, S. A. A Novel Direct Synthesis of the N -Oxides of Quinoline Derivatives. *Bull. Chem. Soc. Jpn.* **2006**, *39* (1), 195–195. https://doi.org/10.1246/bcsj.39.195.

(49) Magnusson, A. O.; Szekrenyi, A.; Joosten, H. J.; et al. NanoDSF as Screening Tool for Enzyme Libraries and Biotechnology Development. *FEBS J.* **2019**, *286* (1), 184–204. https://doi.org/10.1111/febs.14696.

(50) Cobb, S. L.; Murphy, C. D. 19F NMR Applications in Chemical Biology. *J. Fluor. Chem.* **2009**, *130* (2), 132–143. https://doi.org/10.1016/j.jfluchem.2008.11.003.

(51) Dalvit, C. Ligand- and Substrate-Based 19F NMR Screening: Principles and Applications to Drug Discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 243–271. https://doi.org/10.1016/j.pnmrs.2007.07.002.

(52) Rose-Sperling, D.; Tran, M. A.; Lauth, L. M.; et al. 19F NMR as a Versatile Tool to Study Membrane Protein Structure and Dynamics. *Biol. Chem.* **2019**, *400* (10), 1277–1288. https://doi.org/10.1515/hsz-2018-0473.

(53) Hellmich, U. A.; Pfleger, N.; Glaubitz, C. 19F-MAS NMR on Proteorhodopsin: Enhanced Protocol for Site-Specific Labeling for General Application to Membrane Proteins. *Photochem. Photobiol.* **2009**, *85* (2), 535–539. https://doi.org/10.1111/j.1751-1097.2008.00498.x.

(54) Lang, K.; Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **2014**, *114* (9), 4764–4806. https://doi.org/10.1021/cr400355w.

(55) Qianzhu, H.; Abdelkader, E. H.; Herath, I. D.; 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 Sensors* **2022**, *7* (1), 44–49. https://doi.org/10.1021/acssensors.1c02467.

(56) Qianzhu, H.; Abdelkader, E. H.; Herath, I. D.; et al. Site-Specific Incorporation of 7-Fluoro- File://10.93.203.16/Protease/Manuskript_ZIKV_in_solution/Literatur/Chemistry A European J - 2018 - Kenward - Mixed Fluorotryptophan Substitutions at the Same Residue Expand the Versatility.PdfL-Tryptophan into Prote. *ACS Sensors* **2022**, *7* (1), 44–49. https://doi.org/10.1021/acssensors.1c02467.

(57) Kenward, C.; Shin, K.; Rainey, J. K. Mixed Fluorotryptophan Substitutions at the Same Residue Expand the Versatility of 19F Protein NMR Spectroscopy. *Chem. - A Eur. J.* **2018**, *24* (14), 3391–3396. https://doi.org/10.1002/chem.201705638.

(58) Zheng, J. *Biomedical Applications of Biophysics*; Jue, T., Ed.; Humana Press: Totowa, NJ, 2010. https://doi.org/10.1007/978-1-60327-233-9.

(59) Hill, M. E.; Yildiz, M.; Hardy, J. A. 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.

(60) Wu, T.; Yu, J.; Gale-Day, Z.; et al. Three Essential Resources to Improve Differential Scanning Fluorimetry (DSF) Experiments. *bioRxiv* **2020**, 2020.03.22.002543. https://doi.org/10.1101/2020.03.22.002543.

(61) Kroeger, T.; Frieg, B.; Zhang, T.; et al. EDTA Aggregates Induce SYPRO Orange-Based Fluorescence in Thermal Shift Assay. *PLoS One* **2017**, *12* (5), 1–21. https://doi.org/10.1371/journal.pone.0177024.

(62) 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.

(63) Phoo, W. W.; Li, Y.; Zhang, Z.; et al. Structure of the NS2B-NS3 Protease from Zika Virus after Self-Cleavage. *Nat. Commun.* **2016**, *7*, 13410. https://doi.org/10.1038/ncomms13410.

(64) Hammerschmidt, S. J.; Huber, S.; Braun, N. J.; et al. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B / NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm. (Weinheim).* **2022**, e2200518. https://doi.org/10.1002/ardp.202200518.

(65) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and Inexpensive Incorporation of 19F-Tryptophan for Protein NMR Spectroscopy. *Chem. Commun.* **2012**, *48* (86), 10681–10683. https://doi.org/10.1039/c2cc35347d.

(66) Barthels, F.; Hammerschmidt, S. J.; Fischer, T. R.; et al. A Low-Cost 3D-Printable Differential Scanning Fluorometer for Protein and RNA Melting Experiments. *HardwareX* **2022**, *11*, e00256. https://doi.org/10.1016/j.ohx.2022.e00256.

(67) GraphPad Prism 7.0.4 for Windows. GraphPad Software: San Diego, California USA.

Supporting Information: The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and ^{19}F-NMR.

[redacted] Stefan J. Hammerschmidt [a#], [redacted]

[#] authors contributed equally

* corresponding author

[redacted]

[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

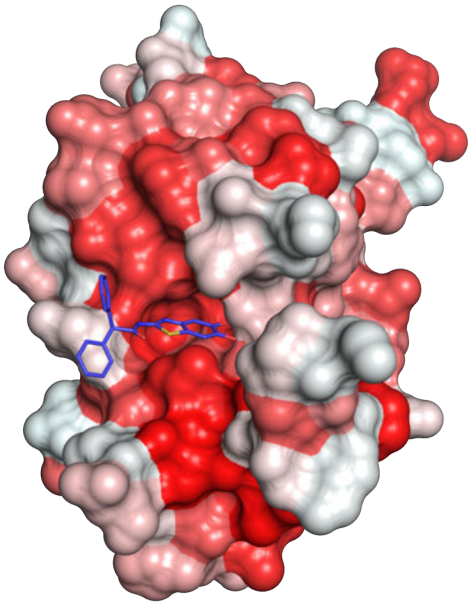


Figure S1. Representation of the surface of the ZIKV protease (PDB: 5GXJ) stained according to hydrophobicity. Deeper red color shades indicate less polarity of the surface.[1] The allosteric ligand $\mathbf{I_a}$ is shown in blue.

1. **Tryptophan Residues in Open and Closed Conformation of the ZIKV Protease**

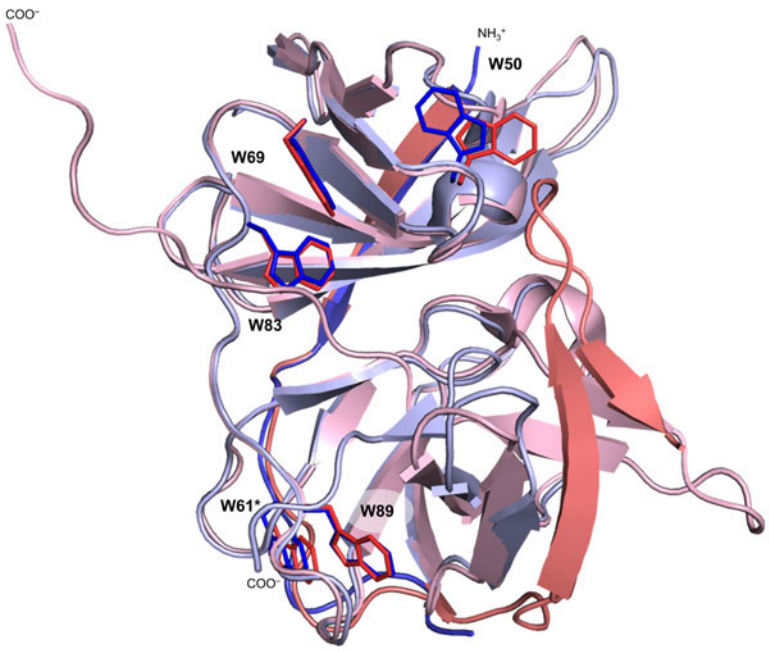


Figure S2. Tryptophan residues in the open (blue, PDB: 5GXJ) and closed conformation (red, PDB: 6Y3B) of the ZIKV protease. Tryptophan residues are highlighted as sticks. NS2B is shown in darker colors; and NS3 in light colors. Residue W61 within NS2B is labeled by an asterisk. W5 at the N-terminus of 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 phenylalanine *via* 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.

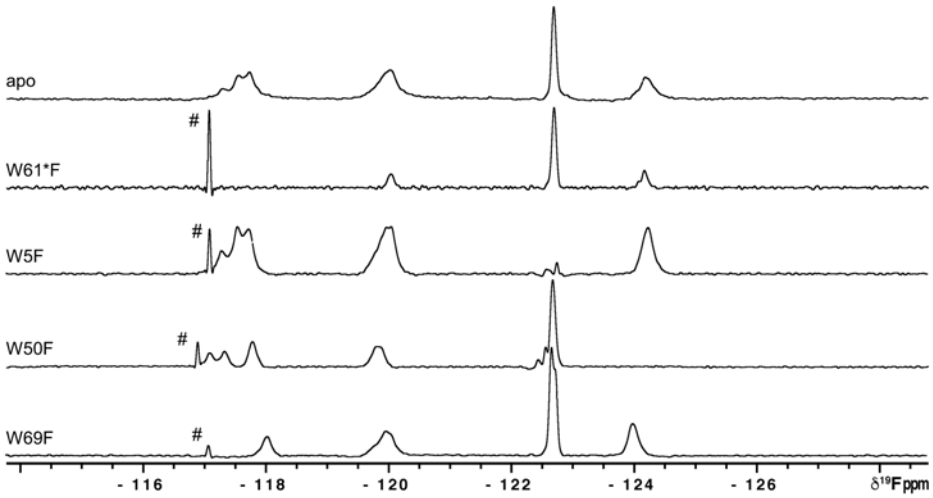


Figure S3. ^{19}F-NMR spectra of the 5F-5Trp labeled W/F mutants of ZIKV protease for assignment purposes. W83F and W89F did not show activity and aberrant shift patterns, so they were excluded from this analysis. # residual solvent peak.

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 ($\mathbf{I_c}$) 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 to 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 individual bursts (individual molecules) occurrences 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 1. 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_{DTT}	361 ± 27	16.0 ± 0.7	0.128 ± 0.005	356 ± 30
$^5ZiPro_{DTT}$	417 ± 16	14.2 ± 0.3	0.113 ± 0.003	272 ± 12
$^5ZiPro^{ATTO\ 488/\ ATTO\ 643}{}_{DTT}$	397 ± 16	7.83 ± 0.18	0.062 ± 0.001	158 ± 7
$^5ZiPro_{TCEP}$	794 ± 115	0.018 ± 0.002	$1.76 \cdot 10^{-3} \pm 1.42 \cdot 10^{-4}$	0.179 ± 0.031
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
$eZiPro_{TCEP}$	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

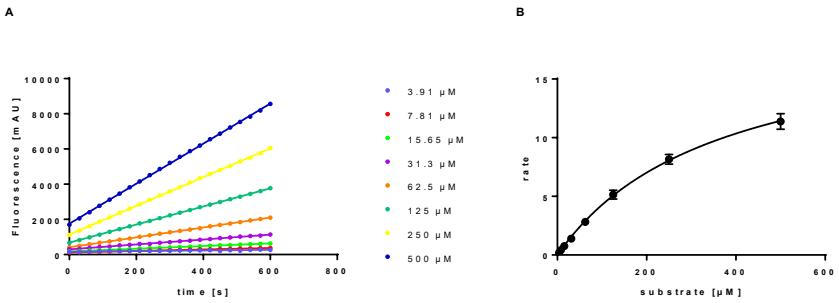


Figure S4. Fluorometric assay of the WT (gZiPro) of the ZIKV protease with 5 mM DTT. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

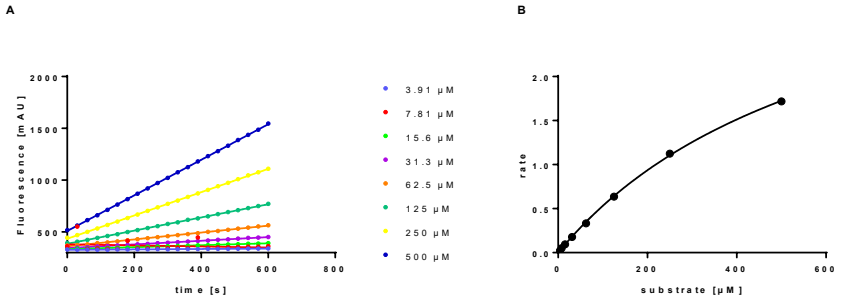


Figure S5. Fluorometric assay of the WT (gZiPro) of the ZIKV protease with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

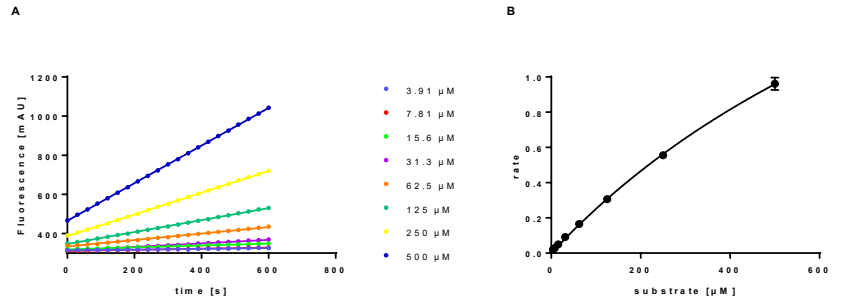


Figure S6. Fluorometric assay of the enzymatically cleavable linked ZIKV protease (eZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

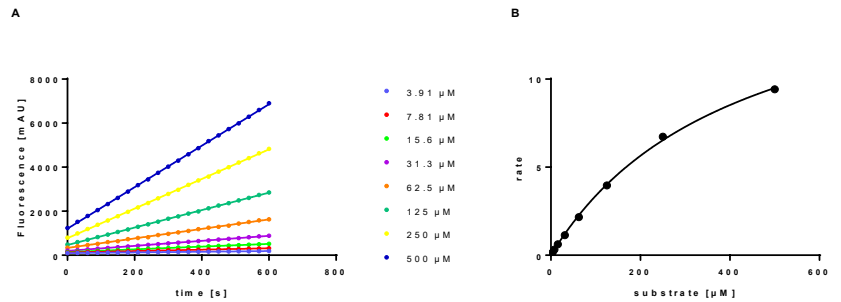


Figure S7. Fluorometric assay of the quintuple mutant of the wildtype of the ZIKV protease (^{5}ZiPro) with 5 mM DTT. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

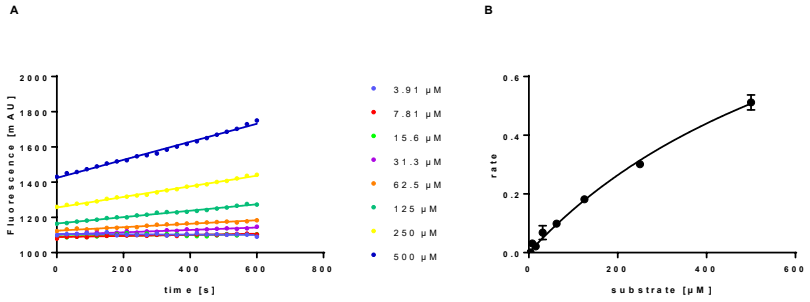


Figure S8. Fluorometric assay of the quintuple mutant of the wildtype of the ZIKV protease (^{5}ZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

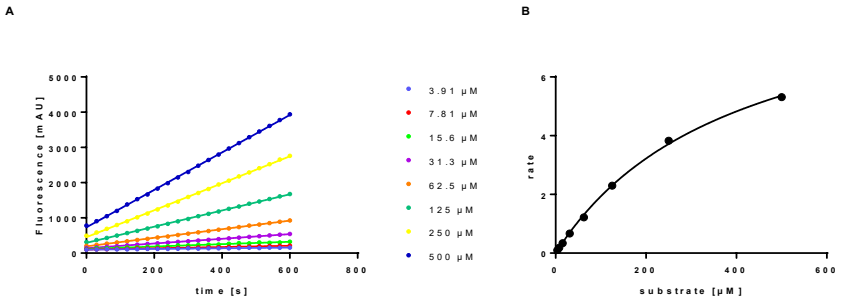


Figure S9. Fluorometric assay of the dye-labeled quintuple mutant of the wildtype of the ZIKV protease (^{5}ZiPro) with 5 mM DTT. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

5F-tryptophan labeled eZiPro

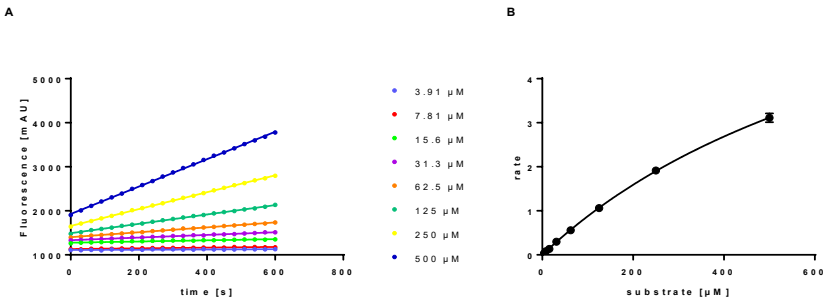


Figure S10. Fluorometric assay of the 5F-tryptophane labeled ZIKV protease (5F-W eZiPro) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

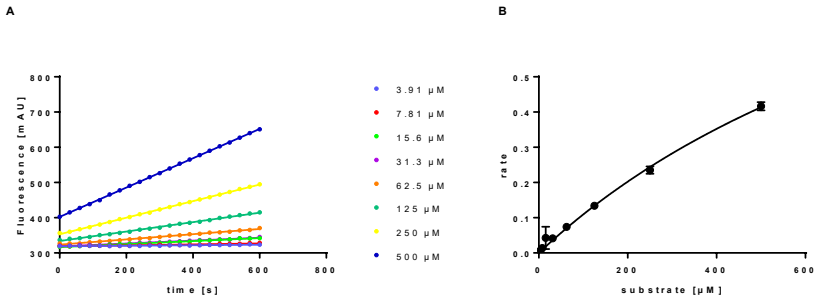


Figure S11. Fluorometric assay of the W61*F mutant of the ZIKV protease (eZiPro W61*F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

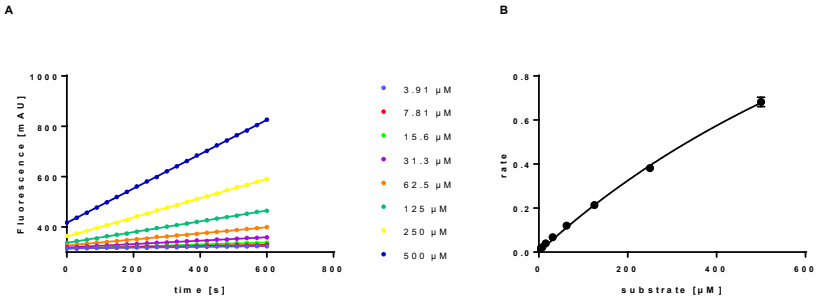


Figure S12. Fluorometric assay of the W5F mutant of the ZIKV protease (eZiPro W5F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

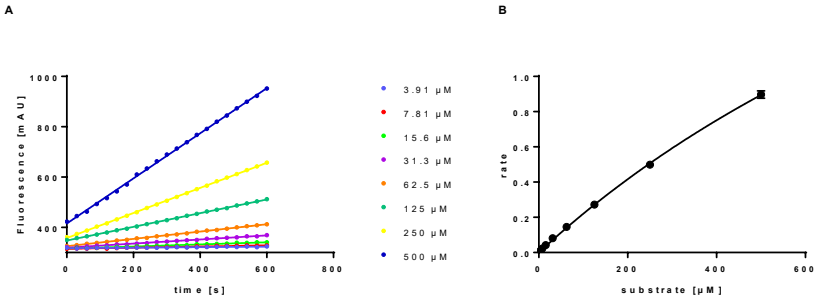


Figure S13. Fluorometric assay of the W50F mutant of the ZIKV protease (eZiPro W50F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

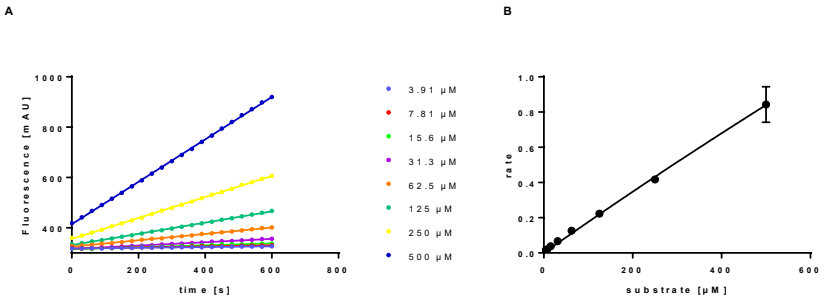


Figure S14. Fluorometric assay of the W69F mutant of the ZIKV protease (eZiPro W69F) with 1 mM TCEP. (**a**) Fluorescence increases over time for different substrate concentrations. (**b**) Fitted curve to determine K_M value.

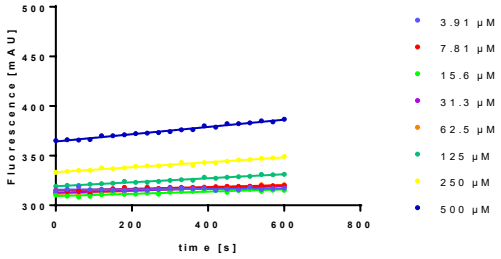


Figure S15. Fluorometric assay of the W83F mutant of the ZIKV protease (eZiPro W83F) with 1 mM TCEP.

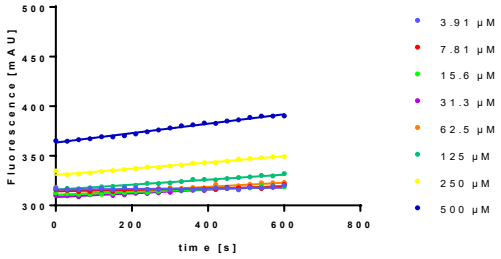


Figure S16. Fluorometric assay of the W89F mutant of the ZIKV protease (eZiPro W89F) with 1 mM TCEP.

REFERENCES

(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. Project 3: Zika Virus NS2B/NS3 Autocleavage Behavior

5.1. *Cis* autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3pro constructs

5.1.1. Context, Project Summary, and Own Contribution

Studying drug interactions requires homogeneous preparations of the isolated target protein. In drug discovery and development targeting ZIKV NS2B/NS3, a minimalistic construct consisting of the hydrophilic cofactor region of NS2B (amino acids 48–95) in conjunction with the N-terminal $NS3_{pro}$ domain (amino acids 1–186) *via* a glycine-rich linker (G_4SG_4) is of particular interest since it is used in most studies.[47,56,240,259–262] For those glycine-linked ZIKV $NS2B_{cf}/NS3_{pro}$ constructs (gZiPro), autocatalytic cleavage in this linker region occurred repeatedly.[33,41,263,264] To overcome this issue, two Arg residues, Arg95* and Arg29, were simultaneously mutated to Ala as they resemble possible cleavage sites for the NS2B/NS3 protease.[37,262] Indeed, this attempt abolished autocleavage, but no further investigations have specified which one of these sites is cleaved. Furthermore, the mechanism of this autocleavage behavior was not investigated. One protease may cleave another (*trans*-cleavage) or can cleave its own autocatalytic site (*cis*-cleavage). Although both Arg residues do not represent the native cleavage site of ZIKV NS2B/NS3, it is of interest to learn more about the behavior of this regularly utilized construct and to derive hints about the behavior of the native full-length NS2B/NS3 protease. Another goal of this publication was to obtain a minimally modified ZIKV NS2B/NS3 construct and to investigate the properties of cleaved or uncleaved proteases. This was achieved by investigating the autocleavage efficiency of the single gZiPro R95*A and R29A mutants. We demonstrated Arg95* to be the only relevant residue for autocleavage. Moreover, we designed an autocleavage assay where the catalytically inactive gZiPro S135A mutant with the intact Arg95* cleavage site was used in high concentrations. A small amount of enzymatically competent gZiPro R95*A mutant with attenuated autocatalytic cleavage was added. Proteolytic cleavage was assessed after incubation by densitometric analysis of Coomassie brilliant blue-stained SDS-PAGE. This absence of *trans*-cleavage demonstrated that autocleavage of the gZiPro exclusively occurs in a *cis* regiment. We recorded melting curves using the intrinsic tryptophane fluorescence and showed that the linker integrity did not significantly impact the thermal stability. However, subtle differences in the molting process occurred between the cleaved and uncleaved states, probably reflecting the possibility of dissociating NS2B of unlinked constructs.

(6) von Hammerstein, F.; Lauth L. M.; **Hammerschmidt, S.**; Wagner, A.; Schirmeister, T.; Hellmich, U. A. *Cis* autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs. *FEBS Lett*. **2019**; 593(16):2204–13. https://doi.org/10.1002/1873-3468.13507

Own contribution: Protein expression and purification, CD-spectroscopy, tryptophan-fluorescence melting curves, SDS-Page analysis, western blotting, enzyme activity assay, writing parts & editing of the manuscript.

Contributions from other authors: Protein expression and purification, cloning of constructs, CD-spectroscopy, SDS-Page analysis, analytical gel filtration, enzyme activity assay, densitometric analysis, CD-spectroscopy, tryptophan-fluorescence melting curves, writing of the original draft of the manuscript.

This work has been published in *FEBS Letters* (impact factor: 3.06).

5.1.2. Publication

The following publication quoted (within " and ") from page 278 to page 292 is exactly the same as the manuscript cited on page 277. "

FEBS Letters

FEBS PRESS science publishing by scientists

COMMUNICATION

Cis autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs

Franziska von Hammerstein[1], Luca M. Lauth[1], Stefan Hammerschmidt[1], Annika Wagner[1,2], Tanja Schirmeister[1] and Ute A. Hellmich[1,2]

1 Institute for Pharmacy and Biochemistry, Johannes Gutenberg University, Mainz, Germany
2 Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe-University, Frankfurt, Germany

Correspondence
U. A. Hellmich, Institute for Pharmacy and Biochemistry, Johannes Gutenberg University, Johann-Joachim-Becherweg 30, 55128 Mainz, Germany
Tel: +49-(0)6131 39-26182
E-mail: u.hellmich@uni-mainz.de

Franziska von Hammerstein and Luca M. Lauth contributed equally

(Received 2 July 2018, revised 4 June 2019, accepted 18 June 2019, available online 9 July 2019)

doi:10.1002/1873-3468.13507

Edited by Jacomine Krijnse-Locker

The flaviviral heterodimeric serine protease NS2B-NS3, consisting of the NS3 protease domain and the NS2B co-factor, is essential for ZIKA virus maturation and replication in cells. For *in vitro* studies a 'linked' construct, where a polyglycine linker connects $NS2B_{CF}$ and $NS3_{pro}$, is often used. This construct undergoes autocatalytic cleavage. Here, we show that linked ZIKV $NS2B_{CF}$-$NS3_{pro}$ is cleaved in *cis* in the $NS2B_{CF}$ exclusively at position R95 and not at the previously proposed alternate cleavage site at residue R29 in the $NS3_{pro}$. Cleavage neither affects protease stability nor activity, despite some observed differences in spectroscopic behavior. This minimally modified construct may thus be useful for future structural and functional studies of the flaviviral protease, for example when testing new inhibitors.

Keywords: autocatalytic cleavage; flavivirus; NS2B-NS3 protease; proteolysisserine; Zika virus

Mosquitoes are the transmitting source of an alarming number of severe diseases. There is a worldwide spread of *Aedes aegypti* due to increased trade, travel behavior, and climate change [1]. Among the diseases spread by *Aedes* mosquitoes are Dengue, Zika, West Nile, and Yellow fever, all of which are caused by single-stranded RNA viruses of the *Flaviviridae* family [2]. Zika fever in particular has gained attention in recent years because it causes severe developmental disorders in newborns [3], and Guillain-Barré syndrome in adults [4]. This is particularly disconcerting since Zika fever was believed for decades to be a nonthreatening disease that leads only to mild flu-like symptoms [5].

In (human) hosts, the flaviviral RNA genome is translated into a viral multi-transmembrane spanning polyprotein of ~ 3000 amino acids. The polyprotein comprises the structural proteins C (capsid), prM (precursor membrane), and E (envelope) as well as seven nonstructural (NS) proteins [6]. NS2B and NS3 form a serine chymotrypsin-like protease, essential for polyprotein processing. NS3 harbors the His51-Asp75-Ser135 catalytic triad while NS2B is wound around the NS3 core domain [7–12]. Since the NS2B-NS3 protease is essential for the progression of the viral replication cycle, it presents an attractive drug target [13,14]. However, drug-interaction or structural studies require homogeneous protein preparations. To date, functional assays with flaviviral NS2B-NS3 proteases have been carried out either with both proteins expressed separately ('unlinked' constructs) or, in the majority of cases, by connecting them *via* a polyglycine linker ('linked' constructs) [15,16]. Of particular relevance for structural studies is a commonly used minimalistic construct consisting of a shortened ZIKV NS2B co-factor

Abbreviations
CD, Circular dichroism; Ni-NTA, Ni^{2+}-nitriloacetic acid; NS, nonstructural; $NS2B_{CF}$, NS2B co-factor; $NS3_{pro}$, NS3 protease domain; prM, precursor membrane; RT, room temperature; SEC, size exclusion chromatography; TEV, tobacco etch virus.

(aa 48-95, $NS2B_{CF}$) in conjunction with the core NS3 protease domain (aa 1–186, $NS3_{pro}$). However, autocatalytic cleavage of the linked $NS2B_{CF}$-$NS3_{pro}$ construct has been observed repeatedly [9,17–19]. To stabilize the construct for structural studies of linked ZIKV protease, the arginine residues at positions $R95^{NS2B}$/$R29^{NS3}$ in regions adhering to an amino acid sequence commonly observed for the cleavage preferences of flaviviral proteases [20] were simultaneously mutated [7,12,16,21]. This successfully abolished autocatalytic cleavage, but without specifying which one (or both) of these sites is being cleaved. More recently, an X-ray structure showing a short $NS2B_{CF}$ C-terminal peptide fragment bound to $NS3_{pro}$ was suggestive of autocatalytic cleavage in *cis* at the NS2B-NS3 junction [17], but this has not yet been conclusively shown in solution.

For future studies, it would be advantageous to obtain a minimally modified ZIKV protease and at the same time gain a better understanding of the properties of cleaved or uncleaved ZIKV protease. Here, we show that $R95^{NS2B}$ is the only autocatalytic cleavage site for the soluble glycine-linked ZIKV protease and that cleavage occurs exclusively in *cis*. Cleavage does not significantly impact protein thermal stability as assessed by tryptophan fluorescence or protein activity, but we find subtle differences in the molten states of cleaved and uncleaved constructs as assessed by circular dichroism and tryptophan fluorescence spectroscopy.

Materials and methods

Ni^{2+}-nitriloacetic acid (Ni-NTA) resin was obtained from Qiagen, Boc-Gly-Arg-Arg-7-amino-4-methylcoumarin from Bachem (Bubendorf BL, Switzerland). All other chemicals were purchased from Carl Roth (Karlsruhe, Germany) of analytical grade or better.

Cloning of constructs

All ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs were designed according to the genome of French Polynesia ZIKV (GenBank ID: KJ776791.1) [22]. Genes were commercially obtained (GenScript and Eurofins Genomics) and cloned into a pET11a vector (linked) or a pACYCDuet-1 vector (unlinked). Both constructs contain an N-terminal hexa-histidine (His_6) tag with a tobacco etch virus (TEV) protease cleavage site, that is, in the linked construct, $NS2B_{CF}$ contains the His_6-tag, while in the unlinked construct, the tag is attached to the $NS3_{pro}$ domain (Fig. 1A). Point mutations were introduced through QuikChange mutagenesis using the Kapa HiFi PCR Kit (KapaBiosystems) with the following primers: (a) $R95A^{NS2B}$: 5′ccgatggctggcggtggtggtagcggc3′ and 5′caccgccagccatcggcggaccatcgtc3′; (b) $R29A^{NS3}$: 5′gacccgtgctctgctgggtagcacccaggtt3′ and 5′cagcagagcacgggtcatcacacggtaaacaccg3′; (c) $S135A^{NS3}$: 5′ggcaccgccggtagcccgattctggacaaatg3′ and 5′ctaccggcggtgccgccggataatccag3′.

Protein expression and purification

Expression was carried out in *Escherichia coli* strain BL21-Gold (DE3) (Agilent Technologies, Santa Clara, CA, USA). Cells were grown in LB medium with the corresponding antibiotic (ampicillin or chloramphenicol) at 37 °C to an OD_{600} of ~ 0.5 and induced with 1 mM isopropyl-D-thiogalactoside. Proteins were expressed at 20 °C for 12–14 h. After harvesting, cells were shock frozen in liquid nitrogen and stored at −20 °C until further use. For purification, cells were resuspended in lysis buffer (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, 0.1% (v/v) Triton X-100, RNase, DNase, lysozyme, 1 mM PMSF, 1 mM DTT, 1 mM benzamidine) and lysed by sonication (Sonoplus; Bandelin, Berlin, Germany). The lysate was cleared by centrifugation and the protein purified by IMAC. After washing (20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole), the protein was eluted (250 mM imidazole). The His_6-tag was removed by TEV cleavage during a dialysis step [20 mM Tris-HCl pH 8, 150 mM NaCl, overnight (oN), 4 °C]. After reverse IMAC, a gelfiltration step (HiLoad 16/60 Superdex 75 or 200 column; GE Healthcare, Chicago, IL, USA) was carried out (50 mM Tris-HCl pH 8, 150 mM NaCl). Throughout all steps, protein concentrations were measured *via* absorbance at 280 nm and the sample purity was assessed *via* SDS/PAGE.

Western blotting

ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs were separated *via* SDS/PAGE (10 μg per lane) and transferred to PVDF membranes (Ø 0.2 μM) by tank blotting. After blocking with non-fat milk, membranes were incubated (oN, 4 °C) with an anti His_6 antibody harboring horseradish peroxidase (1 : 2000; Thermo Fischer Scientific, Waltham, MA, USA). For detection, membranes were incubated with 800 μL Roti-Lumin and chemiluminescence was detected on a Raytest Stella 3200 (Elysia-raytest, Straubenhardt, Germany).

Analytical gelfiltration

For analytical gel filtration, 0.4 mg purified protease in a volume of 100–300 μL was applied to a 24-mL gel filtration column (ENrich SEC 70, Biorad; 50 mM Tris-HCl pH 8, 150 mM NaCl).

Protease activity assay

Fluorometric enzyme assays were performed on a Spark® (Tecan, Männerdorf, Switzerland) Reader using Boc-Gly-Arg-Arg-7-amino-4-methylcoumarin (Boc-GRR-AMC; Bachem) as substrate. Released AMC was excited at a wavelength of 380 nm and detected at 460 nm. The assay was carried out at 25 °C in a final volume of 200 μL in

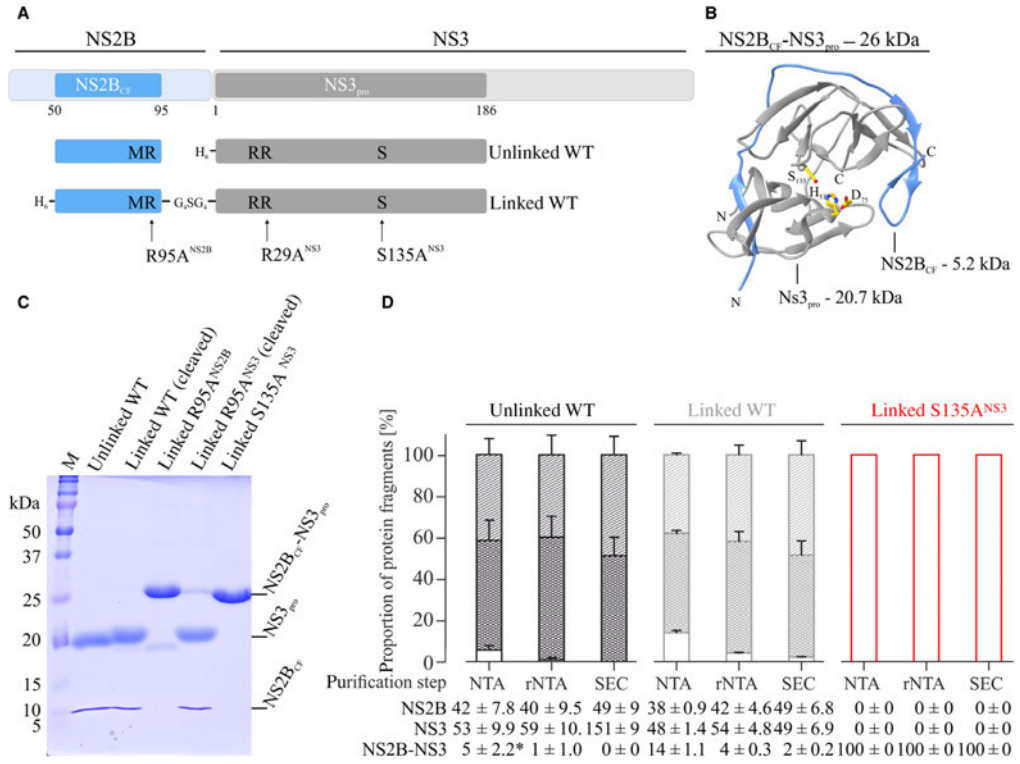


Fig. 1. Autocatalytic cleavage of ZIKV NS2B$_{CF}$-NS3$_{pro}$ constructs. (A) Topology of ZIKV NS2B$_{CF}$-NS3$_{pro}$ constructs used in this study. In the native protein, the co-factor NS2B$_{CF}$ and the protease NS3$_{pro}$ are connected via a polypeptide linker (top). In the constructs typically used for *in vitro* assays, the NS2B$_{CF}$ is shortened to its core ~ 45 amino acids and the NS3$_{pro}$ can be expressed separately (unlinked) or connected through a glycine-linker (G_4SG_4, 'linked'). An N-terminal His_6 tag was used for purification. The active site serine (S135^{NS3}) and the putative cleavage sites (R95^{NS2B} and R29^{NS3}) are indicated. (B) Crystal structure of ZIKV NS2B$_{CF}$-NS3$_{pro}$ (PDB: 5LC0) showing the three-dimensional arrangement of NS3$_{pro}$ (grey) with its co-factor NS2B$_{CF}$ (blue). The protease catalytic triad consisting of Ser135, His51 and Asp75 is shown in gold. (C) Coomassie-stained SDS/PAGE gel of ZIKV NS2B$_{CF}$-NS3$_{pro}$ variants after purification (10 µg per lane) shows two bands corresponding to NS2B$_{CF}$ and NS3$_{pro}$ for all constructs except the linked ZIKV NS2B$_{CF}$-NS3$_{pro}$ S135A^{NS3} (active site mutant) and R95A^{NS2B} variants. (D) During protein purification steps (NTA: Ni-NTA column; rNTA: reverse Ni-NTA column; SEC) densitometric analysis of SDS/PAGE gels shows continuous autocatalytic cleavage for linked WT NS2B$_{CF}$-NS3$_{pro}$, but not for the S135A^{NS3} construct. As expected, unlinked WT ZIKV NS2B$_{CF}$-NS3$_{pro}$ yields equivalent amounts of NS2B$_{CF}$ and NS3$_{pro}$ from the beginning. For the unlinked construct, densitometric analysis of the first purification step indicates the presence of a protein band running at the height of the linked construct (value of 5 ± 2.2% denoted by asterisk). Using Western blots (Fig. S1), we confirmed that this value is based on an impurity and that no linked NS2B$_{CF}$-NS3$_{pro}$ is present.

50 mM Tris-HCl buffer pH 9, with 20% glycerol, and 1 mM Chaps including 5 µL enzyme and 10 µL DMSO. Kinetic parameters were calculated using GRAFIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, West Sussex, UK) (Table 1).

Autocatalytic cleavage assays

Purified ZIKV NS2B$_{CF}$-NS3$_{pro}$ S135A^{NS3} protease was mixed with ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A^{NS2B} in a 7.5 : 1 molar ratio at 37 °C for up to 2 h. For SDS/PAGE analysis, samples were taken at 0, 10, 30, 60, and 120 min (5 µg ZIKV NS2B$_{CF}$-NS3$_{pro}$ S135A^{NS3} and 0.66 µg ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A^{NS2B} per lane).

Densitometric analysis

Densitometric analysis of Coomassie brilliant blue-stained SDS/PAGE gels was carried out using the IMAGEJ software (https://imagej.nih.gov/ij/). Since the Coomassie dye

Table 1. Activity and melting temperatures of ZIKV protease variants. All proteases (except where noted otherwise) were purified at RT. Proteases containing a glycine linker between $NS2B_{CF}$ and $NS3_{pro}$ are denoted as 'linked', even when they undergo autocatalytic cleavage during purification. Kinetic data are derived from a protease assay (25 °C) employing a fluorogenic tripeptide (Boc-GRR-AMC) [20], melting temperatures are based on tryptophan fluorescence measurements (Figs 2B and S2C). For detailed information see Table S1. na, not applicable.

ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs	K_M [μM]	v_{max} [RFU s^{-1}]	k_{cat} [s^{-1}]	$k_{cat}.K_M^{-1}$ [M^{-1} s^{-1}]	T_m [°C]
Unlinked WT	584.3 ± 133.4	9.9 ± 1.1	$2.0 \times 10^{-01} \pm 2.3 \times 10^{-02}$	339.6 ± 86.7	49.4 ± 0.2
Linked WT	557.7 ± 61.5	10.9 ± 0.6	$2.6 \times 10^{-02} \pm 1.4 \times 10^{-03}$	47.2 ± 5.8	48.1 ± 0.2
Linked $R95A^{NS2B}$	367.5 ± 32.5	7.7 ± 0.3	$6.0 \times 10^{-02} \pm 2.2 \times 10^{-03}$	162.4 ± 15.6	48.5 ± 0.1
Linked $R29A^{NS3}$	539.5 ± 61.8	28.2 ± 1.6	$2.0 \times 10^{-01} \pm 1.1 \times 10^{-02}$	365.2 ± 46.5	47.5 ± 0.2
Linked $S135A^{NS3}$	na	na	na	na	50.9 ± 0.1
Linked $R95A^{NS2B}$ (purified at 4 °C)	596.41 ± 30.12	15.17 ± 0.39	$9.8 \times 10^{-02} \pm 0.003$	164.7 ± 9.33	49.1 ± 0.2

specifically stains basic amino acids in polypeptides, densitometric data were normalized according to the quantity of basic amino acids within specific polypeptides [23].

Circular dichroism (CD) spectroscopy

Circular dichroism measurements were carried out on a Jasco-815 CD spectrometer (Jasco, Tokyo, Japan) at protein concentrations below 10 μM in 1.66 mM Tris-HCl and 5 mM NaCl. Scanning speed was set to 50 nm·min^{-1}, bandwidth and scanning intervals were set to 5 nm and 1 nm, respectively. All spectra were obtained from the automatic averaging of five measurements at 15–20 °C.

Tryptophan-fluorescence melting curves

Melting curves from 25–85 °C (5 °C steps with 2 min equilibration) were recorded on a Fluoromax-4 fluorimeter (Horiba Scientific, Kyōto, Japan) at a protein concentration of 2–5 μM. Tris-HCl and NaCl concentrations were adjusted to 1.66 and 5 mM, respectively. At each temperature step, the average wavelength of fluorescence was calculated using Equation 1 where I_n is the intensity at wavelength λ_n and $<\lambda>$ is the averaged emission wavelength to account for the simultaneous changes in both wavelength and intensity during the recoding of melting curves (Fig. S2B). The curve resulting of three melting experiments was fitted using a sigmoidal Boltzmann function.

$$<\lambda> = \frac{\sum(\lambda_n * I_n)}{\sum I_n} \quad (1)$$

Results

In agreement with previous studies [7,15,16,21], we also observed cleavage of linked WT ZIKV $NS2B_{CF}$-$NS3_{pro}$. We therefore set out to identify the cleavage site, whether cleavage occurs in *cis*, that is, an individual protease cleaves its own linker or in *trans*, that is, a protease linker is cleaved by another ZIKV $NS2B_{CF}$-$NS3_{pro}$ protease in the sample and to compare the melting behaviors of cleaved and uncleaved protease constructs.

Autocatalytic cleavage of ZIKV $NS2B_{CF}$-$NS3_{pro}$ occurs exclusively at position R95 in $NS2B_{CF}$

When purifying glycine-linked ZIKV $NS2B_{CF}$-$NS3_{pro}$, cleavage progresses throughout the protein purification and ultimately produces two fragments whose molecular weights agree with those of isolated $NS2B_{CF}$ and $NS3_{pro}$ obtained when purifying the unlinked protease as a comparison (Figs 1 and S1). In principle, cleavage could occur through *E. coli* proteases present during expression and purification. Therefore, we prepared ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ where the active site serine is mutated. Even when purified at room temperature (RT), this mutant does not undergo cleavage (Fig. 1), thus showing that ZIKV protease indeed undergoes autocatalytic cleavage and is not affected by proteases from the expression host.

Because flaviviral NS2B-NS3 proteases generally prefer basic amino acids at the P1 and P2 positions, although with some promiscuity [19,20], two potential ZIKV protease cleavage sites were considered in agreement with the observed fragment sizes on SDS/PAGE (Figs 1C and S1C): a dibasic cleavage site after position R28-$R29^{NS3}$ in the $NS3_{pro}$ core and a second site, M94-$R95^{NS2B}$ at the C-terminal end of the $NS2B_{CF}$. In previous studies, both sites were mutated simultaneously to successfully obtain uncleaved linked ZIKV $NS2B_{CF}$-$NS3_{pro}$ [7,12,16,21]. To establish which one of these sites, or both, are cleaved, we prepared samples of linked protease containing only the single mutations, ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ and $R29A^{NS3}$. While linked $NS2B_{CF}$-$NS3_{pro}$ $R29A^{NS3}$ underwent cleavage similar to the WT protein thus separating the $NS3_{pro}$ domain from the $NS2B_{CF}$, linked $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ was not cleaved (Fig. 1C). Therefore, $R95^{NS2B}$ in the $NS2B_{CF}$ presents the only

relevant ZIKV protease autocatalytic cleavage site when using the glycine-linked construct.

Autocatalytic cleavage does not affect protein structural integrity or activity

Regardless of whether cleavage occurred or not, the interaction of $NS2B_{CF}$ with $NS3_{pro}$ was not compromised, as all probed constructs behaved similarly on a size exclusion chromatography (SEC) column (Fig. 2A). With or without cleavage, $NS3_{pro}$ and $NS2B_{CF}$ remain tightly connected as expected from other studies with flaviviral proteases [16,24]. In addition to the SEC runs, we recorded CD spectra at 15 °C. In agreement with the available crystal and NMR structures of ZIKV $NS2B_{CF}$-$NS3_{pro}$ and with previously published CD spectra [12,15,17,21,24,25], our spectra show large contributions of loops and random coil to the overall protein structure, and are virtually indistinguishable for all protease variants (Fig. 2C).

To test the proteolytic activity of the (mutant) ZIKV $NS2B_{CF}$-$NS3_{pro}$ variants and to exclude that $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ yields inactive protein not capable of undergoing autocatalytic cleavage, we used a classic protease cleavage assay with a fluorophore-labeled tripeptide [20]. All ZIKV serine protease variants display activity with the exception of $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$, which is catalytically silent due to the mutation of the active site serine residue (Table 1). The K_M values as a measure of substrate affinity and the k_{cat} values as a measure of catalytic activity of all tested enzymes were in the same range (e.g., from 300 to 600 μM for the K_M), thus indicating that protease cleavage neither significantly affects protease-substrate affinity nor turnover rates.

Thermal stability of ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs

Autocatalytic cleavage of linked ZIKV $NS2B_{CF}$-$NS3_{pro}$ was not always completed at the end of the purification when carried out at 4 °C, thus yielding potentially heterogeneous protein preparations (data not shown). Hence, we tested whether ZIKV protease was stable enough to be purified at RT and whether this increased autocatalytic cleavage sufficiently to yield homogeneous, cleaved protein. For both, linked and unlinked ZIKV $NS2B_{CF}$-$NS3_{pro}$ WT constructs, two separate bands on SDS/PAGE corresponding to $NS2B_{CF}$ and $NS3_{pro}$ and a single peak during SEC chromatography were observed after RT purification and the proteins had comparable activity (Figs 1C and 2A, Table 1).

To investigate the thermal stability of unlinked/cleaved and linked/uncleaved proteases, we recorded CD spectra at 15 and at 90 °C (Fig. 2C). While all constructs melt, the differences in the spectra for linked/uncleaved and unlinked/cleaved proteases at 90 °C indicate that the linker integrity either affects the melting behavior or the final molten state. Determining the T_m with CD spectroscopy is typically achieved by measuring the molar ellipticity at 222 or 208 nm across a temperature gradient, but since these wavelengths are representative for an α-helical protein, this approach is not feasible for a non-α-helical construct. Instead, we took advantage of the five native tryptophan residues in ZIKV $NS2B_{CF}$-$NS3_{pro}$ to record melting curves *via* tryptophan fluorescence (Figs 2 and S2). All constructs showed a T_m value of ~ 49 °C, thus indicating that the state of the linker does not affect the melting temperature *per se*. Differences in final tryptophan fluorescence intensity and emission wavelength between linked/uncleaved and unlinked/cleaved constructs, however, indicate subtle alterations in the molten protease in dependence of linker integrity.

ZIKV $NS2B_{CF}$-$NS3_{pro}$ protease undergoes autocatalytic cleavage exclusively in *cis*

ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$, which is not cleaved during the purification, shows the same catalytic activity against tripeptides *in vitro* as the WT ZIKV $NS2B_{CF}$-$NS3_{pro}$ (Table 1). The catalytically inactive ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ will carry out neither *cis*- nor *trans*-cleavage, but should in principle be susceptible to hydrolysis by another proteolytically active protease, as long as proteolytic cleavage occurs in *trans*. Therefore, catalytically inactive $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$, the 'substrate', was co-incubated at 37 °C with catalytically active $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$, the 'enzyme' (Fig. 3A). To ensure that the 'enzyme' retains its structural integrity and activity throughout the assay duration, we monitored protein activity at 37 °C for prolonged periods of time (Fig. 3B,D). During this time, no degradation or autocatalytic cleavage products were observed *via* SDS/PAGE (Fig. 3B). Based on experimentally obtained turnover rates (Table 1) the enzyme concentrations were chosen to ensure complete turnover of the entire 'substrate' pool in less than 4 min even when neglecting the increased Stokes radius of the enzyme compared to a peptidic substrate. In agreement with autocatalytic cleavage occurring in *cis*, during the co-incubation of the ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ 'enzyme' with the ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ 'substrate' no cleavage products could be detected.

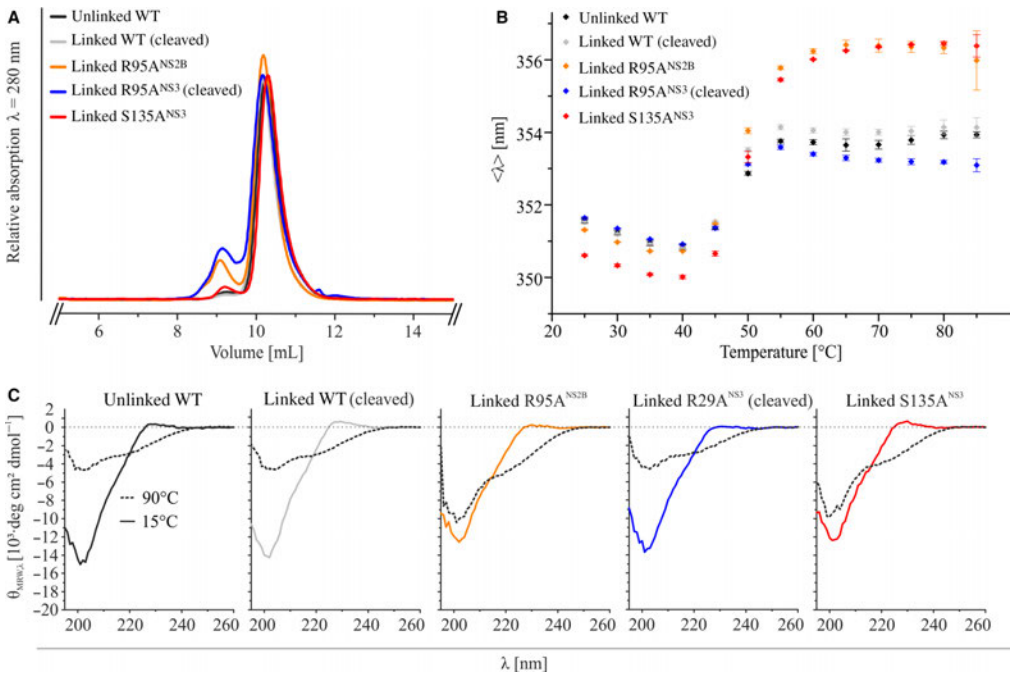


Fig. 2. Structural analyses of ZIKV NS2B$_{CF}$-NS3$_{pro}$ constructs. (A) Analytical SEC of ZIKV NS2B$_{CF}$-NS3$_{pro}$ constructs and (B) Tryptophan fluorescence based melting curves. All protein variants melt at similar temperatures of ~ 49 °C (Table 1). (C) CD spectra indicating that all protein variants are folded correctly and that point mutations and the integrity status of the glycine linker has no influence on overall protein folding and structure. Spectra recorded at 90 °C, depicted as dotted lines, show differences in the melted state of the uncleaved, linked constructs ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A and S135A compared to the cleaved or unlinked constructs.

In combination with our observations during protein purification, where both the linked WT NS2B$_{CF}$-NS3$_{pro}$ as well as the R29A^{NS3} mutant were cleaved into two fragments, our data indicate that ZIKV NS2B$_{CF}$-NS3$_{pro}$ undergoes a single autocatalytic cut at position R95^{NS2B} in *cis*. The introduction of the R95A^{NS2B} single mutant had no effects on ZIKV protease activity and can be used as the minimally modified construct for future functional and structural studies.

Discussion

Autocatalytic cleavage of flaviviral proteases is an important step in the polyprotein maturation process [6] and has been observed in various NS2B-NS3 proteases [9,18,19]. In principle, autocatalytic cleavage can occur in *cis* (each protease cleaves itself), in *trans* (proteases cleaving others) or *via* both mechanisms. For Yellow fever virus NS2B-NS3, dilution experiments with the purified NS2B$_{CF}$-NS3$_{pro}$ have been carried out. Since dilution did not affect cleavage, it was concluded that the observed cleavage activity follows a *cis*-mechanism [9]. For Dengue serotype 2 and West-Nile virus NS2B-NS3, an enzymatic assay similar to our approach here for ZIKV NS2B$_{CF}$-NS3$_{pro}$ was used [18,19] which also showed that these proteases are processed by autocatalytic *cis* cleavage as we observe for the ZIKV enzyme. The sizes of the autocatalytic cleavage products led us and others to conclude that there are two possible cleavage sites in ZIKV NS2B$_{CF}$-NS3$_{pro}$, one at the NS2B$_{CF}$ C-terminal R95^{NS2B} and one at R29^{NS3} in the NS3$_{pro}$ domain. The double mutation simultaneously removing both arginine residues successfully abrogated autocatalytic cleavage [7,12,16,21]. Here, we mutated these sites individually and thus identified the relevant autocatalytic cleavage site to be at position R95^{NS2B} in the C-terminus of the NS2B$_{CF}$. This finding is intriguing for multiple reasons: The ZIKV NS2B$_{CF}$-NS3$_{pro}$ is believed to be

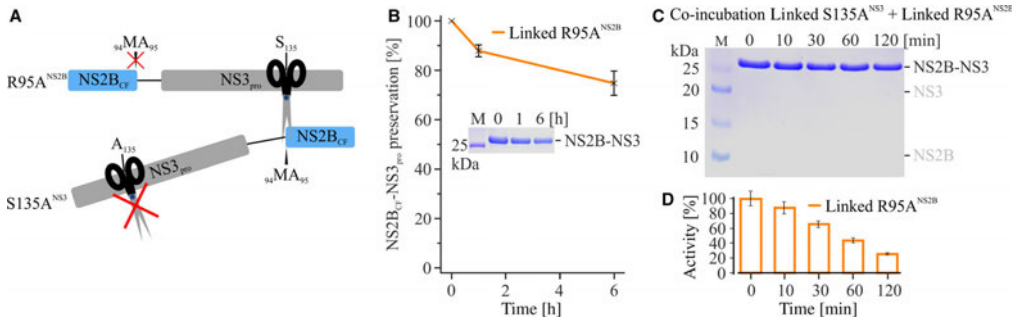


Fig. 3. ZIKV $NS2B_{CF}$-$NS3_{pro}$ autocatalytic cleavage occurs in *cis*. (A) Schematic representation of the experimental approach to investigate whether ZIKV autocatalytic cleavage occurs in *cis* or *trans*. Catalytically active ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ which cannot be cleaved was used as an 'enzyme' and catalytically inactive ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ that maintains the $R95^{NS2B}$ cleavage site used as 'substrate' were mixed. (B) To ensure that ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ is not able to perform autocatalytic cleavage and is stable for the duration of the cleavage assay, it was incubated at 37 °C for up to 6 h. No significant degradation was observed by SDS/PAGE (7.5 μg per lane). (C) SDS/PAGE analysis of the co-incubation of ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ (5 μg) and ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ (0.66 μg) at 37 °C for up to 120 min showed no cleavage of ZIKV $NS2B_{CF}$-$NS3_{pro}$ $S135A^{NS3}$ in the presence of an active protease. (Expected fragment sizes for $NS2B_{CF}$ and $NS3_{pro}$ are indicated in grey). (D) To assess the functional integrity of ZIKV $NS2B_{CF}$-$NS3_{pro}$ $R95A^{NS2B}$ for the duration of the assay, the activity of this protease variant was determined upon incubation at 37 °C *via* a peptide cleavage assay.

specific toward dibasic residues [20]. However, in the frequently used ZIKV $NS2B_{CF}$-$NS3_{pro}$ construct, cleavage occurred at a single basic residue $R95^{NS2B}$ which is preceded by a methionine. In contrast, the second putative autocatalytic cleavage site at $R29^{NS3}$, preceded by an arginine residue and thus adhering more strictly to the classical protease substrate preference sequence, is not cleaved (Figs 1 and 4B). This could be explained by steric reasons. Based on crystal structures of ZIKV $NS2B_{CF}$-$NS3_{pro}$ (e.g., [12]), $R29^{NS3}$ is located at a turn far from the active site and framed by two β-strands potentially limiting local dynamics. Thus, at this position, autocatalytic cleavage in *cis* is sterically hindered. Additionally, the peptide does not seem to have the typically required extended conformation for cleavage in either *cis* or *trans*.

To obtain an X-ray structure of ZIKV $NS2B_{CF}$-$NS3_{pro}$, Phoo *et al*. [17] used a construct with a shortened, native linker sequence instead of the frequently used glycine linker (Fig. 4, PDB: 5GJ4). Here, a cleaved peptide corresponding to the NS2B C-terminus (^{98}TGKR^{101}) was locked in the protease active site. Consequently, cleavage must have occurred at position 101 at some point during purification and/or crystallization. The commonly used glycine-linked construct, however, is abrogated at position $R95^{NS2B}$. Therefore, the structure determined by Phoo *et al*. places the cleavage site six amino acids C-terminal of the cleavage site at position $R95^{NS2B}$ (Fig. 4B). This indicates high intrinsic flexibility in the autocatalytic cleavage peptide recognition and processing complex. Thus, different cleavage patterns may be obtained based on the flaviviral $NS2B_{CF}$-$NS3_{pro}$ constructs used for expression and purification. At this point, no conclusions can be made about the auto-processing behavior of ZIKV NS2B-NS3 *in vivo*, as the physiological linker contains more than 20 additional residues, including basic amino acids (Fig. 4) although it is clear that the NS2B-NS3 protease is essential for viral polyprotein processing and viral maturation [13,14].

Linked ZIKV protease undergoes autocatalytic cleavage during the protein purification (see also e.g. [7,15,16,21]), which may lead to heterogeneous protein preparations and could thus hamper further structural and functional studies. This can be circumvented by purifying the enzyme at RT.

A slight difference in K_M values between linked and unlinked ZIKV $NS2B_{CF}$-$NS3_{pro}$ protease has been reported previously [16]. In our hands, we did not observe any functional differences between the two protein variants larger than the margin of error. However, it should be noted that we used a fluorogenic tripeptide for our cleavage assays, whereas a tetrapeptide was used in the previous study. In addition, we used a single instead of a double mutant to obtain the uncleaved ZIKV protease, namely the $R95A^{NS2B}$ construct instead of a commonly used double mutant at positions $R95^{NS2B}$ and $R29^{NS3}$. Depending on the roles these residues play for global protein dynamics, it is possible that slight functional and/or structural

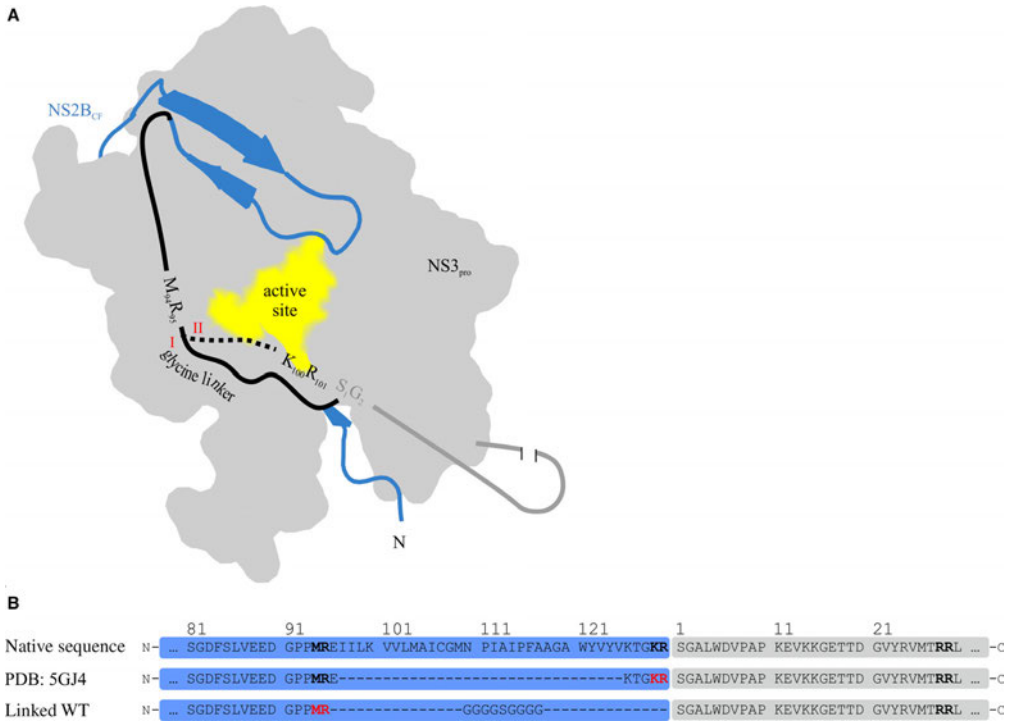


Fig. 4. Construct-dependent cleavage patterns of ZIKV NS2B$_{CF}$-NS3$_{pro}$. (A) In a previous structural study, a linked protease with a shortened native sequence was used instead of the commonly used glycine linker (path I) [17]. Here, the protease underwent autocatalytic cleavage at residue R101^{NS2B} (red letters, sequence alignment) in the NS2B C-terminus (blue). In contrast, the glycine-linked ZIKV NS2B$_{CF}$-NS3$_{pro}$ (path II) undergoes autocatalytic cleavage at position R95^{NS2B} (red letters, sequence alignment). (B) Due to the presence of different linkers, these two cleavage sites are nine amino acids apart from each other.

differences could be observed. Importantly, all constructs (linked, unlinked, and cleaved) showed almost identical melting temperatures of ~ 49 °C, but the linker integrity affected the spectroscopic behavior of the proteins (Figs 2 and S2B). At least for the tryptophan fluorescence spectra, this is potentially explained by the location of the native tryptophan residues (Fig. S2A). Two of these, W61^{NS2B} and W89^{NS3}, are near the NS2B$_{CF}$-NS3$_{pro}$ interface and may thus experience melting differently whether or not the linker is intact.

Together, our observation that glycine-linked ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A^{NS2B} neither shows autocatalytic cleavage during the purification nor when incubated for prolonged periods at RT or even at 37 °C and the finding that ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A^{NS2B}, which behaves as WT in a proteolytic assay, is not capable of cleaving ZIKV NS2B$_{CF}$-NS3$_{pro}$ S135A^{NS3}, indicate that ZIKV protease autocatalytic cleavage occurs exclusively in *cis*. For further studies, cleavage of ZIKV NS2B$_{CF}$-NS3$_{pro}$ can easily be avoided by introducing the single R95A^{NS2B} mutation, constituting a minimal modification of the commonly used construct. Alternatively, homogeneously cleaved protein can be obtained by carrying out the purification at RT.

Acknowledgements

Research was supported by the Carl-Zeiss Foundation (to UAH) and the Centre for Biomolecular Magnetic Resonance (BMRZ) funded by the State of Hesse, Germany. FvH acknowledges a Gutenberg Junior Academy Fellowship and financial support by the Mainz Stufe I program. We acknowledge Nadja Hellmann as well as the entire Schneider group, biochemistry, JGU Mainz, Claudia Janssen and Sabine Maehrlein for fruitful discussions and technical support.

Author contributions

FvH cloned constructs. FvH, LML, SH, AW, and UAH carried out protein purifications. FvH, LML, and SH carried out protease activity assays. FvH, LML, AW, and UAH carried out autocatalytic cleavage assays. LML, AW, and UAH carried out CD measurements and analytical gel filtration. LML and SH carried out Trp fluorescence experiments. FvH and UAH designed experiments. TS and UAH supervised experiments. FvH, LML, SH, and UAH wrote the paper with contributions from all authors.

References

1 Powell JR (2016) Mosquitoes on the move. *Science* **354**, 971–972.

2 Braack L, Gouveia de Almeida AP, Cornel AJ, Swanepoel R and de Jager C (2018) Mosquito-borne arboviruses of African origin. *Parasit Vectors* **11**, 1552.

3 Rasmussen SA, Jamieson DJ, Honein MA and Petersen LR (2016) Zika virus and birth defects-reviewing the evidence for causality. *N Engl J Med* **374**, 1981–1987.

4 Rozé B, Najioullah F, Fergé JL, Dorléans F, Apetse K, Barnay JL, Daudens-Vaysse E, Brouste Y, Césaire R, Fagour L *et al.* (2017) Guillain-Barré syndrome associated with Zika Virus infection in Martinique in 2016: a prospective study. *Clin Infect Dis* **65**, 1462–1468.

5 Maharajan MK, Ranjan A, Chu JF, Foo WL, Chai ZX, Lau EY, Ye HM, Theam XJ and Lok YL (2016) Zika virus infection: current concerns and perspectives. *Clin Rev Allergy Immunol* **51**, 383–394.

6 Lindenbach BD, Thiel HJ and Rice CM (2007) Flaviviridae: The Viruses and their Replication. In Fields Virology (Knipe DM and Howley PM, eds), pp. 1101–1152. Philadelphia, PA: Lippincott-Raven Publishers.

7 Rut W, Zhang L, Kasperkiewicz P, Poreba M, Hilgenfeld R and Drąg M (2017) Extended substrate specificity and first potent irreversible inhibitor/activity-based probe design for Zika virus NS2B-NS3 protease. *Antiviral Res* **139**, 88–94.

8 Chambers TJ, Grakoui A and Rice CM (1991) Processing of the yellow fever virus nonstructural polyprotein. *J Virol* **65**, 6042–6050.

9 Chambers TJ, Weir RC, Grakoui A, McCourt DW, Bazan JF, Fletterick RJ and Rice CM (1990) Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proc Natl Acad Sci USA* **87**, 8898–8902.

10 Falgout B, Miller RH and Lai CJ (1993) Deletion analysis of dengue virus type 4 nonstructural protein NS2B. *J Virol* **67**, 2034–2042.

11 Leung D, Schroder K, White H, Fang NX, Stoermer MJ, Abbenante G, Martin JL, Young PR and Fairlie DP (2001) Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. *J Biol Chem* **276**, 45762–45771.

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

13 Kang C, Keller TH and Luo D (2017) Zika virus protease. *Trends Microbiol* **25**, 797–808.

14 Yao Y, Huo T, Lin YL, Nie S, Wu F, Hua Y, Wu J, Kneubehl AR, Vogt MB, Rico-Hesse R *et al.* (2019) Discovery, X-ray crystallography and antiviral activity of allosteric inhibitors of flavivirus NS2B-NS3 protease. *J Am Chem Soc* **141**, 6832–6863.

15 Li Y, Phoo WW, Loh YR, Zhang Z, Ng EY, Wang W, Keller TH, Luo D and Kang C (2017) Structural characterization of the linked NS2B-NS3 protease of Zika virus. *FEBS Lett* **591**, 2338–2347.

16 Kuiper BD, Slater K, Spellmon N, Holcomb J, Medapureddy P, Muzzarelli KM, Yang Z, Ovadia R, Amblard F, Kovari IA *et al.* (2017) Increased activity of unlinked Zika virus NS2B/NS3 protease compared to linked Zika virus protease. *Biochem Biophys Res Comm* **492**, 668–673.

17 Phoo WW, Li Y, Zhang Z, Lee MY, Loh YR, Tan YB, Ng EY, Lescar J, Kang C and Luo D (2016) Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat Commun* **7**, 13410.

18 Preugschat F, Yao CW and Strauss JH (1990) *In vitro* processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. *J Virol* **64**, 4364–4374.

19 Bera AK, Kuhn RJ and Smith JL (2007) Functional characterization of cis and trans activity of the Flavivirus NS2B-NS3 protease. *J Biol Chem* **282**, 12883–12892.

20 Gruba N, Rodriguez Martinez JI, Grzywa R, Wysocka M, Skoreński M, Burmistrz M, Łęcka M, Lesner A, Sieńczyk M and Pyrć K (2016) Substrate profiling of Zika virus NS2B-NS3 protease. *FEBS Lett* **590**, 3459–3468.

21 Mahawaththa MC, Pearce BJG, Szabo M, Graham B, Klein CD, Nitsche C and Otting G (2017) Solution conformations of a linked construct of the Zika virus NS2B-NS3 protease. *Antiviral Res* **142**, 141–147.

22 Baronti C, Piorkowski G, Charrel RN, Boubis L, Leparc-Goffart I and de Lamballerie X (2014) Complete coding sequence of zika virus from a French polynesia outbreak in 2013. *Genome Announc* **2** , e00500–e00514.

23 Tal M, Silberstein A and Nusser E (1985) Why does Coomassie brilliant blue R interact differently with different proteins? *J Biol Chem* **260**, 9976–9980.

24 Zhang Z, Li Y, Loh YR, Phoo WW, Hung AW, Kang C and Luo D (2016) Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* **354**, 1597–1600.

25 Roy A, Lim L, Srivastava S, Lu Y and Song J (2017) Solution conformations of Zika NS2B-NS3pro and its inhibition by natural products from edible plants. *PLoS ONE* **12**, e0180632.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Measured and calculated data for ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs shown in Table 1.

Fig. S1. Comparison of Western-Blot and Coomassie stained SDS/PAGE gels.

Fig. S2. Tryptophan fluorescence measurements.

Fig. S3. Michaelis-Menten kinetics of ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs with Boc-GRR-AMC substrate.

Supplementary Information

Cis autocatalytic cleavage of glycine linked Zika virus NS2B-NS3 protease constructs

Franziska von Hammerstein[1,#], Luca M. Lauth[1,#], Stefan Hammerschmidt[1], Annika Wagner[1,2], Tanja Schirmeister[1], Ute A. Hellmich[1,2,*]

[1]Institute for Pharmacy and Biochemistry, Johannes Gutenberg University, Mainz, Germany

[2]Centre for Biomolecular Magnetic Resonance (BMRZ), Goethe-University Frankfurt, Germany

[#]these authors contributed equally: FvH, LML

*to whom correspondence should be addressed: u.hellmich@uni-mainz.de

Franziska von Hammerstein, Stefan Hammerschmidt, Tanja Schirmeister

Johannes Gutenberg-University

Institute for Pharmacy and Biochemistry

Staudingerweg 5

55128 Mainz, Germany

Luca M. Lauth, Annika Wagner, Ute A. Hellmich

Johannes Gutenberg-University Mainz

Institute for Pharmacy and Biochemistry

Johann-Joachim-Becherweg 30

55128 Mainz, Germany

Annika Wagner, Ute A. Hellmich

Goethe University Frankfurt

Centre for Biomolecular Magnetic Resonance (BMRZ)

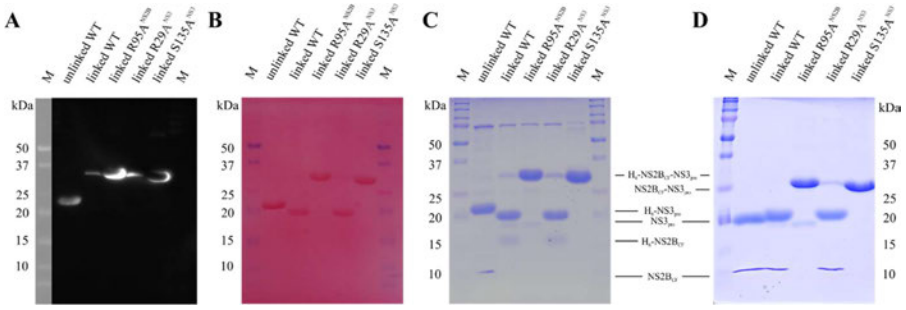
Max von Laue Str. 9

60438 Frankfurt, Germany

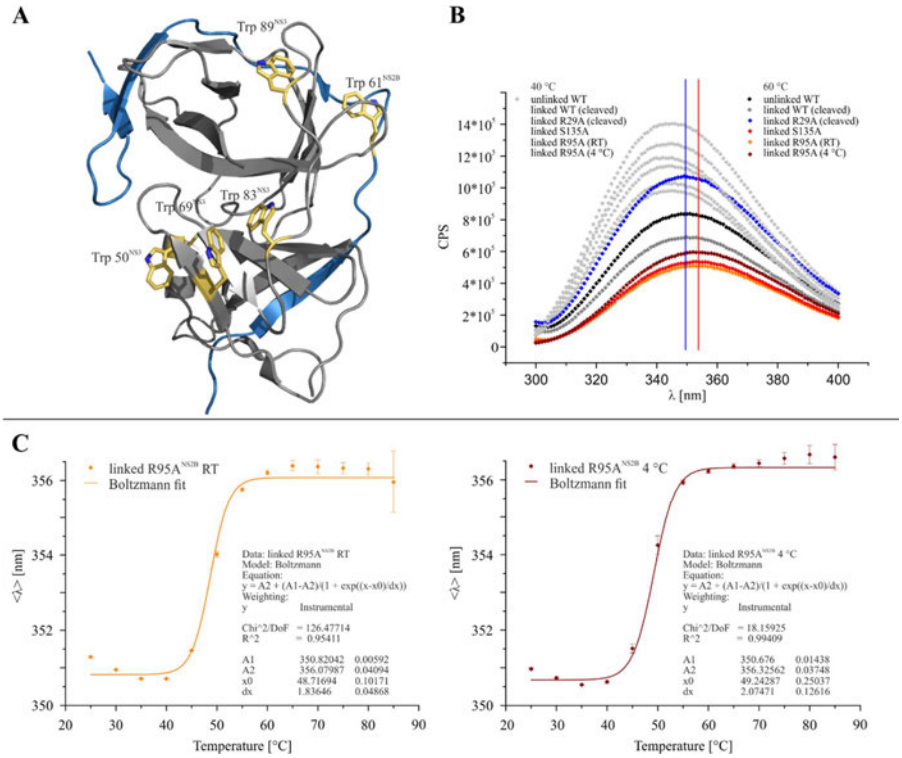
Supplementary table 1: Measured and calculated data for ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs shown in table 1. The v_{max} [$\mu M*s^{-1}$] values were obtained from v_{max} [$RFU*s^{-1}$] by dividing through a previously determined factor of proportionality (f = 3422.6 $RFU*\mu M^{-1}$) based on fluorescence measurements with isolated AMC to correlate relative fluorescence units (RFU) and µM concentrations of substrate. The error (k_{cat}/K_M) was calculated according to equation (S1).

ZIKV $NS2B_{CF}$-$NS3_{pro}$ constructs	c_{enzyme} [µM]	K_M [µM]	σ (K_M)	υ_{max} [RFU s^{-1}]	σ (υ_{max})	υ_{max} [µM s^{-1}]	σ (υ_{max})	k_{cat} [s^{-1}]	σ (k_{cat})	k_{cat}/K_M [M^{-1} s^{-1}]	error (k_{cat}/K_M) [M^{-1} s^{-1}]
unlinked WT	0.015	584.32	133.41	9.85	1.13	2.9E-03	3.3E-04	0.198	0.023	3.4E+02	86.73
linked WT	0.121	557.70	61.46	10.90	0.59	3.2E-03	1.7E-04	0.026	0.001	4.7E+01	5.80
linked R95A	0.038	367.48	32.46	7.74	0.29	2.3E-03	8.5E-05	0.060	0.002	1.6E+02	15.58
linked R29A	0.042	539.46	61.76	28.22	1.58	8.2E-03	4.6E-04	0.197	0.011	3.7E+02	46.51
linked S135A	1.618	X	X	X	X	X	X	X	X	X	X
linked R95A (purified at 4°C)	0.045	596.41	30.12	15.17	0.39	4.4E-03	1.1E-04	0.098	0.003	1.6E+02	9.33

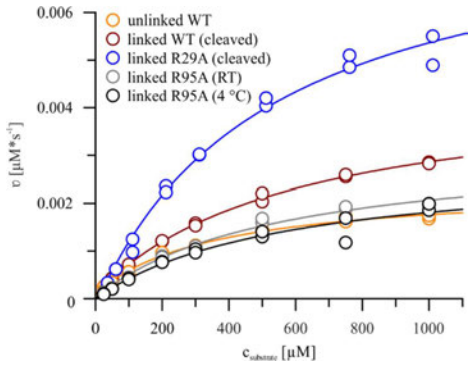
$$error\left(\frac{k_{cat}}{K_M}\right) = \sqrt{\left(\left(\frac{\sigma(k_{cat})}{k_{cat}}\right)^2 + \left(\frac{\sigma(K_M)}{K_M}\right)^2\right)} * \frac{k_{cat}}{K_M} \qquad \text{(S1)}$$



Supplementary Figure 1: ***Comparison of Western-Blot and Coomassie stained SDS-PAGE gels.*** **A:** α-His Western-Blot against the His_6-tagged ZIKV $NS2B_{CF}$-$NS3_{pro}$ after the first purification step (10 µg of Ni-NITA column eluated proteins was loaded). Note the different localization of the His-tag depending on the sample used (compare Fig. 1A). **B:** Ponceau-S stained membrane after blotting shows that $NS2B_{CF}$ and $NS3_{pro}$ cannot be transferred using the same blotting conditions, therefore precluding the use of Western Blots for a densitometric analysis. Instead, a Coomassie brilliant-blue stain based densitometric analysis of autocatalytic cleavage throughout the purification process as shown in Fig. 1D was carried out. **C:** Coomassie brilliant-blue stained control SDS-PAGE gel which was run in parallel to the SDS-PAGE for the Western Blot. **D:** SDS-PAGE (as shown in Fig. 1C) of purified $NS2B_{CF}$-$NS3_{pro}$ constructs (after SEC) for size comparison.



Supplementary Figure 2: ***Tryptophan fluorescence measurements.*** **A:** Crystal structure of ZIKV NS2B$_{CF}$-NS3$_{pro}$ (PDB: 5LC0) showing the three-dimensional arrangement of NS3$_{pro}$ (grey) with its co-factor NS2B$_{CF}$ (blue). The five native tryptophan residues are highlighted in gold. **B:** Representative tryptophan fluorescence emission spectra at 40 °C (grey) and 60 °C (color coded) corresponding to melting curves shown in Fig. 2B displaying differences in bathochromic shift between the different constructs in tryptophan fluorescence spectra. The red and blue vertical lines visualize the emission maximum wavelengths of the linked/uncleaved and unlinked/cleaved constructs at 60 °C, respectively. **C:** Melting curves of linked ZIKV NS2B$_{CF}$-NS3$_{pro}$ R95A purified at RT (left) or 4 °C (right).



Supplementary Figure 3: ***Michaelis-Menten kinetics of ZIKV NS2B$_{CF}$-NS3$_{pro}$ constructs with Boc-GRR-AMC substrate.*** Data points were fitted with Michaelis-Menten equation (S2) using the program GraFit [Version 5.0.13, Erithacus software Limited]. For each ZIKV NS2B$_{CF}$-NS3$_{pro}$ construct data from two independent measurements are shown.

$$v = \frac{v_{max} * c_{substrate}}{K_m + c_{substrate}} \qquad \text{(S2)}$$

“

5.2. Insights into the autocleavage behavior of the ZIKV NS2B/NS3 protease

5.2.1. Context, Project Summary, and Own Contribution

As stated earlier, a shortened construct consisting of the hydrophilic cofactor region of NS2B ($NS2B_{cf}$, amino acids 48–95) is sufficient for the enzymatic activity of NS2B/NS3. A widely applied minimalistic construct, gZiPro, is commonly used in studies for inhibitor development.[47,56,194,200,259–261] This glycine linker, however, was reported to influence the conformational equilibrium of NS2B/NS3, shifting it towards the open conformation,[53,199] thereby reducing catalytic activity.[53,200,250] Bivalently expressed constructs of ZIKV NS2B/NS3 (bZiPro)[47] or enzymatically cleavable linked constructs (eZiPro)[41] do better resemble the native NS2B/NS3. Consequently, the suitability of those constructs for mechanistic studies and inhibitor development was emphasized.[53] To generate eZiPro, the five native C-terminal amino acids of NS2B (KTGKR; amino acids 126–130), representing the native autocleavage site of NS2B/NS3, were introduced to replace the glycine-rich linker, connecting $NS2B_{cf}$ (amino acids 48–96) to the $NS3_{pro}$ domain.[41] Similar constructs were generated with the sequence of DENV4. Thereby, $NS2B_{cf}$ (amino acids 47–96) was tethered to the full-length NS3 by the C-terminal pentapeptide of NS2B (VKTQR, 126–130) as an enzymatically cleavable linker (compare Figure 15).[199] A crystal structure of eZiPro was captured after autocleavage, with the TGKR sequence still residing in the binding pocket of the protease in the closed conformation (PDB-ID: 5GJ4).[41] The prime site of the NS3 binding pocket was unoccupied, indicating that the newly generated NS3 N-terminus was released.[41] This crystal structure indicates *cis* cleavage of the more native linker as well (compare *5.1* Cis *autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3pro constructs*)[41,250] and demonstrates that the C-terminus of NS2B still exhibits high affinity to the ctalytic site and hence could possibly act as inhibitor of the protease. This could be hinting to product-inhibition or autoinhibition mechanisms of high biological relevance, keeping in mind that proteases with largely unselective cleavage after basic residues could damage the host cell and thereby reduce the replication efficiency of the virus. However, there seems to be a distortion by the direct linkage of the TGKR motive to $NS2B_{cf}$ since the P3 Gly and P4 Thr residues do not interact with the $NS3_{pro}$ domain (Figure 3B).[41] In this study, we investigated the autocatalytic behavior of all permutations from basic residues to Ala mutations within the autocatalytic site R95*-EKTGK-R101* of eZiPro by enzymatic activity assays, SDS-PAGE, and protein mass spectroscopy.

In this work, we found the linker integrity and mutations to influence enzyme-kinetic parameters dramatically, as shown by fluorometric enzyme activity assays with the freshly purified enzyme samples. Further, high promiscuity of the protease regarding the sequence of its autocleavage site was observed in protein MS. If mutated, cleavage occurred even after a P1 Ala residue and a new cleavage site at Lys14 in NS3 gained relevance. These compensating mechanisms underline the biological function of autocatalytic cleavage.

For publication of the results, some additional experiments need to be performed. By incubating all 16 catalytically competent mutants and their respective S135A mutants, it can be evaluated whether observed cleavage patterns can result from bacterial enzymes left as impurities in the purified protein samples. If not, the optimal incubation time of the 16 catalytically competent mutants to fully detach NS2B from NS3 can be determined, and subsequent enzymatic activity assays can determine the effects of cleaved linkers on enzyme kinetic parameters. The determined optimal incubation times can be used for further protein MS experiments to achieve a higher resolution for the minor protein species.

(7) **Hammerschmidt, S. J.**; [redacted] Promiscuity in the autocleavage behavior of the Zika virus NS2B/NS3 protease. *Unpublished and not yet submitted manuscript.*

Own contribution: Protein expression and purification, site-directed mutagenesis, SDS-PAGE, densitometric analysis, enzyme activity assay, writing of the original draft & editing of the manuscript.

Contributions from other authors: Protein mass spectroscopy, editing of the manuscript.

5.2.2. Manuscript

The following manuscript, from page 295 to page 310, contains unpublished data.

Promiscuity in the autocleavage behavior of the Zika virus NS2B/NS3 protease

Stefan J. Hammerschmidt[a], [redacted]

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

[b] Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University Jena, Humboldtstrasse 10, 07743 Jena, Germany

[c] Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University, Max von Laue Str. 9, 60438 Frankfurt am Main, Germany

KEYWORDS: Zika Virus, NS2B/NS3 protease, autocatalysis, flavivirus, protein mass spectroscopy

* E-Mail: [redacted]

ABSTRACT: Zika virus replication depends on its heterodimeric serine protease consisting of the NS3 protease domain and the cofactor region of NS2B. For in vitro studies, an uncleavable linked construct is commonly used, although there is a native autocatalysis site between NS2B and NS3. Since this linker was reported to influence the conformational equilibrium and activity of NS2B/NS3, an enzymatically cleavable construct eZiPro, was designed, harboring the native autocleavage site. In this work, we investigate the autocatalytic cleavage of eZiPro to gain information on the biological role of this autocatalysis site by site-directed mutagenesis, SDS-PAGE densitometry, and protein mass spectrometry. Thereby, we report the influence of linker integrity on enzyme kinetic parameters and the high promiscuity of ZIKV NS2B/NS3 regarding cleavage sites in the NS2B/NS3 junction, compensating for introduced mutations.

INTRODUCTION

Zika fever is a tropical disease caused by the eponymous Zika virus (ZIKV), transmitted by mosquitoes of *Aedes aegypti* and *Aedes albopictus* as its main vectors.[1,2] After the ZIKV outbreak in 2016, coinciding with the Olympic games in Brazil and due to its correlation with severe neurological disorders, including in fetuses, the WHO declared ZIKV a public health emergency of international concern.[3–6] Originally found in the tropics and subtropics, these vectors are increasingly found in temperate climate zones due to globalization, international travel, and global warming.[7] Consequently, autochthonous ZIKV infections are rarely but regularly reported in Europe.[8–10] To date, no specific treatment for ZIKV infections is available, limiting the possibilities to merely symptomatic therapies and prevention strategies. On the bright side, several vaccines against ZIKV are currently in clinical trials.[11,12]

ZIKV is a single-stranded (+) sense RNA virus of the Flaviviridae family.[13,14] During the infection process, the viral ~11,000 base genome is translated into a single ~3000 amino acid (aa) precursor polyprotein integrated into the host's endoplasmic reticulum (ER) membrane (Figure 1A). Virus maturation requires proper polyprotein processing. There are cleavage sites exclusively for catalysis by the viral NS3 serine protease ($NS3_{pro}$) and others for processing by host proteases. If the enzymatic activity of $NS3_{pro}$ is suppressed, viral replication is attenuated, qualifying $NS3_{pro}$ as a promising drug target.[15–17] For proper folding and catalytic activity, $NS3_{pro}$ depends on the NS2B cofactor.[16,18] Based on the relative position of $NS2B_{cf}$ to $NS3_{pro}$, the protease can adopt at least two interconverting conformations.[19] In the catalytically competent closed conformation, $NS2B_{cf}$ is tightly wrapped around $NS3_{pro}$. It forms a β-hairpin involved in substrate recognition by contributing $Ser_{81}^{\#}$ and $Asp_{83}^{\#}$ ('#' indicates residues in NS2B) to form parts of the S2 pocket.[20] In the inactive open conformation, the C-terminus of $NS2B_{cf}$ is only loosely bound to NS3 and rather disordered.[21]

ZIKV NS2B/NS3 is a serine protease with His_{51}, Ala_{75}, and Ser_{135} forming the catalytic triad.[22] The substrate preference is for basic residues whereby Arg is preferred over Lys in P1 and Lys over Arg in P2.[23,24] In the S3 pocket, Lys is strongly favored, followed by Arg and Gly, which are also accepted in natural substrates.[24] The S4 pocket is less selective but slightly favoring Arg and Lys over other residues. In the P1' position, small aa like Gly and Ser can be found.[25] Interestingly, an autocatalytic cleavage site between NS2B and NS3 with the sequence KTGKR↓S is present, whose biological relevance remains unclear.[26] However, this site is highly conserved among flaviviruses, indicating biological functions.

Importantly, $NS3_{pro}$ enzymatic activity is also present in truncated constructs retaining only the hydrophilic cofactor region of NS2B ($NS2B_{cf}$, aa 48–95). In a minimalistic construct commonly used in studies for inhibitor development, $NS2B_{cf}$ is tethered to the N-terminal $NS3_{pro}$ domain (amino acids 1–186) by a flexible glycine-rich linker (gZiPro, Figure 1B).[27–33] Although its widespread utilization in research, the presence of this linker was reported to influence the conformational equilibrium of NS2B/NS3 towards the open conformation,[34,35] thereby reducing catalytic activity.[33,34,36] Consequently, some studies advised rather using bivalently expressed constructs of unlinked ZIKV NS2B and NS3 (bZiPro)[28] or enzymatic cleavable linked constructs (eZiPro)[37] as they better resemble the native NS2B/NS3.[34] To generate eZiPro, the five native C-terminal amino acids of NS2B (KTGKR; amino acids 126–130) replace the glycine-rich linker, connecting the $NS2B_{cf}$ region (amino acids 48–96) to the $NS3_{pro}$ domain (Figure 1B).[37] For ZIKV NS2B/NS3 and several closely related flaviviral NS2B/NS3 proteases, glycine-linked constructs were repeatedly reported to undergo autocatalytic cleavage.[22,37–40] To overcome this issue, the two Arg residues, R_{95}# and R_{29}, located in regions adhering to the substrate preferences[23,24] of NS2B/NS3 were mutated to Ala. This measure successfully abolished autocatalytic cleavage and enabled studies with homogenous samples of linked ZIKV NS2B/NS3.[19,33,41] Previously, we showed autocatalytic cleavage of gZiPro exclusively following a cis mechanism (one protease cleaves its own linker).[36] We further demonstrated that only the residue R_{95}# is relevant for the autocatalytic cleavage of this construct.[36] This finding is surprising since R_{95}# is embedded in a PPMR↓G site that should not have considerable cleavage efficiency based on studies using oligopeptides as substrates.[24] This indicates particular promiscuity to substrates when located in the NS2B/NS3 junction, highlighting the biological relevance of this cleavage site. The fact that it is highly conserved among flaviviruses supports this hypothesis.[40,42,43] For the closely related yellow fever virus (YFV) and dengue virus (DENV), cis-cleavage at the NS2B/NS3 junction occurs early in polyprotein processing.[22,38] Remarkably, to date, no function of this site has been reported. Therefore, it is essential to understand the mechanism behind this autocatalytic cleavage and its implications on enzyme activity.

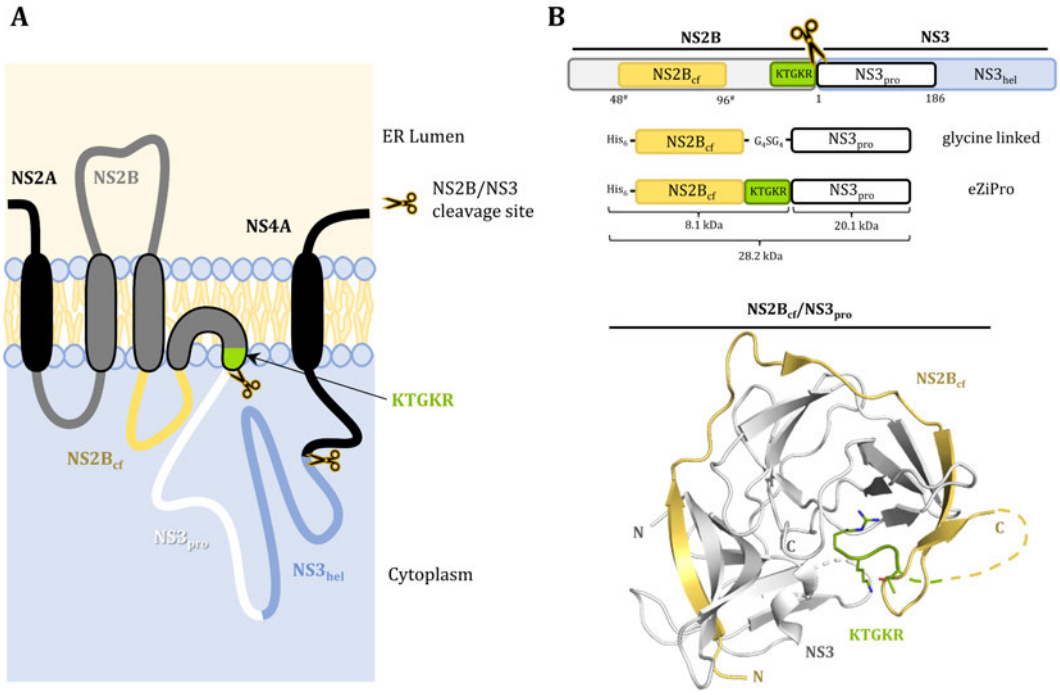


Figure 1. *The expression construct eZiPro.* (**A**) Schematic depiction of NS2B/NS3 as part of the flaviviral precursor polypeptide located at the membrane of the ER. Highlighted are $NS2B_{cf}$ (yellow), the NS2B/NS3 cleavage site (green), $NS3_{pro}$ (white), and $NS3_{hel}$ (blue). The cleavage sites catalyzed by NS2B/NS3 are indicated with yellow scissors. The Graphic was designed following Nitsche et al.[15] (**B**) Top: Linear depiction of full-size NS2B/NS3 and the expression constructs of $NS2B_{cf}$, linked to $NS3_{pro}$ either *via* a glycine-rich linker (gZiPro) or by the KTGKR sequence of the C-terminus of NS2B, yielding the enzymatically cleavable linked construct eZiPro. The molecular sizes of the expected cleavage products of eZiPro are given. Bottom: In the crystal structure of eZiPro (PDB-ID: 5GJ4)[37], only the TGKR residues are resolved, residing in the catalytic site.

This study used eZiPro as a proxy for the native protease. To test for promiscuous cleavage of its $_{94}$#MREKTGKR$_{101}$# site, we mutated the four basic residues within the autocatalytic site to Ala, including all permutations. Then, we analyzed the cleavage patterns and evaluated enzyme kinetics by enzymatic activity assays, densitometric SDS-PAGE analysis, and protein mass spectrometry.

RESULTS and DISCUSSION

Viral proteases with largely unselective cleavage activity could damage the host cell and thereby reduce the replication efficiency of the virus. Hence, one can expect their catalytic activity to be highly regulated. Strategies to achieve this can be high substrate specificity, allosteric sites that are either activated by cofactors or inhibited by reaction products, the necessity of activation of zymogens, auto degradation, or product inhibition. Neither is the ZIKV NS2B/NS3 protease very substrate specific but cleaves unselectively after basic residues,[23,24,36] nor are there reports of other regulatory mechanisms. ZIKV NS2B/NS3 features an allosteric site that can be addressed by small molecules.[44,45] However, no biological role of this site as a regulatory unit addressed by natural ligands has been reported so far. The only autocatalytic activity of ZIKV NS2B/NS3 observed is regarding the NS2B/NS3 junction and the cleavage site between NS3$_{pro}$ and NS3$_{hel}$. The presence or absence of NS3$_{hel}$ did not affect the catalytic activity of DENV proteases but regulates ZIKV NS3$_{hel}$ binding to RNA duplexes.[35,46] Probably due to binding to the active cleft even after cleavage, the NS2B/NS3$_{pro}$ cleavage site KTGKR did reduce the enzymatic activity of eZiPro compared to unlinked constructs and gZiPro.[37]

Densitometric determination of cleavage efficiencies. For densitometric analysis of Coomassie brilliant blue-stained SDS-PAGE, the color densities of the protein bands with molecular sizes <25 kDa were defined as cleavage products and set in relation to the uncut bands with molecular sizes in the range of 25–37 kDa. Upon partial cleavage of the autocatalytic site, three bands would be expected at molecular weights of 27.5–28.5 kDa for the intact eZiPro (NS2B$_{cf}$/NS3$_{pro}$), 20–21 kDa for NS3$_{pro}$ and 7.8–8.3 kDa for NS2B$_{cf}$. No band below 10 kDa corresponding to NS2B$_{cf}$ was resolved in SDS-PAGE. Nonetheless, a rough estimation of the cleavage efficiency could be performed.

In SDS-PAGE, complete cleavage of eZiPro wild-type (WT, RKKR, construct **1**) was observed (Figure S1A). At least three bands in the 10–20 kDa range appeared that could not be attributed to distinct basic residues as cleavage sites. To identify these cleavage products, we first wanted to abolish the autocatalytic cleavage of construct **1**. Based on the preference of ZIKV NS2B/NS3 for Arg in P1, we generated individual Ala mutants for the two Arg residues R$_{95}$# and R$_{101}$# in the $_{94}$#MREKTGKR$_{101}$# site as well as the corresponding double mutant (constructs **2**–**4**).

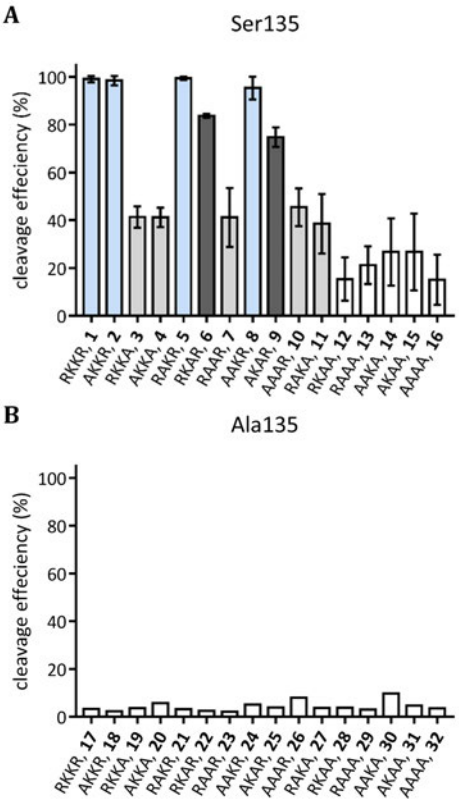


Figure 2. *Densitometric analysis of SDS-PAGE.* (**A**) Cleavage efficiencies (Proportion of the optical density of the bands of fragments of NS2B/NS3 (<25 kDa) to the total density of the bands of the respective lane.) of enzymatically competent eZiPro constructs. Standard deviations of at least triplicates are given as vertical bars. (**B**) The low cleavage efficiencies of active site S135A mutated eZiPro constructs represent low densities at lower molecular weights. Results of single measurements are given. Representative SDS-PAGE can be found in Figure S3A, B. Figure created with GraphPad Prism.[47]

To assess the amount of autocatalytic cleavage during enzyme preparation, densitometric analysis of Coomassie brilliant blue-stained SDS-PAGE was employed. Enzyme purification and subsequent SDS-PAGE with the freshly purified proteases showed construct **1** and the R$_{95}$#A (AKKR, construct **2**) mutant to be fully cleaved. In contrast, R$_{101}$#A (RKKA, construct **3**) and R$_{95}$#A–R$_{101}$#A (AKKA, construct **4**) exhibited a ~40% reduced cleavage efficiency (Figure 3A). Thus, R$_{95}$# seems irrelevant for autocatalytic cleavage of eZiPro which is surprising since it is the only residue responsible for the autocatalytic cleavage gZiPro.[36] This can be either due to a different relative position of R$_{95}$# in eZiPro or the extension of E$_{96}$# since ZIKV NS2B/NS3 prefers small residues in P1. Because still ~40% of **4** was cleaved during enzyme purification, we subsequently mutated the lysine residues of eZiPro to Ala (constructs

5–**7**), including all other Ala permutations (constructs **8**–**16**). Relative cleavage efficiencies of the freshly prepared enzymes were assessed (Figure 3A, Figure S1A).All constructs were also cloned to harbor an $S_{135}A$ mutation (constructs **17**–**32**), abolishing enzymatic activity to exclude that other bacterial enzymes could be accountable for the observed cleavage products (Figure 3B). Indeed, all constructs did not undergo cleavage (<10%). The bands with molecular weights <25 kDa probably reflect impurities.

Almost complete cleavage (>95%) was observed for constructs **1**, **2**, **5** (RAKR), and **8** (AAKR). Since all other constructs lack the two consecutive basic residues K_{100}# and R_{101}#, they represent the most favored cleavage site. The constructs **6** (RKAR) and **9** (AKAR) yield 75–85% cleavage, indicating that R_{101}# is still the favored residue for cleavage as long as at least a Lys is present in P4. All other constructs are auto-processed from 15% to 45%. Even construct **16** (AAAA) seems cleaved, although no basic residues remain in the autocatalytic site. This indicates that the local high linker concentration boosts cleavage efficiency and increases promiscuity.

Enzyme kinetics. To correlate linker integrity (equals 100% minus cleavage efficiency) to enzyme kinetic parameters, fluorometric enzyme activity assays were conducted for all freshly prepared constructs with an intact catalytic triad (Figure S1, Table 1). The Michaelis Menten constant (K_M) is a proxy for substrate affinity to the respective eZiPro construct, and k_{cat} is the turnover rate of substrates. Hence, k_{cat}/K_M can be used to compare the enzymatic activity of different constructs. eZiPro WT (construct **1**) revealed a K_M value of the substrate Boc-GRR-AMC of 205 ± 22 µM, a k_{cat} of $1.92 \cdot 10^{-2} \pm 7 \cdot 10^{-4}$ s^{-1}, resulting in a k_{cat}/K_M of 93.5 ± 10.6 $M^{-1} \cdot s^{-1}$. This enzymatic activity is lower than those of the unlinked and the uncleavable gZiPro R_{95}#A.[36] This finding is in line with the observations of Phoo *et al.* hypothesizing, that after autocatalytic cleavage, the C-terminal end 'TGKR' of $NS2B_{cf}$ decreases the catalytic activity of NS3 in vitro by competing with the substrate.[37]

The k_{cat}/K_M values of the different constructs differed drastically between 28.4 M^{-1} s^{-1} to 2030 M^{-1} s^{-1}. This observation is also reflected by K_M values between 48.7 µM to 1199 µM and k_{cat} values between 0.0109 s^{-1} and 0.186 s^{-1}. The lowest enzymatic activities were found for constructs **12**, **13**, **15**, and **16**, showing the lowest cleavage efficiency. Since cis-autocatalytic cleavage at the NS2B/NS3 junction occurs early in polyprotein processing of YFV and DENV,[22,38] this indicates that cleavage between NS2B and NS3 could be essential to increase catalytic activity. On the other hand, this hypothesis is contradicted by constructs **3**, **4**, and **7** with the highest enzymatic activities (k_{cat}/K_M: 1367 ± 275, 2030 ± 316, and 925 ± 59 M^{-1} s^{-1}, respectively) and only ~41% cleavage efficiency. Although no general trend is observable in kinetics vs. cleavage efficiency plots (Figure S2), these tremendous differences in enzymatic activities of the constructs demonstrate pronounced effects of linker-integrity and hint at possible autoregulatory mechanisms.

Table 1. Enzyme kinetics of the catalytic competent 16 auto-cleavage site permutations.

Nr.	eZiPro constructs[a]	K_M (µM)	v_{max} (nM·s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1} \cdot s^{-1}$)	Cleavage efficiency (%)
1	RKKR	205 ± 22	1.48 ± 0.05	0.0192 ± 0.0007	93.5 ± 10.6	99.1 ± 1.1
2	AKKR	368 ± 42	2.54 ± 0.12	0.0348 ± 0.0016	94.5 ± 11.7	98.5 ± 1.7
3	RKKA	60.7 ± 11.5	3.86 ± 0.26	0.0830 ± 0.0057	1367 ± 275	41.4 ± 3.9
4	AKKA	48.7 ± 7.2	4.24 ± 0.20	0.0989 ± 0.0048	2030 ± 316	41.2 ± 2.9
5	RAKR	133 ± 10	2.11 ± 0.04	0.0349 ± 0.0007	262 ± 20	99.4 ± 0.6
6	RKAR	1112 ± 75	12.4 ± 0.53	0.186 ± 0.008	168 ± 13	83.7 ± 0.7
7	RAAR	93.0 ± 5.6	5.43 ± 0.12	0.0860 ± 0.0019	925 ± 59	41.2 ± 10.1
8	AAKR	1199 ± 161	12.4 ± 0.53	0.186 ± 0.008	155 ± 22	95.4 ± 3.4
9	AKAR	372 ± 29	5.23 ± 0.18	0.0861 ± 0.0029	231 ± 20	74.8 ± 2.9
10	AAAR	386 ± 16	3.92 ± 0.09	0.0645 ± 0.0014	167 ± 8	45.4 ± 6.4
11	RAKA	501 ± 33	2.56 ± 0.08	0.0399 ± 0.0013	79.7 ± 5.8	38.5 ± 8.77
12	RKAA	334 ± 30	0.763 ± 0.03	0.0109 ± 0.0004	32.7 ± 3.2	15.4 ± 7.4
13	RAAA	385 ± 36	0.596 ± 0.03	0.0109 ± 0.005	28.4 ± 2.9	21.2 ± 6.8
14	AAKA	378 ± 21	1.81 ± 0.04	0.0276 ± 0.0007	73.1 ± 4.4	26.7 ± 11.5
15	AKAA	552 ± 42	1.99 ± 0.08	0.0296 ± 0.0012	53.7 ± 4.6	26.7 ± 13.1
16	AAAA	335 ± 28	0.561 ± 0.02	0.00974 ± 0.00036	29.1 ± 2.7	15.1 ± 9.1

a: The eZiPro constructs are denoted by the basic residues of the $_{95}$#REKTGKR$_{101}$# cleavage site (R_{95}#K_{97}#K_{100}#R_{101}#), or the respective Ala mutants are shown.

Incubation experiments. To determine the necessary incubation time to achieve (almost) fully processed proteases, the slow degrading construct **15** (AKAA) was incubated directly after purification in assay buffer at 37 °C under agitation. Samples were drawn at different incubation times, denatured in SDS-Loading buffer for 5 min at 95 °C, and frozen until loaded on an SDS-PAGE (Figure 4A). Interestingly, not only two cleavage products corresponding to $NS3_{pro}$ (20–21 kDa) and $NS2B_{cf}$ (7.8–8.3 kDa) can be found, but several degradation products of $NS3_{pro}$ in the range of 12–20 kDa (Figure 4A, B). The band of the intact eZiPro (red) vanishes over time. The band at ~20 kDa (pale blue) gains intensity but starts to fade after the maximum intensity was reached at 30 h. In line, some smaller bands appear after 10 h (green, pale green), indicating further degradation of $NS3_{pro}$. After 200 h (Figure 4C), eZiPro is nearly fully cleaved. Therefore, 200 h was chosen as the incubation time for further experiments. SDS-PAGE of all enzymatically competent constructs revealed that all constructs except **13** (RAAA) and **16** (AAAA) were fully cleaved (Figure S1C).

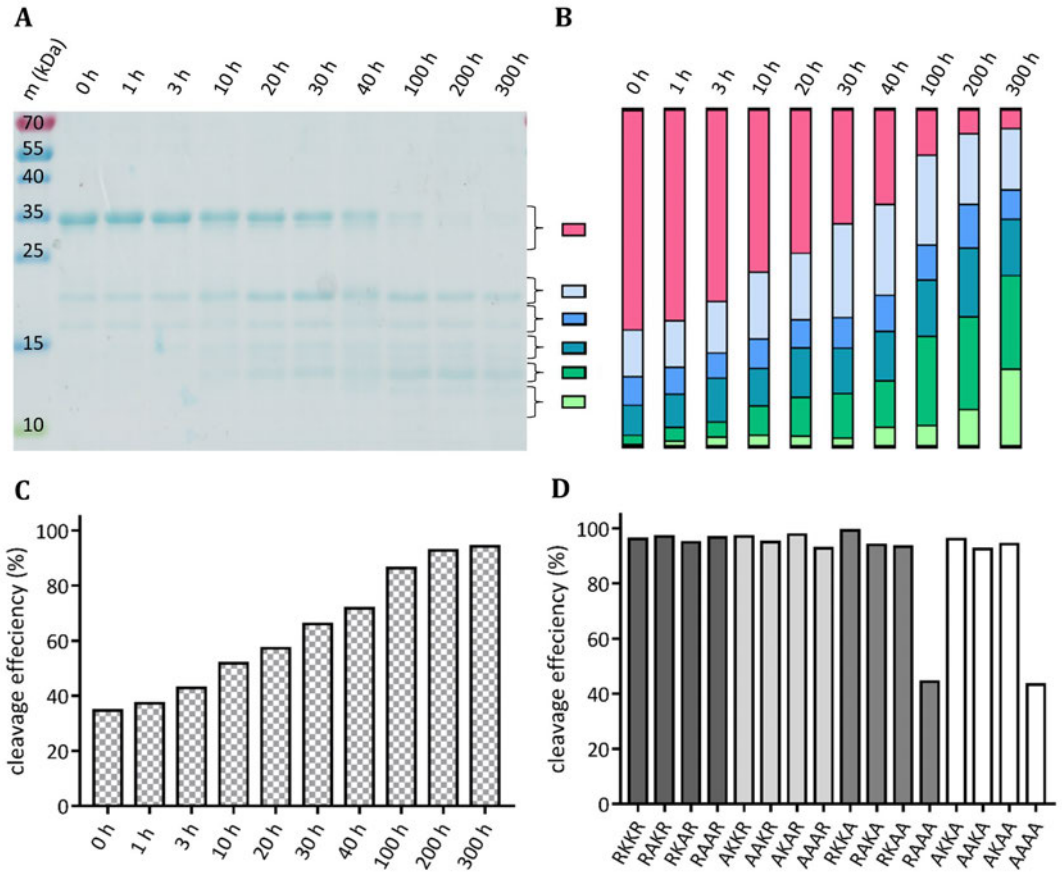


Figure 3. *Incubation of AKAA.* (**A**) Coomassie brilliant blue stained SDS-PAGE of samples drawn at different time points of incubation of AKAA. (**B**) Densitometric analysis of **A**. The fraction of eZiPro with an intact linker (pink) declines, whereas cleavage products at lower molecular weight bands (blue to green) occur. (**C**) Densitometric analysis of **A** shows the cleavage efficiency at different incubation times. At 200 h, almost all eZiPro is cleaved. (**D**) In a single experiment with 200 h incubation, all constructs except RAAA and AAAA are fully cleaved (>90%). Figure created with GraphPad Prism.[47]

The occurrence of sequential degradation of $NS3_{pro}$ could yield catalytically inactive products, thereby explaining the low enzymatic activities of the fully cleaved constructs **1** (RKKR), **2** (AKKR), and **5** (RAKR) with cleavage above 98%. Therefore, we excluded those constructs from the analysis plots and found a good correlation between cleavage efficiency and k_{cat} (R^2 = 0.861, Figure 5). This strongly hints at higher turnover rates after autocatalytic linker cleavage. The lack of correlation with K_M and k_{cat}/K_M values could result from different affinities of the intact linker or their cleavage fragments to the catalytic cleft, competing with substrate binding.

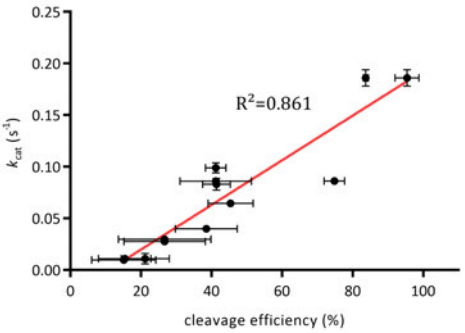


Figure 4. Correlation between linker integrity and k_{cat}. The Turnover rate k_{cat} is plotted against the cleavage efficiency of enzymatic competent autocleavage site permutations. Constructs **1**, **2**, and **5** are excluded from this analysis. Figure created with GraphPad Prism.[47]

A crystal structure (PDB-ID: 5GJ4) was captured after autocleavage with the linker still residing in the catalytic site (Figure 2).[37] This indicates that the protease undergoes autocleavage in a cis mechanism, as we demonstrated for gZiPro.[36,37] In this structure, the P3 Gly and P4 Thr residues are distorted hence not interacting with the $NS3_{pro}$ domain in this crystal structure. This is probably due to the artificially shortened nature of eZiPro, preventing interactions with the S3 and S4 pockets. Hence, in the native construct with an extended linker sequence, the auto-inhibitory effect of this linker after autocleavage could be even higher. This observation hints at possible product inhibition or autoinhibition mechanisms regulating enzymatic activity.

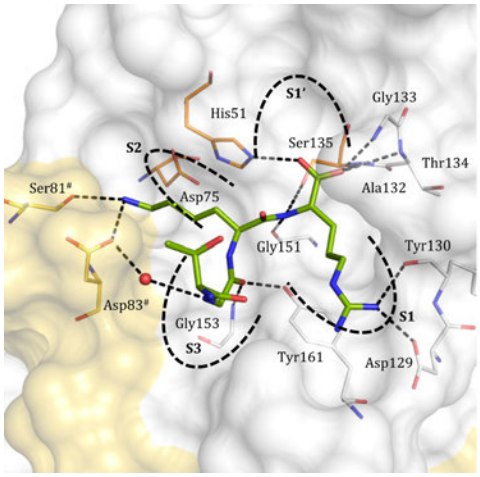


Figure 5. *Binding mode of the TGKR-fragment of the autocleavage site of the ZIKV NS2B/NS3 protease.* The active site of native linked ZIKV NS2B/NS3 (eZiPro) in the closed conformation after self-cleavage (PDB-ID: 5GJ4).[37] The newly generated C-terminus TGKR (green sticks) is still bound to the catalytic site. NS3 is depicted as a white surface, and NS2B is colored yellow. Polar interactions are shown as black dashed lines. For a clear view, only those amino acids that form polar interactions are displayed as lines. Amino acids of the catalytic triad are shown as orange sticks and amino acids in NS2B are indicated by '#'. Binding sub-pockets are displayed according to Schechter&Berger.[48] Figure created using PyMOL.[49]

Protein Mass Spectrometry. Since many cleavage products occurred in SDS-PAGE of the freshly purified eZiPro constructs that could not be identified, protein mass spectrometry was used to determine the exact products and their respective cleavage sites. Noteworthy, adducts of the polypeptides were found together with Na (22.0 Da) and as the *N*-gluconoyl-adducts (178.2 Da) of the hexahistidine (His_6) tag harboring species as occasionally observed for *N*-terminally His_6 tagged proteins with a vicinal Ser residue, expressed in *Escherichia coli* (*E. coli*).[50,51] This modification could hence be used to identify fragments of the intact N-terminus. Several findings can be deduced from the obtained MS spectra. (i) Several constructs (**4**, and **11**–**16**) revealed cleavage of the His_6-Tag directly after the sixth His residue. This observation could be of interest for experiments with recombinant NS2B/NS3 proteases relying on the immobilization of the protease *via* its His_6 tag. (ii) Despite R_{95}# cleavage being relevant for the cleavage of glycine-linked constructs, no construct revealed cleavage at this residue. These results are in accordance with the cleavage efficiencies of constructs **2**–**4** and could be due to E_{96}# that is special too demanding to be placed in the S1' pocket. (iii) For constructs **1**, **2**, **5**, and **8** with intact K_{100}# and R_{101}#, no other cleavage products were detected in protein-MS without further incubation, highlighting the cleavage preference for these residues. (iv) When K_{100}# or R_{101}# are mutated to Ala (K_{100}#A: **6**, **7**, **9**, and **10**; R_{101}#A: **3**, **4**,

and **11**), cleavage still occurs after residue 101#. This is surprising since Lys should be preferred in P1 over Ala.[23,24] Furthermore an additional cleavage occurs at K_{14} in $NS3_{pro}$, embedded in a $_{11}$EVKK↓G_{15} site. (v) Promiscuity of the cleavage site preference in the NS2B/NS3 junction is highlighted by constructs **4**, **11,** and **15**, where cleavage occurred at $_{95}$#AEKT↓GKA_{101}#, $_{95}$#REATGK↓A_{101}# and $_{95}$#AEKT↓GAA_{101}#, respectively. (vi) Construct **16** did not reveal cleavage at the autocleavage site but at the His_6 tag and K_{14}.

CONCLUSION

In this work, we investigated the autocleavage behavior of $NS2B_{cf}/NS3_{pro}$ by analyzing eZiPro with various linker mutations. Thereby enzyme kinetic parameters differed highly among the freshly prepared constructs. A correlation of linker integrity with enzymatic turnover rates k_{cat} was found, indicating that cleavage of this linker leads to higher enzymatic activities required for further polyprotein processing. This is in line with studies that showed autocatalytic cleavage of NS2B/NS3 occurs early in polyprotein-processing of the closely related YFV and DENV.[22,38] The cleavage of this linker seems to be driven by its position near the active site cleft and the resulting high local concentrations yielding high promiscuity of the autocatalysis regarding the linker sequence. If the most favored residue, R_{101}#, is mutated, cleavage still occurs after A_{101}#, and a new cleavage site in $NS3_{pro}$ ($_{11}$EVKK↓G_{15}) compensates for these mutations. In the densitometric analysis of SDS-PAGE and protein-MS, R_{95}# proved irrelevant for autocatalytic cleavage, contrasting earlier observations for gZiPro, probably due to the addition of E_{96}#, which is too large to be placed in the S1' pocket. Furthermore, we observed successive degradation of $NS3_{pro}$ upon extended incubation periods that could serve as a self-limiting mechanism. With this work, we contribute to a more detailed understanding of the autocatalytic cleavage behavior of ZIKV NS2B/NS3 that could also be transferred to other flaviviral proteases featuring this highly conserved cleavage site.

EXPERIMENTAL SECTION

Reagents. If not stated elsewhere, all chemicals were purchased from Sigma-Aldrich Chemie GmbH (Munich, Germany), Thermo Fisher Scientific Inc. (Waltham, MA; USA), and Carl Roth GmbH + Co. KG (Karlsruhe, Germany). LB-media was purchased from ForMedium Ltd. (Swaffham, UK). Primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA, USA).

Expression constructs and Site-directed Mutagenesis. The generation of the 16 permutations of basic residues in the cleavage site of ZIKV NS2B/NS3 and their respective active site S_{135}A mutants was performed based on a pET-11a vector harboring the sequence of the French Polynesia ZIKV strain (GenBank ID: KJ776791.2, $NS2B_{cf}$: amino acids 1423–1467 and $NS3_{pro}$: amino acids 1503–1688)[52] as used before[44]. This construct was mutated to replace the glycine-rich linker with a Glu (amino acid 1468) and the native cleavage site KTGKR of NS2B/NS3 (aa 1498–1502) to resemble the previously reported eZiPro denoted construct.[37] The construct contains an N-terminal hexahistidine (His_6) tag for purification, followed by a tobacco etch virus protease (TEV)-cleavage site. The mutants were generated with successive site-directed mutagenesis using the Kapa HiFi PCR kit (Kapa Biosystems Inc. Woburn, MA, USA). Since the mutated basic residues are in the vicinity, the primers for one mutation must be designed multiple times to match the already introduced mutations (Table S1).

Recombinant Protein Expression and Preparation. The enzymatically cleavable linked ZIKV NS2B/NS3 protease mutants (eZiPro) were expressed and purified as described previously.[53,54] Briefly, the pET-11a vector was transformed into competent *Escherichia coli* (*E.* coli) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). Cultures were grown in LB medium containing 100 mg·L^{-1} ampicillin at 37 °C until an optical density (OD_{600}) of ~0.8 was reached. Expression was induced by adding 1 mM isopropyl-β-D-thiogalactoside (IPTG) and conducted at 18 °C for ~20 h. Cell pellets were harvested by centrifugation at 9000 rpm at 4 °C for 15 min. After resuspending 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), cells were lysed by 10 cycles à 45 s of sonication at 60% power (Sonoplus HD 2200; Bandelin, Berlin, Germany). Centrifugation at 20,000 rpm at 4 °C for 1 h cleared the crude lysate. The His_6 tagged NS2B/NS3 was isolated from the supernatant 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). A linear gradient of wash buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole) and elution buffer with 250 mM imidazole was applied. The eluted fractions containing NS2B/NS3 were pooled, 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 (20 mM Tris-HCl, pH 8.0, 150 mM NaCl). Fractions containing NS2B/NS3 were pooled, concentrated, flash-frozen in liquid nitrogen, and stored at –80 °C until further use. SDS-PAGE confirmed the purity and identity of NS2B/NS3 with Coomassie blue staining and subsequent fluorometric enzyme activity assays.

Fluorometric Enzyme Activity Assay. Fluorometric enzyme activity assays were performed as described previously.[54] Briefly, each well of white flat-bottom 96-well plates (Greiner bio-one, Kremsmünster, Austria) contained 200 µL reaction mixture composed of 5 µL of NS2B/NS3 (final concentration 50 nM) in assay buffer (50 mM Tris-HCl pH 9.0, 1 mM CHAPS) with 5 µL substrate (final concentration 100 µM Boc-GRR-AMC, Bachem, Bubendorf, Switzerland) in DMSO and 10 µL of DMSO. The fluorescence of released AMC was recorded at λ_{ex} = 380 nm and λ_{em} = 460 nm wavelengths in 30 s increments for 10 min on a Spark 10M plate reader (Tecan, Männerdorf, Switzerland).

***K*M determination.** The Michaelis-Menten constants K_M of the Boc-GRR-AMC substrate to NS2B/NS3 was determined using eight final substrate concentrations (1000, 800, 600, 400, 300, 200, 100, 50 µM, Figure S1). K_M was calculated using GRAFIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, West Sussex, UK)[55] by fitting to the Michaelis-Menten equation (1)

$$v = \frac{V_{max} \cdot [S]}{K_M + [S]} \quad (1)$$

with v (ΔF/min) as the substrate hydrolysis rate V_{max} as the maximum slope of the dose-response curve, and the substrate concentration [S]. Kinetic parameters were calculated using a previously determined value of 3422.6 RFU·µM−1 for the AMC fluorescence for this experimental setup.[36]

Densitometric analysis. Densitometric analysis of Coomassie brilliant blue stained SDS-PAGE gels was carried out using the ImageJ software (https://imagej.net/ij/).[56]

Protein Mass Spectrometry Markus

ASSOCIATED CONTENT

SUPPORTING INFORMATION

Supplementary Figures, Supplementary Tables. Additional References.

AUTHOR INFORMATION

Corresponding Author

* To whom correspondence should be addressed. [redacted]

Author Contributions

SJH: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. [redacted] Data curation, Formal analysis, Methodology, Validation, Writing – review & editing. [redacted] Supervision, Writing – review & editing. [redacted]: Conceptualization, Project administration, Supervision, Writing – review & editing.

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Funding Sources

Notes

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.

ACKNOWLEDGMENTS

We thank [redacted] ([redacted]) for cloning the eZiPro constructs **1**–**4**.

ABBREVIATIONS

PDB, RCSB protein data bank; cpd, compound; K_i, inhibition constant; TRIS, tris(hydroxymethyl)aminomethane; HEPES, 4-(2-hydroxyethyl)- 1-piperazineethanesulfonic acid; K_d, dissociation constant; IC_{50}, the concentration at half-maximal inhibition; His_6, hexahistidine; *E. coli, Escherichia coli*; OD_{600}, optical density at λ = 600 nm; IPTG, isopropyl-*β*-D-thiogalactopyranosid; FPLC, fast protein liquid chromatography; DTT, dithiothreitol; IMAC, immobilized metal affinity chromatography; Ni-NTA, nickel nitrilotriacetic acid; CV, column volumes; MWCO, molecular weight cut-off; AEX, anion exchange; K_M Michaelis-Menten constant.

REFERENCES

(1) Hayes, E. B. Zika Virus Outside Africa. *Emerg. Infect. Dis.* **2009**, *15* (9), 1347–1350. https://doi.org/10.3201/eid1509.090442.

(2) Powell, J. R. Mosquitoes on the Move. *Science (80-.).* **2016**, *354* (6315), 971–972. https://doi.org/10.1126/science.aal1717.

(3) Cao-Lormeau, V.-M.; Blake, A.; Mons, S.; Lastère, S.; Roche, C.; Vanhomwegen, J.; Dub, T.; Baudouin, L.; Teissier, A.; Larre, P.; Vial, A.-L.; Decam, C.; Choumet, V.; Halstead, S. K.; Willison, H. J.; Musset, L.; Manuguerra, J.-C.; Despres, P.; Fournier, E.; Mallet, H.-P.; Musso, D.; Fontanet, A.; Neil, J.; Ghawché, F. Guillain-Barré Syndrome Outbreak Associated with Zika Virus Infection in French Polynesia: A Case-Control Study. *Lancet* **2016**, *387* (10027), 1531–1539. https://doi.org/10.1016/S0140-6736(16)00562-6.

(4) 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.

(5) Chan, J. F. W.; Choi, G. K. Y.; Yip, C. C. Y.; Cheng, V. C. C.; Yuen, K.-Y. Zika Fever and Congenital Zika Syndrome: An Unexpected Emerging Arboviral Disease. *J. Infect.* **2016**, *72* (5), 507–524. https://doi.org/10.1016/j.jinf.2016.02.011.

(6) WHO. The History of Zika Virus. **2016**, *7. Februar*.

(7) Angelo, K. M.; Stoney, R. J.; Brun-Cottan, G.; Leder, K.; Grobusch, M. P.; Hochberg, N.; Kuhn, S.; Bottieau, E.; Schlagenhauf, P.; Chen, L.; Hynes, N. A.; Perez, C. P.; Mockenhaupt, F. P.; Molina, I.; Crespillo-Andújar, C.; Malvy, D.; Caumes, E.; Plourde, P.; Shaw, M.; McCarthy, A. E.; Piper-Jenks, N.; Connor, B. A.; Hamer, D. H.; Wilder-Smith, A. Zika among International Travellers Presenting to GeoSentinel Sites, 2012–2019: Implications for Clinical Practice. *J. Travel Med.* **2020**, *27* (4), 1–9. https://doi.org/10.1093/jtm/taaa061.

(8) Giron, S.; Franke, F.; Decoppet, A.; Cadiou, B.; Travaglini, T.; Thirion, L.; Durand, G.; Jeannin, C.; L'Ambert, G.; Grard, G.; Noël, H.; Fournet, N.; Auzet-Caillaud, M.; Zandotti, C.; Aboukaïs, S.; Chaud, P.; Guedj, S.; Hamouda, L.; Naudot, X.; Ovize, A.; Lazarus, C.; de Valk, H.; Paty, M.-C.; Leparc-Goffart, I. Vector-Borne Transmission of Zika Virus in Europe, Southern France, August 2019. *Eurosurveillance* **2019**, *24* (45), 1–4. https://doi.org/10.2807/1560-7917.ES.2019.24.45.1900655.

(9) Nuñez, A. I.; Talavera, S.; Birnberg, L.; Rivas, R.; Pujol, N.; Verdún, M.; Aranda, C.; Berdugo, M.; Busquets, N. Evidence of Zika Virus Horizontal and Vertical Transmission in Aedes Albopictus from Spain but Not Infectious Virus in Saliva of the Progeny. *Emerg. Microbes Infect.* **2020**, *9* (1), 2236–2244. https://doi.org/10.1080/22221751.2020.1830718.

(10) Vermeulen, T. D.; Reimerink, J.; Reusken, C.; Giron, S.; de Vries, P. J. Autochthonous Dengue in Two Dutch Tourists Visiting Département Var, Southern France, July 2020. *Eurosurveillance* **2020**, *25* (39), 13–15. https://doi.org/10.2807/1560-7917.ES.2020.25.39.2001670.

(11) Fernandez, E.; Diamond, M. S. Vaccination Strategies against Zika Virus. *Curr. Opin. Virol.* **2017**, *23*, 59–67. https://doi.org/10.1016/j.coviro.2017.03.006.

(12) Hazlewood, J. E.; Tang, B.; Yan, K.; Rawle, D. J.; Harrison, J. J.; Hall, R. A.; Hobson-Peters, J.; Suhrbier, A. The Chimeric

Binjari-Zika Vaccine Provides Long-Term Protection against ZIKA Virus Challenge. *Vaccines* **2022**, *10* (1), 85. https://doi.org/10.3390/vaccines10010085.

(13) Braack, L.; Gouveia de Almeida, A. P.; Cornel, A. J.; Swanepoel, R.; de Jager, C. Mosquito-Borne Arboviruses of African Origin: Review of Key Viruses and Vectors. *Parasit. Vectors* **2018**, *11* (1), 29. https://doi.org/10.1186/s13071-017-2559-9.

(14) Simmonds, P.; Becher, P.; Bukh, J.; Gould, E. A.; Meyers, G.; Monath, T.; Muerhoff, S.; Pletnev, A.; Rico-Hesse, R.; Smith, D. B.; Stapleton, J. T. ICTV Virus Taxonomy Profile: Flaviviridae. *J. Gen. Virol.* **2017**, *98* (1), 2–3. https://doi.org/10.1099/jgv.0.000672.

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

(16) Lindenbach, B. D.; Thiel, H.-J.; Rice, C. M. Flaviviridae: The Viruses and Their Replication. *Fields Virol.* **2007**, 1101–1151.

(17) Voss, S.; Nitsche, C. Inhibitors of the Zika Virus Protease NS2B-NS3. *Bioorg. Med. Chem. Lett.* **2020**, *30* (5), 126965. https://doi.org/10.1016/j.bmcl.2020.126965.

(18) Falgout, B.; Pethel, M.; Zhang, Y. M.; Lai, C. J. Both Nonstructural Proteins NS2B and NS3 Are Required for the Proteolytic Processing of Dengue Virus Nonstructural Proteins. *J. Virol.* **1991**, *65* (5), 2467–2475. https://doi.org/10.1128/jvi.65.5.2467-2475.1991.

(19) Mahawaththa, M. C.; Pearce, B. J. G.; Szabo, M.; Graham, B.; Klein, C. D.; Nitsche, C.; Otting, G. Solution Conformations of a Linked Construct of the Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *142*, 141–147. https://doi.org/10.1016/j.antiviral.2017.03.011.

(20) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. *Science (80-.).* **2016**, *353* (6298), 503–505. https://doi.org/10.1126/science.aag2419.

(21) 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* (4), 372–373. https://doi.org/10.1038/nsmb1073.

(22) Chambers, T. J.; Weir, R. C.; Grakoui, A.; McCourt, D. W.; Bazan, J. F.; Fletterick, R. J.; Rice, C. M. Evidence That the N-Terminal Domain of Nonstructural Protein NS3 from Yellow Fever Virus Is a Serine Protease Responsible for Site-Specific Cleavages in the Viral Polyprotein. *Proc. Natl. Acad. Sci.* **1990**, *87* (22), 8898–8902. https://doi.org/10.1073/pnas.87.22.8898.

(23) Gruba, N.; Rodriguez Martinez, J. I.; Grzywa, R.; Wysocka, M.; Skoreński, M.; Burmistrz, M.; Łęcka, M.; Lesner, A.; Sieńczyk, M.; Pyrć, K. Substrate Profiling of Zika Virus NS2B-NS3 Protease. *FEBS Lett.* **2016**, *590* (20), 3459–3468. https://doi.org/10.1002/1873-3468.12443.

(24) Rut, W.; Zhang, L.; Kasperkiewicz, P.; Poreba, M.; Hilgenfeld, R.; Drąg, M. Extended Substrate Specificity and First Potent Irreversible Inhibitor/Activity-Based Probe Design for Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *139*, 88–94. https://doi.org/10.1016/j.antiviral.2016.12.018.

(25) Li, J.; Lim, S. P.; Beer, D.; Patel, V.; Wen, D.; Tumanut, C.; Tully, D. C.; Williams, J. A.; Jiricek, J.; Priestle, J. P.; Harris, J. L.; Vasudevan, S. G. Functional Profiling of Recombinant NS3 Proteases from All Four Serotypes of Dengue Virus Using Tetrapeptide and Octapeptide Substrate Libraries. *J. Biol. Chem.* **2005**, *280* (31), 28766–28774. https://doi.org/10.1074/jbc.M500588200.

(26) Behnam, M. A. M.; Nitsche, C.; Boldescu, V.; Klein, C. D. The Medicinal Chemistry of Dengue Virus. *J. Med. Chem.* **2016**, *59* (12), 5622–5649. https://doi.org/10.1021/acs.jmedchem.5b01653.

(27) Huber, S.; Braun, N. J.; Schmacke, L. C.; Quek, J. P.; Murra, R.; Bender, D.; Hildt, E.; Luo, D.; Heine, A.; Steinmetzer, T. Structure-Based Optimization and Characterization of Macrocyclic Zika Virus NS2B-NS3 Protease Inhibitors. *J. Med. Chem.* **2022**, *65* (9), 6555–6572. https://doi.org/10.1021/acs.jmedchem.1c01860.

(28) Phoo, W. W.; Zhang, Z.; Wirawan, M.; Chew, E. J. C.; Chew, A. B. L.; Kouretova, J.; Steinmetzer, T.; Luo, D. Structures of Zika Virus NS2B-NS3 Protease in Complex with Peptidomimetic Inhibitors. *Antiviral Res.* **2018**, *160*, 17–24. https://doi.org/10.1016/j.antiviral.2018.10.006.

(29) Kumar, A.; Liang, B.; Aarthy, M.; Singh, S. K.; Garg, N.; Mysorekar, I. U.; Giri, R. Hydroxychloroquine Inhibits Zika Virus NS2B-NS3 Protease. *ACS Omega* **2018**, *3* (12), 18132–18141. https://doi.org/10.1021/acsomega.8b01002.

(30) Nitsche, C.; Zhang, L.; Weigel, L. F.; Schilz, J.; Graf, D.; Bartenschlager, R.; Hilgenfeld, R.; Klein, C. D. 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.

(31) Yang, Y.; Cao, L.; Gao, H.; Wu, Y.; Wang, Y.; Fang, F.; Lan, T.; Lou, Z.; Rao, Y. Discovery, Optimization, and Target Identification of Novel Potent Broad-Spectrum Antiviral Inhibitors. *J. Med. Chem.* **2019**, *62* (8), 4056–4073. https://doi.org/10.1021/acs.jmedchem.9b00091.

(32) Li, Y.; Phoo, W. W.; Loh, Y. R.; Zhang, Z.; Ng, E. Y.; Wang, W.; Keller, T. H.; Luo, D.; Kang, C. Structural Characterization of the Linked <scp>NS</Scp> 2B- <scp>NS</Scp> 3 Protease of Zika Virus. *FEBS Lett.* **2017**, *591* (15), 2338–2347. https://doi.org/10.1002/1873-3468.12741.

(33) Kuiper, B. D.; Slater, K.; Spellmon, N.; Holcomb, J.; Medapureddy, P.; Muzzarelli, K. M.; Yang, Z.; Ovadia, R.; Amblard, F.; Kovari, I. A.; Schinazi, R. F.; Kovari, L. C. Increased Activity of Unlinked Zika Virus NS2B/NS3 Protease Compared to Linked Zika Virus Protease. *Biochem. Biophys. Res. Commun.* **2017**, *492* (4), 668–673. https://doi.org/10.1016/j.bbrc.2017.03.108.

(34) Hill, M. E.; Yildiz, M.; Hardy, J. A. 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.

(35) 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. https://doi.org/10.1016/j.antiviral.2020.104900.

(36) von Hammerstein, F.; Lauth, L. M.; Hammerschmidt, S.; Wagner, A.; Schirmeister, T.; Hellmich, U. A. 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.

(37) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D. Structure of the NS2B-NS3 Protease from Zika Virus after Self-Cleavage. *Nat. Commun.* **2016**, *7*, 13410. https://doi.org/10.1038/ncomms13410.

(38) Preugschat, F.; Yao, C. W.; Strauss, J. H. 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.

(39) Bera, A. K.; Kuhn, R. J.; Smith, J. L. Functional Characterization of Cis and Trans Activity of the Flavivirus NS2B-NS3 Protease. *J. Biol. Chem.* **2007**, *282* (17), 12883–12892. https://doi.org/10.1074/jbc.M611318200.

(40) Bera, A. K.; Kuhn, R. J.; Smith, J. L. Functional Characterization of Cis and Trans Activity of the Flavivirus NS2B-NS3 Protease. *J. Biol. Chem.* **2007**, *282* (17), 12883–12892. https://doi.org/10.1074/jbc.M611318200.

(41) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. *Science (80-.).* **2016**, *353* (6298), 503–505. https://doi.org/10.1126/science.aag2419.

(42) Falgout, B.; Pethel, M.; Zhang, Y. M.; Lai, C. J. Both Nonstructural Proteins NS2B and NS3 Are Required for the Proteolytic Processing of Dengue Virus Nonstructural Proteins. *J. Virol.* **1991**, *65* (5), 2467–2475. https://doi.org/10.1128/jvi.65.5.2467-2475.1991.

(43) Chambers, T. J.; Grakoui, A.; Rice, C. M. Processing of the Yellow Fever Virus Nonstructural Polyprotein: A Catalytically Active NS3 Proteinase Domain and NS2B Are Required for Cleavages at Dibasic Sites. *J. Virol.* **1991**, *65* (11), 6042–6050. https://doi.org/10.1128/jvi.65.11.6042-6050.1991.

(44) Millies, B.; von Hammerstein, F.; Gellert, A.; Hammerschmidt, S.; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; Kersten, C.; Kiefer, W.; Bodem, J.; Hildt, E.; Windbergs, M.; Hellmich, U. A.; Schirmeister, T. 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.

(45) Santos, N. P.; Santos, L. H.; Torquato Quezado de Magalhães, M.; Lei, J.; Hilgenfeld, R.; Salgado Ferreira, R.; Bleicher, L. Characterization of an Allosteric Pocket in Zika Virus NS2B-NS3 Protease. *J. Chem. Inf. Model.* **2022**, *62* (4), 945–957. https://doi.org/10.1021/acs.jcim.1c01326.

(46) Cao, X.; Liu, K.; Yan, S.; Li, S.; Li, Y.; Jin, T.; Liu, S. Mechanical Regulation of the Helicase Activity of Zika Virus NS3. *Biophys. J.* **2022**, *121* (24), 4900–4908. https://doi.org/10.1016/j.bpj.2022.07.030.

(47) GraphPad Prism. GraphPad Software: San Diego, California USA 2014. www.graphpad.com.

(48) Harper, E.; Berger, A. On the Size of the Active Site in Proteases: Pronase. *Biochem. Biophys. Res. Commun.* **1972**, *46* (5), 1956–1960. https://doi.org/10.1016/0006-291X(72)90076-9.

(49) The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography*. Schrödinger 2002, pp 82–92. http://www.pymol.org.

(50) Gardner, A.; Sweitzer, T. D.; Taylor, A. H.; Patel, P. Methods for Preventing Gluconoylation of Proteins. WO 2004/078936 A2, 2004.

(51) Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; LeMotte, P. K.; McColl, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous α-N-6-Phosphogluconoylation of a "His Tag" InEscherichia Coli:The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **1999**, *267* (1), 169–184. https://doi.org/10.1006/abio.1998.2990.

(52) Baronti, C.; Piorkowski, G.; Charrel, R. N.; 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), e00500-14. https://doi.org/10.1128/genomeA.00500-14.

(53) Hammerschmidt, S. J.; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B/NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm. (Weinheim).* **2023**, *356* (4), e2200518. https://doi.org/10.1002/ardp.202200518.

(54) Maus, H.; Barthels, F.; Hammerschmidt, S. J.; Kopp, K.; Millies, B.; Gellert, A.; Ruggieri, A.; Schirmeister, T. SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B/NS3 Proteases. *Bioorganic Med. Chem.* **2021**, *47* (August), 116392. https://doi.org/10.1016/j.bmc.2021.116392.

(55) Robin J. Leatherbarrow. GraFit 6. Erithacus Software Limited: East Grinstead, West Sussex, U.K. 2007. http://erithacus.com/grafit/.

(56) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9* (7), 671–675. https://doi.org/10.1038/nmeth.2089.

Promiscuity in the autocleavage behavior of the Zika virus NS2B/NS3 protease.

Stefan J. Hammerschmidt[a], [redacted]

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

[b] Institute of Organic Chemistry & Macromolecular Chemistry (IOMC), Friedrich Schiller University, Jena, Germany

KEYWORDS: Zika Virus, NS2B/NS3 protease, autocatalysis, protein mass spectroscopy

* E-Mail: [redacted]

CONTENT

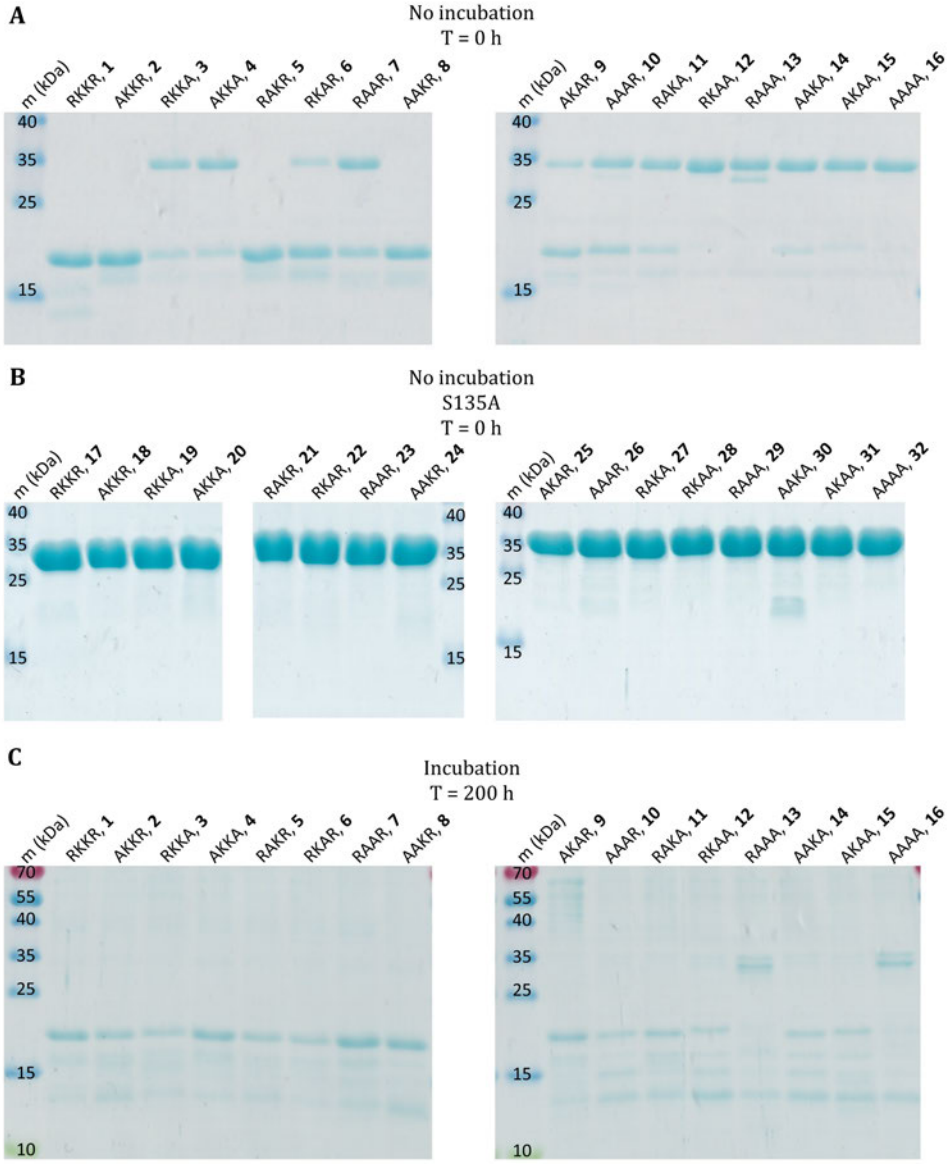
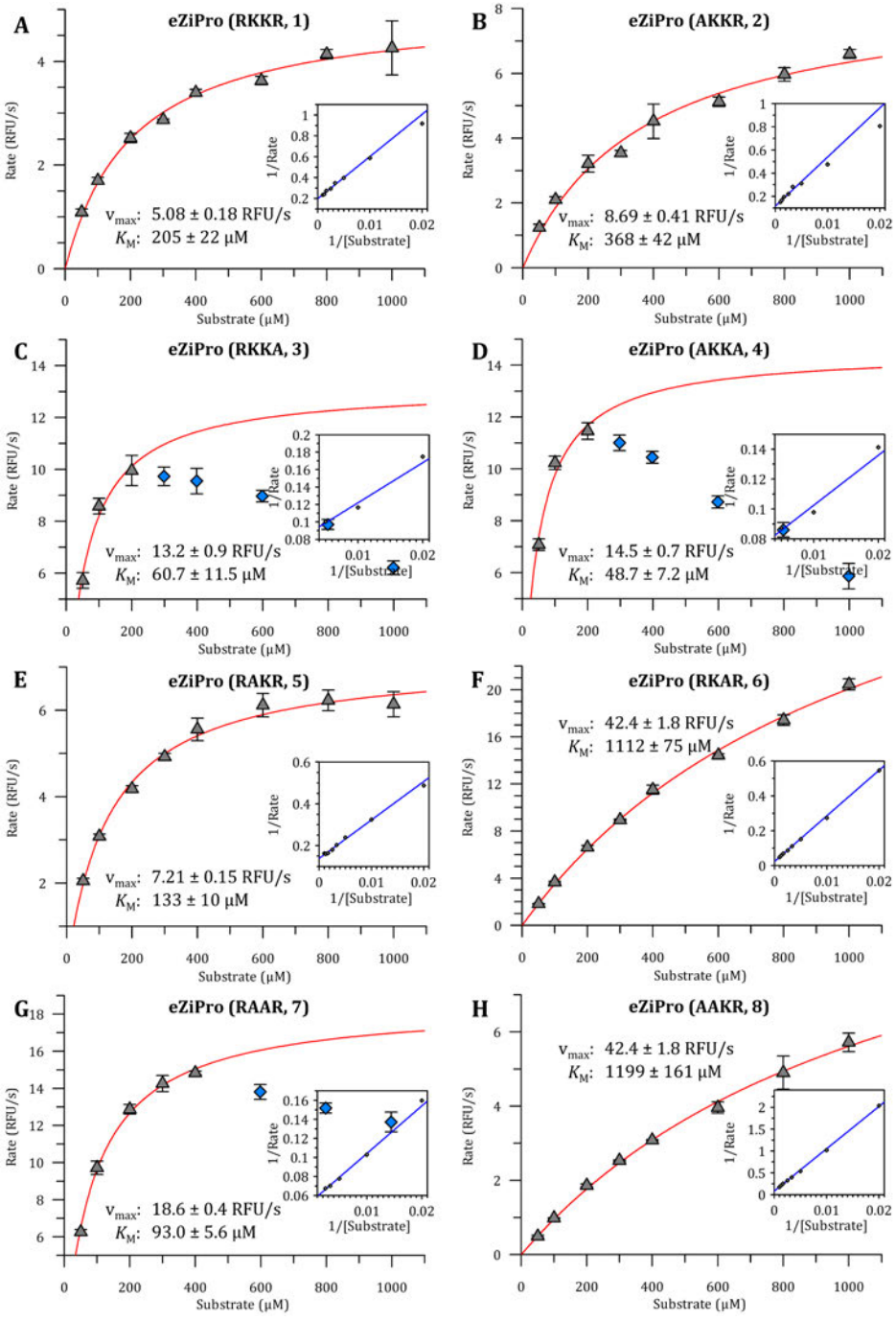


Figure S1. *Exemplarily SDS-PAGE for densitometric analysis.* (**A**) SDS-PAGE of all enzymatically active constructs and (**B**) their respective S135A mutants. After protein purification under similar conditions regarding time and temperature, the eZiPro constructs were aliquoted and flash-frozen so that the cleavage efficiencies of all constructs could be determined by densitometric analysis. (**C**) SDS-PAGE of all enzymatically active constructs after 200 h incubation in assay buffer under agitation at 37 °C shows nearly fully enzymatic cleavage except for RAAA and AAAA constructs. As a marker (m), the PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo Fisher Scientific, Waltham, MA, USA), was used.


A
eZiPro (RKKR, 1)
v_{max}: 5.08 ± 0.18 RFU/s
K_M: 205 ± 22 µM
B
eZiPro (AKKR, 2)
v_{max}: 8.69 ± 0.41 RFU/s
K_M: 368 ± 42 µM
C
eZiPro (RKKA, 3)
v_{max}: 13.2 ± 0.9 RFU/s
K_M: 60.7 ± 11.5 µM
D
eZiPro (AKKA, 4)
v_{max}: 14.5 ± 0.7 RFU/s
K_M: 48.7 ± 7.2 µM
E
eZiPro (RAKR, 5)
v_{max}: 7.21 ± 0.15 RFU/s
K_M: 133 ± 10 µM
F
eZiPro (RKAR, 6)
v_{max}: 42.4 ± 1.8 RFU/s
K_M: 1112 ± 75 µM
G
eZiPro (RAAR, 7)
v_{max}: 18.6 ± 0.4 RFU/s
K_M: 93.0 ± 5.6 µM
H
eZiPro (AAKR, 8)
v_{max}: 42.4 ± 1.8 RFU/s
K_M: 1199 ± 161 µM
Rate (RFU/s)
Substrate (µM)
1/Rate
1/[Substrate]

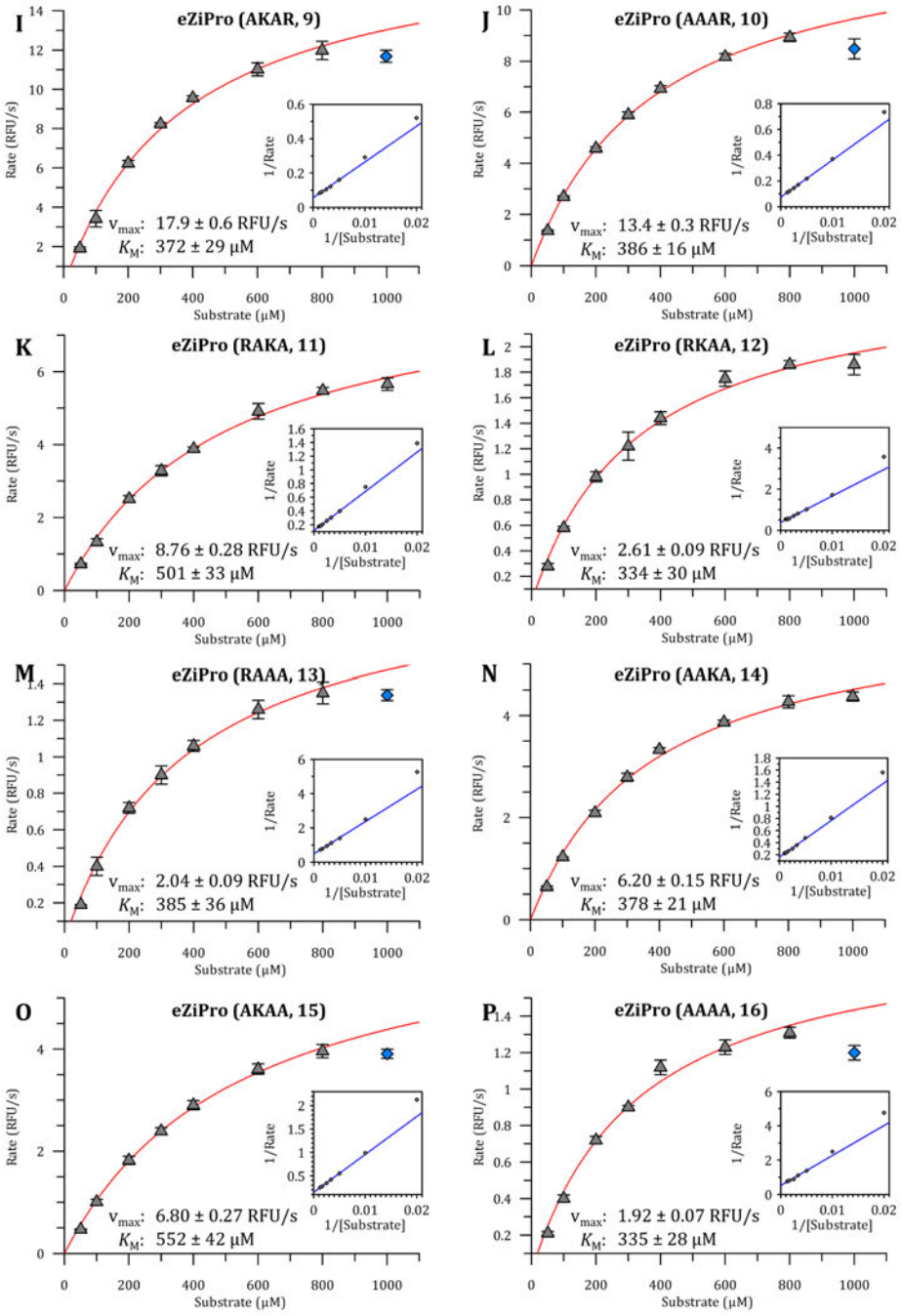


Figure S2. *K_M-determinations of eZiPro and Boc-GRR-AMC.* (**A**–**P**) The 16 cleavage site permutations are indicated in brackets. Thereby, only the basic residues of the REKTGKR cleavage site (R95#K97#K100#R101#) or the respective Ala mutants are shown. Data of technical triplicates are given as gray triangles. Errors are indicated as black bars. The K_M fit is shown as a red line. Since substrate inhibition occurred by several constructs, data for high substrate concentrations laying outside the fitting curve with their error bars are excluded from evaluation and given as blue rhombuses. In the smaller secondary graph, the respective linearized Lineweaver-Burk plot[1] is presented. Figure generated with GRAPHIT (Version 5.0.13; Erithacus Software Limited, East Grinstead, UK).[2]

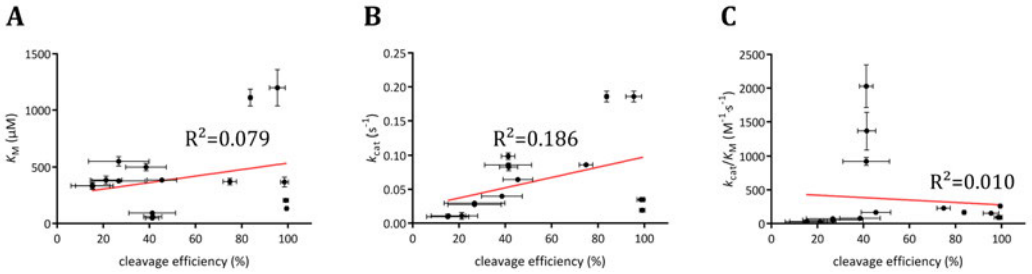


Figure S3. *Enzyme kinetic parameters K_M **(A)**, k_{cat} **(B)**, and k_{cat}/K_M **(C)** plotted against the cleavage efficiency.* Figure created with GraphPad Prism.[3]

Table S2. Primers used for site directed mutagenesis.

Mutation	Direction	Template mutations	Sequence 5' to 3'
R_{95}#A	for	RKKR	CCGCCGATGGCTGAAAAAACCGGCAAACG
	rev	RKXX	GTTTTTTCAGCCATCGGCGGACCATCGTCTTC
R_{101}#A	for	RKKR	ACCGGCAAAGCGAGCGGCGCGCTGTGG
	rev	RKKR	CCGCGCCGCTCGCTTTGCCGGTTTTTTCACGC
K_{97}#A	for	XKKR	ATGCGTGAAGCGACCGGCAAACGCAGCG
	for	AKKR	GGCTGAAGCGACCGGCAAACGCAGCG
	for	RKKA	ATGCGTGAAGCGACCGGCAAAGCGAGCG
	for	AKKA	ATGGCTGAAGCGACCGGCAAAGCGAGCG
	rev	AKAA	CTGAAGCGACCGGCGCGGCGAGCG
	rev	RKXX	CCGGTCGCTTCACGCATCGGCGGACC
K_{101}#A	for	AKXX	CCGGTCGCTTCAGCCATCGGCGGACC
	for	XXKR	CGGCGCGCGCAGCGGCGCGCTG
	for	XXKA	AAACCGGCGCAGCGAGCGGCG
	rev	RKKR	CGCTGCGCGCGCCGGTTTTTTCACGCATC
	rev	XAKR	GCCGCTGCGCGCGCCGGTCGCTTC
	rev	AKKR	GCTGCGCGCGCCGGTTTTTTCAGCCATCG
	rev	RKKA	CGCTCGCTGCGCCGGTTTTTTCAC
	rev	AKKA	GCTCGCTGCGCCGGTTTTTTCAGC
	rev	XAKA	GCTCGCTGCGCCGGTCGCTTC
S_{135}A	for	XXXX	GGTACGGCTGGCTCCCCGATCGTC
	rev	XXXX	GAGCCAGCCGTACCCGGAGAGAAATC

For clarity, the bases of the introduced Ala mutation are colored red, the bases of residues already mutated to Ala are yellow and the bases of native basic residues of the cleavage site are colored green. The sequence of the template's basic aa R_{95}#, K_{97}#, K_{100}#, and R_{101}# are given in brackets. (X can be any aa; aa in NS2B are indicated with a '#')

REFERENCES

(1) Lineweaver, H.; Burk, D. The Determination of Enzyme Dissociation Constants. *J. Am. Chem. Soc.* **1934**, *56* (3), 658–666. https://doi.org/10.1021/ja01318a036.

(2) Robin J. Leatherbarrow. GraFit 6. Erithacus Software Limited: East Grinstead, West Sussex, U.K. 2007. http://erithacus.com/grafit/.

(3) GraphPad Prism. GraphPad Software: San Diego, California USA 2014. www.graphpad.com.

6. Project 4: Cation Interactions with Sortase A

6.1. Interactions of *Staphylococcus aureus* Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors

6.1.1. Context, Project Summary, and Own Contribution

In this project, the influences of 45 cationic metal ligands on *Staphylococcus aureus* SrtA have been investigated using a toolset of several biophysical methods such as ITC, MST, X-ray co-crystallography, FRET-based transpeptidase assays, and a Tb^{3+}-FRET binding assay. Thereby, we identified activating and inhibiting, metal cations, and such exhibiting a mixed mechanism. Through thorough investigations, several false positive hits and nonspecific binders could be identified as quenchers of the FRET signal, thiophilic binders of the active site Cys184 or protein aggregating ions. ITC allowed us to determine several metal cations' thermodynamic binding profiles and K_d values. Thereby, we revealed that lanthanides bind to adventitious sites since binding isotherms could be fitted with sequential binding sites. ITC experiments of the cations against SrtA that was preincubated with an active site-directed compound demonstrated cooperativity of the SrtA activating cations Ca^{2+} and Mg^{2+}. This indicates allosteric activation of SrtA by binding to the Ca^{2+} binding site. MST showed aggregation of SrtA under high concentrations of the lanthanides, probably leading to inhibition in activity assays. In Tb^{3+}-FRET assays, Trp and Tyr residues get excited with light. Through FRET, the fluorescence of Tb^{3+} can be recorded when bound in proximity. This experiment revealed higher fluorescence of Tb^{3+} with increasing concentrations of SrtA. This fluorescence can be reduced when Ca^{2+} is added in excess. This indicates the binding of Tb^{3+} to the Ca^{2+} binding site and subsequent displacement. In X-ray crystallography, co-crystals were obtained in complex with a covalent peptidomimetic inhibitor of SrtA and with either $CaCl_2$, $MgCl_2$, $MnCl_2$, $ZrOCl_2$, or $LaCl_3$. However, only the structure of the Mn^{2+} complexed SrtA could be solved yet. Mn^{2+} was coordinated in the Ca^{2+} binding site, inducing a conformation of SrtA like that of the Ca^{2+} coordinated structure (PDB-ID: 2KID).[102]

For publication of the results, some additional experiments need to be performed. SrtA mutants Q172W and F103W, harboring an additional Trp near the Ca^{2+} binding site, were already cloned. They could increase the intensity of the Tb^{3+} fluorescence in Tb^{3+}-FRET assays, validating our results. Ti^{3+} and Ge^{3+} are known for their antistaphylococcal properties and biofilm reduction.[265–267] Hence, they should be tested in enzyme activity assays to evaluate if they act as inhibitors of SrtA. To complement our DSF data (not included in the manuscript since all assayed cations showed a similar shifted T_m), nanoDSF experiments with and without active-site inhibitors should be performed. More ITC experiments of the activating cations Pd^{2+}, Ba^{2+}, and Mn^{2+} should be conducted to evaluate cooperativity with active site-directed ligands. Since conditions for the co-crystallization of SrtA were established in this work, these experiments should be performed in collaboration with a specialized working group.

(8) **Hammerschmidt, S. J.**; [redacted] Interactions of Staphylococcus Aureus Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors. *Unpublished and not yet submitted manuscript.*

Own contribution: Establishment of enzyme expression and purification, mutagenesis of SrtA active site mutants, Gibson Assembly of SrtA dimers, enzyme activity assays, ITC experiments, MST experiments, DSF experiments, X-ray crystallography, writing of the original draft & editing of the manuscript.

Contributions from other authors: Solving the crystal structure of SrtA in complex with cpd **3** and Mn^{2+}, Tb^{3+}-FRET assay, Synthesis of cpds **2** and **3**, Synthesis of the Lanthanide-Chlorides, Writing of the Experimental Section “Chemistry”.

6.1.2. Manuscript

The following manuscript, from page 313 to page 369, contains unpublished data.

Interactions of Staphylococcus Aureus Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors.

Stefan J. Hammerschmidt [a], [redacted]

[a] Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University of Mainz, Staudinger Weg 5, 55128 Mainz, Germany

[b] Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Hans Dieter Hüsch Weg 15, 55128 Mainz, Germany

KEYWORDS: Staphylococcus aureus, sortase A, metal cation, allostery, x-ray structure, isothermal titration calorimetry, anti-virulence

ABSTRACT

In recent years, the rapid spread of multiple antibiotic resistances among Gram-positive bacteria, including *Staphylococcus aureus*, is responsible for several thousands of deaths annually and poses a major challenge to modern healthcare. Sortase A is a bacterial transpeptidase crucial for biofilm formation and virulence. For catalytic activity, sortase A relies on Ca^{2+} as a cofactor, allosterically activating sortase by pre-organizing an intrinsically disordered loop. To date, only active site inhibitors of SrtA have been reported that are facing several hindrances, such as the rigidification of flexible loop upon binding, reverse protonation of the catalytic cysteine, and high local concentrations of the competing natural substrates. Therefore, to provide more detailed information on the interactions and effects of cationic cofactors on SrtA conformation and catalytic activity, we investigated a set of 47 metal cations by competitive fluorometric enzyme activity assays. We used microscale thermophoresis and differential scanning fluorimetry to gain additional information on binding behavior and determined the thermodynamic binding profiles with isothermal titration calorimetry. Further, we performed co-crystallization experiments with the active site inhibitors **2** and **3** as reference compounds in the presence of different cations. Within this work, we expand knowledge of the interaction of SrtA with its cofactor Ca^{2+} as well as other activating and inhibiting ligands. Our research successfully identified several new inhibitory cations for SrtA activity and elucidated the mechanism behind their inhibitory activity. Cocrystals with different activating cations revealed binding to the Ca^{2+} binding site, allowing SrtA to adopt the active conformation. Taken together, the findings of this work support fundamental research as a basis for developing new inhibitors for SrtA. Thereby, it contributes to overcoming the growing resistance of biofilm-forming pathogenic bacteria by incorporation of SrtA inhibiting cations in the alloys of implants and other medical instruments.

INTRODUCTION

The Role of SrtA in Staphylococcal Virulence. The rapid emergence of multiple antibiotic-resistant claims tens of thousands of annual deaths.[1] Drug-resistant strains cause nearly 3 million infections in the USA, with 35,000 fatalities.[2] Among them, methicillin-resistant *staphylococcus aureus* (MRSA) is responsible for 324,000 hospitalizations and 10,600 deaths, leading to 1.7 billion dollars in healthcare costs.[2] A major causative for the high resistance development observed in *Staphylococcus aureus* (*S. aureus*) is its ability to grow in 3-dimensional communities, called biofilm.[3,4] As a membrane-anchored surface transpeptidase, sortase A (SrtA) catalyzes the anchoring reaction of several microbial surface components recognizing adhesive matrix molecules (MSCRAMMs).[5] SrtA catalyzes a two-step reaction: First, the proteolytic cleavage of the so-called sorting signal LPXT↓G, in which the catalytic Cys184 forms an acyl intermediate, which is later connected to the

pentaglycine motive of the membrane-anchored lipid II, a ubiquitous building block of the peptidoglycan matrix.[5] Biofilms harbor several virulence factors which are anchored to the periplasmic matrix by SrtA,[6] enabling adhesion to host tissues and medical instruments,[6–8] higher antibiotic resistances,[4] immune escape,[9] host cell-damaging,[10] nutrient extraction[11,12] and intracellular communication.[13,14] Hence, SrtA represents a promising target for developing anti-virulence drugs.[15]

SrtA as Target for Inhibitor Development. Inhibitors of SrtA can potentially reduce biofilm formation[16] but do not impede bacterial growth like conventional antibiotics.[5,17] Thus, the lower selective pressure they impose could prevent the development of resistances.[18,19] So far, all small molecule inhibitors of SrtA were reported to bind in its active site.[18,20,21] Development of those inhibitors, however, is challenged by three major features of SrtA, protecting it against active site-directed inhibitors.[18] (i) Since the SrtA binding site does not feature a pronounced oxyanion hole as it can be found in related proteases, the catalytically active form is the zwitterionic state of Cys184 and His120.[22,23] The catalytic Cys184 is reversely protonated, decreasing the population of the nucleophilic thiolate below 0.1%.[23,24] Hence, the Cys184 is less reactive than cysteines in related proteases and challenging to target with protease-specific electrophilic warheads.[23,25,26] (ii) The activity of SrtA is regulated by the intrinsically disordered region (IDR) of the $\beta6/\beta7$ loop, which undergoes strong rigidization upon binding the Ca^{2+} cofactor and the sorting signal LPXTG.[27,28] Thus, inhibitor binding is likely to be impeded by a high entropic penalty that is reflected in the high K_M of the substrate of ~5 mM.[29] (iii) To compensate for this low enzymatic activity due to the reverse protonation and the high K_M value of the sorting signal, concentrations of both the LPXTG substrates and the lipid II acceptors are high in the local environment of SrtA. Therefore, the binding of active site-directed inhibitors is less efficient through high competition.[30]

Allosteric Activation of SrtA by Calcium. The calcium-binding site is formed by Gly105, Glu108, and Asp112 of the $\beta3/\beta4$ loop and Glu171 of the IDR $\beta6/\beta7$ loop.[22] The apo-SrtA has a disordered $\beta6/\beta7$ loop that is highly flexible in the open state. This is probably due to the electrostatic repulsion of the anionic Ca^{2+} binding residues. The dynamic $\beta7/\beta8$ loop is closed (Figure 1A).[31]

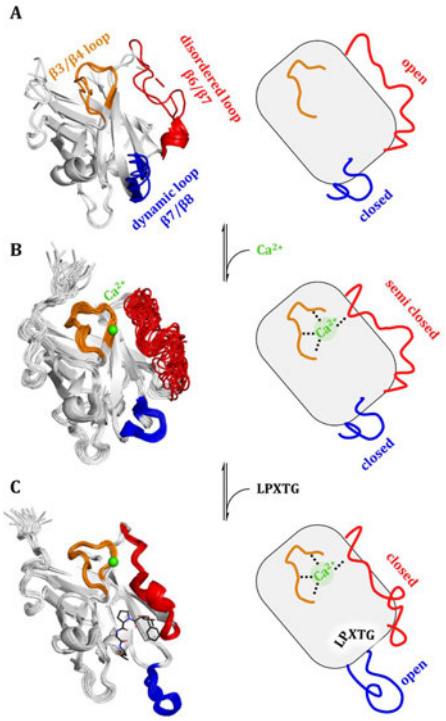


Figure 1. *Cooperativity of Ca^{2+} and the sorting signal.* Left: X-ray (**A**) or nuclear magnetic resonance (NMR) ensembles (**B**, **D**) of SrtA (white) in different complexed states. Right: The respective schematic presentation of the disordered (red) and the dynamic loop's (blue) conformations. (**A**) The three subunits of the asymmetric unit of the apo-SrtA show an open conformation of the disordered and a closed conformation of the dynamic loop in the absence of ligands (RCSB protein data bank (PDB)-ID[34]: 1T2P).[31]. (**B**) A Calcium ion (green) coordinates the $\beta3/\beta4$ loop (orange) and the C-terminal part of the disordered loop, leading to a partly closed conformation, pre-organizing it for binding of the sorting signal. The N-terminal region stays highly flexible (PDB-ID: 1IJA).[22] (**C**) In the complexed state with Ca^{2+} and the sorting signal, the disordered loop adopts a rigid closed conformation. The dynamic loop is displaced by the sorting signal and stretches, forming a binding interface for a second substrate (PDB-ID: 2KID),[27] The figure was created using Pymol.[35]

Ca^{2+} binding to this site enhances enzymatic activity by allosteric preorganization of the $\beta6/\beta7$ loop for binding the sorting signal LPXTG (Figure 1B).[27,28,32,33] In the holo-form, complexed with Ca^{2+} and the sorting signal, the disordered loop adopts a well-defined conformation. The sorting signal displaces the dynamic loop that opens to form a binding pocket for the pentaglycine acceptor(Figure 1C).[27] The rigidification of SrtA upon Ca^{2+} binding is also in line with an increase of the melting

temperature (T_m) of 3.4 K in differential scanning calorimetry (DSC).[33]

The Relevance of SrtA in Bioengineering. Modulating the activity of SrtA is not only of interest because of its promising role as an anti-virulence target to suppress biofilm formation, but the versatility of the transpeptidase functionality is also frequently utilized in biochemical engineering. Since the catalytic activity of SrtA does not rely on its N-terminal membrane anchoring, truncated $SrtA_{\Delta 59}$ is well soluble and widely used in biochemical applications and *in vitro* experiments.[22,36] The so-called *sortagging* can be used to label proteins terminally by either adding a C-terminal LPXTG functionality or an N-terminal oligo-glycine tag.[37,38] Furthermore, SrtA was used to perform backbone cyclization reactions of peptides and proteins,[39−41] for protein-to-protein fusion products,[42] immobilization of biocatalyst onto a solid surface,[43] preparation of complex glycoconjugates,[44] antibody–drug conjugation,[45,46] and in vivo protein modification.[47] However, the applicability of SrtA-mediated ligations is limited by the reliance on its Ca^{2+} cofactor. Hence, the reactions must be performed in media that do not chelate or precipitate calcium ions, such as the commonly used phosphate and carbonate buffers as well as EDTA and EGTA. Attempts to overcome these issues pursued the generation of Ca^{2+} independent mutants of SrtA by replacing some of the Ca^{2+} chelating Glu with Lys, enabling interaction to the $\beta 6/\beta 7$ loop.[48,49] Although some activity could be achieved, the most suitable SrtA variant only yielded ∼80% of the wild-type's (WT) activity.[49]

Metal Cations as Antibacterial Agents. Due to progressing antibiotic resistance, maintaining effective anti-microbial treatment options urges new therapeutic strategies to combat these threats. Especially biofilm-forming pathogens in chronic infections can often not be adequately treated and complicate wound healing.[50] These pathogens are even more dramatic as they attach to medical devices such as catheters, endotracheal tubes, and implants.[8,51] Consequently, surface metal ions recently gained attention for their anti-microbial effects.[52−57] They found several applications as anti-microbial agents. They are utilized in alloys or coatings for implants[58] or catheters,[59] in wound dressings,[60] or liquid hand-washing formulations, and can be used to impregnate clothes.[61,62] However, it was demonstrated that anti-microbial activity is not a property of a metal ion directed against all biofilm-forming bacteria in the same way. There are some disturbing differences between *Pseudomonas aeruginosa* (*P. aeruginosa*), *S. aureus,* and *E. coli*.[63] Bacteria in biofilms exhibit pronounced resistance to heavy metal ions either due to the physical barrier that biofilms impose or mechanisms that convert them to less toxic forms by chelation or reduction.[64] However, in the case of *S. aureus*, some ions were found to act even better against bacteria in biofilms compared to their planktonic form.[63−65] Ions that are found effective against the biofilm formation of *S. aureus* are Ti^{3+}, Ni^{2+}, Cu^{2+}, Ag^{+}, and Ga^{3+}.[58,63,66] However, data about Ag^{+} efficacy against *S. aureus* biofilm formation is inconclusive.[63] Interestingly, Ni^{2+} was only effective against *S. aureus* biofilm formation, possibly indicating a SrtA-mediated mechanism.[63]

Cation Interactions with SrtA. Given the hindrances of developing active site-directed inhibitors, it seems convincing to actively search for other options for suppressing its function. So far, to our knowledge, neither allosteric nor cofactor competitive inhibitors of SrtA have been reported. Although some metal ions were assayed to find activators of SrtA, the residual activity of SrtA without Ca^{2+} was reported to be only 8-fold or 3-fold lower than with the addition of 2 mM or 5 mM of Ca^{2+}, respectively.[22,36] The study investigating 2 mM of cations found an 8-fold increase of activity in the presence of Ca^{2+} assayed the eight cations K^{+}, Mg^{2+}, Ca^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, $Zn^{2+}$and Cd^{2+} only found Mg^{2+}, Mn^{2+} to achieve ∼50% of the activity in the presence of Ca^{2+}. In contrast, all other ions failed to stimulate SrtA activity.[22] This contrasts the study with 5 mM of the eight metal ions Mg^{2+}, Ca^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}, reporting a 3-fold increase of activity with Ca^{2+}. The ∼50% activity in the presence of Mg^{2+} and Mn^{2+} was reproducible, whereas all other cations seemed to lower the residual activity of SrtA and act inhibitory. Just stating this finding, no validation or thorough characterization of the mechanism of the inhibitory cations was performed.[36] Since these studies used Ca^{2+} in their purification buffers and did not describe further purification[22] or stated that assays were performed "without additional metal ion (control)",[36] we assume that residual Ca^{2+} from protein purification could be blamed for the relatively high enzymatic activities in the calcium-free control measurements. This led us to hypothesize that ions may bind the allosteric binding site, thereby replacing Ca^{2+} without inducing a proper folding of SrtA.

Hence, this work of extensively investigating SrtA-cation interactions could benefit the scientific fields of bioengineering and modern healthcare. Activating ions could replace calcium ions in industrial sortagging processes, especially when calcium-free reaction buffers are needed. Given the hindrances in inhibitor development, inhibiting ions could be of high interest concerning applications in healthcare since they could provide the foundation for designing new medical alloys or coatings for medicinal products and implants and the utilization of new metal ions in surface disinfectants. Additionally, a deeper mechanistic understanding and structural information of the allosteric modulation of SrtA activity with cations could lead to new strategies for SrtA-targeting anti-virulence drugs. In this regard, allosteric SrtA inhibitors or cation-chelating substrates of SrtA could be designed, releasing inhibitory ions upon proteolytic cleavage. To provide information on a broad spectrum of SrtA-cation interactions and to derive prerequisites of activating or inhibiting ions, we chose to perform FRET-based SrtA activity and inhibition assays to evaluate a set of 47 metal chlorides. Then we selected promising candidates for an in-depth analysis of their mode of action by using intrinsic tryptophane fluorescence for differential scanning fluorimetry (nanoDSF), microscale thermophoresis (MST), isothermal titration calorimetry (ITC) and co-crystallization experiments.

RESULTS AND DISCUSSION

Calcium-Free Activity Assay of Sortase A. All assayed cations were used as chloride salts for better comparability of our data. Previous work tested the effects of a smaller set of cations in Ca^{2+}-free assays. However, the activity of Ca^{2+}-free SrtA was much higher compared to our assay (Figure 2).[22,36] This leads to the conclusion that Ca^{2+} was retained in the protein samples from purification. However, inhibitory cations competing with the Ca^{2+} binding site within loops $\beta3/\beta4$ and $\beta6/\beta7$ could not certainly be detected since no Ca^{2+} competitive assay or further investigations were conducted.[22,36] The Activity of SrtA is strongly dependent on environmental pH.[36] Thus, lowered pH by adding Lewis acids could lead to inactive SrtA, thereby disguising the specific effects of these cations on SrtA activity. To overcome this issue, pH measurements of the final assay buffer (with 5 mM Tris•HCl as buffering agent) were conducted and revealed drastically decreased pH upon adding 10 mM of the metal chlorides (Table S2). Since several chloride salts represent strong Lewis acids,[67] the cations were only used at 5 mM concentration. The initial assay buffer was adjusted to 50 mM Tris•HCl to better compensate for the proton release. In line with previous studies, where SrtA enzymatic activity was found to increase 8-fold upon adding 2 mM of Ca^{2+} or 3-fold in 5 mM Ca^{2+}, we found SrtA activity to increase 20-fold in the presence of 5 mM Ca^{2+} compared to our calcium-free conditions (Figure 2A). No rebuffering steps of the protease are described in the aforementioned studies. Further, Ca^{2+} was present during the final protein purification steps. Hence, the reported smaller activating effects of Ca^{2+} could be explained by a higher SrtA activity without further addition of Ca^{2+}, as residual Ca^{2+} could be present from the size exclusion/storage buffers.[22,36] The finding that activities in the presence of Co^{2+}, Cu^{2+}, Fe^{2+}, Ni^{2+}, and Zn^{2+} were reported to be slightly lower than the negative control without additional cations further supports this hypothesis.[36] Herein, SrtA was eluted from size exclusion chromatography (SEC) in a calcium-free buffer. Hence, almost no enzymatic activity could be observed in our negative control without adding cations, and this is similar to the activity in the presence of 10 mM EDTA (Figure 2A). The tested alkaline metals Li^+, Na^+, K^+, Rb^+, and Cs^+ did not influence SrtA activity (Figure 2A). Na^+ was present in each measurement, as the assay buffer contains 150 mM NaCl. The first new activating cations could be identified among the alkaline earth metals (Mg2+, Ca2+, Sr2+, and Ba^{2+}, Figure 2A). Besides Ca^{2+} (reference, 100%), both Sr^{2+} and Ba^{2+} showed relative activating effects of 25% and 14%, respectively. In contrast to previous studies[22,36] addition of Mg^{2+} only led to a minor increase in activity (Figure 2A). In the set of assayed transition metals (Sc^{3+}, Y^{3+}, Ti^{3+}, Zr^{4+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Ru^{3+}, Co^{2+}, Ni^{2+}, Pd^{2+}, Cu^+, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}), only Mn^{2+} revealed notably activating effects of 42% in line with previous measurements (Figure 2B).[22,36] In the metal group (Al^{3+}, Ga^{3+}, In^{3+}, Sn^{2+}, Pb^{2+}, Sb^{3+}, and Bi^{3+}), both Sn^{2+} and Pb^{2+} activated SrtA to 14% and 15% compared to Ca^{2+}, respectively (Figure 2C). In the group of tested lanthanides (La^{3+}, Ce^{3+}, Nd^{3+}, Sm^{3+}, Eu^{3+}, Gd^{3+}, Tb^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, Yb^{3+}, and Lu^{3+}), 5 mM La^{3+} yielded a relative activity of 111% and was even more activating than the natural ligand Ca^{2+}. The lanthanides Ce^{3+}, Pr^{3+}, and Nd^{3+}, activated SrtA with 35, 40, and 25%, respectively. All lanthanides with higher atomic numbers did not stimulate SrtA activity (Figure 2D). It can be assumed, that SrtA activating metal cation address the calcium binding site.

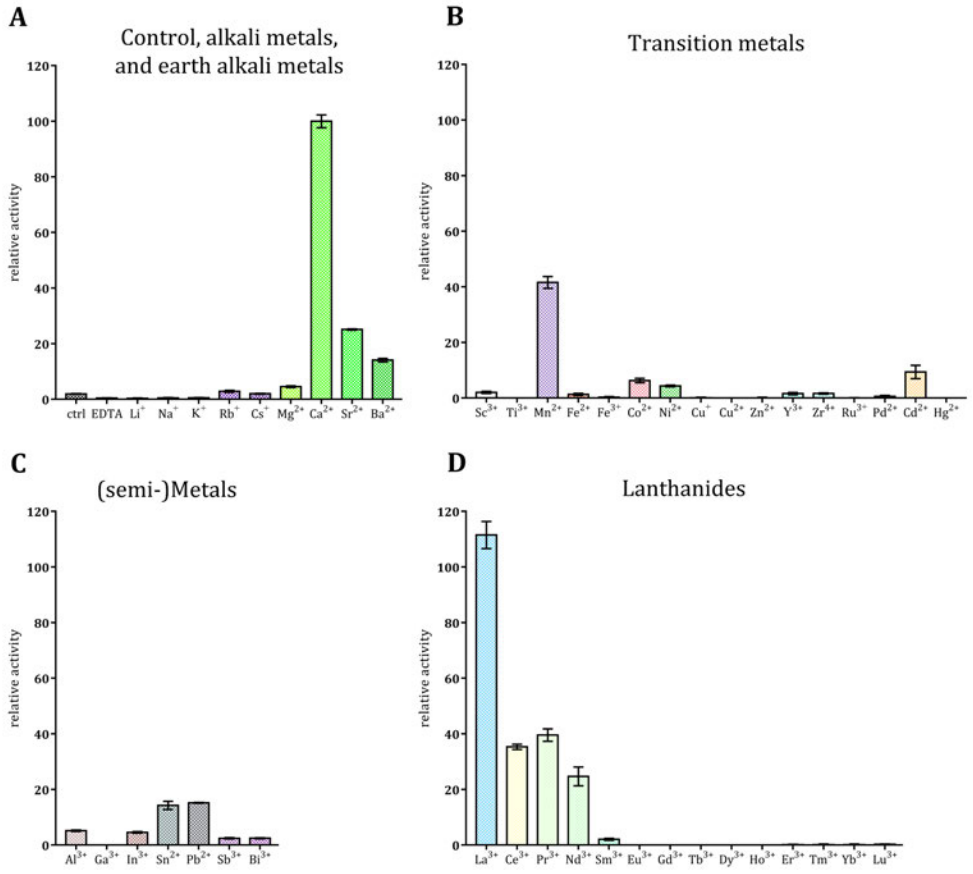


Figure 2. *Enzymatic activities of wild-type SrtA in the presence of 5 mM of respective cation chlorides.* Activities are normalized to the activity in the presence of 5 mM Ca^{2+} ions (100%). Relative activity values are indicated as means of at least technical triplicates ± standard deviation. Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[68] **A** Relative activities of control measurements (without additional cations, with 5 mM EDTA), alkaline and earth-alkaline metals. **B** Relative activities of transition metal cations. **C** Relative activities of (semi-) metals. **D** Relative activities of lanthanide ions. The figure was created using GraphPad Prism 7.0.4.[69] For tabularized data, see Table S3.

Calcium-Competitive Activity Assay of Sortase A. After identifying new activating cations as cofactors for SrtA, we sought to elucidate further effects of the cations that only displayed minor to no effects in the enzyme activity assays. Therefore, we employed a Ca^{2+}-competitive assay, where the effects of 5 mM of the respective cations on SrtA activity were tested in the presence of equimolar concentrations of Ca^{2+}. By these means, we can determine inhibitory effects on SrtA, possibly due to the displacement of Ca^{2+}. Contrary, cations that do neither influence SrtA activity in the presence nor absence of Ca^{2+} are unlikely to interact with SrtA. Contrary to the literature, reporting SrtA activity to reach its maximum at 10–50 mM Ca^{2+},[36] the activity of 10 mM Ca^{2+} was not altered to that of 5 mM Ca^{2+}. In the remaining group of alkaline and earth-alkaline metals, only Ba^{2+} slightly inhibits SrtA activity of ~15% (Figure 3A). Since both Sr^{2+} and Ba^{2+} were found to simulate SrtA activity in the Ca^{2+}-free assay but to a lower extent than Ca^{2+}, it can be assumed that they either bind with less affinity or act as partial agonists of SrtA. Thereby, Ba^{2+} seems to displace some Ca^{2+} ions, leading to observed inhibition of SrtA. In the group of transition metals, several inhibitory cations were found. In line with Ba^{2+}, the activating cations Mn^{2+}, Co^{2+}, Ni^{2+}, and Cd^{2+} seem to inhibit SrtA to a low extent (<30%). Thus, the observed inhibitory effects may result from the replacement of bound Ca^{2+} by cations acting as partial agonists (Figure 3B). Several other transition metals inhibited SrtA activity more potently.

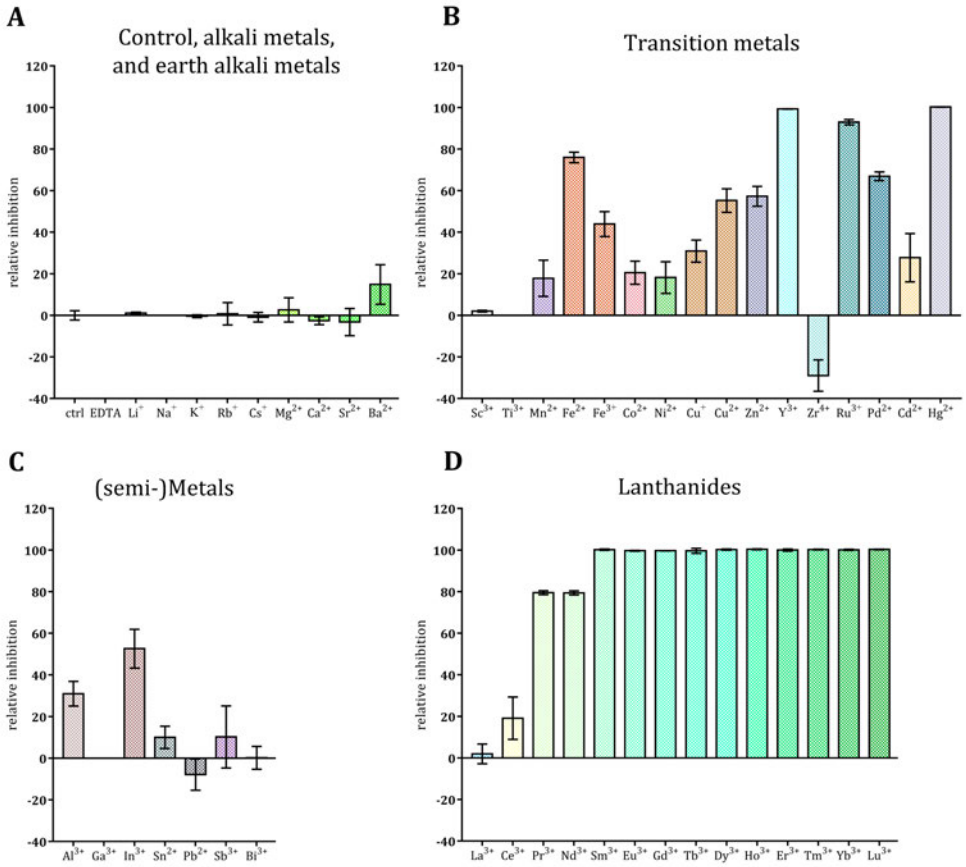


Figure 3. *Enzymatic activities of wild-type SrtA in the presence of 5 mM of respective cation chlorides.* Activities are normalized to the activity in the presence of 5 mM Ca^{2+} ions (100%). Relative activity values are indicated as means of at least technical triplicates ± standard deviation. Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[68] **A** Relative activities of control measurements (without additional cations, with 5 mM EDTA), alkaline and earth-alkaline metals. **B** Relative activities of transition metal cations. **C** Relative activities of (semi-)metals. **D** Relative activities of lanthanide ions. The figure was created using GraphPad Prism 7.0.4.[69] For tabularized data, see Table S3.

Moderate inhibition (30%< inhibition <70%) was observed for Fe^{3+}, Cu^{+}, Cu^{2+}, Zn^{2+}, and Pd^{2+}. Strong inhibition was yielded with Sc^{3+}, Fe^{2+}, Y^{3+}, Ru^{3+}, and Hg^{2+}. From the (semi-)metal group, only Al^{3+} and In^{3+} inhibit SrtA moderately (Figure 3C). The lanthanides inhibit SrtA quite contrary to their activating behavior. La^{3+} does not exhibit inhibitory potential, whereas Ce^{3+} is a weak inhibitor, both Pr^{3+}, and Nd^{3+} are strong inhibitors reaching 79% inhibition, and all other lanthanides fully inhibit SrtA activity (Figure 3D). Strong inhibition of ions that do not activate SrtA in Ca^{2+}-free assays should be interpreted carefully, as several unspecific effects could explain these observations. Noteworthy, Zr^{4+} additionally enhanced the activity of SrtA in the presence of 5 mM Ca^{2+} by 29%. Since Zr^{4+} displayed no activating effects in the Ca^{2+}-free assay, the mechanism behind its Ca^{2+}-depended synergistic effect remains elusive.

Since 5 mM Ru^{3+} ions led to dark-colored solutions, we investigated whether it falsely mimics inhibition by quenching the fluorescence signal. We also included Sc^{3+}, Y^{3+}, La^{3+}, and the lanthanides since they exhibit complex spectroscopic behavior used to determine phases of the diffracted X-rays and in bioanalytical assays.[70–74] Therefore, Fluorescence measurements of varying concentrations of the fluorophore anthranilic acid (Abz) either with or without 5 mM Ru^{3+} (Figure S1) revealed a quenching factor (QF) of ~200 for Ru^{3+}. This demonstrates that Ru^{3+} highly quenches Abz

fluorescence, thereby mimicking inhibition of SrtA. All other tested cations showed QF between 0.5 and 2.4 (Table S3). It is possible that Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+} do not inhibit SrtA by competing with Ca^{2+} binding, as they were found to bind thiol groups due to their high thiophilicities and, therefore, may inactivate cysteine proteases.[75–78] To confirm this hypothesis of binding of these cations to the active site Cys184, ITC measurements with the inactive C184S and C184A mutants were performed.

Microscale Thermophoresis. MST experiments were performed for Ca^{2+}, Mn^{2+}, Sc^{3+}, and the lanthanides La^{3+}, Sm^{3+}, and Tb^{3+}. Since the activating cations Ca^{2+}, Mn^{2+}, and La^{3+} did not show any thermophoretic shift in 1:1 dilution series starting from a final concentration of 400 µM (data not shown), we performed a binding check for Sc^{3+}, La^{3+}, Sm^{3+}, and Tb^{3+} (Figure 4). At 5 mM, all cations seemed to induce a thermophoretic shift suitable for K_d determinations. However, bumby thermophoresis curves were observed for Tb^{3+} and Sc^{3+}. Bumby curves are indicative of ligand-induced protein aggregation. Hence, we took advantage of this and evaluated the experiments with dilution series of the lanthanides not for K_d values but to determine the lowest concentration that leads to protein aggregation. We observed aggregation at 25 µM for La^{3+} and 12.5 µM for Sm^{3+}, and 1.56 µM for Sc^{3+}. Surprisingly, the activating La^{3+} causes aggregation at these low concentrations, although high activation of SrtA could be observed at 5 mM La^{3+}. Furthermore, Sc^{3+}, which neither inhibits nor activates SrtA, also causes protein aggregation. This demonstrates that aggregation does not necessarily inhibit the SrtA activity. On the other hand, the inhibiting Sm^{3+} aggregates SrtA at lower concentrations than La^{3+}.

Isothermal Titration Calorimetry. ITC experiments were conducted for selected metal cations based on their effects on SrtA activity to characterize their binding profiles further. As activating ions, Mg^{2+}, Ca^{2+}, Mn^{2+}, La^{3+}, and Ce^{3+} were selected. Fe^{2+}, Zn^{2+}, Ru^{3+}, Hg^{2+}, Dy^{3+}, and Tb^{3} were selected for their inhibitory activity. Sc^{3+} was used as the negative control, and Nd^{3+} and Zr^{4+} for their mixed behavior. From the tested cations, Fe^{2+}, Hg^{2+}, Ru^{3+} Sc^{2+}, and Zr^{4+} achieved no binding signal (data not shown), indicating unspecific effects adjacent to the Ca^{2+} binding site, which is in line with the assay data for Sc^{2+} and with the hypothesized unspecific acting of the thiophilic Hg^{2+} and the fluorescence quenching Ru^{3+}. From the set of activating ions, Mg^{2+}, Ca^{2+}, Mn^{2+}, La^{3+}, and Ce^{3+} were found to bind to SrtA (Table 1).

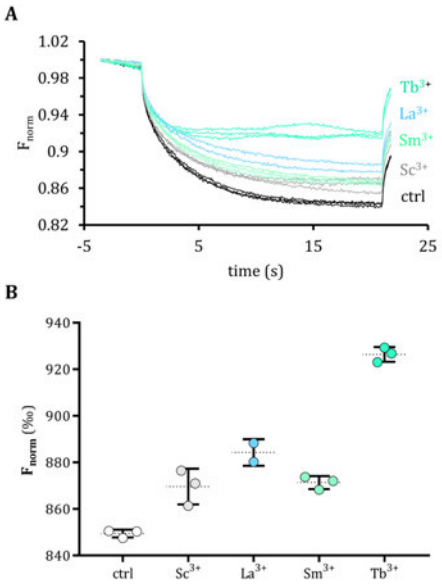


Figure 4. *MST binding check.* (**A**) The thermophoresis curves for trivalent metal-cations at 5 mM show shifts compared to the ctrl measurement. Bumpy curves of Tb^{3+} indicate protein aggregation. (**B**) The means of the normalized fluorescence (F_{norm}), evaluated between 4 s and 5 s of the laser-on time, are given with their standard deviations. The figure was created using Microsoft Excel and GraphPad Prism 7.0.4.[69]

Table 1. Thermodynamic binding profiles of metal ions to SrtA WT.

Ion	K_d (μM)			ΔG° (kJ·mol^{-1})			ΔH° (kJ·mol^{-1})			−TΔS° (kJ·mol^{-1})			N		
Mg^{2+}	276 ± 19			−20.4 ± 0.7			−13.2 ± 0.6			−19.4 ± 0.4			1.00[a]		
Ca^{2+}	79.3 ± 17.1			−23.4 ± 0.2			−26.5 ± 6.1			3.1 ± 4.1			1.06 ± 0.10		
Mn^{2+}	108 ± 63			−22.9 ± 1.0			−2.8 ± 1.4			−20.1 ± 1.8			1.00[a]		
Zn^{2+}	0.0204 ± 0.0088	2.89 ± 0.07		−45.1 ± 2.5	−31.6 ± 0.1		−29.7 ± 0.3	−34.7 ± 0.2		−15.4 ± 1.6	3.1 ± 0.6		0.61 ± 0.01	1.85 ± 0.02	
La^{3+}	0.248 ± 0.073	5.77 ± 0.50		−38.0 ± 1.3	−30.1 ± 1.0		1.9 ± 0.4	6.9 ± 0.1		−39.9 ± 1.4	−37.0 ± 1.0		0.70 ± 0.04	3.02 ± 0.05	
Ce^{3+}	3.11 ± 0.36			−31.5 ± 0.5			2.5 ± 0.5			−5.7 ± 0.9			1.80 ± 0.02		
Nd^{3+}	2.4 ± 0.28			−32.1 ± 0.5			7.4 ± 0.2			−39.5 ± 0.2			2.50 ± 0.02		
Tb^{3+}	24.1 ± 14.4	35.2 ± 20.9	182 ± 109	−26.4 ± 0.6	−25.5 ± 0.6	−21.5 ± 0.7	8.5 ± 3.1	39.8 ± 20.7	−6.8 ± 16.9	−34.9 ± 0.5	−65.3 ± 5.6	−14.7 ± 5.9	n.a.	n.a.	n.a.
Dy^{3+}	3.31 ± 1.31			−31.8 ± 1.4			−0.9 ± 0.1			−30.9 ± 1.6			1.34 ± 0.06		

[a] Affinity was too low to allow the determination of stoichiometry. Hence, N was manually set to be 1.0 since Mg^{2+} and Mn^{2+} are neither strong thiophile, oxophile, nor strong reducing agents.[75,79] n.a.: not applicable. The Isotherm was fitted as three sequential binding sites; hence, no binding stoichiometry was obtained.

The K_d of Ca^{2+} was determined to be 79 ± 17 μM which is in good accordance with the literature (K_d: 55 ± 7 μM).[32] The Stoichiometry is near one what was expected based on solution NMR structures showing one Ca^{2+} to be coordinated in the calcium-binding site.[27] The less activating ions and Mg^{2+} Mn^{2+} also bind to SrtA. Whereas Mg^{2+} binds less affine (K_d: 276 ± 19 μM), the released heat of Mn^{2+} (ΔH°: −2.8 ± 1.4 kJ·mol^{-1}) was too low to be adequately fitted, resulting in larger errors (K_d: 108 ± 63 μM). For the thermogram of the titration of Zn^{2+} to SrtA WT, two specific binding events were observed at stoichiometries of N_1: 0.6 and N_2: 1.9 (Table 1, Figure S2F). Interestingly, binding affinities were higher than those of Ca^{2+} with 20 ± 9 nM and 2.9 ± 0.1 μM, respectively. In the case of La^{3+} and the lanthanides, more sequential binding events could be observed that are more or less pronounced and could not be fitted in each case (Table 1, Table S4, Figure S2F). For the tested lanthanides, the determined K_d of the binding event with the highest affinity were all below those of Ca^{2+} (La^{3+}: 0.25 ± 0.07 μM, Ce^{3+}: 3.1 ± 0.4 μM, Nd^{3+}:2.4 ± 0.3 μM, Tb^{3+}: 24 ± 14 μM, and Dy^{3+}: 3.3 ± 1.3 μM). The thermodynamic binding profiles of the activating ions Mg^{2+} and Ca^{2+}, are enthalpy driven, whereas the lanthanides all have a dominant negative entropic contribution to the binding energy (Figure 5).

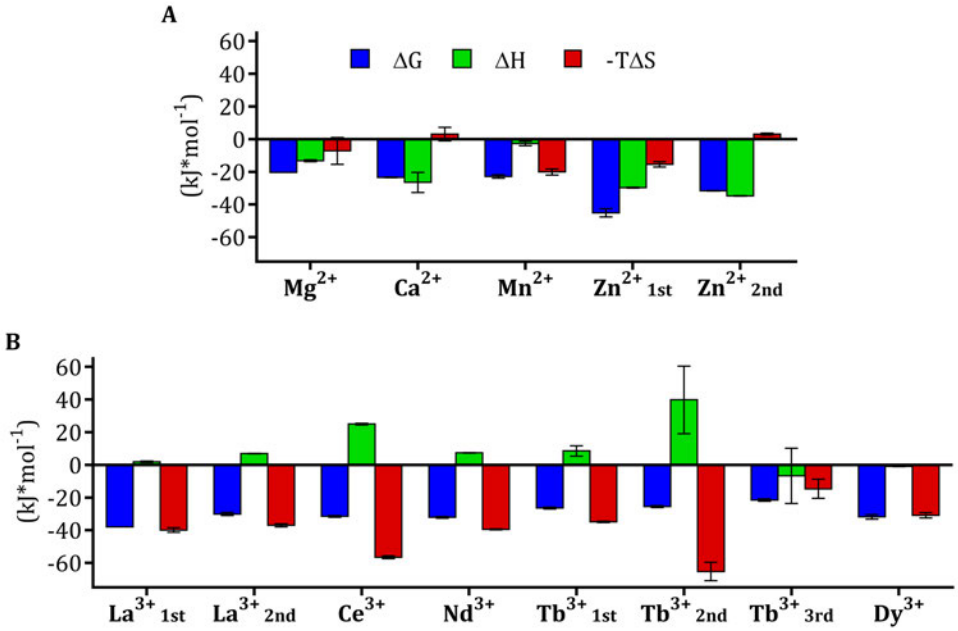


Figure 5. *Signature blots of metal ions vs. SrtA displaying the enthalpic (green) and temperature-dependent entropic (red) contributions to the overall binding energy (blue)*. Errors are indicated as black bars. The figure was created using GraphPad Prism 7.0.4.[69]

Mg^{2+}, Ca^{2+} and Zn^{2+} binding to SrtA WT in presence of the covalent reversible inhibitor cpd **3** was performed to get information on cooperativity. The K_d of Ca^{2+} was reduced 4-fold to 19 ± 0.6 µM. However, in presence of the small molecule inhibitor **3**, the K_d was less reduced than reported in literature where larger thioester analogue of SrtA with a much larger peptide was generated (SrtA-QALPECG-NH_2).[32] Interestingly, ΔH° of Ca^{2+}-binding was improved. In contrast, a high entropic penalty was determined (with cpd **3**: ΔH°: −56 ± 1, −TΔS°: 29 ± 3, Table 1, S4). This fits the assumption that the simultaneous binding of Ca^{2+} and the LPXTG binding motif of SrtA leads to the pronounced rigidification of the flexible loop, which is not observed with solely bound LPXTG or Ca^{2+} (Figure 1). Similar effects were observed with Mg^{2+}; a 3-fold reduced K_d of 97 ± 14 µM was yielded (Table S4). In contrast, the K_d of both binding events of Zn^{2+} were increased by a factor of two upon adding cpd **3**, indicating negative cooperativity between the substrate-mimetic inhibitor and Zn^{2+} which is in line with the SrtA inhibiting properties of Zn^{2+}. Due to the high thiophilicity of Zn^{2+}, it is possible that binding to the active site Cys184 occurs.[75] Therefore, the active site mutants of SrtA replacing the catalytic Cys by Ser (SrtA C184S) or Ala (SrtA C184A) were generated, expressed, and purified. Enzymatic assays confirmed that these mutants did not possess catalytic activity (Figure S3). ITC experiments conducted with both mutants did not show the first binding event. Hence, it can be concluded that the first binding event of Zn^{2+} was indeed binding to the active site Cys184. Since binding to the active site Cys leads to inhibition of SrtA, it remains unclear whether the binding of Zn^{2+} to the Ca^{2+} binding site is activating or inhibiting. ITC experiments with the lanthanides to the active site mutants did not systemically alter their binding behavior (Table S4, Figures S2, S4).

Concentration Dependency. Next, the effects of different concentrations of the lanthanide ions were analyzed in the enzymatic activity assays. Therefore, dilution series of the SrtA stimulating ions La^{3+}, Ce^{3+}, Pr^{3+}, and Nd^{3+} were added to the reaction mixture without Ca^{2+}. The optimal concentration was 5 mM for La^{3+} and 1.5 mM for Ce^{3+}, Pr^{3+}, and Nd^{3+} (Figure 6). Notably, La^{3+} reached even higher enzymatic activity than the 5 mM Ca^{2+} reference, and Ce^{3+} reached similar activities as Ca^{2+}. Higher concentrations reduced the enzymatic activity.

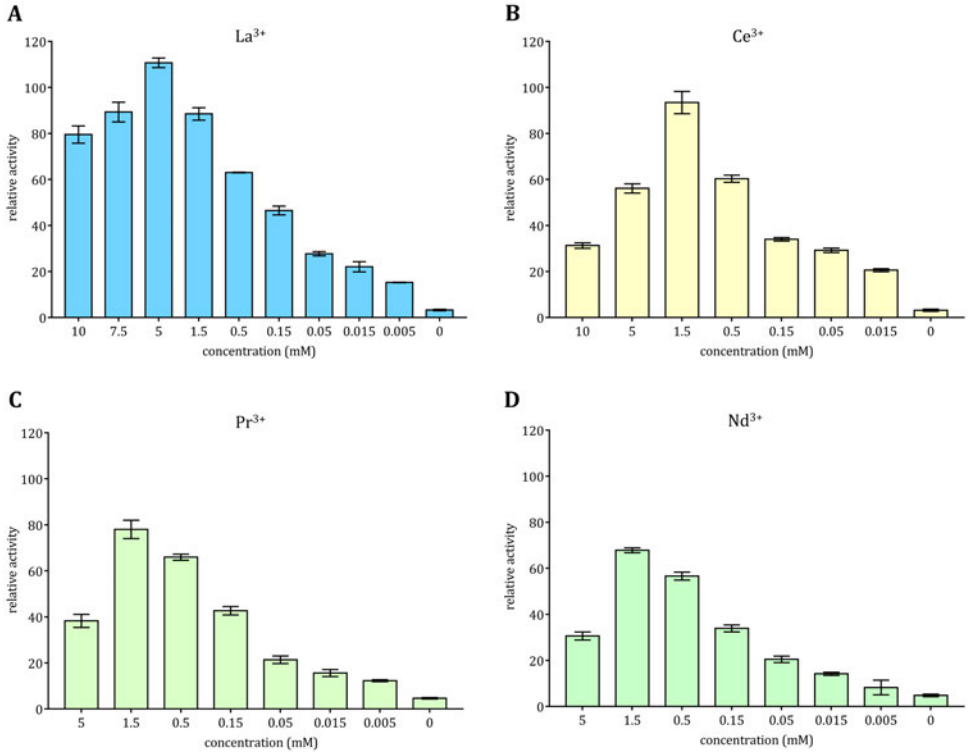


Figure 6. *Concentration-dependent activation of SrtA by lanthanides La^{3+}, Ce^{3+}, Pr^{3+}, and Nd^{3+} (**A–D**)*. The figure was created using GraphPad Prism 7.0.4.[69] Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[68]

This prompted us to also test the inhibitory lanthanides as well as Sc^{3+} and Y^{3+} for their stimulating activity at lower concentrations (Figure S5). All tested ions were found to activate SrtA at lower concentrations but with reduced intrinsic activity with higher atomic numbers. The maximal activity being shifted to lower concentrations for Tm3+, Yb3+, and Lu3+ suggests that lanthanides can bind to the Ca^{2+} binding site and act as partial agonists of SrtA. Sc^{3+} and Y^{3+} acted similarly to lanthanides with maximal activities at 1.5 mM. This finding is in line with other studies finding lanthanides to compete and party substitute for Ca^{2+} functions.[80,81]Next, the concentration-dependent effects of the highly inhibitory lanthanides and Sc3+ and Y3+ on SrtA activity in the presence of 5 mM Ca^{2+} and 1 mM Ca^{2+} was investigated (Figures S6, S7). In 5 mM Ca^{2+}, several lanthanides exhibited synergistic effects of Ca^{2+} on SrtA activity. Particularly 150 µM Ho^{3+} achieved an ~1.49-fold increased activity compared to 5 mM Ca^{2+} alone, followed by Tm^{3+}, Lu^{3+}, Dy^{3+}, and Gd^{3+} (1.43, 1.37, 1.34, and 1.34-fold activity, respectively). Interestingly, this synergistic effect was less pronounced in 1 mM Ca^{2+}.

Terbium-FRET Assay. The gain in enzymatic activity upon lanthanide addition suggests binding to the Ca^{2+} binding site. However, ITC experiments revealed multiple binding events. Therefore, biophysical evidence for the binding of lanthanides to the Ca^{2+} binding site is needed. To this means, we employed a Terbium-FRET assay similar to previous works.[81] The absorption maxima of Tb^{3+} are at 292 nm and 362 nm, and a local emission maximum is at 544 nm (Figure S8B, C). When tryptophane and tyrosine residues are excited (280–282 nm), their emission (348 and 303 nm) allows FRET to occur to Tb^{3+} in proximity. In SrtA, Tyr153 is the closest FRET partner with 10.9 Å (Figure S8A). Hence, the Tb^{3+} emission at 544 nm can be used to measure the amount of bound Tb^{3+} directly. In this time-resolved (tr-)FRET experiment, the emission signal is detected with a delay upon excitation in contrast to conventional steady-state fluorescence.

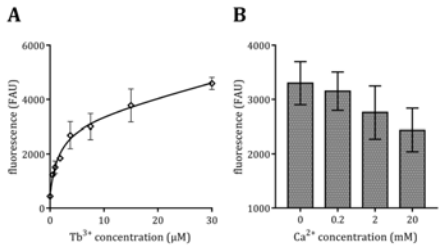


Figure 7. *Tb^{3+}-FRET assay against SrtA.* (**A**) The binding curve of a Tb^{3+} dilution series allows the determination of a K_d of 1.70 ± 0.76 µM. (**B**) Increasing concentrations of Ca^{2+} at a constant Tb^{3+} concentration led to lower Tb^{3+}-FRET signals, demonstrating displacement of Tb^{3+} from the Ca^{2+} binding site.

Since the excitation light pulse is shorter than the decay time of the fluorescence signal from the long-life Tb^{3+}, this allows for the separation of the FRET signal from the scattered light of the excitation. Lanthanides have long fluorescence lifetimes (in the range of milliseconds) due to their low absorption coefficients and slow emission rates. From measurements of the Tb^{3+} dilution series with SrtA, the K_d of Tb^{3+} was determined to be 1.70 ± 0.76 µM (Figure 7A), similar to the K_d values of La^{3+}, Ce^{3+}, Nd^{3+},.and Dy^{3+}, derived from ITC. The lowest K_d value of Tb^{3+} from ITC was 8-fold higher. However, this discrepancy can arise from fitting three binding events and relatively high errors. In the Ca^{2+} competitive experiment, Ca^{2+} is added in a dilution series to displace Tb^{3+} from the binding site, thereby increasing the distance from FRET partners hence reducing the FRET signal. The results demonstrate that Tb^{3+} is displaced by Ca^{2+}, proving the binding of Tb^{3+} to the Ca^{2+} binding site (Figure 7B).

Protein Crystallization. Co-crystallization experiments were performed for SrtA WT in the apo-state, and complexed only by cationic ligands ($MgCl_2$, $CaCl_2$, $MnSO_4$, $ZnCl_2$, $ZrOCl_2$, La^{3+}, Sm^{3+}, or Tb^{3+}), by cpds **2** or **3**, or both. Exclusively crystals with both cpd **3** and a metal ion (Ca^{2+}, $MgCl_2$, Mn^{2+}, $ZrOCl_2$, and La^{3+}) were obtained. Based on previous findings that the cooperative cofactor and substrate binding lead to a more structured intrinsically disordered loop of SrtA[28], it is conclusive that a better-defined and less flexible conformation of SrtA favors crystallization. In a previous study, we showed cpd **2** to form a disulfide bridge with the active site Cys184 of SrtA, thereby only transferring

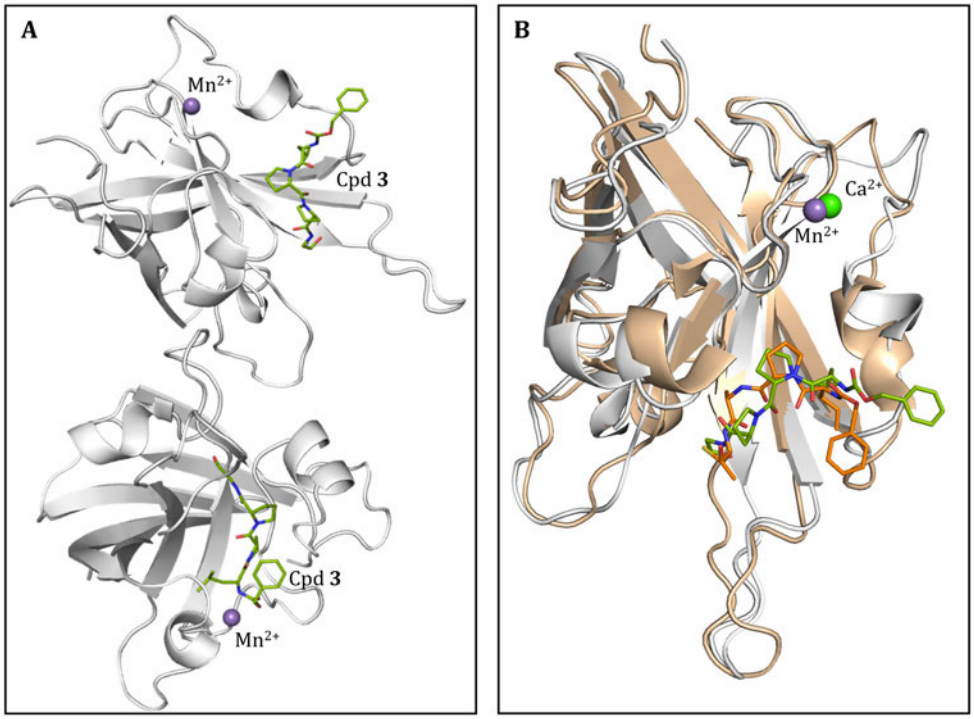


Figure 8. *X-ray structure of SrtA in complex with Mn^{2+} (violet sphere) and cpd **3** (green sticks).* (**A**) In the asymmetric unit, SrtA (white cartoon) assembles as a dimer. (**B**) Superimposed structure of fully complexed SrtA (pale orange cartoon) with Ca^{2+} (bright green sphere) and a substrate-mimetic compound (orange sticks) with our structure. The figure was created using Pymol.[35]

a methanethiol group.[20] Although completely attenuating SrtA activity, we hypothesize that this small group cannot rigidize SrtA sufficiently to enhance crystallization. From the obtained crystals, only the structure of SrtA in the presence of $MnCl_2$, diffracting to 3.0 Å could be solved. SrtA crystallized as a dimer (Figure 7A) with an overall structure resembling the fully complex structure with PDB-ID 2KID. In our structure, an Mn^{2+} ion placed in the Ca^{2+} binding site, cpd **3** is covalently bound to the active site Cys184 and mimics the sorting signal (Figure 7).

SrtA Dimerization. Based on previous findings, SrtA forms homodimers in vitro that are 6–8-fold more active than the monomer. Since the K_d of ~55 μM for dimerization at the interaction interface formed by Asn123, Lys137, and Tyr143 is low,[82,83] we came up with a different strategy to achieve higher SrtA activities. *Via* Gibson assembly, we sought to connect the C- and N-termini of SrtA with the flexible glycine-rich linker G_4SG_4 to enhance dimer formation. However, the primers self-assembled so that two different constructs were obtained. One construct was connected with a tetraglycine linker, and one with a very long linker, consisting of $G_4SG_5SG_2SG_5SG_5SG_2SG_2SG_5SG_2SG_5SG_2SG_4$. Both mutants could be expressed and purified using our standard protocol. Activity assays at 5 μM of the SrtA mutants revealed activities as high as 200% compared to the monomeric SrtA WT (Figure S3). Hence, the activity was neither increased nor decreased upon dimerization. In the case of the tetraglycine linked SrtA dimer, the linker is too short to allow dimerization of the linked monomers, as observed in the crystal structures. Nevertheless, the local concentration of SrtA should be increased, at least resulting in a slightly higher dimer population. The long-linked SrtA dimer should allow the dimerization of the two monomers. However, although no increased activity could be obtained, the two constructs retained full activity and could probably be used when SrtA function is desired, and higher molecular weights can be exploited for the separation of SrtA after the reaction.

CONCLUSION

In this work, we were dedicated to elucidating the influence of a set of 47 cations on SrtA activity (Table S1). Using a FRET based activity assay, some activating cations were identified. La^{3+} even outperformed the activating properties of Ca^{2+} with a relative activity of 111%, followed by Mn^{2+}, yielding an activity of 42%. In Ca^{2+} competitive assays, several ions were found to inhibit SrtA activity. Further investigations of the binding mechanism revealed the most potent inhibitors, Hg^{2+}, Y^{3+}, Zn^{2+}, Ru^{3+}, and several lanthanides, to act *via* unspecific mechanisms. Thorough investigations revealed Ru^{3+} as a quencher for the assay readout, Hg^{2+}, and Zn^{2+} to likely bind to the active site Cys184. Sc^{3+}, Y^{3+}, and the lanthanides are likely to aggregate SrtA at higher concentrations, as shown by MST, probably leading to the observed inhibition. ITC experiments allowed the determination of thermodynamic binding profiles of selected cations and revealed multiple binding sites for the tested lanthanides. Noteworthy, the K_d values for the activating cations Mg^{2+}, Ca^{2+}, Mn^{2+}, and La^{3+} were in line with the activating effects. La^{3+} and the other lanthanides showed K_d values in the low micromolar range, whereas the natural cofactor Ca^{2+} binds with a K_d of ~75 μM. Both the affinity of Mg^{2+} and Ca^{2+} was decreased when SrtA was incubated with cpd **3**, highlighting the cooperativity of SrtA activation cations with the sorting signal. Zn^{2+} showed two binding events, one with a K_d of ~20 nM and one with ~2.9 μM. The high affine binding event vanished when the active site Cys184 was mutated to Ser or Ala, suggesting binding to Cys184 due to the high thiophilicity of Zn^{2+}. It is not uncommon for lanthanides to compete with Ca^{2+} for its binding sites with high affinities,[84,85] thereby functionally substituting Ca^{2+} or inhibiting protein function.[80] Also, binding of lanthanides at adventitious sites was reported as in the case of *Bacillus subtilus* PyrR where Sm^{3+} could bind in the oligomerization interface of the protein[86] or, in the case of Yb^{3+}, binding to only two carboxylate side chains of cadherin NCD1.[87] Previously, Ca^{2+}-displacement from the amoebic protein *Eh*CaBP by the lanthanides Tm^{3+} and Yb^{3+} was fitted by a model of four sequential binding sites.[85] Therefore, we investigated the concentration-dependent effects of lanthanides on the SrtA activity and found all of them to partly activate SrtA before abolishing SrtA activity at higher concentrations. To provide evidence for the binding of lanthanides to the Ca^{2+} binding site, we employed a Tb^{3+}-FRET assay, demonstrating the binding of Tb^{3+} to SrtA and displacement of Tb^{3+} by Ca^{2+}, proving binding to the Ca^{2+} binding site. X-ray co-crystallization was used to solve the structure of cpd **3**, and Mn^{2+} bound SrtA in a conformation similar to that of Ca^{2+} bound SrtA.

Previous studies investigated several metal cations for their activity in preventing biofilm formation. Among them, Cu^{2+}, Ag^{+}, Ti^{3+}, Ga^{3+}, Al^{3+}, Ni^{2+}, and Zn^{2+} were reported to eradicate established biofilms of *S.*

aureus.[63,88] This study suggests that Cu^{2+} and Zn^{2+} could act as inhibitors of SrtA, probably by binding the catalytic Cys184. Ni^{2+} and Al^{3+} also inhibited SrtA, but to a lower extent. Thus, they were not further investigated. Ag^{+} was unsuitable for our experiments as it forms hydroxides and precipitates in our experimental conditions but could also act as a thiophile.[52] In line with previous studies reporting certain metals to enhance the antibacterial spectrum of antibiotics,[89,90] our findings can contribute to developing new anti-microbial materials by releasing cations that inhibit SrtA activity. To further increase the selectivity against SrtA, one could think of metal chelators in the LPXTG motif loaded with SrtA-inhibiting cations. The X could be replaced, and the desired aa, such as histidine residues, could flank the sorting signal. Upon cleavage by SrtA, the ions would be released. Additionally, La^{3+} could prove valuable for bioengineering purposes by increasing the SrtA activity.

Material and methods

Site Directed Mutagenesis of SrtA C184A and C184S Mutants and Dimerized SrtA. Mutations of the WT SrtA constructs active site cysteine (Cys184) were introduced through Quikchange mutagenesis using polymerase chain reaction (PCR) with the Kapa HiFi PCR kit (KapaBiosystems) according to the manufactorers instructions using the following primers: (i) C184A: 5′-CATTAATTACT**GCT**GATGATTACAATGAAAAGACAGGC-3′ and 5′-CATTGTAATCATC**AGC**AGTAATTAATGTTAATTGTTTATCTTTAC-3′, (ii) C184S: 5′-CATTAATTACT**TCT**GATGATTACAATGAAAAGACAGGC-3′ and 5′-ATTGTAATCATC**AGA**AGTAATTAATGTTAATTGTTTATCTTTAC-3′ (Integrated DNA Technologies, Coralville, USA). Dimerization of SrtA via a flexible glycine rich linker (G4SG4) was attempted by Gibson assembly using following primers (iii) vector backbone: 5′-TACAGAAGTCAAACTCGAGCACCACCACCAC-3′ and 5′-ACCGCCGCTACCGCCACCGCCCTCGAGTTTGACTTCTGTAGCTACAAAG-3′, (iV) insert 5′-GGCGGTAGCGGCGGTGGCGGTATGCAAGCTAAACCTCAAATTCCG-3′ and 5′-CTCGAGTTTGACTTCTGTAGCTACAAAG-3′. The success of PCR was assessed after Dpn1 digestion (3 h, 37 °C) by Agarose Gel electrophoresis (GelRed® staining Biotium Inc., CA, USA). The introduction of the desired mutation was confirmed with Sanger sequencing performed by StarSEQ GmbH (Mainz, Germany).

Recombinant Sortase A Expression and Purification. Expression of the *S. aureus* SrtA was mainly performed as described previously.[20,92] A pET-23b expression construct, containing the codon optimized sequence for SrtA (GenBank ID: MBE7584187.1) with an uncleavable C-terminal hexahistidine (His_6) tag was received from the group of Prof. Schwarzer (University of Tübingen, Germany).[20] For protein expression, it was transformed in competent ($CaCl_2$-method) *E. coli* strain BL21-Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA) and grown in a 100 µg/mL ampicillin containing LB medium under agitation at 37 °C to an OD_{600} of ~0.7. Expression was induced by adding 1 mM isopropyl-D-thiogalactoside (IPTG) and incubating under agitation for 16 h at 20 °C. After harvesting, pelleted cells were resuspended in lysis buffer (20 mM Tris•HCl pH 6.9, 300 mM NaCl, 0.1% Triton X_{-100}, RNase, DNase, lysozyme) and lysed by multiple cycles of sonication (Sonoplus, Bandelin, Berlin, Germany). The lysate was cleared by centrifugation (45 min at 15 krpm). The protein was purified from the supernatant by immobilized metal affinity chromatography (IMAC) on a HisTrap HP 5 mL column (GE Healthcare, Chicago, Illinois), washed with 20 column volumes (CV) HisTrap buffer A (20 mM Tris•HCl pH 7.2, 300 mM NaCl, 30 mM imidazole) and eluted with 20 CV of a linear gradient 0-100% with HisTrap buffer B (20 mM Tris•HCl pH 7.2, 500 mM NaCl, 300 mM imidazole). Eluted proteins were subsequently subjected to a gel-filtration step (HiLoad 16/600 Superdex 75 pg column, GE Healthcare) and eluted in storage buffer (20 mM Tris•HCl pH 7.50, 150 mM NaCl) either with or without 5 mM $CaCl_2$ as required for different downstream measurements and indicated below. Purified proteins were concentrated, shock frozen in liquid nitrogen and stored at –80 °C until further usage. Throughout all steps, protein concentrations were measured *via* absorbance at 280 nm using a NanoDrop™ 8000 Spectrophotometer (Thermo Scientific™ Waltham, Massachusetts). Sample purity was assessed *via* Coomassie brilliant blue stained SDS-PAGE.

SrtA Activation Assay. Transpeptidation efficacy of isolated SrtA was performed in vitro, as described previously.[93,94] Briefly, assay buffer (50 mM Tris•HCl pH 7.50, 150 mM NaCl) was supplemented with the FRET-pair substrate Abz-LPETG-Dap(Dnp)-OH at 25 µM (Genscript, New Jersey, NY, USA) and 0.5 mM tetraglycine (Sigma Aldrich, St.

Louis, MO, USA) as an acceptor for the transpeptidase reaction. The FRET-pair substrate was added from a stock solution in DMSO into black flat-bottom 96-well plates (Greiner Bio-One, Kremsmünster, Austria). Cations were added from 20 mM Stock, resulting in 5 mM. For better comparability, all assayed cations were used as chloride salts (Table S1). Negative controls were performed without adding $CaCl_2$. The reaction was initiated by adding SrtA at a final concentration of 1 µM and monitored for 30 min at 25 °C in an Infinite M200 Pro plate reader with λ_{ex} 320 nm and λ_{em} 430 nm (Tecan, Männedorf, Switzerland). Assays were carried out in technical triplicates. The enzyme kinetics were analyzed as described previously.[95] The competitive assay was performed similarly to the activation assay but with additional 5 mM $CaCl_2$ in each reaction. The control measurement was mock treatment with assay buffer instead of further cation chloride enriched assay buffer. Quenching control measurements were performed with Abz (1.25, 2.5, 5, and 10 µM) instead of the FRET-pair substrate and 5 mM of the cations. The slopes of the linear regression of the resulting Abz-concentration vs. fluorescence curves with and without added metal ions were divided to calculate the QF of the respective cations.

Microscale Thermophoresis. A Monolith NT.115 Pico instrument (NanoTemper Technologies, Munich, Germany) was utilized to perform MST experiments. Purified His_6-tagged SrtA was labeled with the Monolith Protein Labeling Kit RED-NHS (NanoTemper Technologies) according to the manufacturer's instructions. MST sample preparation was conducted as described by the manufacturer's standard protocols. In brief, the labeled SrtA was used in a final concentration of 5 nM diluted in MST buffer (50 mM Tris, 150 mM NaCl, 5 mM CaCl2, 0.1% (v/v) Tween-20, 10% DMSO, pH 7.5). Addition of 5 µM unlabeled SrtA diminished adhesion to the capillaries. For the initial screening of SrtA binding, metal salts were used at a final concentration of 5 mM. K_d measurements were attempted with 1:1 dilution series of the cations starting with final concentrations of 400 µM. Measurements were performed in triplets at 25 °C, at 40% MST-power, using the Pico - RED excitation color and Monolith NT.115 Capillaries (NanoTemper Technologies). The obtained data were processed using the MO.Affinity Analysis software Version 2.3 (NanoTemper Technologies). K_d measurements were performed with a 1:1 dilution series in 8 capillaries, starting with 400 µM of the respective metal ion as the highest final concentration.

Isothermal Titration Calorimetry. All ITC experiments were performed in technical triplicates on a MicroCal PEAQ-ITC (Malvern Panalytical, Kassel, Deutschland). Briefly, SrtA (SrtA WT, SrtA C184A, and SrtA C184S) were subjected to buffer exchange into ITC buffer (50 mM 2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethansulfonic acid (HEPES) pH 7.5, 150 mM NaCl) with Amicon® Ultra 10 K centrifugal filters (Millipore, Billerica, USA). The first four flowthroughs were discarded, followed by four centrifugation steps, where the flowthrough was collected. After determining protein concentration using absorbance at 280 nm on a NanoDrop™ 2000 spectrophotometer (Thermo Scientific™ Waltham, Massachusetts), and the absorption coefficient calculated with the ExPASy ProtParam tool[96], SrtA concentration was adjusted with flowthrough to 50 µM. For titrations against SrtA in the presence of an inhibitor, cpd **3** was preincubated with SrtA, and the cation chlorides were dissolved in flowthrough to 1 mM. Titration experiments were performed with 19 injections at 25 °C. Control experiments (buffer vs. buffer; SrtA vs. buffer; buffer vs. cation) were performed for each experiment and subtracted from the raw data.

Tb^{3+}-FRET Assay. All experiments were conducted in black flat-bottom 96-well plates (Greiner Bio-One, Kremsmünster, Austria) using an Infinite M200 Pro plate reader (Tecan, Männedorf, Switzerland) with λ_{ex} 280 nm. Emission was recorded at λ_{em} 545 nm with a delay of 100 µs and an integration time of 2000 µs. SrtA WT was used at a final concentration of 5 µM in assay buffer (50 mM Tris•HCl pH 7.50, 150 mM NaCl). To determine the K_d of Tb^{3+} binding to the Ca^{2+} binding site, Tb^{3+} was added in a 1:1 dilution series of 30 µM, 15 mM, 7.5 µM, 3.8 µM, 1.9 µM, 0.94 µM, 0.47 µM, and 0 µM. Since no correction was performed for the fluorescence of Tb^{3+} without SrtA, the K_d was fitted as specific binding with an unspecific linear offset using GraphPad Prism 7.0.4.[69] For subsequent displacement experiments with Ca^{2+}, Tb^{3+} was used at 10 µM for sufficient saturation. Ca^{2+} was added in concentrations of 20 mM, 2 mM, 0.2 mM, and 0 µM.

Protein Crystallization. According to previously published conditions, an initial 24 well grid screening was prepared, containing 0.1 M NaCl, 25 mM 2-(*N*-morpholino)ethanesulfonic acid (MES), varying in pH (6.0, 6.1, 6.35, 6.5) and ammonium sulfate concentration (2.9–3.4 M).[31] Both SrtA purified with or without $CaCl_2$, stored at –

80 °C, thawed on ice, and concentrated in Amicon® Ultra 10 K centrifugal filters (Millipore, Billerica, USA) to >60 mg/mL. The flowthrough was collected. The concentration was determined using absorbance at 280 nm and an absorption coefficient calculated with the ExPASy ProtParam tool[96] on a NanoDrop™ 2000 spectralphotometer (Thermo Scientific™ Waltham, MA, USA). For co-crystallization experiments, inhibitor **3** was added to SrtA from a 50 mM stock in DMSO in a molar 1 to 1.1 excess (3.09 mM) and incubated for 1 h on ice before adjusting protein concentration to 50 mg/mL with flowthrough. This resulted in four protein samples: (i) apo-SrtA, (ii) SrtA + Ca^{2+}, (iii) SrtA + cpd **3**, (iv) SrtA + Ca^{2+} + cpd **3**. For hanging drop vapor diffusion crystallization, 700 µL precipitant solution was placed in the reservoirs of 24 well comboplates (Greiner Bio-One, Kremsmünster, Austria). Protein solutions were mixed with the precipitant solution in a 1 µL to 1 µL ratio on a glass cover slide and placed tightly on the corresponding well. Crystals were grown for 1 week at rt. In all attempts to crystallize apo-protein SrtA, no crystal could be obtained. Instead, crystals grew preferably in simultaneous presence of both Ca^{2+} and an active site inhibitor, whereas the absence of one of the ligands led to gelatinous phase separation (Figure S9). After initial hits, further 12 well grid screens were prepared, containing 0.1 M NaCl, 25 mM MES with varying pH (5.8, 5.9, 6.0, and 6.1) and ammonium sulfate concentration (2.8, 2.9, and 3.0 M) with additional 20 mM of either $CaCl_2$, $ZnCl_2$, $MnSO_4$, $MgCl_2$, $ZrOCl_2$, La^{3+}, Sm^{3+}, or Tb^{3+} and with either cpd **2** or **3**. After adjusting the pH, a white flocculent precipitate of the Zr^{4+} containing solutions formed and was removed by centrifugation. In the case of Zn^{2+}, precipitation occurred after mixing with protein solution at higher pH values. Only crystals with cpd **3** were obtained in presence of $MgCl_2$, $ZrOCl_2$, and La^{3+} in 5–14 d at room temperature.

Data Collection and Structure Refinement. The harvested crystals were preincubated for ~30 s in the respective precipitant buffer with an additional 16% (v/v) ethylene glycol before flash freezing in a 100 KN2 cryo stream at the Institute of Molecular Physiology (JGU, Mainz, Germany). X-ray diffraction data were collected using the beamline of a AXS Microstar-H generator (Bruker Nano GmbH, Berlin, Germany) and a MAR345 image-plate detector distance of 350 mm and a wavelength of 1.5417 Å in 1°-steps after 7 min exposition time. Data were processed with XDS and CCP4.[97,98] Crystals belonging to the 173, P 63 space group, with a = b = 146.209 Å, c = 47.675 Å. One dimer is present in the asymmetric unit. For structure determination, a previously determined structure of SrtA (PDB: 1T2W) was used as a model structure for molecular replacement using PhaserMR of the phenix software package (https://phenix-online.org/) after the removal of the ligand and all water molecules.[99,100] Final refinement was performed with WinCoot.[101]

Chemistry. In general, cations were either purchased as chloride salt or converted into their chlorides by precipitation as hydroxide with NaOH, decanting the supernatant and dissolving in diluted HCl. Lanthanides were purchased from Nova Elements s.a.s. (Palermo, Italy) with >99.95% purity in the native state and dissolved in diluted HCl. The chlorides were obtained after lyophilization of the solutions (Table S1). Cpds **2** and **3** were synthesized as described previously (see SI Inhibitor synthesis).

Abbreviations

methicillin-resistant *staphylococcus aureus* (MRSA), *Staphylococcus aureus* (*S. aureus*), Sortase A (SrtA), microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), intrinsically disordered region (IDR), temperature (T_m), differential scanning calorimetry (DSC), nuclear magnetic resonance (NMR), *Pseudomonas aeruginosa* (*P. aeruginosa*), RCSB protein data bank (PDB), wild-type (WT); differential scanning fluorimetry (nanoDSF), microscale thermophoresis (MST), isothermal titration calorimetry (ITC), size exclusion chromatography (SEC); Corey–Pauling–Koltun (CPK), anthranilic acid (Abz), quenching factor (QF), time-resolved (tr), polymerase chain reaction (PCR), isopropyl-D-thiogalactoside (IPTG), immobilized metal affinity chromatography (IMAC), column volumes (CV), 2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethansulfonic acid (HEPES), 2-(*N*-morpholino)ethanesulfonic acid (MES), 2,3-diaminopropionicacid (Dap), 2,4-dinitrophenol (Dnp)

Acknowledgements

We thank [redacted] and [redacted] for their contributions to the project during their bachelor thesis.

Literature

1. Klevens, R. M. *et al.* Invasive methicillin-resistant Staphylococcus aureus infections in the United States. *J. Am. Med. Assoc.* **298**, 1763–1771 (2007).

2. Zhen, X. *et al. Antibiotic resistance threats in the United States, 2019*. *Anti-microbial Resistance & Infection Control* vol. 10 https://aricjournal.biomedcentral.com/articles/10.1186/s13756-020-00872-w (2019).

3. Høiby, N., Bjarnsholt, T., Givskov, M., Molin, S. & Ciofu, O. Antibiotic resistance of bacterial biofilms. *Int. J. Antimicrob. Agents* **35**, 322–332 (2010).

4. Flemming, H. C. *et al.* Biofilms: An emergent form of bacterial life. *Nat. Rev. Microbiol.* **14**, 563–575 (2016).

5. Mazmanian, S. K., Liu, G., Ton-That, H. & Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science (80-.).* **285**, 760–763 (1999).

6. Schneewind, O. & Missiakas, D. Sortases, surface proteins, and their roles in Staphylococcus Aureus disease and vaccine development. *Protein Secret. Bact.* **7**, 173–188 (2019).

7. Tsompanidou, E. *et al.* The Sortase A Substrates FnbpA, FnbpB, ClfA and ClfB Antagonize Colony Spreading of Staphylococcus aureus. *PLoS One* **7**, 1–7 (2012).

8. Koo, H., Allan, R. N., Howlin, R. P., Stoodley, P. & Hall-Stoodley, L. Targeting microbial biofilms: current and prospective therapeutic strategies. *Nat. Rev. Microbiol.* **15**, 740–755 (2017).

9. Falugi, F., Kim, H. K., Missiakas, D. M. & Schneewind, O. Role of Protein A in the Evasion of Host Adaptive Immune Responses by Staphylococcus aureus. *MBio* **4**, (2013).

10. Wardenburg, J. B., Patel, R. J. & Schneewind, O. Surface proteins and exotoxins are required for the pathogenesis of Staphylococcus aureus pneumonia. *Infect. Immun.* **75**, 1040–1044 (2007).

11. Pishchany, G. *et al.* IsdB-dependent hemoglobin binding is required for acquisition of heme by Staphylococcus aureus. *J. Infect. Dis.* **209**, 1764–1772 (2014).

12. Muryoi, N. *et al.* Demonstration of the iron-regulated surface determinant (Isd) heme transfer pathway in Staphylococcus aureus. *J. Biol. Chem.* **283**, 28125–28136 (2008).

13. Novick, R. P. & Geisinger, E. Quorum Sensing in Staphylococci. *Annu. Rev. Genet.* **42**, 541–564 (2008).

14. Hughes, D. T. & Sperandio, V. Inter-kingdom signalling: communication between bacteria and their hosts. *Nat. Rev. Microbiol.* **6**, 111–120 (2008).

15. Cascioferro, S., Totsika, M. & Schillaci, D. Sortase A: An ideal target for anti-virulence drug development. *Microb. Pathog.* **77**, 105–112 (2014).

16. Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E. & Schneewind, O. Staphylococcus aureus sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5510–5515 (2000).

17. Mu, D., Xiang, H., Dong, H., Wang, D. & Wang, T. Isovitexin, a potential candidate inhibitor of sortase a of staphylococcus aureus USA300. *J. Microbiol. Biotechnol.* **28**, 1426–1432 (2018).

18. Cascioferro, S. *et al.* Sortase A Inhibitors: Recent Advances and Future Perspectives. *J. Med. Chem.* **58**, 9108–9123 (2015).

19. Rasko, D. A. & Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nat. Rev. Drug Discov.* **9**, 117–128 (2010).

20. Barthels, F. *et al.* Asymmetric Disulfanylbenzamides as Irreversible and Selective Inhibitors of Staphylococcus aureus Sortase A. *ChemMedChem* **15**, 839–850 (2020).

21. Jaudzems, K. *et al.* Targeting Bacterial Sortase A with Covalent Inhibitors: 27 New Starting Points for Structure-Based Hit-to-Lead Optimization. *ACS Infect. Dis.* **6**, 186–194 (2020).

22. Ilangovan, U., Ton-That, H., Iwahara, J., Schneewind, O. & Clubb, R. T. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of Staphylococcus aureus. *Proc. Natl. Acad. Sci.* **98**, 6056–6061 (2001).

23. Frankel, B. A., Kruger, R. G., Robinson, D. E., Kelleher, N. L. & McCafferty, D. G. Staphylococcus aureus sortase transpeptidase SrtA: Insight into the kinetic mechanism and evidence for a reverse protonation catalytic mechanism. *Biochemistry* **44**, 11188–11200 (2005).

24. Connolly, K. M. *et al.* Sortase from Staphylococcus aureus Does Not Contain a Thiolate-Imidazolium Ion Pair in Its Active Site. *J. Biol. Chem.* **278**, 34061–34065 (2003).

25. Kruger, R. G., Barkallah, S., Frankel, B. A. & McCafferty, D. G. Inhibition of the

Staphylococcus aureus sortase transpeptidase SrtA by phosphinic peptidomimetics. *Bioorganic Med. Chem.* **12**, 3723–3729 (2004).

26. Scott, C. J. *et al.* Irreversible inhibition of the bacterial cysteine protease-transpeptidase sortase (SrtA) by substrate-derived affinity labels. *Biochem. J.* **366**, 953–958 (2002).

27. Suree, N. *et al.* The structure of the Staphylococcus aureus sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J. Biol. Chem.* **284**, 24465–24477 (2009).

28. Pang, X. & Zhou, H. X. Disorder-to-Order Transition of an Active-Site Loop Mediates the Allosteric Activation of Sortase A. *Biophys. J.* **109**, 1706–1715 (2015).

29. Kruger, R. G., Dostal, P. & McCafferty, D. G. Development of a high-performance liquid chromatography assay and revision of kinetic parameters for the Staphylococcus aureus sortase transpeptidase SrtA. *Anal. Biochem.* **326**, 42–48 (2004).

30. Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of Staphylococcus aureus at the LPXTG motif. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 12424–12429 (1999).

31. Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O. & Narayana, S. V. L. Crystal structures of Staphylococcus aureus Sortase A and its substrate complex. *J. Biol. Chem.* **279**, 31383–31389 (2004).

32. Wang, X., Chen, J. L., Otting, G. & Su, X. C. Conversion of an amide to a high-energy thioester by Staphylococcus aureus sortase A is powered by variable binding affinity for calcium. *Sci. Rep.* **8**, 2–10 (2018).

33. Ugur, I. *et al.* Ca 2+ binding induced sequential allosteric activation of sortase A: An example for iontriggered conformational selection. *PLoS One* **13**, 1–19 (2018).

34. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

35. The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography* vol. 40 82–92 (2002).

36. Wu, Z., Hong, H., Zhao, X. & Wang, X. Efficient expression of sortase A from Staphylococcus aureus in Escherichia coli and its enzymatic characterizations. *Bioresour. Bioprocess.* **4**, (2017).

37. Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E. & Ploegh, H. L. Sortagging: A versatile method for protein labeling. *Nat. Chem. Biol.* **3**, 707–708 (2007).

38. Popp, M. W. L. & Ploegh, H. L. Making and breaking peptide bonds: Protein engineering using sortase. *Angew. Chemie - Int. Ed.* **50**, 5024–5032 (2011).

39. Wu, Z., Guo, X. & Guo, Z. Sortase A-catalyzed peptide cyclization for the synthesis of macrocyclic peptides and glycopeptides. *Chem. Commun.* **47**, 9218 (2011).

40. Van'T Hof, W., Maňásková, S. H., Veerman, E. C. I. & Bolscher, J. G. M. Sortase-mediated backbone cyclization of proteins and peptides. *Biol. Chem.* **396**, 283–293 (2015).

41. Popp, M. W., Dougan, S. K., Chuang, T. Y., Spooner, E. & Ploegh, H. L. Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3169–3174 (2011).

42. Witte, M. D. *et al.* Preparation of unnatural N-to-N and C-to-C protein fusions. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11993–11998 (2012).

43. Chan, L. *et al.* Covalent attachment of proteins to solid supports and surfaces via sortase-mediated ligation. *PLoS One* **2**, 1–5 (2007).

44. Guo, X., Wang, Q., Swarts, B. M. & Guo, Z. Sortase-catalyzed peptide-glycosylphosphatidylinositol analogue ligation. *J. Am. Chem. Soc.* **131**, 9878–9879 (2009).

45. Beerli, R. R., Hell, T., Merkel, A. S. & Grawunder, U. Sortase enzyme-mediated generation of site-specifically conjugated antibody drug conjugates with high In Vitro and In Vivo potency. *PLoS One* **10**, 1–17 (2015).

46. Voloshchuk, N., Liang, D. & Liang, J. Sortase A Mediated Protein Modifications and Peptide Conjugations. *Curr. Drug Discov. Technol.* **12**, 205–213 (2016).

47. Glasgow, J. E., Salit, M. L. & Cochran, J. R. In Vivo Site-Specific Protein Tagging with Diverse Amines Using an Engineered Sortase Variant. *J. Am. Chem. Soc.* **138**, 7496–7499 (2016).

48. Hirakawa, H., Ishikawa, S. & Nagamune, T. Design of Ca 2+-independent Staphylococcus aureus sortase A mutants. *Biotechnol. Bioeng.* **109**, 2955–2961 (2012).

49. Jeong, H. J., Abhiraman, G. C., Story, C. M., Ingram, J. R. & Dougan, S. K. Generation of Ca2+-independent sortase A mutants with enhanced activity for protein and cell surface labeling. *PLoS One* **12**, 1–15 (2017).

50. Percival, S. L. *et al.* A review of the scientific evidence for biofilms in wounds. *Wound Repair Regen.* **20**, 647–657 (2012).

51. Thorarinsdottir, H. R., Kander, T., Holmberg, A., Petronis, S. & Klarin, B. Biofilm formation on three different endotracheal tubes: A prospective clinical trial. *Crit. Care* **24**, 1–12 (2020).

52. Lemire, J. A., Harrison, J. J. & Turner, R. J. Anti-microbial activity of metals: Mechanisms, molecular targets and applications. *Nat. Rev. Microbiol.* **11**, 371–384 (2013).

53. Rzhepishevska, O. *et al.* The antibacterial activity of Ga3+ is influenced by ligand complexation as well as the bacterial carbon source. *Antimicrob. Agents Chemother.* **55**, 5568–5580 (2011).

54. Middaugh, J. *et al.* Aluminum triggers decreased aconitase activity via Fe-S cluster disruption and the overexpression of isocitrate dehydrogenase and isocitrate lyase: A metabolic network mediating cellular survival. *J. Biol. Chem.* **280**, 3159–3165 (2005).

55. Macomber, L. & Imlay, J. A. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8344–8349 (2009).

56. Holt, K. B. & Bard, A. J. Interaction of silver(I) ions with the respiratory chain of Escherichia coli: An electrochemical and scanning electrochemical microscopy study of the anti-microbial mechanism of micromolar Ag. *Biochemistry* **44**, 13214–13223 (2005).

57. Calderón, I. L. *et al.* Tellurite-mediated disabling of [4Fe-4S] clusters of Escherichia coli dehydratases. *Microbiology* **155**, 1840–1846 (2009).

58. Patenge, N. *et al.* Evaluation of anti-microbial effects of novel implant materials by testing the prevention of biofilm formation using a simple small scale medium-throughput growth inhibition assay. *Biofouling* **28**, 267–277 (2012).

59. Ewald, A., Glückermann, S. K., Thull, R. & Gbureck, U. Anti-microbial titanium/silver PVD coatings on titanium. *Biomed. Eng. Online* **5**, 1–10 (2006).

60. Kostenko, V., Lyczak, J., Turner, K. & Martinuzzi, R. J. Impact of silver-containing wound dressings on bacterial biofilm viability and susceptibility to antibiotics during prolonged treatment. *Antimicrob. Agents Chemother.* **54**, 5120–5131 (2010).

61. Gant, V. A. *et al.* Three novel highly charged copper-based biocides: Safety and efficacy against healthcare-associated organisms. *J. Antimicrob. Chemother.* **60**, 294–299 (2007).

62. Borkow, G., Zatcoff, R. C. & Gabbay, J. Reducing the risk of skin pathologies in diabetics by using copper impregnated socks. *Med. Hypotheses* **73**, 883–886 (2009).

63. Gugala, N., Lemire, J. A. & Turner, R. J. The efficacy of different anti-microbial metals at preventing the formation of, and eradicating bacterial biofilms of pathogenic indicator strains. *J. Antibiot. (Tokyo).* **70**, 775–780 (2017).

64. Harrison, J. J., Turner, R. J. & Ceri, H. Persister cells, the biofilm matrix and tolerance to metal cations in biofilm and planktonic Pseudomonas aeruginosa. *Environ. Microbiol.* **7**, 981–994 (2005).

65. Harrison, J. J., Ceri, H., Stremick, C. A. & Turner, R. J. Biofilm susceptibility to metal toxicity. *Environ. Microbiol.* **6**, 1220–1227 (2004).

66. Wang, M. & Tang, T. Surface treatment strategies to combat implant-related infection from the beginning. *J. Orthop. Transl.* **17**, 42–54 (2019).

67. Lepetit, C., Maraval, V., Canac, Y. & Chauvin, R. On the nature of the dative bond: Coordination to metals and beyond. The carbon case. *Coord. Chem. Rev.* **308**, 59–75 (2016).

68. Willighagen, E. L. Processing CML conventions in Java. *Internet J. Chem.* **4**, 1–9 (2001).

69. Software, G. Graphpad prism for windows. (2014).

70. Victor, T. W. *et al.* Lanthanide-Binding Tags for 3D X-ray Imaging of Proteins in Cells at Nanoscale Resolution. *J. Am. Chem. Soc.* **142**, 2145–2149 (2020).

71. Bünzli, J. C. G. Lanthanide Luminescence: From a Mystery to Rationalization, Understanding, and Applications. *Handb. Phys. Chem. Rare Earths* **50**, 141–176 (2016).

72. Ungur, L., Szabo, B., Alothman, Z. A., Al-Kahtani, A. A. S. & Chibotaru, L. F.

Mechanisms of Luminescence in Lanthanide Complexes: A Crucial Role of Metal-Ligand Covalency. *Inorg. Chem.* **61**, 5972–5976 (2022).

73. Eliseeva, S. V. & Bünzli, J. C. G. Lanthanide luminescence for functional materials and bio-sciences. *Chem. Soc. Rev.* **39**, 189–227 (2010).

74. Lei, L. *et al.* Manipulation of time-dependent multicolour evolution of X-ray excited afterglow in lanthanide-doped fluoride nanoparticles. *Nat. Commun.* **13**, 1–11 (2022).

75. Kepp, K. P. A quantitative scale of oxophilicity and thiophilicity. *Inorg. Chem.* **55**, 9461–9470 (2016).

76. Yee, K. K. *et al.* Effective mercury sorption by thiol-laced metal-organic frameworks: In strong acid and the vapor phase. *J. Am. Chem. Soc.* **135**, 7795–7798 (2013).

77. Hsu, J. T. A. *et al.* Evaluation of metal-conjugated compounds as inhibitors of 3CL protease of SARS-CoV. *FEBS Lett.* **574**, 116–120 (2004).

78. Basak, A. *et al.* Enzymic characterization in vitro of recombinant proprotein convertase PC4. *Biochem. J.* **343**, 29–37 (1999).

79. Bard, A. J., Parsons, R. & Jordan, J. *Standard Potentials in Aqueous Solution*. (Routledge, 2017). doi:10.1201/9780203738764.

80. Martin, R. B. & Richardson, F. S. Lanthanides as probes for calcium in biological systems. *Q. Rev. Biophys.* **12**, 181–209 (1979).

81. Brayshaw, L. L., Smith, R. C. G., Badaoui, M., Irving, J. A. & Price, S. R. Lanthanides compete with calcium for binding to cadherins and inhibit cadherin-mediated cell adhesion. *Metallomics* **11**, 914–924 (2019).

82. Zhu, J. *et al.* Single mutation on the surface of Staphylococcus aureus sortase A can disrupt its dimerization. *Biochemistry* **47**, 1667–1674 (2008).

83. Lu, C. *et al.* Staphylococcus aureus Sortase A Exists as a Dimeric Protein In Vitro. *Biochemistry* **46**, 9346–9354 (2007).

84. Pidcock, E. & Moore, G. R. Structural characteristics of protein binding sites for calcium and lanthanide ions. *J. Biol. Inorg. Chem.* **6**, 479–489 (2001).

85. Mustafi, S. M., Mukherjee, S., Chary, K. V. R., Del Bianco, C. & Luchinat, C. Energetics and mechanism of Ca2+ displacement by lanthanides in a calcium binding protein. *Biochemistry* **43**, 9320–9331 (2004).

86. Tomchick, D. R., Turner, R. J., Switzer, R. L. & Smith, J. L. Adaptation of an enzyme to regulatory function: Structure of Bacillus subtilis PyrR, a pyr RNA-binding attenuation protein and uracil phosphoribosyltransferase. *Structure* **6**, 337–350 (1998).

87. Moore, J. D., Skinner, M. A., Swatman, D. R., Hawkins, A. R. & Brown, K. A. Reactivation of 3-Dehydroquinate Synthase by Lanthanide Cations. *J. Am. Chem. Soc.* **120**, 7105–7106 (1998).

88. Wu, C. *et al.* Zinc as an agent for the prevention of biofilm formation by pathogenic bacteria. *J. Appl. Microbiol.* **115**, 30–40 (2013).

89. Harrison, J. J. *et al.* Copper and quaternary ammonium cations exert synergistic bactericidal and antibiofilm activity against Pseudomonas aeruginosa. *Antimicrob. Agents Chemother.* **52**, 2870–2881 (2008).

90. Morones-Ramirez, J. R., Winkler, J. A., Spina, C. S. & Collins, J. J. Silver Enhances Antibiotic Activity Against Gram-Negative Bacteria. *Sci. Transl. Med.* **5**, 1–7 (2013).

91. Chitambar, C. R. & Narasimhan, J. Targeting Iron-Dependent DNA Synthesis with Gallium and Transferrin-Gallium. *Pathobiology* **59**, 3–10 (1991).

92. Schmohl, L., Bierlmeier, J., Gerth, F., Freund, C. & Schwarzer, D. Engineering sortase A by screening a second-generation library using phage display. *J. Pept. Sci.* **23**, 631–635 (2017).

93. Li, N. *et al.* An enzyme-mediated protein-fragment complementation assay for substrate screening of sortase A. *Biochemical and Biophysical Research Communications* vol. 486 (Elsevier Ltd, 2017).

94. Schmohl, L. *et al.* Identification of sortase substrates by specificity profiling. *Bioorganic Med. Chem.* **25**, 5002–5007 (2017).

95. Schirmeister, T. *et al.* Quantum Chemical-Based Protocol for the Rational Design of Covalent Inhibitors. *J. Am. Chem. Soc.* **138**, 8332–8335 (2016).

96. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* 571–607 (2005). doi:10.1385/1-59259-890-0:067.

97. Collaborative Computational Project, N. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763 (1994).

98. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).

99. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).

100. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 861–877 (2019).

101. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

Interactions of staphylococcus aureus Sortase A with various Cations: Basis for the design of novel inhibitors.

Stefan J. Hammerschmidt [a], [redacted]

[a] Institute for Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University of Mainz, Staudinger Weg 5, 55128 Mainz, Germany

[b] Institute of Molecular Physiology, Johannes Gutenberg-University of Mainz, Hans Dieter Hüsch Weg 15, 55128 Mainz, Germany

KEYWORDS: Staphylococcus aureus, sortase A, metal cation, allostery, x-ray structure, isothermal titration calorimetry, anti-virulence

CONTENT

Table S1. Used cation chlorides with annotations concerning their handling.

Salt	Preparation	Comments
LiCl		
NaCl		present in reaction buffers
KCl		
RbCl		
CsCl	From Cs_2CO_3 with HCl	
$MgCl_2$		
$CaCl_2$		
$SrCl_2$		
$BaCl_2$		
$ScCl_3$	From Sc triflate with NaOH precipitation and HCl	
$TiCl_3$		Still needs to be performed
$MnCl_2$		daily fresh prepared
$FeCl_2$		Not fully soluable
$FeCl_3$		Not fully soluable
$CoCl_2$		
$NiCl_2$	From $NiSO_4$ with NaOH precipitation and HCl	
CuCl		Not fully soluable, daily fresh prepared
$CuCl_2$		
$ZnCl_2$		
YCl_3	From native Y with HCl	
$ZrOCl_2$		
$RuCl_3$		Deep black color
$PdCl_2$		Not fully soluable
$CdCl_2$	From $Cd(NO_3)_2$ with NaOH precipitation and HCl	
$HgCl_2$		
$AlCl_3$		
$GaCl_3$		Still needs to be performed
$InCl_3$	From native In with HCl	
$SnCl_2$		
$PbCl_2$		
$SbCl_3$		Not fully soluable
$BiCl_3$		
$LaCl_3$	From native La with HCl	
$CeCl_3$	From native Ce with HCl	
$PrCl_3$	From native Pr with HCl	
$NdCl_3$	From native Nd with HCl	
$SmCl_3$	From native Sm with HCl	
$EuCl_3$	From native Eu with HCl	
$GdCl_3$	From native Gd with HCl	
$TbCl_3$	From native Tb with HCl	
$DyCl_3$	From native Dy with HCl	
$HoCl_3$	From native Ho with HCl	
$ErCl_3$	From native Er with HCl	

$TmCl_3$	From native Tm with HCl
$YbCl_3$	From native Yb with HCl
$LuCl_3$	From native Lu with HCl
NH_4Cl	

Table S2. pH-Values of $RuCl_3$ in assay buffer**.** $RuCl_3$ was used in different concentrations. Since the pH of the assay buffer A breaks with 2 mM $RuCl_3$ the Tris•HCl content was increased to 50 mM in Assay buffer B to maintain the pH even with 5 mM of $RuCl_3$.

$RuCl_3$ concentration (mM)	Assay buffer A 5 mM Tris•HCl (pH)	Assay buffer B 50 mM Tris•HCl (pH)
0.0032	7.69	
0.016	7.66	
0.08	7.68	
0.4	7.61	
2	6.88	
5		7.49
10	2.34	
50	0.94	

Table S3. Enzyme assay data. Values are given as mean of at least technical triplicates and ± errors.

Cation	relative activity	relative inhibition	Quenching factors
ctrl	1.98±0.10	0.00±2.21	1.00±0.00
EDTA	0.42±0.11	n.d.	n.d.
LiI	0.34±0.15	1.09±0.44	n.d.
NaI	0.50±0.12	n.d.	n.d.
KI	0.55±0.11	-0.42±0.69	n.d.
RbI	2.89±0.35	0.76±5.38	0.95±0.00
CsI	2.02±0.16	-0.86±2.31	n.d.
MgII	4.55±0.29	2.64±5.81	n.d.
CaII	100±2	-2.57±1.87	n.d.
SrII	25.1±0.3	-3.20±6.53	n.d.
BaII	14.1±0.6	14.9±9.5	n.d.
ScIII	2.03±0.38	92.0±1.4	0.54±0.00
TiIII			
MnII	41.6±2.1	17.8±8.7	n.d.
FeII	1.31±0.37	76.0±2.6	n.d.
FeIII	0.27±0.25	43.9±6.0	n.d.
CoII	6.29±0.78	20.5±5.6	n.d.
NiII	4.33±0.31	18.2±7.6	n.d.
CuI	0.10±0.19	30.9±5.3	n.d.
CuII	-0.29±0.27	55.3±5.7	n.d.
ZnII	-0.08±0.34	57.3±4.8	n.d.

YIII	1.59±0.44	99.3±0.1	0.92±0.00
ZrIV	1.63±0.23	-29.0±7.5	n.d.
RuIII	-0.02±0.16	92.9±1.3	204±38
PdII	0.57±0.38	66.9±2.1	n.d.
CdII	9.38±2.37	27.8±11.6	n.d.
HgII	-0.24±0.24	100±0	n.d.
AlIII	5.14±0.25	31.0±5.9	n.d.
GaIII			
InIII	4.57±0.27	52.6±9.3	n.d.
SnII	14.2±1.5	10.0±5.3	n.d.
PbII	15.2±0.1	-7.88±7.5	n.d.
SbIII	2.43±0.24	10.2±14.9	n.d.
BiIII	2.44±0.18	0.18±5.49	n.d.
LaIII	111±5	1.96±4.77	0.79±0.00
CeIII	35.3±0.9	19.1±10.2	n.d.
PrIII	39.6±2.3	79.5±0.9	1.73±0.02
NdIII	24.7±3.4	79.4±1.0	1.75±0.01
SmIII	2.07±0.35	100±0	1.92±0.02
EuIII	-1.63±0.24	99.7±0.2	1.90±0.01
GdIII	-1.57±0.33	99.7±0.1	1.76±0.01
TbIII	-0.23±0.21	99.6±1.2	1.52±0.01
DyIII	-0.41±0.08	100±0	1.75±0.00
HoIII	-0.51±0.14	100±0	2.36±0.01
ErIII	0.11±0.19	100±1	1.83±0.01
TmIII	0.15±0.23	100±0	1.67±0.01
YbIII	0.15±0.23	100±0	1.87±0.02
LuIII	0.27±0.09	100±0	0.95±0.00
NH4I	2.70±1.05	-6.41±5.70	

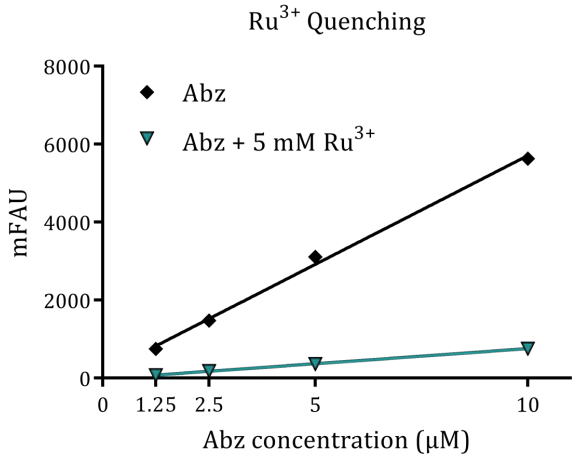
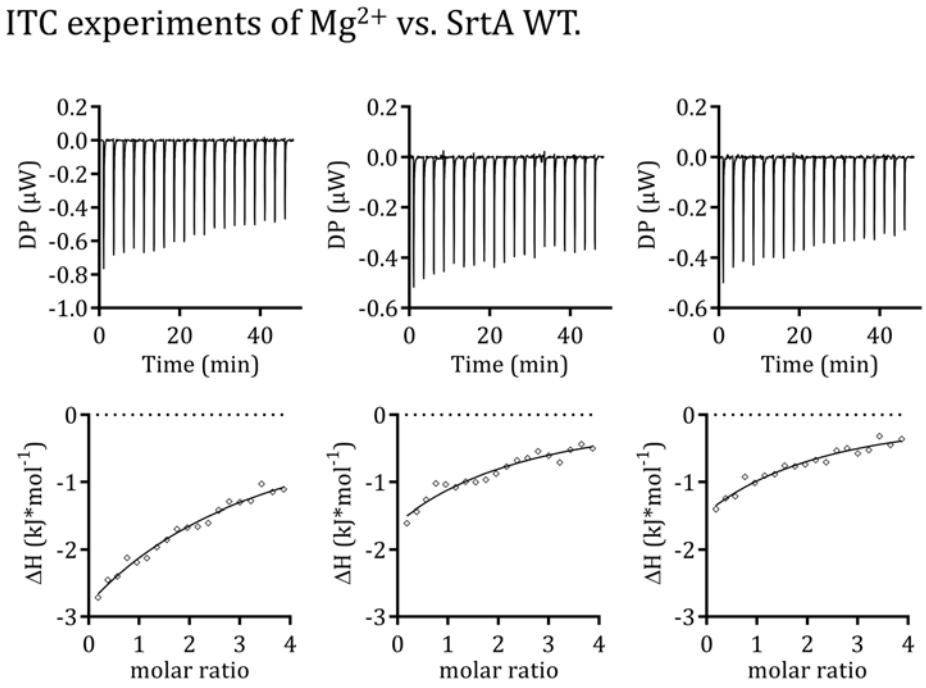
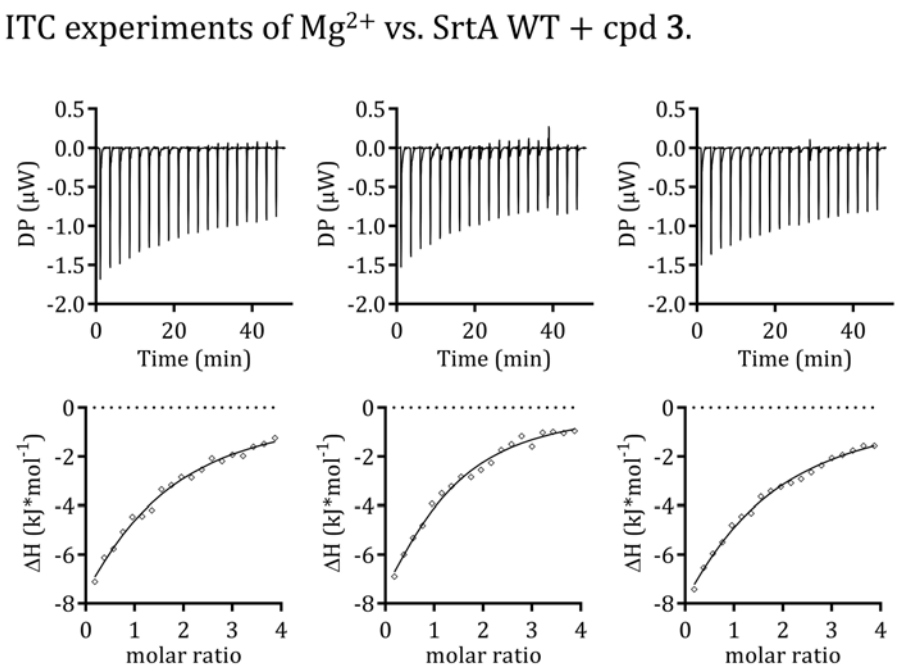


Figure S1. *Quenching of the Abz fluorophore by Ru^{3+} ions.* The fluorescence of differing Abz concentrations (1.25 M, 2.5 M, 5 M, 10 M) was monitored in standard assay conditions with and without 5 mM $RuCl_3$ with λ_{ex}: 320 nM and λ_{em}: 420 nM. Measurements were performed in technical triplicates; standard errors are indicated as bars (too small to be displayed).[1]



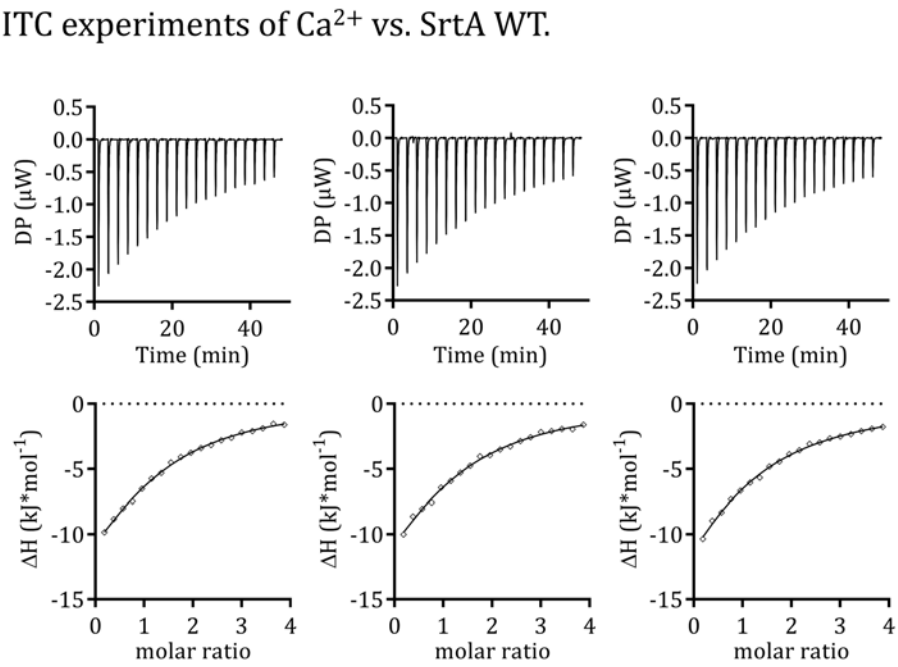
SrtA_WT_vs_Mg_1	SrtA_WT_vs_Mg_2	SrtA_WT_vs_Mg_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.00	N (sites) = 1.00	N (sites) = 1.00
KD (M) = 386e-6 ± 30.2e-6	KD (M) = 248e-6 ± 43.1e-6	KD (M) = 194e-6 ± 23.0e-6
ΔH (kJ/mol) = -23.8 ± 1.36	ΔH (kJ/mol) = -8.96 ± 1.03	ΔH (kJ/mol) = -6.84 ± 0.503
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0
ΔG (kJ/mol) = -19.5	ΔG (kJ/mol) = -20.6	ΔG (kJ/mol) = -21.2
-TΔS (kJ/mol) = 4.35	-TΔS (kJ/mol) = -11.6	-TΔS (kJ/mol) = -14.4

Figure S2A. Thermograms and isotherms of Mg^{2+} vs. SrtA WT.



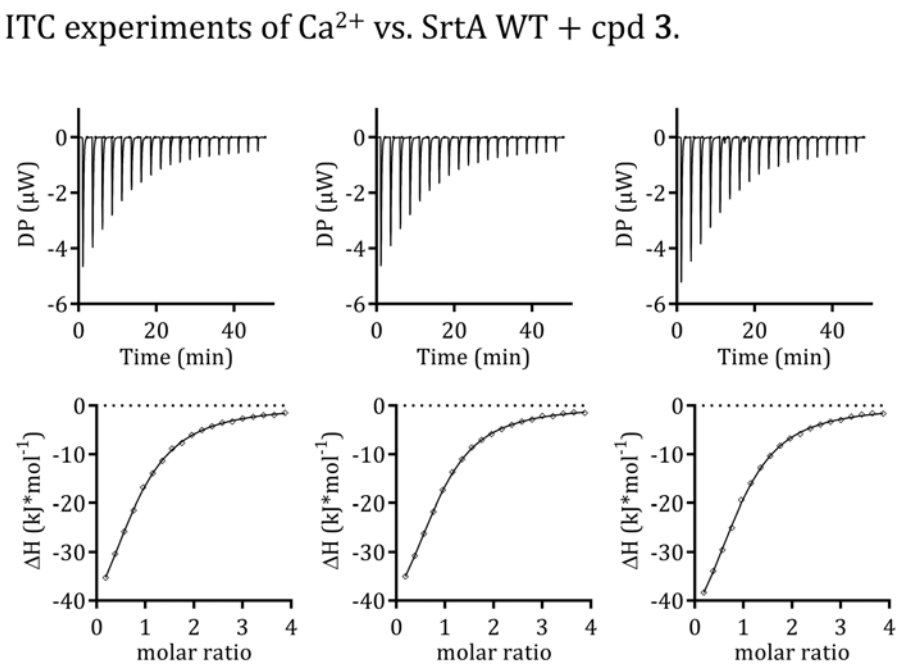
SrtA_WT+cpd3_vs_Mg_1	SrtA_WT+cpd3_vs_Mg_2	SrtA_WT+cpd3_vs_Mg_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.00	N (sites) = 1.00	N (sites) = 1.00
KD (M) = 109e-6 ± 30.5e-6	KD (M) = 61.0e-6 ± 19.2e-6	KD (M) = 121e-6 ± 25.7e-6
ΔH (kJ/mol) = -23.0 ± 5.51	ΔH (kJ/mol) = -15.7 ± 3.41	ΔH (kJ/mol) = -25.9 ± 4.85
Offset (kJ/mol) = -0.506 ± 0.418	Offset (kJ/mol) = -1.14 ± 0.351	Offset (kJ/mol) = -0.384 ± 0.346
ΔG (kJ/mol) = -22.6	ΔG (kJ/mol) = -24.1	ΔG (kJ/mol) = -22.4
-TΔS (kJ/mol) = 0.375	-TΔS (kJ/mol) = -8.38	-TΔS (kJ/mol) = 3.49

Figure S2B. Thermograms and isotherms of Mg^{2+} vs. SrtA WT + cpd **3**.



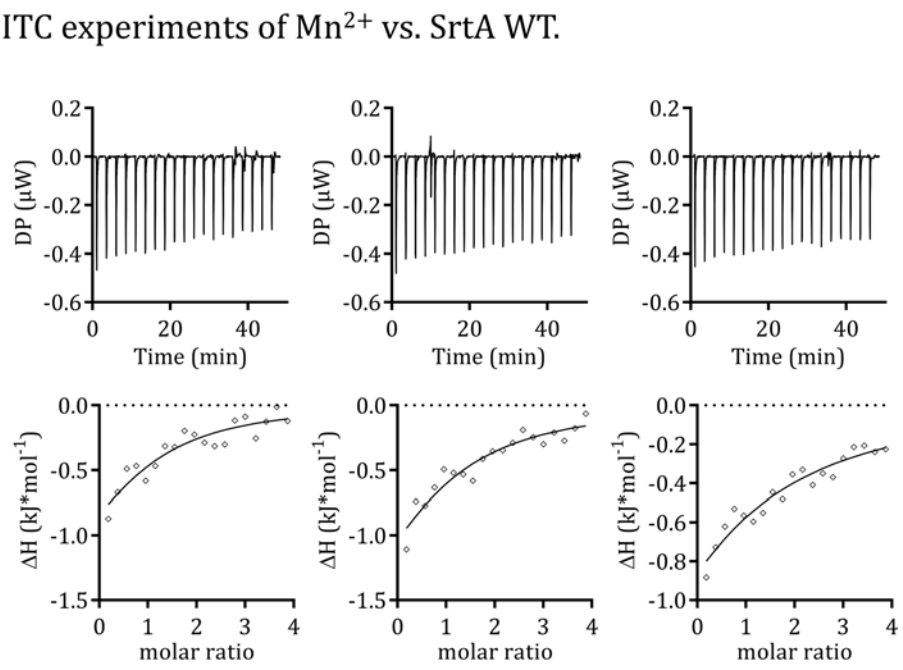
SrtA_WT_vs_Ca_1	SrtA_WT_vs_Ca_2	SrtA_WT_vs_Ca_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.14 ± 0.105	N (sites) = 1.10 ± 0.181	N (sites) = 0.929 ± 0.226
KD (M) = 69.9e-6 ± 19.5e-6	KD (M) = 77.1e-6 ± 32.6e-6	KD (M) = 91.0e-6 ± 34.5e-6
ΔH (kJ/mol) = -22.9 ± 5.46	ΔH (kJ/mol) = -24.8 ± 9.77	ΔH (kJ/mol) = -31.8 ± 14.6
Offset (kJ/mol) = -0.667 ± 0.350	Offset (kJ/mol) = -0.603 ± 0.533	Offset (kJ/mol) = -0.511 ± 0.467
ΔG (kJ/mol) = -23.7	ΔG (kJ/mol) = -23.5	ΔG (kJ/mol) = -23.1
-TΔS (kJ/mol) = -0.822	-TΔS (kJ/mol) = 1.26	-TΔS (kJ/mol) = 8.75

Figure S2C. Thermograms and isotherms of Ca^{2+} vs. SrtA WT.



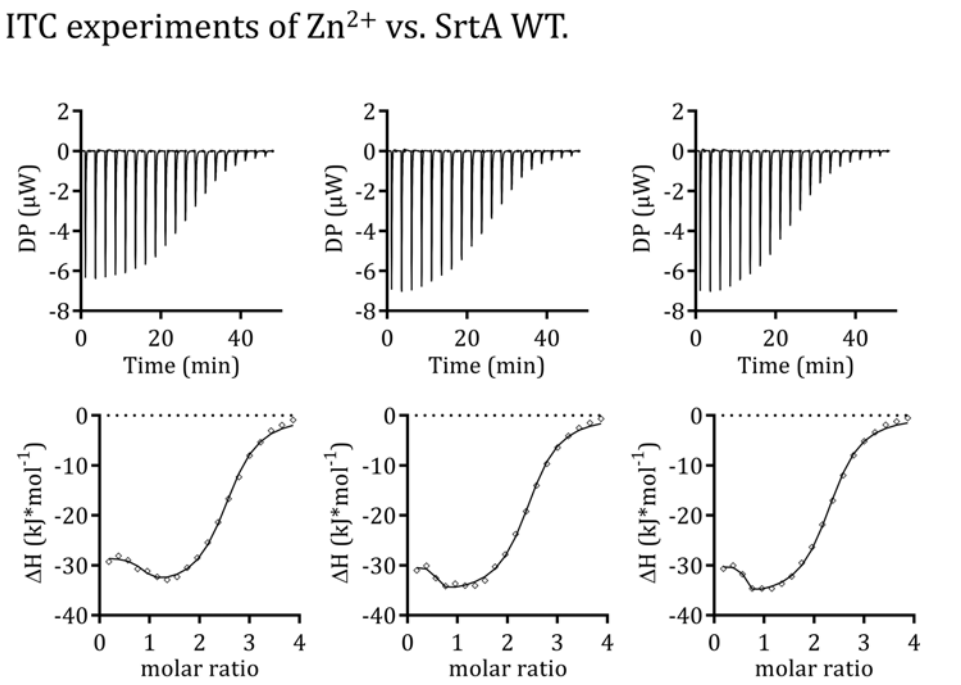
SrtA_WT+cpd3_vs_Ca_1	SrtA_WT+cpd3_vs_Ca_2	SrtA_WT+cpd3_vs_Ca_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 0.794 ± 1.3e-2	N (sites) = 0.845 ± 7.9e-3	N (sites) = 0.898 ± 1.4e-2
KD (M) = 22.6e-6 ± 1.23e-6	KD (M) = 18.4e-6 ± 702e-9	KD (M) = 18.5e-6 ± 1.25e-6
ΔH (kJ/mol) = -58.9 ± 1.88	ΔH (kJ/mol) = -53.1 ± 1.00	ΔH (kJ/mol) = -57.0 ± 1.84
Offset (kJ/mol) = 4.2e-2 ± 0.149	Offset (kJ/mol) = -0.215 ± 0.102	Offset (kJ/mol) = -0.224 ± 0.207
ΔG (kJ/mol) = -26.5	ΔG (kJ/mol) = -27.1	ΔG (kJ/mol) = -27.1
-TΔS (kJ/mol) = 32.3	-TΔS (kJ/mol) = 26.0	-TΔS (kJ/mol) = 29.9

Figure S2D. Thermograms and isotherms of Ca^{2+} vs. SrtA WT + cpd **3**.



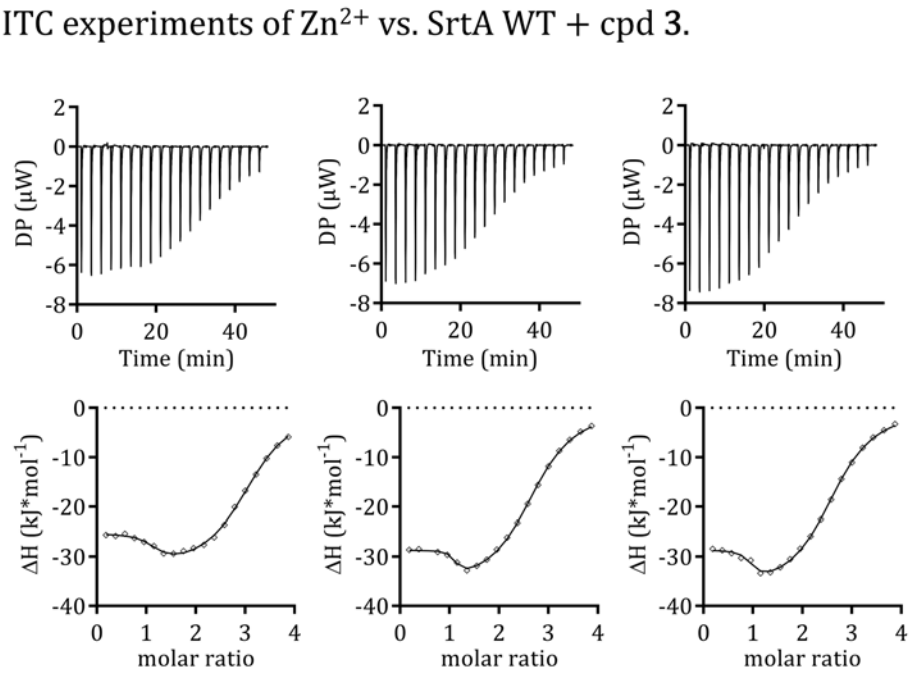
SrtA_WT_vs_Mn_1	SrtA_WT_vs_Mn_2	SrtA_WT_vs_Mn_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.00	N (sites) = 1.00	N (sites) = 1.00
KD (M) = 64.8e-6 ± 79.8e-6	KD (M) = 83.8e-6 ± 80.8e-6	KD (M) = 175e-6 ± 150e-6
ΔH (kJ/mol) = -1.84 ± 1.60	ΔH (kJ/mol) = -2.66 ± 2.00	ΔH (kJ/mol) = -3.74 ± 3.14
Offset (kJ/mol) = -0.369 ± 0.160	Offset (kJ/mol) = -0.268 ± 0.175	Offset (kJ/mol) = -0.235 ± 0.175
ΔG (kJ/mol) = -23.9	ΔG (kJ/mol) = -23.3	ΔG (kJ/mol) = -21.5
-TΔS (kJ/mol) = -22.1	-TΔS (kJ/mol) = -20.6	-TΔS (kJ/mol) = -17.7

Figure S2E. Thermograms and isotherms of Mn^{2+} vs. SrtA WT.



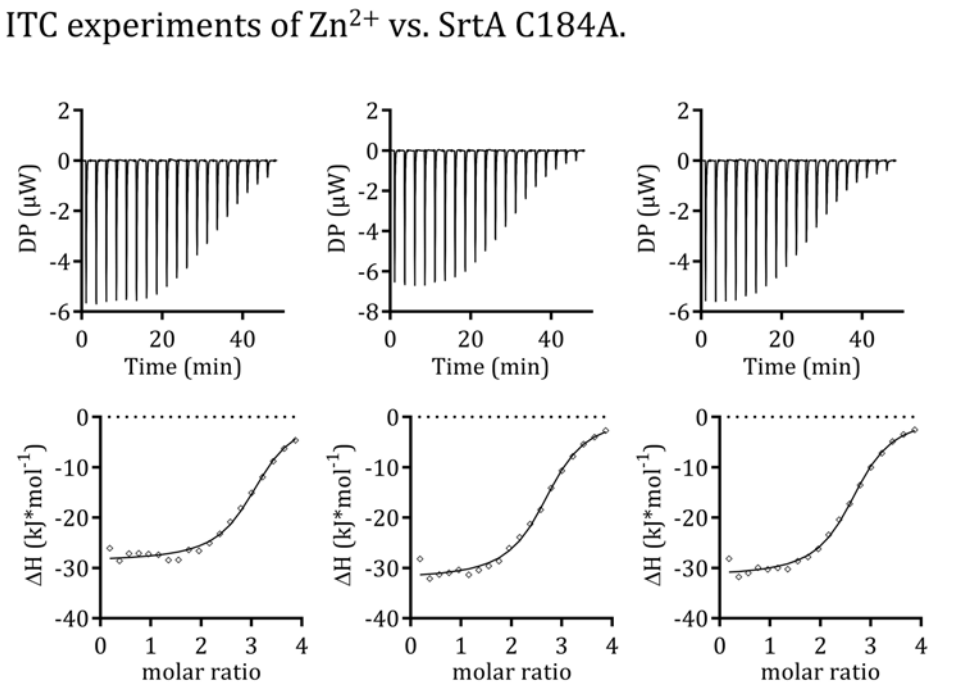
SrtA_WT_vs_Zn_1	SrtA_WT_vs_Zn_2	SrtA_WT_vs_Zn_3
Model: Two Sets of Sites	Model: Two Sets of Sites	Model: Two Sets of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N1 (sites) = 1.79 ± 2.8e-2	N1 (sites) = 1.94 ± 3.4e-2	N1 (sites) = 1.81 ± 2.5e-2
KD1 (M) = 3.08e-6 ± 37.3e-9	KD1 (M) = 2.68e-6 ± 183e-9	KD1 (M) = 2.91e-6 ± 118e-9
ΔH1 (kJ/mol) = -33.8 ± 0.358	ΔH1 (kJ/mol) = -35.0 ± 0.324	ΔH1 (kJ/mol) = -35.4 ± 0.283
N2 (sites) = 0.797 ± 2.5e-2	N2 (sites) = 0.497 ± 3.1e-2	N2 (sites) = 0.532 ± 1.5e-2
KD2 (M) = 47.4e-9 ± 22.5e-9	KD2 (M) = 9.27e-9 ± 12.8e-9	KD2 (M) = 4.56e-9 ± 5.15e-9
ΔH2 (kJ/mol) = -28.4 ± 0.438	ΔH2 (kJ/mol) = -30.4 ± 0.593	ΔH2 (kJ/mol) = -30.3 ± 0.422
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2F. Thermograms and isotherms of Zn^{2+} vs. SrtA WT.



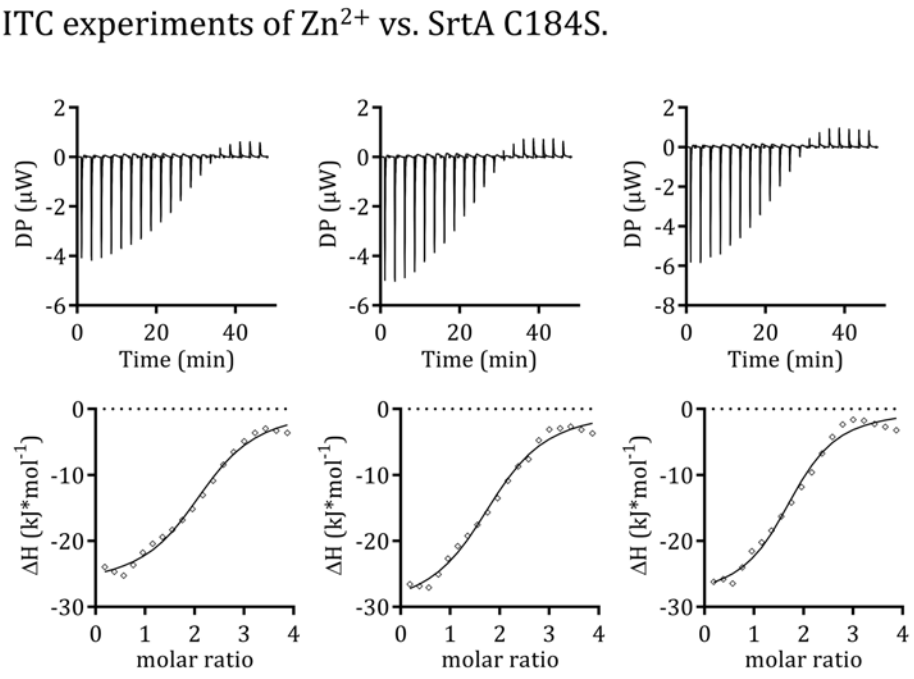
SrtA_WT+cpd3_vs_Zn_1	SrtA_WT+cpd3_vs_Zn_2	SrtA_WT+cpd3_vs_Zn_3
Model: Two Sets of Sites	Model: Two Sets of Sites	Model: Two Sets of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N1 (sites) = 1.08 ± 7.3e-3	N1 (sites) = 1.03 ± 4.6e-3	N1 (sites) = 1.86 ± 1.6e-2
KD1 (M) = 94.0e-9 ± 6.64e-9	KD1 (M) = 21.0e-9 ± 11.7e-9	KD1 (M) = 5.91e-6 ± 42.5e-9
ΔH1 (kJ/mol) = -25.3 ± 0.149	ΔH1 (kJ/mol) = -28.7 ± 0.159	ΔH1 (kJ/mol) = -34.8 ± 0.205
N2 (sites) = 2.10 ± 1.3e-2	N2 (sites) = 1.80 ± 7.9e-3	N2 (sites) = 0.887 ± 1.1e-2
KD2 (M) = 5.32e-6 ± 118e-9	KD2 (M) = 6.17e-6 ± 22.6e-9	KD2 (M) = 49.6e-9 ± 7.52e-9
ΔH2 (kJ/mol) = -31.3 ± 0.161	ΔH2 (kJ/mol) = -33.7 ± 0.152	ΔH2 (kJ/mol) = -28.7 ± 0.199
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2G. Thermograms and isotherms of Zn^{2+} vs. SrtA WT + cpd **3**.



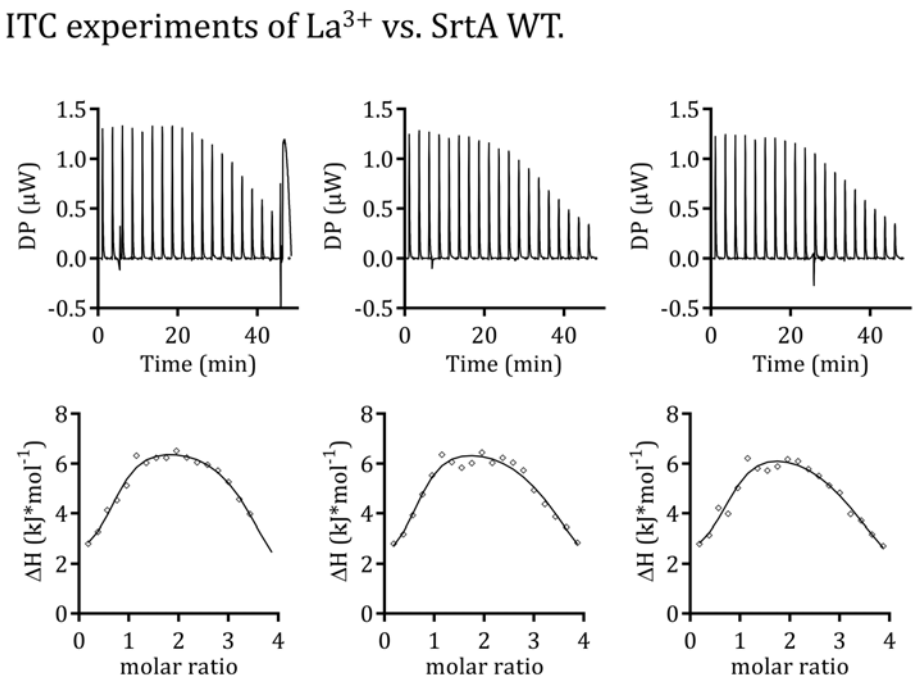
SrtA_C184A_vs_Zn_1	SrtA_C184A_vs_Zn_2	SrtA_C184A_vs_Zn_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 3.04 ± 2.8e-2	N (sites) = 2.68 ± 2.7e-2	N (sites) = 2.64 ± 2.4e-2
KD (M) = 2.53e-6 ± 371e-9	KD (M) = 2.91e-6 ± 403e-9	KD (M) = 2.83e-6 ± 343e-9
ΔH (kJ/mol) = -28.6 ± 0.368	ΔH (kJ/mol) = -32.1 ± 0.457	ΔH (kJ/mol) = -31.6 ± 0.395
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0
ΔG (kJ/mol) = -32.0	ΔG (kJ/mol) = -31.6	ΔG (kJ/mol) = -31.7
-TΔS (kJ/mol) = -3.39	-TΔS (kJ/mol) = 0.454	-TΔS (kJ/mol) = -0.151

Figure S2H. Thermograms and isotherms of Zn^{2+} vs. SrtA C184A.



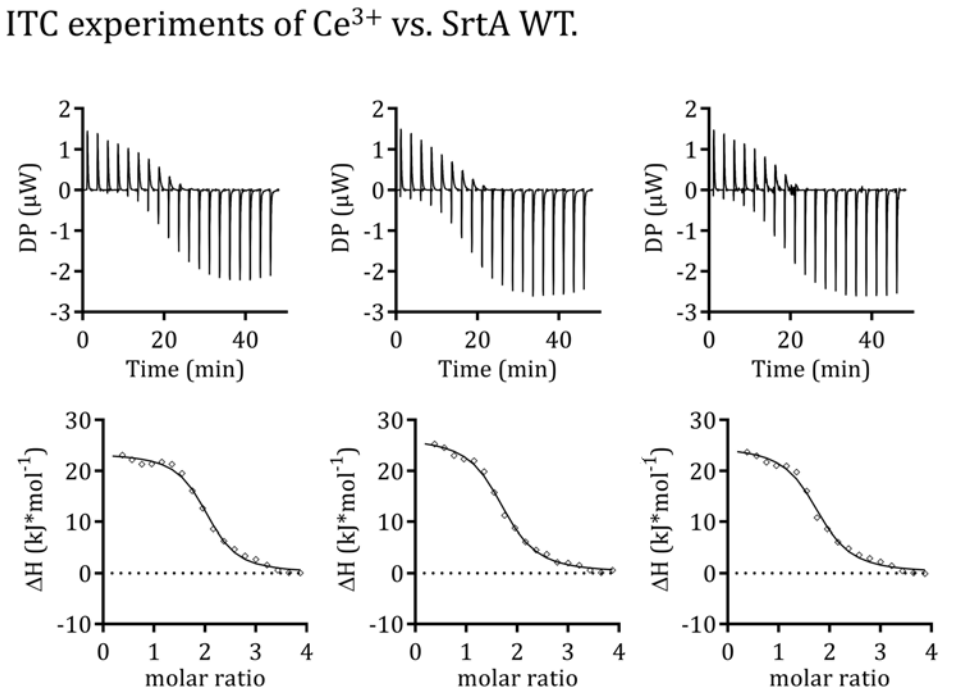
SrtA_C184S_vs_Zn_1	SrtA_C184S_vs_Zn_2	SrtA_C184S_vs_Zn_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 2.20 ± 8.2e-2	N (sites) = 1.90 ± 7.4e-2	N (sites) = 1.78 ± 6.7e-2
KD (M) = 8.44e-6 ± 2.10e-6	KD (M) = 10.1e-6 ± 2.69e-6	KD (M) = 7.23e-6 ± 2.16e-6
ΔH (kJ/mol) = -26.8 ± 1.83	ΔH (kJ/mol) = -30.4 ± 2.24	ΔH (kJ/mol) = -28.7 ± 2.05
Offset (kJ/mol) = 9.01 ± 1.06	Offset (kJ/mol) = 9.00 ± 1.04	Offset (kJ/mol) = 8.39 ± 0.942
ΔG (kJ/mol) = -29.0	ΔG (kJ/mol) = -28.5	ΔG (kJ/mol) = -29.4
-TΔS (kJ/mol) = -2.25	-TΔS (kJ/mol) = 1.89	-TΔS (kJ/mol) = -0.717

Figure S2I. Thermograms and isotherms of Zn^{2+} vs. SrtA C184S.



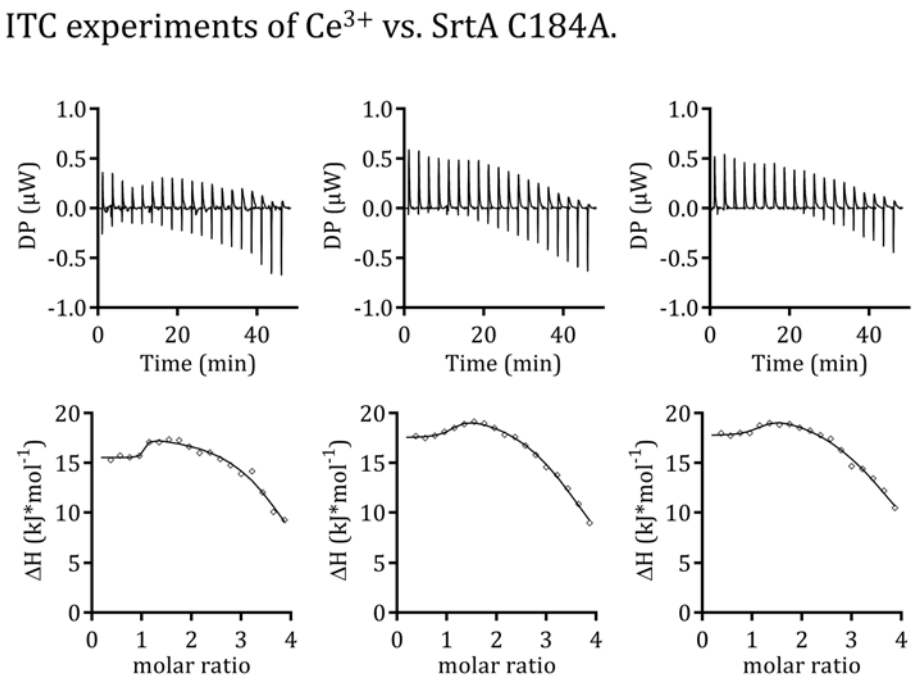
SrtA_WT_vs_La_1	SrtA_WT_vs_La_2	SrtA_WT_vs_La_3
Model: Two Sets of Sites	Model: Two Sets of Sites	Model: Two Sets of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N1 (sites) = 0.706 ± 3.1e-2	N1 (sites) = 0.653 ± 7.7e-2	N1 (sites) = 0.744 ± 9.0e-2
KD1 (M) = 109e-9 ± 29.0e-9	KD1 (M) = 258e-9 ± 150e-9	KD1 (M) = 378e-9 ± 157e-9
ΔH1 (kJ/mol) = 2.15 ± 0.333	ΔH1 (kJ/mol) = 1.63 ± 0.882	ΔH1 (kJ/mol) = 1.90 ± 0.781
N2 (sites) = 2.91 ± 5.1e-2	N2 (sites) = 3.13 ± 0.102	N2 (sites) = 3.03 ± 0.133
KD2 (M) = 3.15e-6 ± 107e-9	KD2 (M) = 6.03e-6 ± 890e-9	KD2 (M) = 8.13e-6 ± 1.19e-6
ΔH2 (kJ/mol) = 6.83 ± 0.113	ΔH2 (kJ/mol) = 6.99 ± 0.232	ΔH2 (kJ/mol) = 6.90 ± 0.228
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2J. Thermograms and isotherms of La^{3+} vs. SrtA WT.



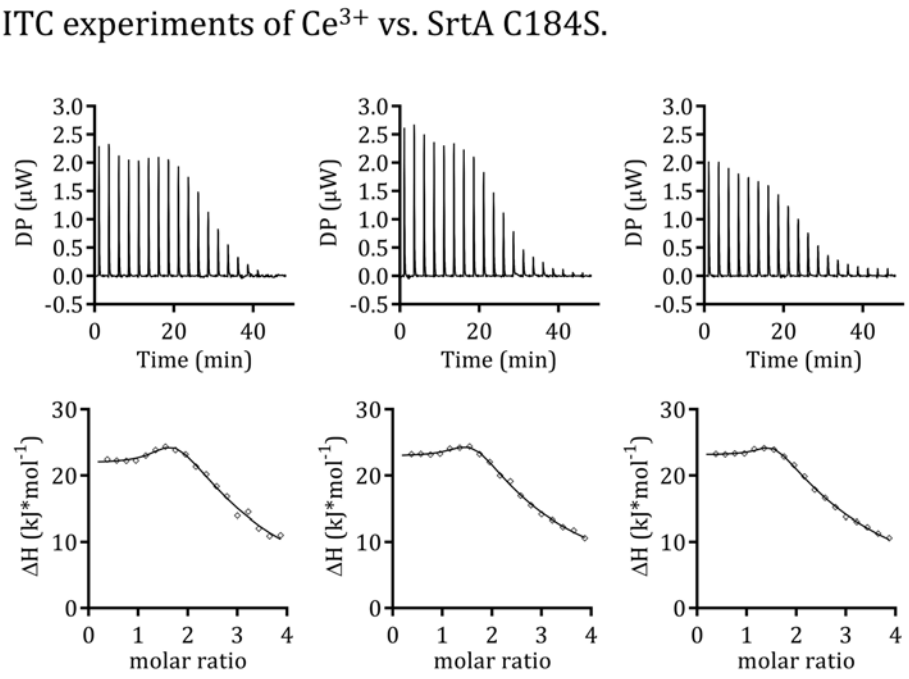
SrtA_WT_vs_Ce_1	SrtA_WT_vs_Ce_2	SrtA_WT_vs_Ce_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.99 ± 3.0e-2	N (sites) = 1.69 ± 2.5e-2	N (sites) = 1.73 ± 3.6e-2
KD (M) = 2.33e-6 ± 470e-9	KD (M) = 3.61e-6 ± 576e-9	KD (M) = 3.40e-6 ± 789e-9
ΔH (kJ/mol) = 23.5 ± 0.791	ΔH (kJ/mol) = 26.6 ± 0.814	ΔH (kJ/mol) = 24.8 ± 1.08
Offset (kJ/mol) = -5.73 ± 0.448	Offset (kJ/mol) = -7.38 ± 0.358	Offset (kJ/mol) = -6.40 ± 0.496
ΔG (kJ/mol) = -32.2	ΔG (kJ/mol) = -31.1	ΔG (kJ/mol) = -31.2
-TΔS (kJ/mol) = -55.7	-TΔS (kJ/mol) = -57.7	-TΔS (kJ/mol) = -56.1

Figure S2K. Thermograms and isotherms of Ce^{3+} vs. SrtA WT.



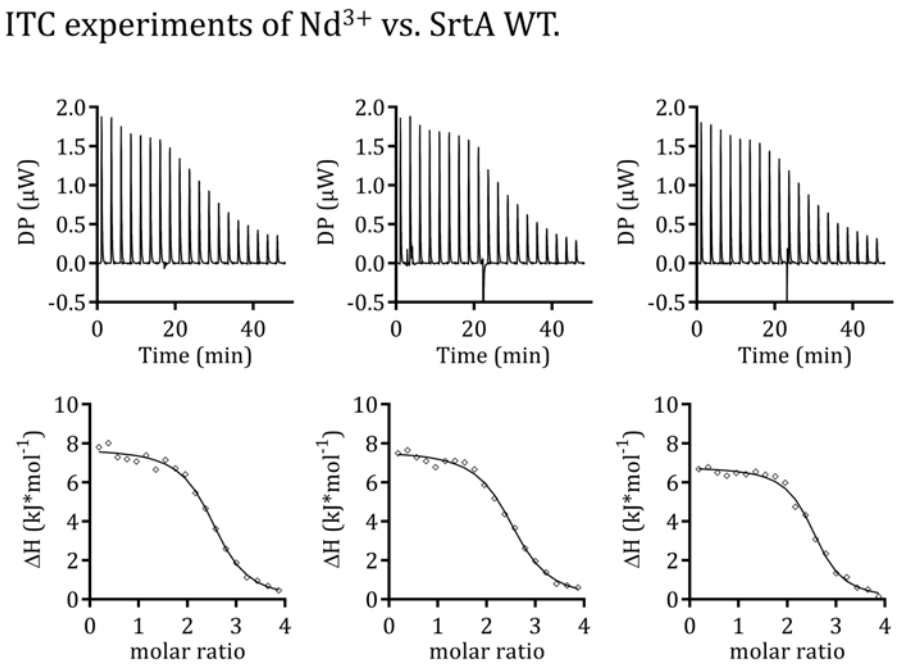
SrtA_C184A_vs_Ce_1	SrtA_C184A_vs_Ce_2	SrtA_C184A_vs_Ce_3
Model: Two Sets of Sites	Model: Two Sets of Sites	Model: Two Sets of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N1 (sites) = 0.968 ± 2.6e-3	N1 (sites) = 1.06 ± 9.3e-3	N1 (sites) = 1.10 ± 7.4e-2
KD1 (M) = 1.55e-9 ± 1.08e-9	KD1 (M) = 68.8e-9 ± 6.73e-9	KD1 (M) = 115e-9 ± 24.2e-9
ΔH1 (kJ/mol) = 15.5 ± 0.311	ΔH1 (kJ/mol) = 17.5 ± 0.115	ΔH1 (kJ/mol) = 17.8 ± 0.392
N2 (sites) = 3.12 ± 4.0e-2	N2 (sites) = 2.99 ± 1.5e-2	N2 (sites) = 3.28 ± 8.5e-2
KD2 (M) = 6.00e-6 ± 2.96e-9	KD2 (M) = 9.42e-6 ± 245e-9	KD2 (M) = 12.4e-6 ± 129e-9
ΔH2 (kJ/mol) = 17.4 ± 0.152	ΔH2 (kJ/mol) = 19.7 ± 9.8e-2	ΔH2 (kJ/mol) = 19.8 ± 0.184
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2L. Thermograms and isotherms of Ce^{3+} vs. SrtA C184A.



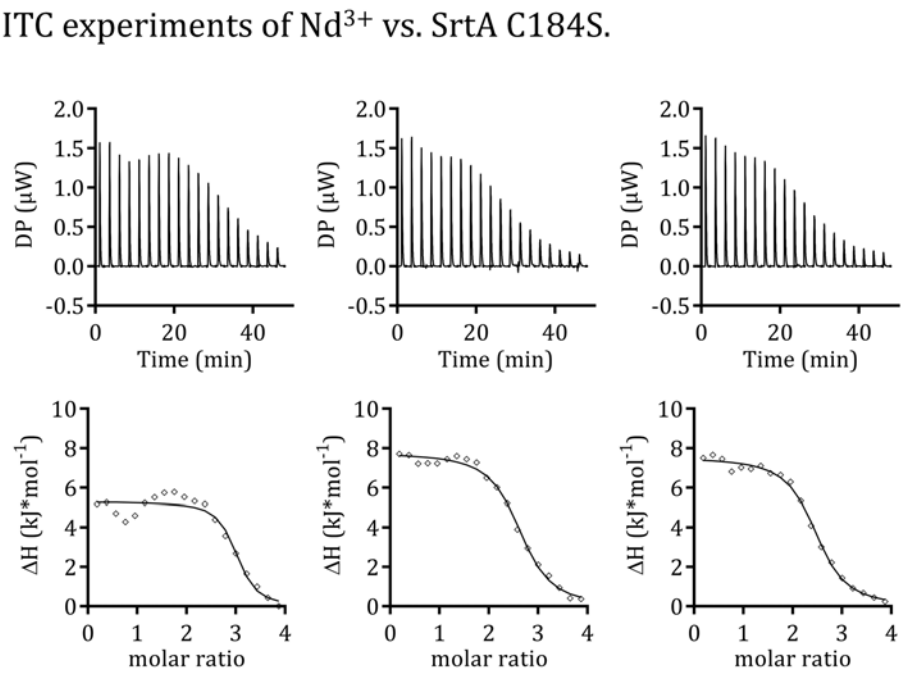
SrtA_C184S_vs_Ce_1	SrtA_C184S_vs_Ce_2	SrtA_C184S_vs_Ce_3
Model: Two Sets of Sites	Model: Two Sets of Sites	Model: Two Sets of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N1 (sites) = 1.45 ± 8.6e-2	N1 (sites) = 5.15 ± 0.183	N1 (sites) = 4.55 ± 9.7e-2
KD1 (M) = 1.43e-6 ± 306e-9	KD1 (M) = 390e-6 ± 31.3e-6	KD1 (M) = 229e-6 ± 12.5e-6
ΔH1 (kJ/mol) = 22.3 ± 0.373	ΔH1 (kJ/mol) = 28.9 ± 0.565	ΔH1 (kJ/mol) = 27.3 ± 0.332
N2 (sites) = 3.39 ± 0.241	N2 (sites) = 1.38 ± 4.8e-2	N2 (sites) = 1.32 ± 3.2e-2
KD2 (M) = 119e-6 ± 22.6e-6	KD2 (M) = 3.36e-6 ± 305e-9	KD2 (M) = 1.27e-6 ± 48.5e-9
ΔH2 (kJ/mol) = 29.0 ± 0.735	ΔH2 (kJ/mol) = 24.0 ± 0.201	ΔH2 (kJ/mol) = 23.5 ± 0.118
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2M. Thermograms and isotherms of Ce^{3+} vs. SrtA C184S.



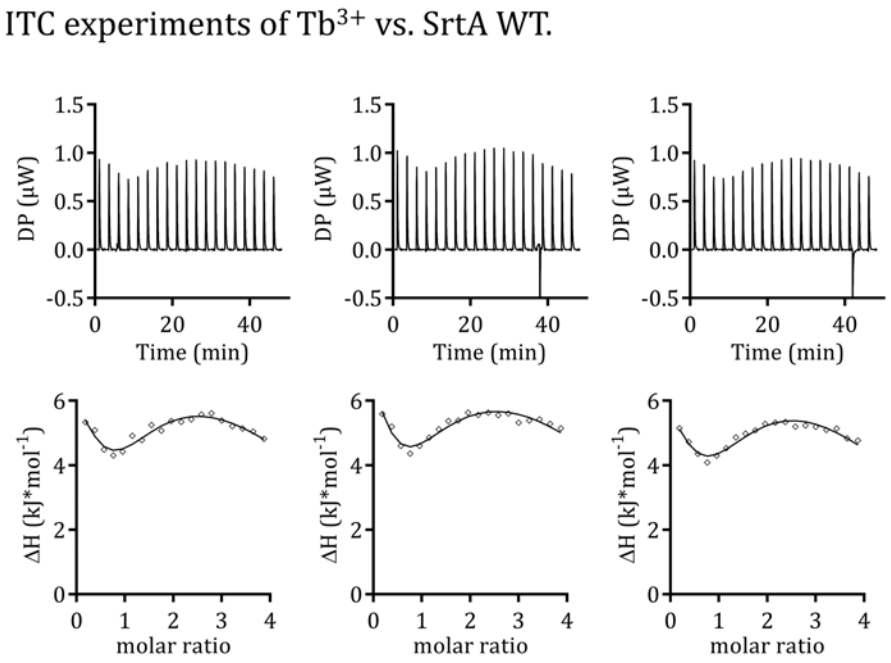
SrtA_WT_vs_Nd_1	SrtA_WT_vs_Nd_2	SrtA_WT_vs_Nd_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 2.50 ± 4.5e-2	N (sites) = 2.50 ± 4.2e-2	N (sites) = 2.50 ± 3.4e-2
KD (M) = 2.47e-6 ± 527e-9	KD (M) = 2.95e-6 ± 531e-9	KD (M) = 1.79e-6 ± 353e-9
ΔH (kJ/mol) = 7.72 ± 0.329	ΔH (kJ/mol) = 7.62 ± 0.294	ΔH (kJ/mol) = 6.79 ± 0.236
Offset (kJ/mol) = 1.67 ± 0.243	Offset (kJ/mol) = 1.51 ± 0.215	Offset (kJ/mol) = 1.81 ± 0.176
ΔG (kJ/mol) = -32.0	ΔG (kJ/mol) = -31.6	ΔG (kJ/mol) = -32.8
-TΔS (kJ/mol) = -39.8	-TΔS (kJ/mol) = -39.2	-TΔS (kJ/mol) = -39.6

Figure S2M. Thermograms and isotherms of Nd^{3+} vs. SrtA WT.



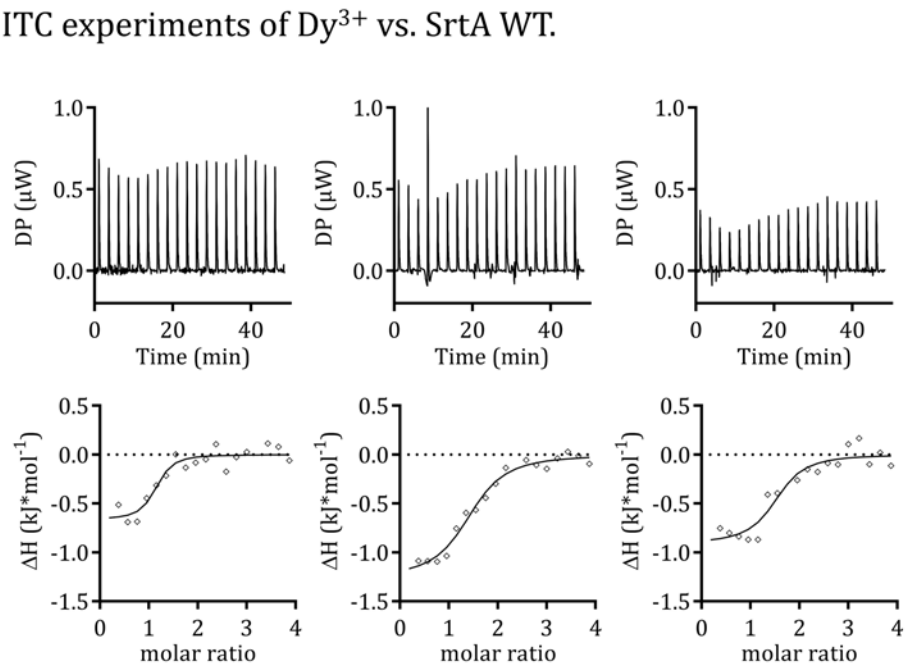
SrtA_C184S_vs_Nd_1	SrtA_C184S_vs_Nd_2	SrtA_C184S_vs_Nd_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 2.93 ± 7.9e-2	N (sites) = 2.58 ± 3.9e-2	N (sites) = 2.42 ± 3.1e-2
KD (M) = 633e-9 ± 408e-9	KD (M) = 1.92e-6 ± 387e-9	KD (M) = 1.95e-6 ± 342e-9
ΔH (kJ/mol) = 5.29 ± 0.545	ΔH (kJ/mol) = 7.74 ± 0.297	ΔH (kJ/mol) = 7.51 ± 0.232
Offset (kJ/mol) = 7.61 ± 0.467	Offset (kJ/mol) = 6.39 ± 0.227	Offset (kJ/mol) = 6.47 ± 0.168
ΔG (kJ/mol) = -35.4	ΔG (kJ/mol) = -32.7	ΔG (kJ/mol) = -32.6
-TΔS (kJ/mol) = -40.7	-TΔS (kJ/mol) = -40.4	-TΔS (kJ/mol) = -40.1

Figure S2N. Thermograms and isotherms of Nd^{3+} vs. SrtA C184S.



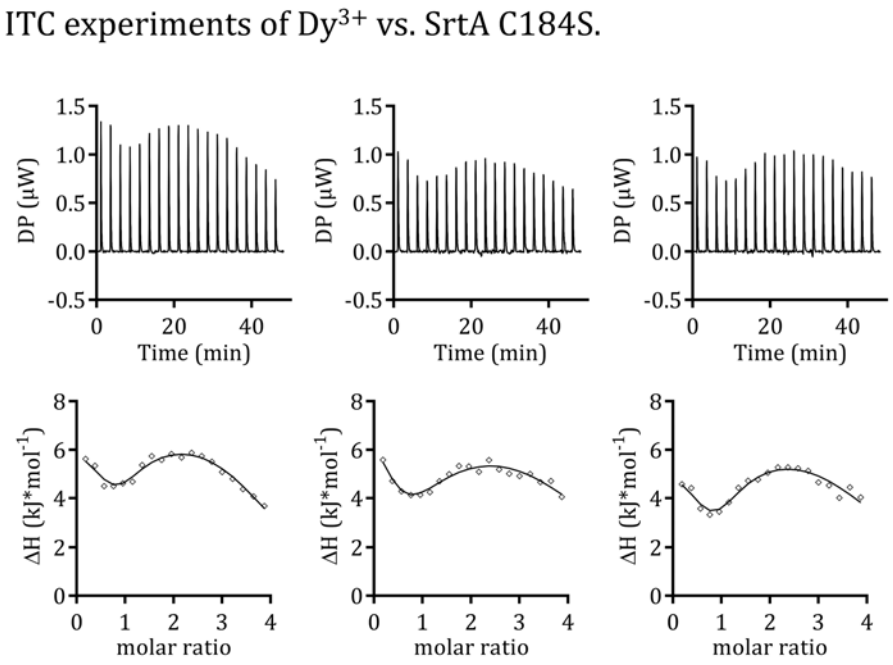
SrtA_WT_vs_Tb_1	SrtA_WT_vs_Tb_2	SrtA_WT_vs_Tb_3
Model: Sequential Binding Sites	Model: Sequential Binding Sites	Model: Sequential Binding Sites
Number of Sites: 3.00	Number of Sites: 3.00	Number of Sites: 3.00
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
KD1 (M) = 20.9e-6 ± 11.6e-6	KD1 (M) = 32.6e-6 ± 28.0e-6	KD1 (M) = 18.9e-6 ± 5.14e-6
ΔH1 (kJ/mol) = 8.04 ± 4.17	ΔH1 (kJ/mol) = 10.0 ± 8.21	ΔH1 (kJ/mol) = 7.46 ± 1.84
KD2 (M) = 209e-6 ± 529e-6	KD2 (M) = 222e-6 ± 710e-6	KD2 (M) = 115e-6 ± 6.97e-10
ΔH2 (kJ/mol) = -3.63 ± 36.8	ΔH2 (kJ/mol) = -14.3 ± 34.3	ΔH2 (kJ/mol) = -2.35 ± 6.81
KD3 (M) = 32.5e-6 ± 66.4e-6	KD3 (M) = 26.5e-6 ± 68.2e-6	KD3 (M) = 46.7e-6 ± 8.12e-6
ΔH3 (kJ/mol) = 37.0 ± 46.9	ΔH3 (kJ/mol) = 47.0 ± 40.6	ΔH3 (kJ/mol) = 35.3 ± 3.89
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2O. Thermograms and isotherms of Tb^{3+} vs. SrtA WT.



SrtA_WT_vs_Dy_1	SrtA_WT_vs_Dy_2	SrtA_WT_vs_Dy_3
Model: One Set of Sites	Model: One Set of Sites	Model: One Set of Sites
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
N (sites) = 1.44 ± 6.5e-2	N (sites) = 1.44 ± 6.5e-2	N (sites) = 1.50 ± 0.142
KD (M) = 5.14e-6 ± 2.22e-6	KD (M) = 5.14e-6 ± 2.22e-6	KD (M) = 2.94e-6 ± 2.86e-6
ΔH (kJ/mol) = -1.26 ± 0.122	ΔH (kJ/mol) = -1.26 ± 0.122	ΔH (kJ/mol) = -0.910 ± 0.155
Offset (kJ/mol) = 3.15 ± 4.0e-2	Offset (kJ/mol) = 3.15 ± 4.0e-2	Offset (kJ/mol) = 1.89 ± 6.2e-2
ΔG (kJ/mol) = -30.2	ΔG (kJ/mol) = -30.2	ΔG (kJ/mol) = -31.6
-TΔS (kJ/mol) = -29.0	-TΔS (kJ/mol) = -29.0	-TΔS (kJ/mol) = -30.7

Figure S2P. Thermograms and isotherms of Dy^{3+} vs. SrtA WT.



SrtA_C184S_vs_Dy_1	SrtA_C184S_vs_Dy_2	SrtA_C184S_vs_Dy_3
Model: Sequential Binding Sites	Model: Sequential Binding Sites	Model: Sequential Binding Sites
Number of Sites: 3.00	Number of Sites: 3.00	Number of Sites: 3.00
[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6	[Cell] (M) = 50.0e-6
[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3	[Syr] (M) = 1.00e-3
KD1 (M) = 3.85e-6 ± 693e-9	KD1 (M) = 12.5e-6 ± 5.86e-6	KD1 (M) = 5.17e-6 ± 1.45e-6
ΔH1 (kJ/mol) = 6.12 ± 0.903	ΔH1 (kJ/mol) = 7.35 ± 3.19	ΔH1 (kJ/mol) = 5.22 ± 1.18
KD2 (M) = 150e-6 ± 125e-6	KD2 (M) = 167e-6 ± 301e-6	KD2 (M) = 370e-6 ± 1.15e-3
ΔH2 (kJ/mol) = -10.9 ± 13.3	ΔH2 (kJ/mol) = -18.1 ± 15.3	ΔH2 (kJ/mol) = -30.5 ± 57.2
KD3 (M) = 5.07e-6 ± 3.37e-6	KD3 (M) = 11.2e-6 ± 15.3e-6	KD3 (M) = 4.32e-6 ± 12.3e-6
ΔH3 (kJ/mol) = 30.5 ± 13.2	ΔH3 (kJ/mol) = 40.7 ± 15.2	ΔH3 (kJ/mol) = 51.0 ± 54.9
Offset (kJ/mol) = 0	Offset (kJ/mol) = 0	Offset (kJ/mol) = 0

Figure S2Q. Thermograms and isotherms of Dy^{3+} vs. SrtA C184S.

Table S4. Thermodynamic binding profiles of metal ions to SrtA in presence of the covalent-reversible inhibitor 3 and the SrtA active site mutans SrtA C164A and C164S.

Ion	Target	K_d (μM)			ΔG° (kJ·mol^{-1})			ΔH° (kJ·mol^{-1})			−TΔS° (kJ·mol^{-1})			N		
Mg^{2+}	SrtA + cpd **3**	97.0 ± 14.8			−23.0 ± 0.8			−21.5 ± 2.7			−1.5 ± 5.0			1.00[a]		
Ca^{2+}	SrtA + cpd **3**	19.8 ± 0.6			−26.9 ± 0.3			−56.3 ± 0.9			29.4 ± 2.6			0.85 ± 0.01		
Zn^{2+}	SrtA + cpd **3**	0.0549 ± 0.0051	5.80 ± 0.04		−41.9 ± 1.6	−29.9 ± 0.1		−27.6 ± 0.1	−33.3 ± 0.1		−14.3 ± 0.9	3.4 ± 1.6		0.52 ± 0.10	1.93 ± 0.02	
Zn^{2+}	SrtA C184A	2.76 ± 0.22			−31.8 ± 0.2			−30.8 ± 0.2			−10.3 ± 1.7			2.79 ± 0.02		
Zn^{2+}	SrtA C184S	8.59 ± 1.35			−29.0 ± 0.4			−28.6 ± 1.2			−35.9 ± 1.7			1.96 ± 0.04		
Ce^{3+}	SrtA C184A	0.0618 ± 0.0084	9.27 ± 0.09		−43.6 ± 4.8	−28.8 ± 0.7		16.9 ± 0.2	19.0 ± 0.1		−60.6 ± 3.8	−47.8 ± 0.4		1.04 ± 0.02	3.13 ± 0.03	
Ce^{3+}	SrtA C184S	2.02 ± 0.14	246 ± 14		−32.8 ± 1.1	−20.9 ± 1.2		23.3 ± 0.1	28.4 ± 0.3		−56.0 ± 0.8	−49.3 ± 1.5		1.38 ± 0.03	4.36 ± 0.11	
Nd^{3+}	SrtA C184S	1.50 ± 0.22			−33.6 ± 1.3			6.9 ± 0.2			−40.4 ± 0.2			2.64 ± 0.03		
Dy^{3+}	SrtA C184S	6.86 ± 4.34	7.17 ± 4.69	229 ± 144	−29.7 ± 1.0	−29.7 ± 1.2	−21.0 ± 1.0	40.7 ± 19.5	6.2 ± 1.2	−19.8 ± 20.2	−70.5 ± 8.6	−36.0 ± 0.8	−1.2 ± 9.0	n.a.	n.a.	n.a.

[a] Affinity was too low to allow the determination of the stoichiometry. Hence, N was manually set to be 1.0 since Mg^{2+} is neither a strong thiophile, oxophile nor a strong reducing agent.[2,3]
n.a.: not applicable. The Isotherm was fitted as 3 sequential binding sites; hence, no binding stoichiometry was obtained.

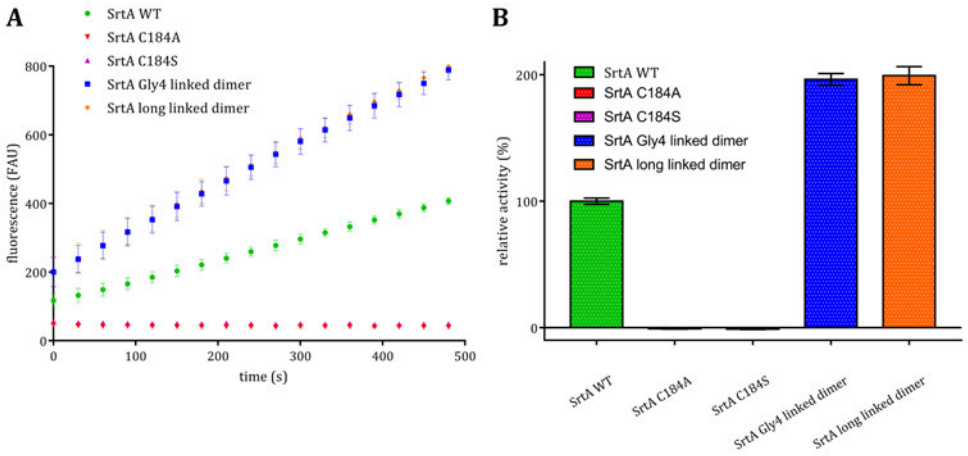


Figure S3. *Activity of SrtA constructs used in this work.* (**A**) Fluorescence-time traces and (**B**) Relative activities.

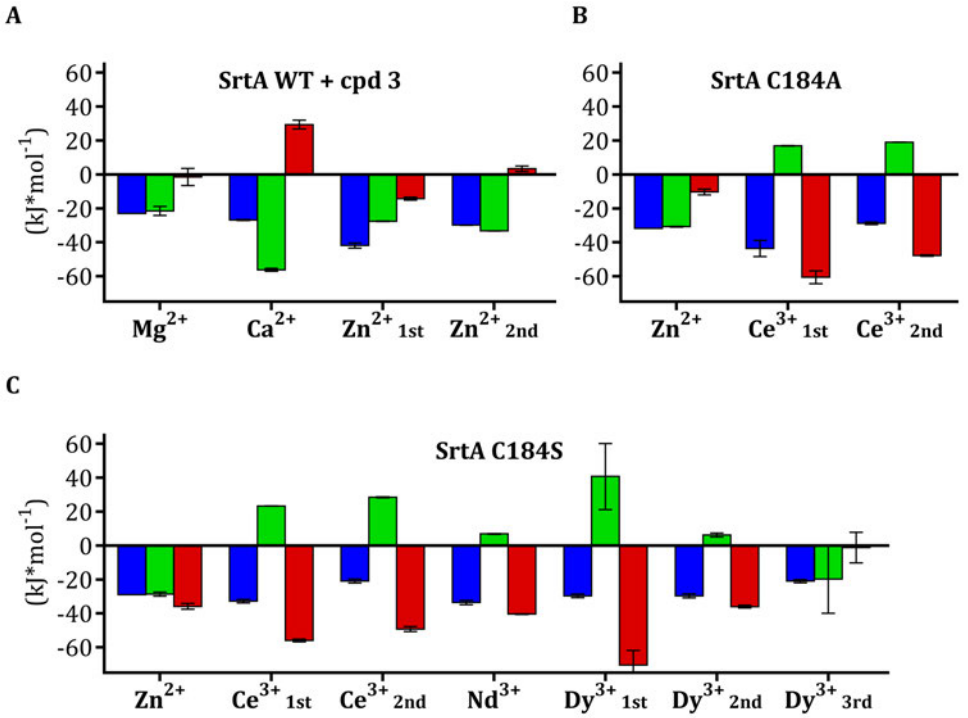
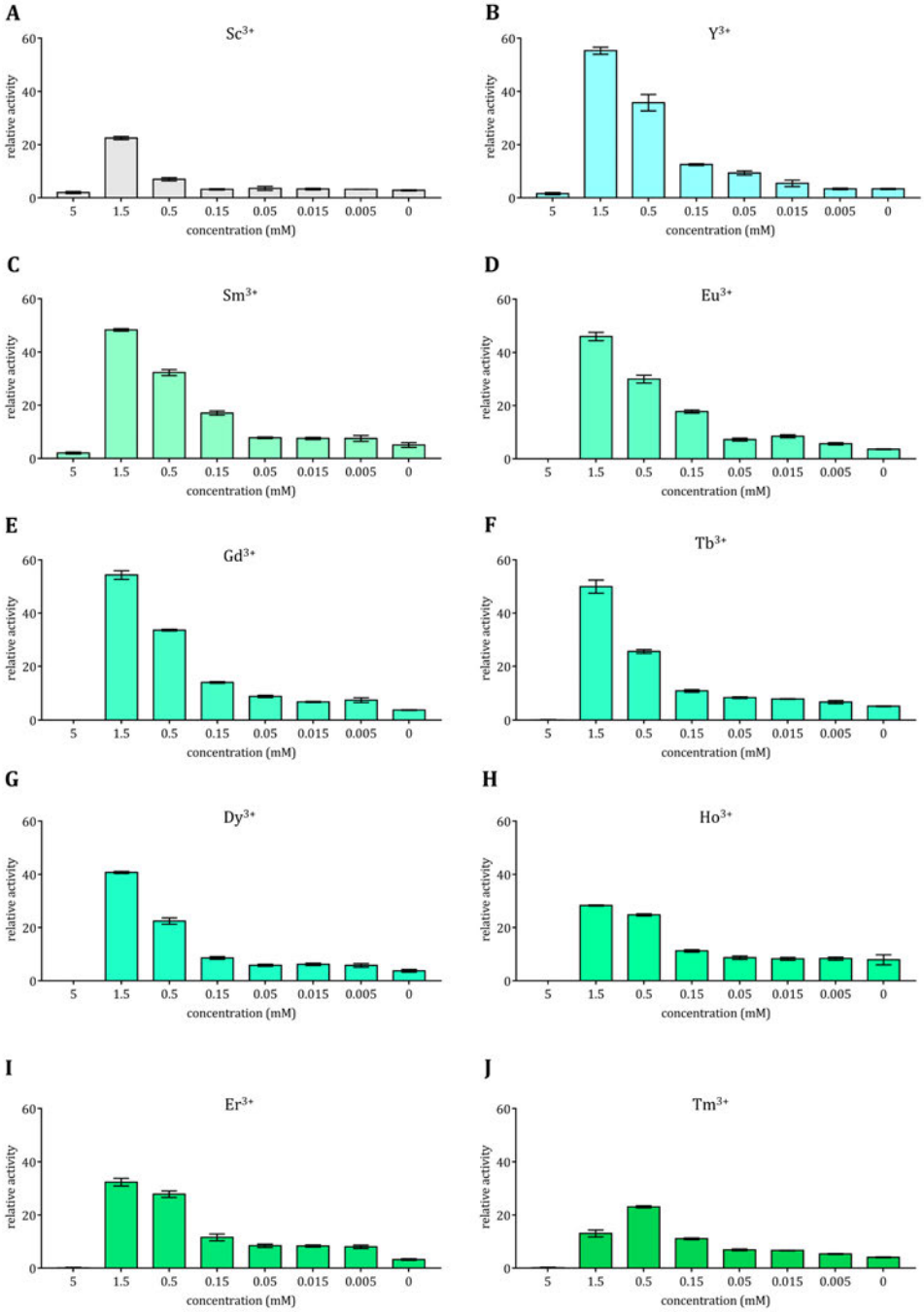


Figure S4. *Signature blots of metal ions vs. SrtA + cpd **3** and the active site mutants.* (**A**) and the active site mutants SrtA C164A (**B**) and SrtA C164S (**C**). displaying enthalpic (green) and temperature dependent entropic (red) contributions to the overall binding energy (blue). Errors are indicated as black bars. Figure was created using GraphPad Prism 7.0.4.[4]


A
Sc3+
B
Y3+
C
Sm3+
D
Eu3+
E
Gd3+
F
Tb3+
G
Dy3+
H
Ho3+
I
Er3+
J
Tm3+
relative activity
0
20
40
60
5
1.5
0.5
0.15
0.05
0.015
0.005
0
concentration (mM)

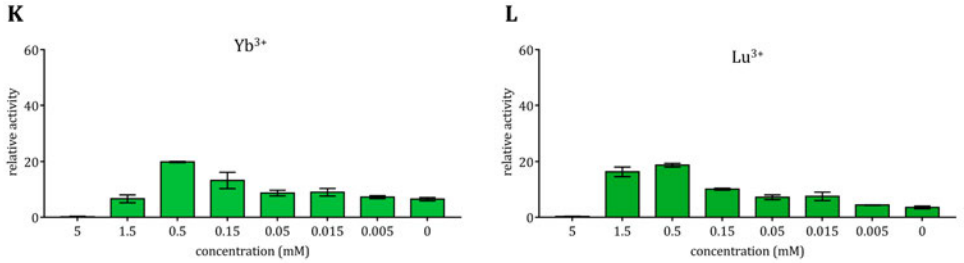
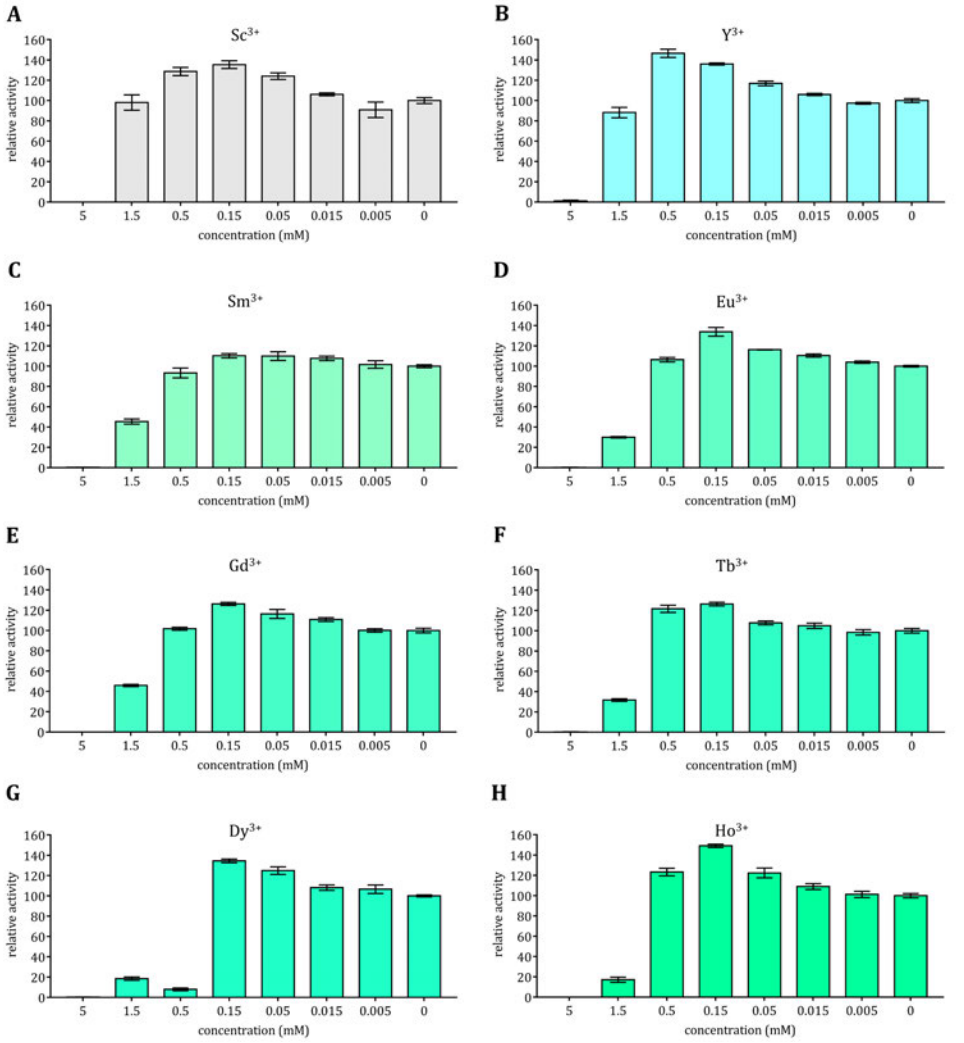


Figure S5. C*oncentration-dependent activation of SrtA by Sc^{3+}, Y^{3+}, Sm^{3+}, Eu^{3+}, Gd^{3+}, Tb^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, Yb^{3+}, and Lu^{3+} (**A–K**).* Figure was created using GraphPad Prism 7.0.4.[4] Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[5]



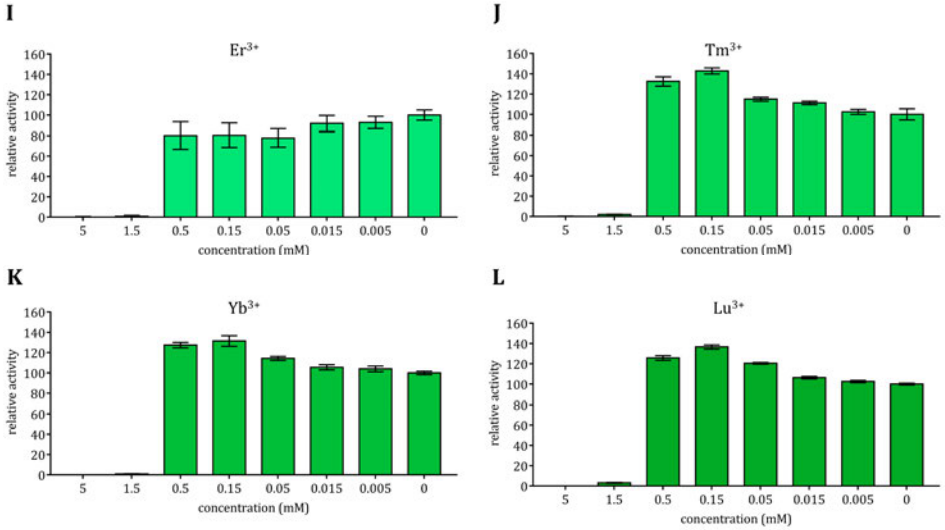
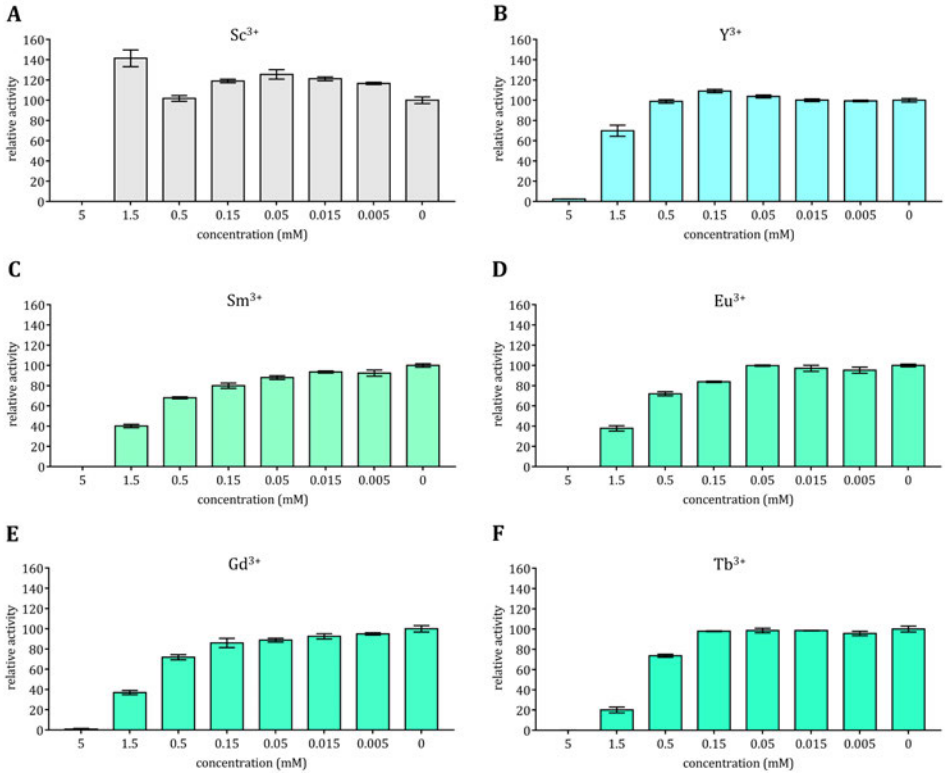


Figure S6. *Concentration-dependent effects of Sc^{3+}, Y^{3+}, Sm^{3+}, Eu^{3+}, Gd^{3+}, Tb^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, Yb^{3+}, and Lu^{3+} **(A–K)** on SrtA activity in presence of 5 mM Ca^{2+}.* Figure was created using GraphPad Prism 7.0.4.[4] Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[5]



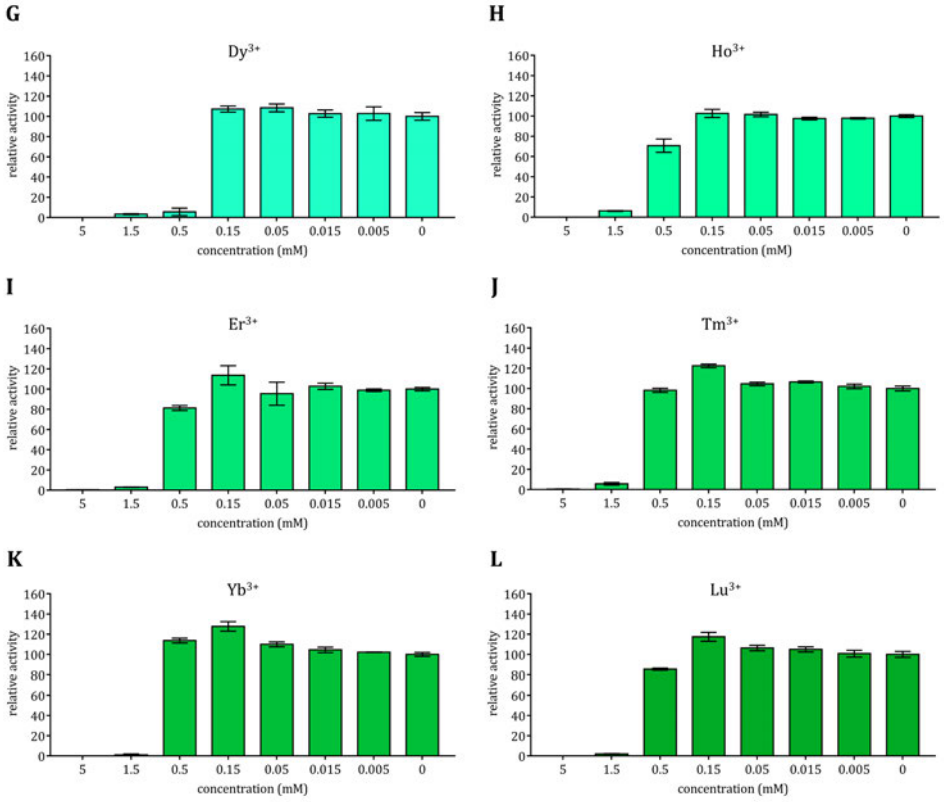


Figure S7. *Concentration-dependent effects of Sc^{3+}, Y^{3+}, Sm^{3+}, Eu^{3+}, Gd^{3+}, Tb^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, Yb^{3+}, and Lu^{3+} (**A**–**K**) on SrtA activity in presence of 1 mM Ca^{2+}.* Figure was created using GraphPad Prism 7.0.4.[4] Metal ions are Corey–Pauling–Koltun (CPK) colored according to the Jmol Atomic Color Scheme.[5]

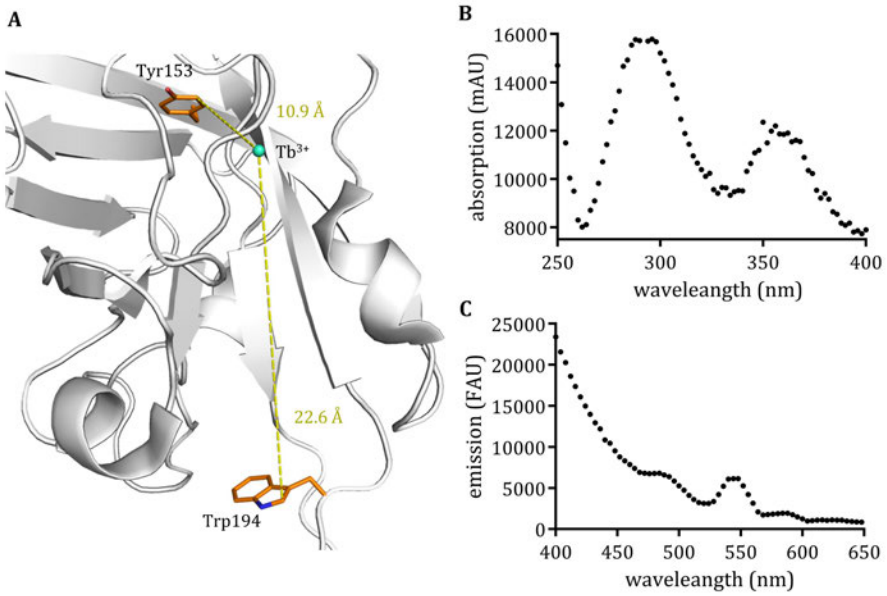


Figure S8. *Tb^{3+}-FRET assay.* (**A**) Relative positions of Tb^{3+} in the Ca^{2+} binding site of SrtA (PDB-ID 2KID). **B** absorption spectrum and **C** emission spectrum of Tb^{3+}.

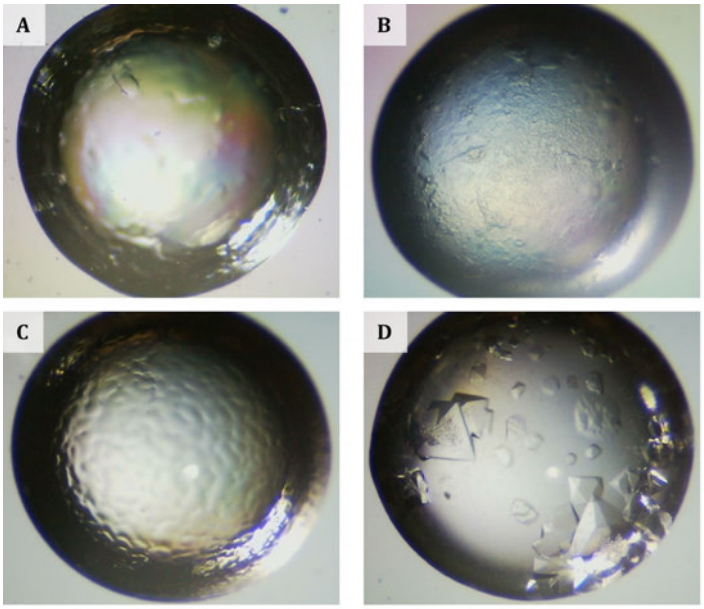
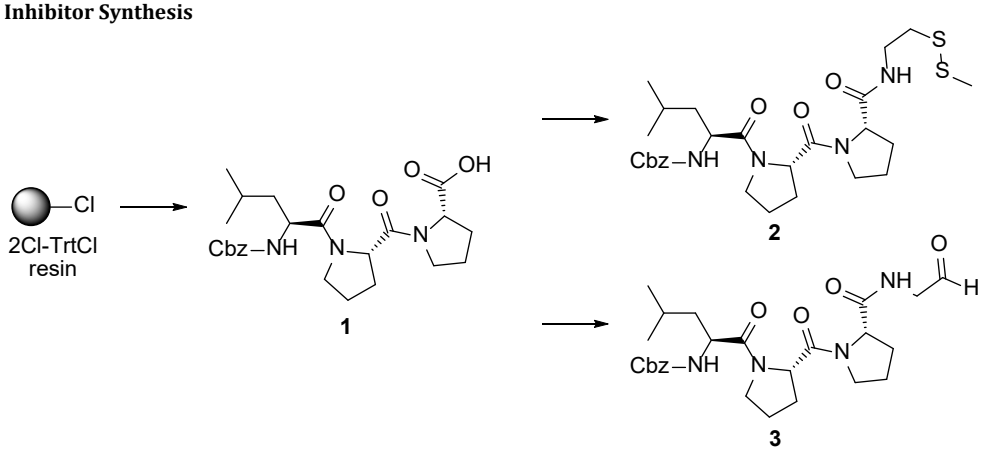


Figure S9. *Representative photographed crystallization droplets.* Image was recorded after 2 weeks at rt with a Microscope Wi-Fi CMOS C-mount Camera 5MP mounted on a Trinocular Stereomicroscope with 7x–45x zoom (Molecular Dimensions, Maumee, OH, USA) with reservoir solution containing 0.1 M NaCl, 25 mM MES pH 6.0, 2.9 M $(NH_4)_2SO_4$. (**A**) SrtA without any ligands showed phase separation as well as **C** SrtA with Ca^{2+}. (**B**) Addition of cpd. **3** prior to crystallization led to more gelatinous structures. (**D**) Only the combination of both Ca^{2+} and cpd. **3** led to SrtA crystals with square bipyramidal structures after 5 d.

Inhibitor Synthesis



Scheme SSXXX1. Synthesis of inhibitors **2** and **3**.

((Benzyloxy)carbonyl)-L-leucyl-L-prolyl-L-proline (1)

The title compound was synthesized by solid-phase synthesis using 2Cl-TrtCl resin (1.6 mmol/g, Novabiochem) and standard peptide coupling procedures.[6] ((Benzyloxy)carbonyl)-L-leucyl-L-prolyl-L-proline was obtained as a colorless solid (830 mg, 1.81 mmol, 78%). m.p.: 88–94 °C; ^{1}H NMR (300 MHz, DMSO-d_6): δ = 12.43 (s, 1H), 7.49–7.23 (m, 6H), 5.00 (s, 2H), 4.65–4.55 (m, 1H), 4.37–4.14 (m, 2H), 3.79–3.59 (m, 2H), 3.57–3.41 (m, 2H), 2.69 (s, 1H), 2.24–2.06 (m, 2H), 2.02–1.59 (m, 7H), 1.51–1.26 (m, 2H), 0.89 (d, *J* = 6.6 Hz, 6H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 173.3, 170.2, 169.7, 156.0, 137.0, 128.3, 127.8, 127.7, 65.3, 58.4, 57.4, 50.7, 46.6, 46.2, 39.6, 38.2, 28.4, 27.5, 24.5, 24.4, 24.0, 23.1, 21.3.

Benzyl ((*S*)-4-methyl-1-((*S*)-2-((*S*)-2-((2-(methyldisulfanyl)ethyl)carbamoyl)pyrrolidine-1-carbonyl)pyrrolidin-1-yl)-1-oxopentan-2-yl)carbamate (2, FB308)

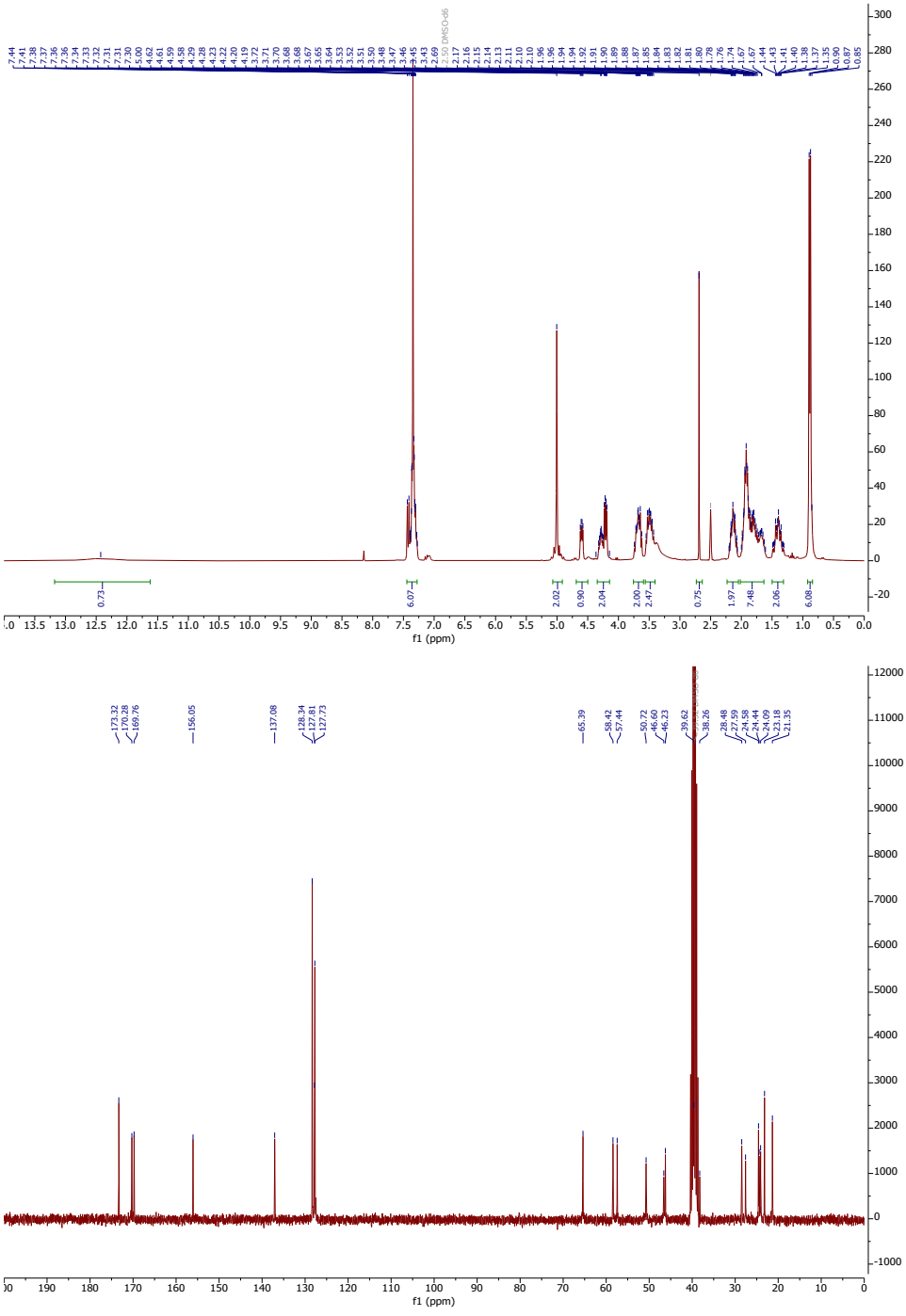
((Benzyloxy)carbonyl)-L-leucyl-L-prolyl-L-proline (92 mg, 0.20 mmol), 2-(methyldisulfanyl)ethan-1-amine (30 mg, 0.24 mmol), TBTU (71 mg, 0.22 mmol), and DIPEA (139 µL, 0.80 mmol) were suspended in ethyl acetate (10 mL) and stirred for 24 h at room temperature. The organic phase was extract with $NaHCO_3$ sat., HCl (1 M), and purified by column chromatography (DCM/MeOH 10:1) to yield the title compound as a colorless solid (80 mg, 0.14 mmol, 71%). m.p.: 85–92 °C; ^{1}H NMR (300 MHz, DMSO-d_6): δ = 7.92 (t, *J* = 5.6 Hz, 1H), 7.39–7.26 (m, 6H), 5.00 (s, 2H), 4.63–4.52 (m, 1H), 4.45–4.15 (m, 2H), 3.78–3.59 (m, 2H), 3.31–3.15 (m, 1H), 2.89–2.64 (m, 3H), 2.45–2.31 (m, 4H), 2.29–1.57 (m, 10H), 1.49–1.25 (m, 3H), 0.89 (d, *J* = 5.2 Hz, 6H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 171.6, 170.2, 169.9, 156.0, 137.0, 128.3, 127.8, 127.7, 127.6, 65.3, 59.3, 57.4, 50.6, 46.5, 37.8, 36.4, 29.1, 27.8, 24.4, 24.0, 23.1, 22.6, 21.3; ESI-MS *m/z*: $[M+Na]^+$ 587.24; purity (HPLC) = 99.3%.

Benzyl ((*S*)-4-methyl-1-oxo-1-((*S*)-2-((*S*)-2-((2-oxoethyl)carbamoyl)pyrrolidine-1-carbonyl)pyrrolidin-1-yl)pentan-2-yl)carbamate (3, FB313)

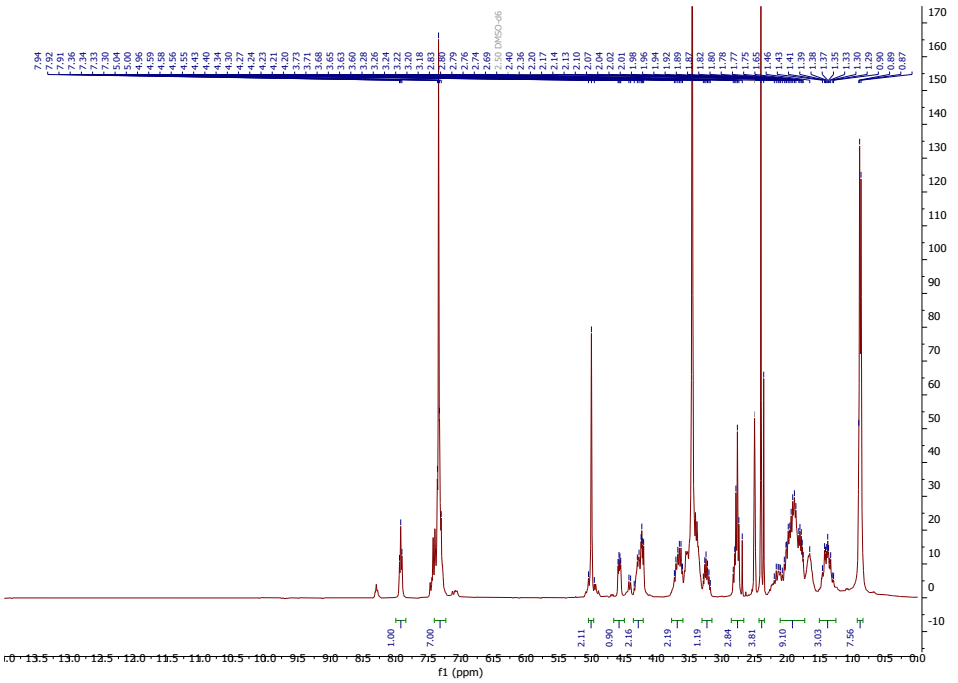
((Benzyloxy)carbonyl)-L-leucyl-L-prolyl-L-proline (92 mg, 0.20 mmol), 2,2-dimethoxy-1-aminoethane (26 µL, 0.24 mmol), TBTU (71 mg, 0.22 mmol), and DIPEA (139 µL, 0.80 mmol) were suspended in ethyl acetate (10 mL) and stirred for 24 h at room temperature. The organic phase was extract with $NaHCO_3$ sat., HCl (1 M), and purified by column chromatography (DCM/MeOH 10:1) to yield the title compound as a colorless solid (69 mg, 0.13 mmol, 69%). m.p.: 95–100 °C; ^{1}H NMR (300 MHz, DMSO-d_6): δ = 9.40 (s, 1H), 7.46–7.25 (m, 7H), 5.00 (s, 2H), 4.77–4.52 (m, 1H), 4.42–4.20 (m, 2H), 3.99–3.37 (m, 5H), 3.27–2.85 (m, 1H), 2.25–1.58 (m, 10H), 1.50–1.24 (m, 2H), 0.88 (d, *J* = 6.5 Hz, 6H); ^{13}C NMR (75 MHz, DMSO-d_6): δ = 199.9,

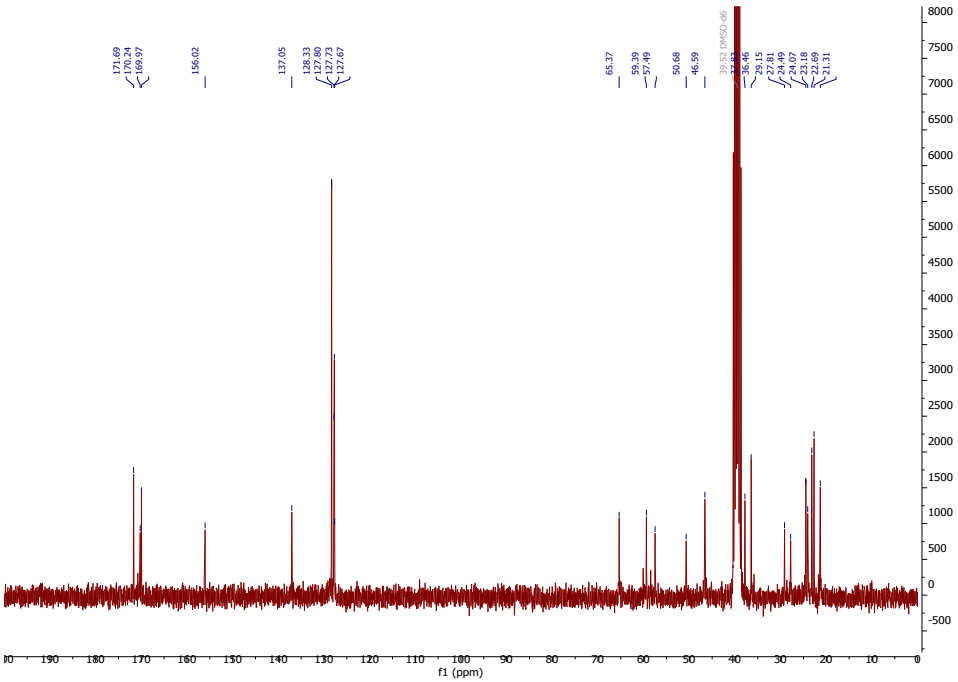
172.5, 171.6, 170.2, 170.0, 169.9, 156.0, 137.0, 128.3, 127.7, 127.7, 88.6, 65.3, 59.1, 57.4, 50.6, 49.1, 46.6, 45.7, 29.2, 27.7, 24.4, 24.0, 23.1, 21.3; ESI-MS *m/z*: $[M+H]^+$ 501.49; purity (HPLC) = 97.6%.

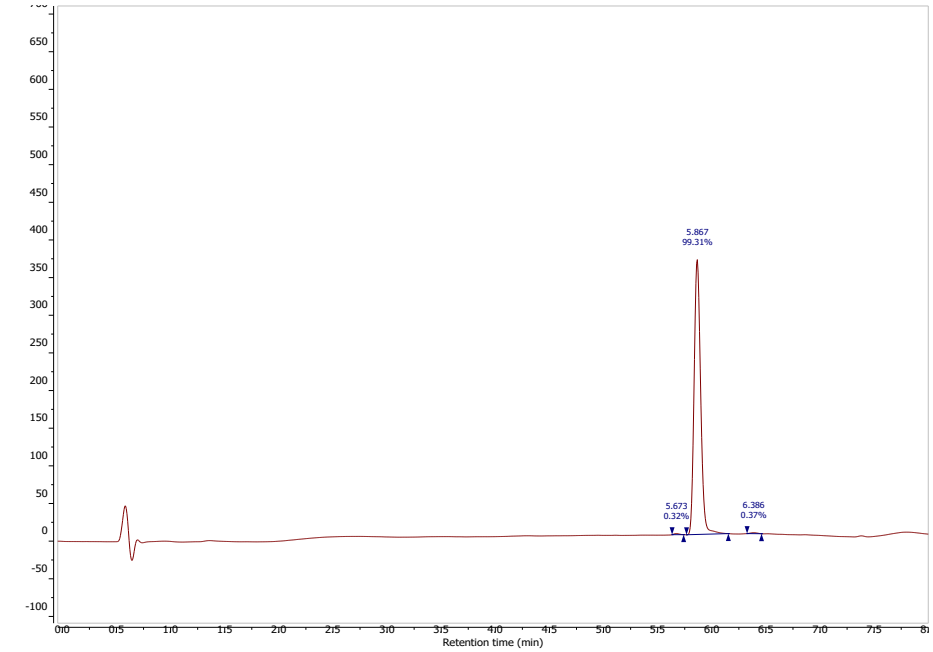
Cpd. 1

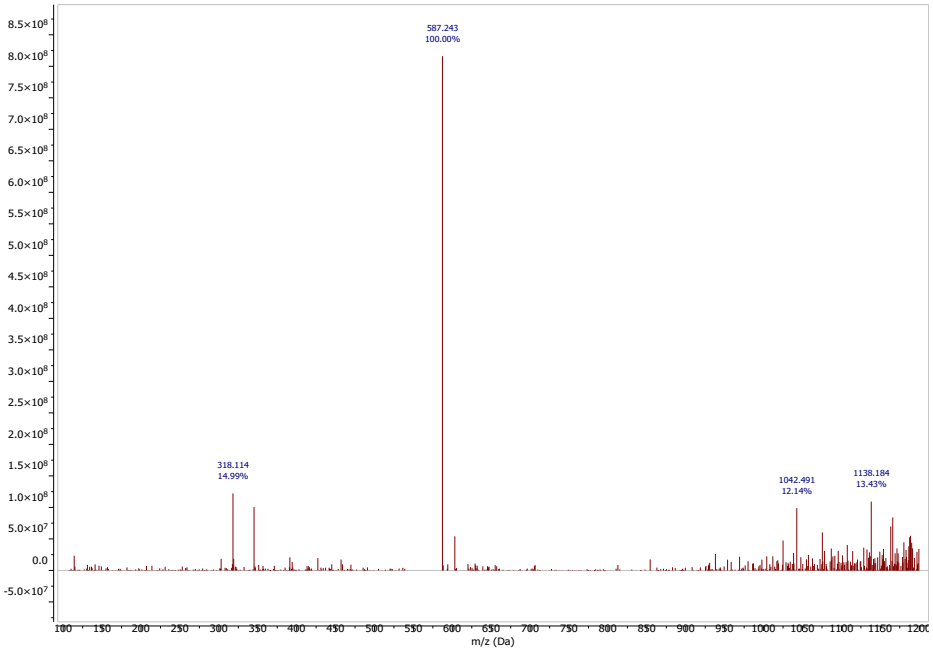


Cpd. 2

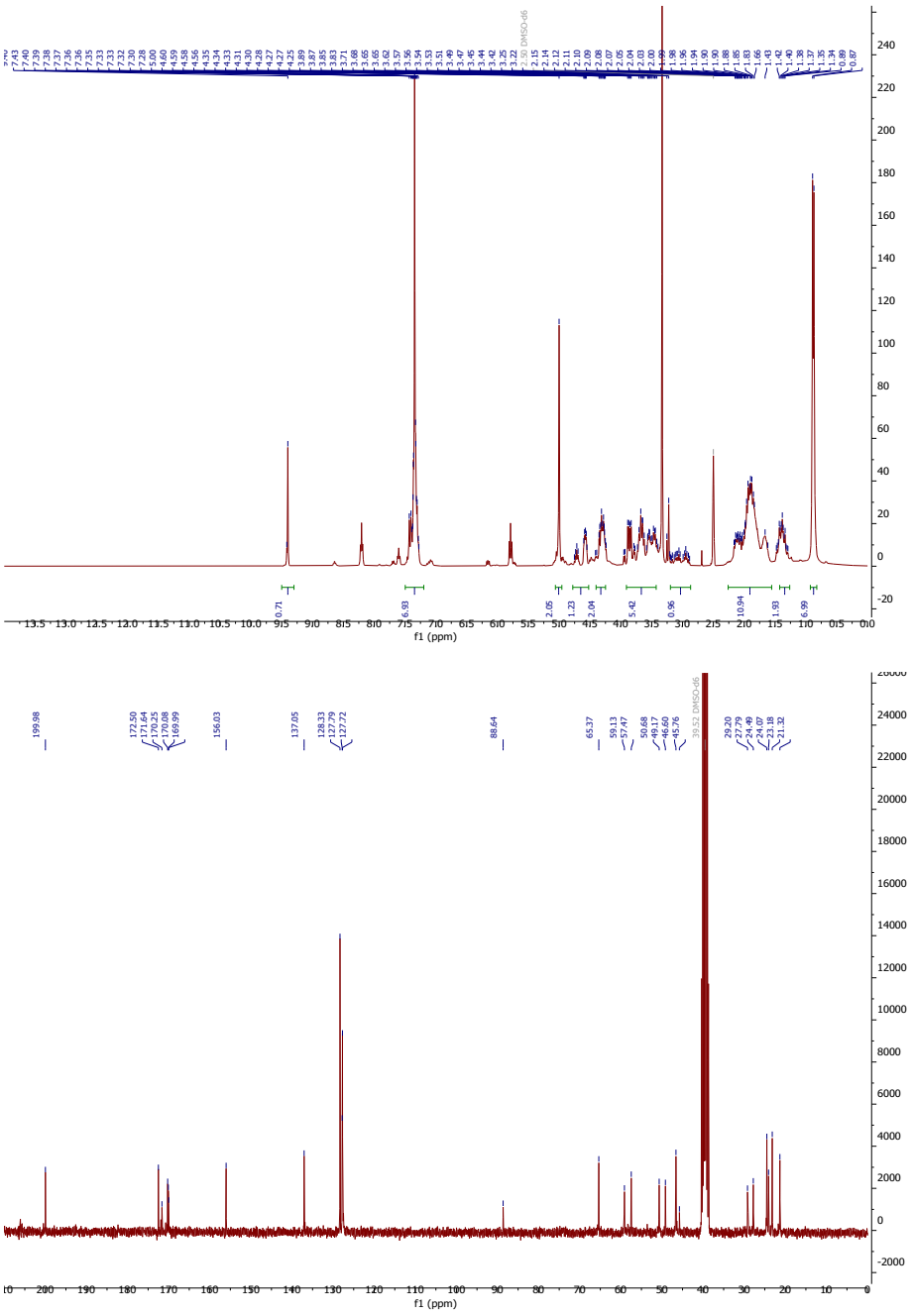


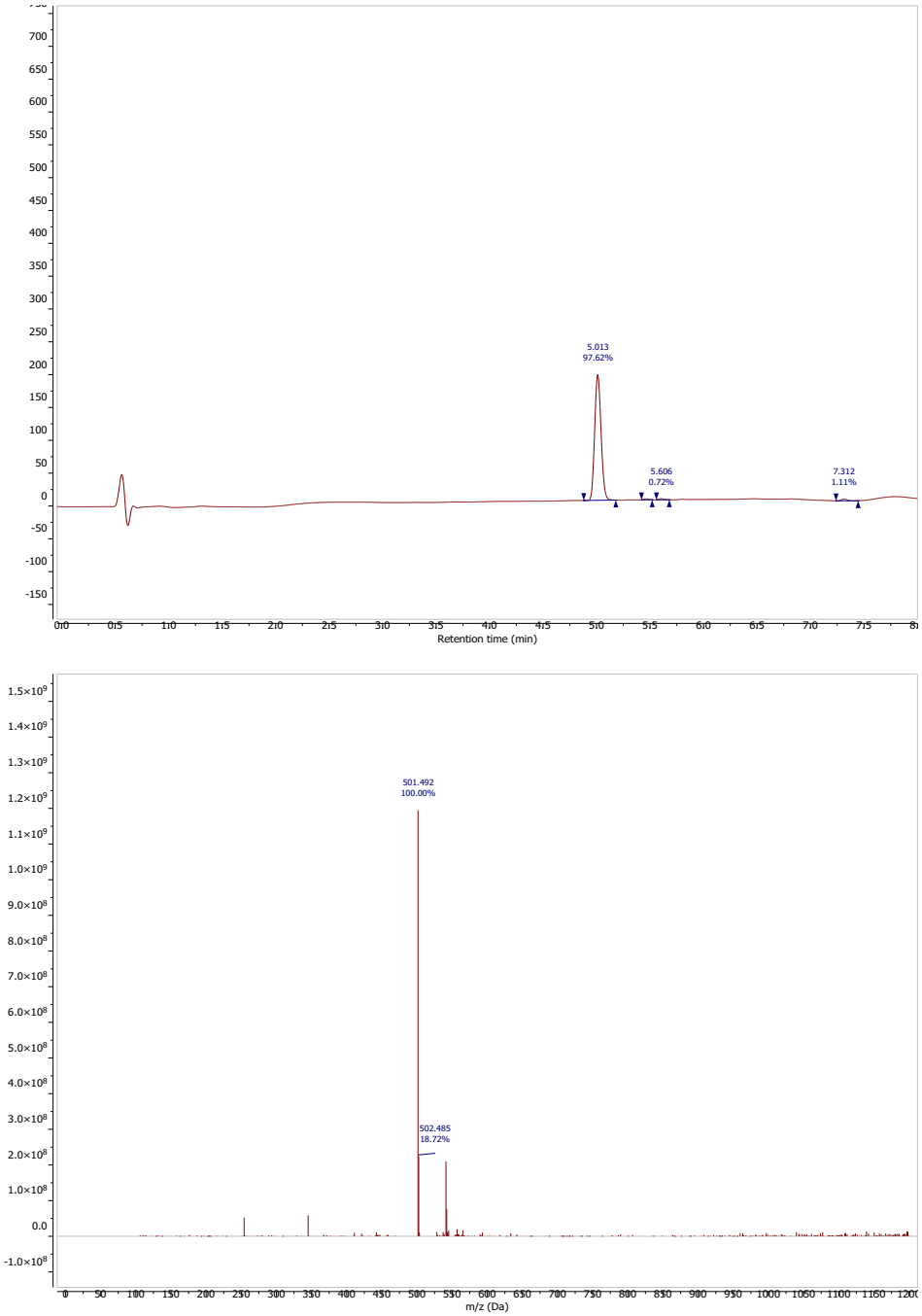



5.867
99.31%
5.673
0.32%
6.386
0.37%
Retention time (min)


587.243
100.00%
318.114
14.99%
1042.491
12.14%
1138.184
13.43%
m/z (Da)

Cpd. 3





Literature

1. GraphPad Prism 7.0.4 for Windows.

2. Kepp, K. P. A quantitative scale of oxophilicity and thiophilicity. *Inorg. Chem.* **55**, 9461–9470 (2016).

3. Bard, A. J., Parsons, R. & Jordan, J. *Standard Potentials in Aqueous Solution*. (Routledge, 2017). doi:10.1201/9780203738764.

4. Software, G. Graphpad prism for windows. (2014).

5. Willighagen, E. L. Processing CML conventions in Java. *Internet J. Chem.* **4**, 1–9 (2001).

6. Kreutzer, A. G., Salveson, P. J., Hyunjun, Y. & Guaglianone, G. *Standard practices for Fmoc-based solid-phase peptide synthesis in the Nowick laboratory*. vol. 1.7.2 http://www.chem.uci.edu/~jsnowick/groupweb/files/Standard_practices_for_Fmoc_based_solid_phase_peptide_synthesis_in_the_Nowick_Laboratory_V_1point6.pdf (2020).

7. Conclusions and Outlook

In the course of this doctoral thesis, four main projects were addressed.

Project 1. In the first project, the entropic optimization of ZIKV and MT-SP1 inhibitors was investigated. The results demonstrated that predictions of entropic binding contributions are still inaccurate and that commonly accepted dogmas must be taken cautiously. The higher affinity of macrocyclic of ZIKV NS2B/NS3 inhibitors was not due to an entropic benefit achieved by reducing the necessity of ligand preorganization. Conversely, the macrocycle experienced an entropic penalty compared to some linear reference cpds, which retained higher flexibility in the complexed state (*3.1 Thermodynamic characterization of a macrocyclic Zika virus NS2B/NS3 protease inhibitor and its acyclic analogs*). On the other hand, the introduction of higher ligand symmetry of MT-SP1 inhibitors to enable multiple indistinguishable binding modes yielded an entropic benefit that could be accurately predicted by calculating the resulting statistical benefit. Although not pronounced, such symmetry effects are neglected mainly in commonly used prediction models and literature. Hence, to our knowledge, this work was the first to quantify the entropic benefit of higher ligand symmetry by ITC (*3.2 Improving Binding Entropy by Introducing Higher Ligand Symmetry*). Studies with other sets of enzymes and inhibitors are crucial for final assessments of these two strategies to achieve inhibitors with improved binding affinities.

To our knowledge, no educational literature directed at post-graduates in medicinal chemistry and related disciplines is available teaching advanced ITC methodology. Hence, we designed teaching experiments covering workflows to determine the number of protons transferred upon ligand binding (n_{proton}) and the respective buffer ionization correction. Furthermore, these experiments demonstrate the determination of binding-induced changes in isobaric heat capacities (ΔC_P) and the utilization of displacement experiments for analyzing ligands displaying too-low or too-high affinities for direct titration experiments. Notably, these experiments are already utilized for the training of students in pharmacy. To guide these advanced evaluation processes, we implemented the openly accessible web server *ITCcalc* (*3.3 Advanced Isothermal Titration Calorimetry for Medicinal Chemists with ITCcalc*). Although this work covers some ITC-related gaps in educational papers, there are still some missing spots that could be regarded as further advanced methodologies, such as kinetic evaluations of ITC experiments,[163,166,268] or the characterization of enzyme kinetics by using the recently implemented technique 2D-ITC.[269,270]

Project 2. The second project addressed the binding mode of allosteric inhibitors of flaviviral NS2B/NS3 proteases. The binding conformation was assessed by docking experiments against open and closed crystal structures of DENV and ZIKV NS2B/NS3 proteases. ROC curves of both conformational pairs revealed a slightly better discriminating of binders from nonbinders for the open conformation. The allosteric binding site was mapped by Cys-hopping and subsequent maleimide coupling to block the binding site (*4.2 Proline-Based Allosteric Inhibitors of Zika and Dengue Virus NS2B/NS3 Proteases*). In succeeding attempts to determine the allosteric mode of action, several methodologies like EPR, ^{19}F-NMR, (nano)DSF, and smFRET were successfully used to sense the conformational transitions of the protease. However, undisputable assignment of the respective conformation could only be provided by EPR and smFRET, determining the distances between two conformation-sensitive labels (*4.3 The effects of allosteric and competitive inhibitors on ZIKV protease conformational dynamics explored through smFRET, nanoDSF, DSF, and* ^{19}F-NMR).[241,256,271–273] Although the binding conformation of allosteric inhibitors could be pinpointed as the open conformation, several crucial pieces of information regarding the allosteric mode of inhibition are still missing. Whether the open conformation is stabilized by an induced fit or a conformational selection mechanism could be deconvoluted from smFRET experiments on immobilized enzymes, allowing us to determine the residence times of the respective conformations in the presence and absence of different inhibitors. The precise binding mode can be obtained by co-crystallization. Since this work is the first to report bZiPro crystals obtained in the apo-state, these preliminary results enable soaking experiments with allosteric ligands in crystals expected to be in the open conformation (*1.2.2 Project 2: Chasing the Binding Conformation of NS2B/NS3 Protease Inhibitors*).

Project 3. This doctoral thesis's third project concerned the autocatalytic cleavage behavior of ZIKV NS2B/NS3. In the first article, an artificial glycine-linked construct was examined. Employing an autocleavage assay, we demonstrated this cleavage process to occur exclusively in *cis*. We showed that only one point mutation R95*A is required to attenuate autocatalytic cleavage and that Arg29* only plays neglectable roles. (*5.1 Cis autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3pro constructs*). As a result, we yielded a new minimally modified construct suitable for further studies. In the follow-up manuscript, the autocleavage behavior of a closer-to-native construct, harboring the native autocleavage site connecting $NS2B_{cf}$ and $NS3_{pro}$ (eZiPro), was investigated. We mutated all basic residues in this R95*-EKTGKR101* site to Ala, including all 16 Ala-permutations and their respective active site S135A mutants. By densitometric analysis of SDS-PAGE, we evaluated the cleavage efficiencies of the constructs and correlated linker integrity with enzymatic activity, revealing tremendous influences on K_M and k_{cat} values. Cleavage of the linker increased k_{cat} but had mixed effects on K_M. Protein mass spectrometry determined the respective cleavage fragments and demonstrated the preferred autocatalytic cleavage after

Arg101*. If this site was subsequently mutated to Ala, other cleavage sites gained relevance, possibly pointing at evasive mechanisms to maintain autocatalytic cleavage. Prior publication of this manuscript, further experiments need to be performed to determine binding affinities of the cleavage sites after catalysis to reveal possible auto-inhibitory effects, explaining the mixed impact on K_M values (*5.2 Insights into the autocleavage behavior of the ZIKV NS2B/NS3 protease*).

Project 4. The last project was dedicated to SrtA and its modulation using different cationic metal ions to compete for binding to the allosteric Ca^{2+} binding site. Therefore, a set of 47 cations was evaluated by enzyme activity assays, Ca^{2+} competitive assays, DSF, MST, ITC, and X-ray crystallography. Several inhibitory and activating ions were identified with even higher binding affinities than Ca^{2+}. Particularly lanthanides were found to bind SrtA tightly. La^{3+} activated SrtA to a greater extent than Ca^{2+}. Other lanthanides modulated or strongly inhibited SrtA activity, probably due to unspecific binding, distinct from the Ca^{2+} site. This finding could be concluded from multiple binding events in ITC experiments. ITC experiments in the presence of covalent inhibitors miming the LPXT sorting signal of SrtA revealed the cooperativity of activating cations as observed earlier for Ca^{2+}.[156] X-ray crystallography yielded crystals of SrtA in complex with Ca^{2+}, Mg^{2+}, Mn^{2+}, and La^{3+}. However, only one structure in complex with Mn^{2+} has been solved so far (*6.1 Interactions of Staphylococcus* aureus *Sortase A with Various Cations: Basis for the Design of Innovative Inhibitors*). Those lanthanides and other inhibiting ions could prove valuable as anti-microbial additives in alloys and coatings of medical devices. One could also think of ion-chelating peptides harboring the LPXTG sorting signal to achieve controlled drug delivery, thus reducing the side effects of toxic cations. After processing by SrtA, they could release inhibitory cations.

Altogether, this work utilized a toolset of biophysical methods to describe ligand binding to pathologically relevant proteases. New insights into general attempts to improve binding entropy and essential information about protease-specific inhibition modes that will benefit future scientists in drug development were gained.

8. Bibliography

(1) Mitra, A.; Mawson, A. Neglected Tropical Diseases: Epidemiology and Global Burden. *Trop. Med. Infect. Dis.* **2017**, *2* (3), 36. https://doi.org/10.3390/tropicalmed2030036.

(2) Hayes, E. B. Zika Virus Outside Africa. *Emerg. Infect. Dis.* **2009**, *15* (9), 1347–1350. https://doi.org/10.3201/eid1509.090442.

(3) Powell, J. R. Mosquitoes on the Move. *Science (80-.).* **2016**, *354* (6315), 971–972. https://doi.org/10.1126/science.aal1717.

(4) Collins, M. H.; Metz, S. W. Progress and Works in Progress: Update on Flavivirus Vaccine Development. *Clin. Ther.* **2017**, *39* (8), 1519–1536. https://doi.org/10.1016/j.clinthera.2017.07.001.

(5) WHO. Zika Virus Disease – India. **2021**, *14 October*.

(6) PAHO/WHO. Zika Suspected and Confirmed Cases Reported by Countries and Territories in the Americas Cumulative Cases, 2015-2017. *Paho/Who* **2017**, No. May 2017, 1.

(7) Bhatt, S.; Gething, P. W.; Brady, O. J.; Messina, J. P.; Farlow, A. W.; Moyes, C. L.; Drake, J. M.; Brownstein, J. S.; Hoen, A. G.; Sankoh, O.; Myers, M. F.; George, D. B.; Jaenisch, T.; Wint, G. R. W.; Simmons, C. P.; Scott, T. W.; Farrar, J. J.; Hay, S. I. The Global Distribution and Burden of Dengue. *Nature* **2013**, *496* (7446), 504–507. https://doi.org/10.1038/nature12060.

(8) Brady, O. J.; Gething, P. W.; Bhatt, S.; Messina, J. P.; Brownstein, J. S.; Hoen, A. G.; Moyes, C. L.; Farlow, A. W.; Scott, T. W.; Hay, S. I. Refining the Global Spatial Limits of Dengue Virus Transmission by Evidence-Based Consensus. *PLoS Negl. Trop. Dis.* **2012**, *6* (8), e1760. https://doi.org/10.1371/journal.pntd.0001760.

(9) Cao-Lormeau, V.-M.; Blake, A.; Mons, S.; Lastère, S.; Roche, C.; Vanhomwegen, J.; Dub, T.; Baudouin, L.; Teissier, A.; Larre, P.; Vial, A.-L.; Decam, C.; Choumet, V.; Halstead, S. K.; Willison, H. J.; Musset, L.; Manuguerra, J.-C.; Despres, P.; Fournier, E.; Mallet, H.-P.; Musso, D.; Fontanet, A.; Neil, J.; Ghawché, F. Guillain-Barré Syndrome Outbreak Associated with Zika Virus Infection in French Polynesia: A Case-Control Study. *Lancet* **2016**, *387* (10027), 1531–1539. https://doi.org/10.1016/S0140-6736(16)00562-6.

(10) 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.

(11) Chan, J. F. W.; Choi, G. K. Y.; Yip, C. C. Y.; Cheng, V. C. C.; Yuen, K.-Y. Zika Fever and Congenital Zika Syndrome: An Unexpected Emerging Arboviral Disease. *J. Infect.* **2016**, *72* (5), 507–524. https://doi.org/10.1016/j.jinf.2016.02.011.

(12) WHO. The History of Zika Virus. **2016**, *7. Februar*.

(13) WHO. *Dengue and severe dengue*. Fact-sheet. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.

(14) Guzman, M. G.; Halstead, S. B.; Artsob, H.; Buchy, P.; Farrar, J.; Gubler, D. J.; Hunsperger, E.; Kroeger, A.; Margolis, H. S.; Martínez, E.; Nathan, M. B.; Pelegrino, J. L.; Simmons, C.; Yoksan, S.; Peeling, R. W. Dengue: A Continuing Global Threat. *Nat. Rev. Microbiol.* **2010**, *8* (S12), S7–S16. https://doi.org/10.1038/nrmicro2460.

(15) Katzelnick, L. C.; Gresh, L.; Halloran, M. E.; Mercado, J. C.; Kuan, G.; Gordon, A.; Balmaseda, A.; Harris, E. Antibody-Dependent Enhancement of Severe Dengue Disease in Humans. *Science (80-.).* **2017**, *358* (6365), 929–932. https://doi.org/10.1126/science.aan6836.

(16) Halstead, S. B. Dengvaxia Sensitizes Seronegatives to Vaccine Enhanced Disease Regardless of Age. *Vaccine* **2017**, *35* (47), 6355–6358. https://doi.org/10.1016/j.vaccine.2017.09.089.

(17) Villar, L.; Dayan, G. H.; Arredondo-García, J. L.; Rivera, D. M.; Cunha, R.; Deseda, C.; Reynales, H.; Costa, M. S.; Morales-Ramírez, J. O.; Carrasquilla, G.; Rey, L. C.; Dietze, R.; Luz, K.; Rivas, E.; Miranda Montoya, M. C.; Cortés Supelano, M.; Zambrano, B.; Langevin, E.; Boaz, M.; Tornieporth, N.; Saville, M.; Noriega, F. Efficacy of a Tetravalent Dengue Vaccine in Children in Latin America. *N. Engl. J. Med.* **2015**, *372* (2), 113–123. https://doi.org/10.1056/NEJMoa1411037.

(18) Mustafa, M. S.; 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.

(19) Fernandez, E.; Diamond, M. S. Vaccination Strategies against Zika Virus. *Curr. Opin. Virol.* **2017**, *23*, 59–67. https://doi.org/10.1016/j.coviro.2017.03.006.

(20) Hazlewood, J. E.; Tang, B.; Yan, K.; Rawle, D. J.; Harrison, J. J.; Hall, R. A.; Hobson-Peters, J.; Suhrbier, A. The Chimeric Binjari-Zika Vaccine Provides Long-Term Protection against ZIKA Virus Challenge. *Vaccines* **2022**, *10* (1), 85. https://doi.org/10.3390/vaccines10010085.

(21) Torresi, J.; Ebert, G.; Pellegrini, M. Vaccines Licensed and in Clinical Trials for the Prevention of Dengue. *Hum. Vaccin. Immunother.* **2017**, *13* (5), 1059–1072. https://doi.org/10.1080/21645515.2016.1261770.

(22) Braack, L.; Gouveia de Almeida, A. P.; Cornel, A. J.; Swanepoel, R.; de Jager, C. Mosquito-Borne Arboviruses of African Origin: Review of Key Viruses and Vectors. *Parasit. Vectors* **2018**, *11* (1), 29. https://doi.org/10.1186/s13071-017-2559-9.

(23) Simmonds, P.; Becher, P.; Bukh, J.; Gould, E. A.; Meyers, G.; Monath, T.; Muerhoff, S.; Pletnev, A.; Rico-Hesse, R.; Smith, D. B.; Stapleton, J. T. ICTV Virus Taxonomy Profile: Flaviviridae. *J. Gen. Virol.* **2017**, *98* (1), 2–3. https://doi.org/10.1099/jgv.0.000672.

(24) Falgout, B.; Pethel, M.; Zhang, Y. M.; Lai, C. J. Both Nonstructural Proteins NS2B and NS3 Are Required for the Proteolytic Processing of Dengue Virus Nonstructural Proteins. *J. Virol.* **1991**, *65* (5), 2467–2475. https://doi.org/10.1128/jvi.65.5.2467-2475.1991.

(25) Lindenbach, B. D.; Thiel, H.-J.; Rice, C. M. Flaviviridae: The Viruses and Their Replication. *Fields Virol.* **2007**, 1101–1151.

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

(27) 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), e0134823. https://doi.org/10.1371/journal.pone.0134823.

(28) Yusof, R.; Clum, S.; Wetzel, M.; Murthy, H. M. K.; 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.

(29) Shannon, A. E.; Chappell, K. J.; Stoermer, M. J.; Chow, S. Y.; Kok, W. M.; Fairlie, D. P.; Young, P. R. Simultaneous Uncoupled Expression and Purification of the Dengue Virus NS3 Protease and NS2B Co-Factor Domain. *Protein Expr. Purif.* **2016**, *119*, 124–129. https://doi.org/10.1016/j.pep.2015.11.022.

(30) Noinaj, N.; Easley, N. C.; Oke, M.; Mizuno, N.; Gumbart, J.; Boura, E.; Steere, A. N.; Zak, O.; Aisen, P.; Tajkhorshid, E.; Evans, R. W.; Gorringe, A. R.; Mason, A. B.; Steven, A. C.; Buchanan, S. K. Structural Basis for Iron Piracy by Pathogenic Neisseria. *Nature* **2012**, *483* (7387), 53–58. https://doi.org/10.1038/nature10823.

(31) Voss, S.; Nitsche, C. Inhibitors of the Zika Virus Protease NS2B-NS3. *Bioorg. Med. Chem. Lett.* **2020**, *30* (5), 126965. https://doi.org/10.1016/j.bmcl.2020.126965.

(32) Rawlings, N. D. MEROPS: The Peptidase Database. *Nucleic Acids Res.* **2004**, *32* (90001), 160D – 164. https://doi.org/10.1093/nar/gkh071.

(33) Chambers, T. J.; Weir, R. C.; Grakoui, A.; McCourt, D. W.; Bazan, J. F.; Fletterick, R. J.; Rice, C. M. Evidence That the N-Terminal Domain of Nonstructural Protein NS3 from Yellow Fever Virus Is a Serine Protease Responsible for Site-Specific Cleavages in the Viral Polyprotein. *Proc. Natl. Acad. Sci.* **1990**, *87* (22), 8898–8902. https://doi.org/10.1073/pnas.87.22.8898.

(34) 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* (4), 372–373. https://doi.org/10.1038/nsmb1073.

(35) Kang, C.; Gayen, S.; Wang, W.; Severin, R.; Chen, A. S.; Lim, H. A.; Chia, C. S. B.; Schüller, A.; Doan, D. N. P.; Poulsen, A.; Hill, J.; Vasudevan, S. G.; Keller, T. H. Exploring the Binding of Peptidic West Nile Virus NS2B–NS3 Protease Inhibitors by NMR. *Antiviral Res.* **2013**, *97* (2), 137–144. https://doi.org/10.1016/j.antiviral.2012.11.008.

(36) Poulsen, A.; Kang, C.; Keller, T. Drug Design For Flavivirus Proteases: What Are We Missing? *Curr. Pharm. Des.* **2014**, *20* (21), 3422–3427. https://doi.org/10.2174/13816128113199990633.

(37) Gruba, N.; Rodriguez Martinez, J. I.; Grzywa, R.; Wysocka, M.; Skoreński, M.; Burmistrz, M.; Łęcka, M.; Lesner, A.; Sieńczyk, M.; Pyrć, K. Substrate Profiling of Zika Virus NS2B-NS3 Protease. *FEBS Lett.* **2016**, *590* (20), 3459–3468. https://doi.org/10.1002/1873-3468.12443.

(38) Rut, W.; Zhang, L.; Kasperkiewicz, P.; Poreba, M.; Hilgenfeld, R.; Drąg, M. Extended Substrate Specificity and First Potent Irreversible Inhibitor/Activity-Based Probe Design for Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *139*, 88–94. https://doi.org/10.1016/j.antiviral.2016.12.018.

(39) Li, J.; Lim, S. P.; Beer, D.; Patel, V.; Wen, D.; Tumanut, C.; Tully, D. C.; Williams, J. A.; Jiricek, J.; Priestle, J. P.; Harris, J. L.; Vasudevan, S. G. Functional Profiling of Recombinant NS3 Proteases from All Four Serotypes of Dengue Virus Using Tetrapeptide and Octapeptide Substrate Libraries. *J. Biol. Chem.* **2005**, *280* (31), 28766–28774. https://doi.org/10.1074/jbc.M500588200.

(40) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(41) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D. Structure of the NS2B-NS3 Protease from Zika Virus after Self-Cleavage. *Nat. Commun.* **2016**, *7*, 13410. https://doi.org/10.1038/ncomms13410.

(42) Ren, J.; Lee, H.; Kotak, A.; Johnson, M. E. MD Simulations Reveal Alternate Conformations of the Oxyanion Hole in the Zika Virus NS2B/NS3 Protease. *Proteins Struct. Funct. Bioinforma.* **2020**, *88* (2), 345–354. https://doi.org/10.1002/prot.25809.

(43) Harper, E.; Berger, A. On the Size of the Active Site in Proteases: Pronase. *Biochem. Biophys. Res. Commun.* **1972**, *46* (5), 1956–1960. https://doi.org/10.1016/0006-291X(72)90076-9.

(44) The PyMOL Molecular Graphics System. *CCP4 Newsletter On Protein Crystallography*. Schrödinger 2002, pp 82–92. http://www.pymol.org.

(45) Falgout, B.; Miller, R. H.; Lai, C. J. Deletion Analysis of Dengue Virus Type 4 Nonstructural Protein NS2B: Identification of a Domain Required for NS2B-NS3 Protease Activity. *J. Virol.* **1993**, *67* (4), 2034–2042. https://doi.org/10.1128/jvi.67.4.2034-2042.1993.

(46) 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* (21), 9753–9771. https://doi.org/10.1021/acs.jmedchem.9b01196.

(47) Phoo, W. W.; Zhang, Z.; Wirawan, M.; Chew, E. J. C.; Chew, A. B. L.; Kouretova, J.; Steinmetzer, T.; Luo, D. Structures of Zika Virus NS2B-NS3 Protease in Complex with Peptidomimetic Inhibitors. *Antiviral Res.* **2018**, *160*, 17–24. https://doi.org/10.1016/j.antiviral.2018.10.006.

(48) Barthels, F.; Schirmeister, T.; Kersten, C. BANΔIT: B'-Factor Analysis for Drug Design and Structural Biology. *Mol. Inform.* **2021**, *40* (1), 2000144. https://doi.org/10.1002/minf.202000144.

(49) Carugo, O. B -Factor Accuracy in Protein Crystal Structures. *Acta Crystallogr. Sect. D Struct. Biol.* **2022**, *78* (1), 69–74. https://doi.org/10.1107/S2059798321011736.

(50) Kim, Y. M.; Gayen, S.; Kang, C.; Joy, J.; Huang, Q.; Chen, A. S.; Wee, J. L. K.; Ang, M. J. Y.; Lim, H. A.; Hung, A. W.; Li, R.; Noble, C. G.; Lee, L. T.; Yip, A.; Wang, Q.-Y.; Chia, C. S. B.; Hill, J.; Shi, P.-Y.; Keller, T. H. NMR Analysis of a Novel Enzymatically Active Unlinked Dengue NS2B-NS3 Protease Complex. *J. Biol. Chem.* **2013**, *288* (18), 12891–12900. https://doi.org/10.1074/jbc.M112.442723.

(51) Götz, C.; Hinze, G.; Gellert, A.; Maus, H.; von Hammerstein, F.; Hammerschmidt, S. J.; Lauth, L. M.; Hellmich, U. A.; Schirmeister, T.; Basché, T. 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.

(52) 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* (5), e1006411. https://doi.org/10.1371/journal.ppat.1006411.

(53) Hill, M. E.; Yildiz, M.; Hardy, J. A. 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.

(54) 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* (1), 438–446. https://doi.org/10.1128/JVI.06225-11.

(55) da Silva-Júnior, E. F.; de Araújo-Júnior, J. X. Peptide Derivatives as Inhibitors of NS2B-NS3 Protease from Dengue, West Nile, and Zika Flaviviruses. *Bioorg. Med. Chem.* **2019**, *27* (18), 3963–3978. https://doi.org/10.1016/j.bmc.2019.07.038.

(56) Huber, S.; Braun, N. J.; Schmacke, L. C.; Quek, J. P.; Murra, R.; Bender, D.; Hildt, E.; Luo, D.; Heine, A.; Steinmetzer, T. Structure-Based Optimization and Characterization of Macrocyclic Zika Virus NS2B-NS3 Protease Inhibitors. *J. Med. Chem.* **2022**, *65* (9), 6555–6572. https://doi.org/10.1021/acs.jmedchem.1c01860.

(57) Braun, N. J.; Quek, J. P.; Huber, S.; Kouretova, J.; Rogge, D.; Lang-Henkel, H.; Cheong, E. Z. K.; Chew, B. L. A.; Heine, A.; Luo, D.; Steinmetzer, T. Structure-Based Macrocyclization of Substrate Analogue NS2B-NS3 Protease Inhibitors of Zika, West Nile and Dengue Viruses. *ChemMedChem* **2020**, *15* (15), 1439–1452. https://doi.org/10.1002/cmdc.202000237.

(58) Nitsche, C.; Onagi, H.; Quek, J.-P.; Otting, G.; Luo, D.; Huber, T. Biocompatible Macrocyclization between Cysteine and 2-Cyanopyridine Generates Stable Peptide Inhibitors. *Org. Lett.* **2019**, *21* (12), 4709–4712. https://doi.org/10.1021/acs.orglett.9b01545.

(59) Behnam, M. A. M.; Graf, D.; Bartenschlager, R.; Zlotos, D. P.; Klein, C. D. 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.

(60) Lim, H. A.; Joy, J.; Hill, J.; San Brian Chia, C. Novel Agmatine and Agmatine-like Peptidomimetic Inhibitors of the West Nile Virus NS2B/NS3 Serine Protease. *Eur. J. Med. Chem.* **2011**, *46* (7), 3130–3134. https://doi.org/10.1016/j.ejmech.2011.04.055.

(61) Stoermer, M. J.; Chappell, K. J.; Liebscher, S.; Jensen, C. M.; Gan, C. H.; Gupta, P. K.; Xu, W.-J.; Young, P. R.; Fairlie, D. P. Potent Cationic Inhibitors of West Nile Virus NS2B/NS3 Protease With Serum Stability, Cell Permeability and Antiviral Activity. *J. Med. Chem.* **2008**, *51* (18), 5714–5721. https://doi.org/10.1021/jm800503y.

(62) Chu, J. J. H.; Lee, R. C. H.; Ang, M. J. Y.; Wang, W.-L.; Lim, H. A.; Wee, J. L. K.; Joy, J.; Hill, J.; Brian Chia, C. S. Antiviral Activities of 15 Dengue NS2B-NS3 Protease Inhibitors Using a Human Cell-Based Viral Quantification Assay. *Antiviral Res.* **2015**, *118*, 68–74. https://doi.org/10.1016/j.antiviral.2015.03.010.

(63) Su, X.-C.; Ozawa, K.; Qi, R.; Vasudevan, S. G.; Lim, S. P.; Otting, G. NMR Analysis of the Dynamic Exchange of the NS2B Cofactor between Open and Closed Conformations of the West Nile Virus NS2B-NS3 Protease. *PLoS Negl. Trop. Dis.* **2009**, *3* (12), e561. https://doi.org/10.1371/journal.pntd.0000561.

(64) Behnam, M. A. M.; Klein, C. D. P. Conformational Selection in the Flaviviral NS2B-NS3 Protease. *Biochimie* **2020**, *174*, 117–125. https://doi.org/10.1016/j.biochi.2020.04.014.

(65) Li, Z.; Brecher, M.; Deng, Y.-Q.; Zhang, J.; Sakamuru, S.; Liu, B.; Huang, R.; Koetzner, C. A.; Allen, C. A.; Jones, S. A.; Chen, H.; Zhang, N.-N.; Tian, M.; Gao, F.; Lin, Q.; Banavali, N.; Zhou, J.; Boles, N.; Xia, M.; Kramer, L. D.; Qin, C.-F.; Li, H. Existing Drugs as Broad-Spectrum and Potent Inhibitors for Zika Virus by Targeting NS2B-NS3 Interaction. *Cell Res.* **2017**, *27* (8), 1046–1064. https://doi.org/10.1038/cr.2017.88.

(66) Hooper, J. D.; Clements, J. A.; Quigley, J. P.; Antalis, T. M. Type II Transmembrane Serine Proteases. Insights into an Emerging Class of Cell Surface Proteolytic Enzymes. *J. Biol. Chem.* **2001**, *276* (2), 857–860. https://doi.org/10.1074/jbc.R000020200.

(67) Szabo, R.; Bugge, T. H. Membrane-Anchored Serine Proteases as Regulators of Epithelial Function. *Biochem. Soc. Trans.* **2020**, *48* (2), 517–528. https://doi.org/10.1042/BST20190675.

(68) Netzel-Arnett, S.; Hooper, J. D.; Szabo, R.; Madison, E. L.; Quigley, J. P.; Bugge, T. H.; Antalis, T. M. Membrane Anchored Serine Proteases: A Rapidly Expanding Group of Cell Surface Proteolytic Enzymes with Potential Roles in Cancer. *Cancer Metastasis Rev.* **2003**, *22* (2–3), 237–258. https://doi.org/10.1023/A:1023003616848.

(69) Martin, C. E.; List, K. Cell Surface–Anchored Serine Proteases in Cancer Progression and Metastasis. *Cancer Metastasis Rev.* **2019**, *38* (3), 357–387. https://doi.org/10.1007/s10555-019-09811-7.

(70) Lin, C.-Y.; Anders, J.; Johnson, M.; Sang, Q. A.; Dickson, R. B. Molecular Cloning of CDNA for Matriptase, a Matrix-Degrading Serine Protease with Trypsin-like Activity. *J. Biol. Chem.* **1999**, *274* (26), 18231–18236. https://doi.org/10.1074/jbc.274.26.18231.

(71) Ding, K.-F. Effect of SNC19/ST14 Gene Overexpression on Invasion of Colorectal Cancer Cells. *World J. Gastroenterol.* **2005**, *11* (36), 5651. https://doi.org/10.3748/wjg.v11.i36.5651.

(72) List, K. Matriptase: A Culprit in Cancer? *Futur. Oncol.* **2009**, *5* (1), 97–104. https://doi.org/10.2217/14796694.5.1.97.

(73) Lee, S.-L.; Dickson, R. B.; Lin, C.-Y. Activation of Hepatocyte Growth Factor and Urokinase/Plasminogen Activator by Matriptase, an Epithelial Membrane Serine Protease. *J. Biol. Chem.* **2000**, *275* (47), 36720–36725. https://doi.org/10.1074/jbc.M007802200.

(74) Steinmetzer, T.; Schweinitz, A.; Stürzebecher, A.; Dönnecke, D.; Uhland, K.; Schuster, O.; Steinmetzer, P.; Müller, F.; Friedrich, R.; Than, M. E.; Bode, W.; Stürzebecher, J. Secondary Amides of Sulfonylated 3-Amidinophenylalanine. New Potent and Selective Inhibitors of Matriptase. *J. Med. Chem.* **2006**, *49* (14), 4116–4126. https://doi.org/10.1021/jm051272l.

(75) Hammami, M.; Rühmann, E.; Maurer, E.; Heine, A.; Gütschow, M.; Klebe, G.; Steinmetzer, T. New 3-Amidinophenylalanine-Derived Inhibitors of Matriptase. *Medchemcomm* **2012**, *3* (7), 807. https://doi.org/10.1039/c2md20074k.

(76) List, K.; Haudenschild, C. C.; Szabo, R.; Chen, W.; Wahl, S. M.; Swaim, W.; Engelholm, L. H.; Behrendt, N.; Bugge, T. H. Matriptase/MT-SP1 Is Required for Postnatal Survival, Epidermal Barrier Function, Hair Follicle Development, and Thymic Homeostasis. *Oncogene* **2002**, *21* (23), 3765–3779. https://doi.org/10.1038/sj.onc.1205502.

(77) Wilkinson, D. J.; Habgood, A.; Lamb, H. K.; Thompson, P.; Hawkins, A. R.; Désilets, A.; Leduc, R.; Steinmetzer, T.; Hammami, M.; Lee, M. S.; Craik, C. S.; Watson, S.; Lin, H.; Milner, J. M.; Rowan, A. D. Matriptase Induction of Metalloproteinase-Dependent Aggrecanolysis In Vitro and In Vivo: Promotion of Osteoarthritic Cartilage Damage by Multiple Mechanisms. *Arthritis Rheumatol.* **2017**, *69* (8), 1601–1611. https://doi.org/10.1002/art.40133.

(78) Fuentes-Prior, P. Priming of SARS-CoV-2 S Protein by Several Membrane-Bound Serine Proteinases Could Explain Enhanced Viral Infectivity and Systemic COVID-19 Infection. *J. Biol. Chem.* **2021**, *296*, 100135. https://doi.org/10.1074/jbc.REV120.015980.

(79) Bost, P.; Giladi, A.; Liu, Y.; Bendjelal, Y.; Xu, G.; David, E.; Blecher-Gonen, R.; Cohen, M.; Medaglia, C.; Li, H.; Deczkowska, A.; Zhang, S.; Schwikowski, B.; Zhang, Z.; Amit, I. Host-Viral Infection Maps Reveal Signatures of Severe COVID-19 Patients. *Cell* **2020**, *181* (7), 1475-1488.e12. https://doi.org/10.1016/j.cell.2020.05.006.

(80) List, K.; Kosa, P.; Szabo, R.; Bey, A. L.; Wang, C. B.; Molinolo, A.; Bugge, T. H. Epithelial Integrity Is Maintained by a Matriptase-Dependent Proteolytic Pathway. *Am. J. Pathol.* **2009**, *175* (4), 1453–1463. https://doi.org/10.2353/ajpath.2009.090240.

(81) Kilpatrick, L. M.; Harris, R. L.; Owen, K. A.; Bass, R.; Ghorayeb, C.; Bar-Or, A.; Ellis, V. Initiation of Plasminogen Activation on the Surface of Monocytes Expressing the Type II Transmembrane Serine Protease Matriptase. *Blood* **2006**, *108* (8), 2616–2623. https://doi.org/10.1182/blood-2006-02-001073.

(82) Le Gall, S. M.; Szabo, R.; Lee, M.; Kirchhofer, D.; Craik, C. S.; Bugge, T. H.; Camerer, E. Matriptase Activation Connects Tissue Factor–Dependent Coagulation Initiation to Epithelial Proteolysis and Signaling. *Blood* **2016**, *127* (25), 3260–3269. https://doi.org/10.1182/blood-2015-11-683110.

(83) Klok, F. A.; Kruip, M. J. H. A.; van der Meer, N. J. M.; Arbous, M. S.; Gommers, D.; Kant, K. M.; Kaptein, F. H. J.; van Paassen, J.; Stals, M. A. M.; Huisman, M. V.; Endeman, H. Confirmation of the High Cumulative Incidence of Thrombotic Complications in Critically Ill ICU Patients with COVID-19: An

Updated Analysis. *Thromb. Res.* **2020**, *191* (April), 148–150. https://doi.org/10.1016/j.thromres.2020.04.041.

(84) Kojima, K.; Inouye, K. Activation of Matriptase Zymogen. *J. Biochem.* **2011**, *150* (2), 123–125. https://doi.org/10.1093/jb/mvr075.

(85) Takeuchi, T.; Harris, J. L.; Huang, W.; Yan, K. W.; Coughlin, S. R.; Craik, C. S. Cellular Localization of Membrane-Type Serine Protease 1 and Identification of Protease-Activated Receptor-2 and Single-Chain Urokinase-Type Plasminogen Activator as Substrates. *J. Biol. Chem.* **2000**, *275* (34), 26333–26342. https://doi.org/10.1074/jbc.M002941200.

(86) Damalanka, V. C.; Han, Z.; Karmakar, P.; O'Donoghue, A. J.; La Greca, F.; Kim, T.; Pant, S. M.; Helander, J.; Klefström, J.; Craik, C. S.; Janetka, J. W. Discovery of Selective Matriptase and Hepsin Serine Protease Inhibitors: Useful Chemical Tools for Cancer Cell Biology. *J. Med. Chem.* **2019**, *62* (2), 480–490. https://doi.org/10.1021/acs.jmedchem.8b01536.

(87) Yuan, C.; Chen, L.; Meehan, E. J.; Daly, N.; Craik, D. J.; Huang, M.; Ngo, J. C. Structure of Catalytic Domain of Matriptase in Complex with Sunflower Trypsin Inhibitor-1. *BMC Struct. Biol.* **2011**, *11* (1), 30. https://doi.org/10.1186/1472-6807-11-30.

(88) Sisay, M. T.; Steinmetzer, T.; Stirnberg, M.; Maurer, E.; Hammami, M.; Bajorath, J.; Gütschow, M. Identification of the First Low-Molecular-Weight Inhibitors of Matriptase-2. *J. Med. Chem.* **2010**, *53* (15), 5523–5535. https://doi.org/10.1021/jm100183e.

(89) Friedrich, R.; Fuentes-Prior, P.; Ong, E.; Coombs, G.; Hunter, M.; Oehler, R.; Pierson, D.; Gonzalez, R.; Huber, R.; Bode, W.; Madison, E. L. Catalytic Domain Structures of MT-SP1/Matriptase, a Matrix-Degrading Transmembrane Serine Proteinase. *J. Biol. Chem.* **2002**, *277* (3), 2160–2168. https://doi.org/10.1074/jbc.M109830200.

(90) Béliveau, F.; Désilets, A.; Leduc, R. Probing the Substrate Specificities of Matriptase, Matriptase-2, Hepsin and DESC1 with Internally Quenched Fluorescent Peptides. *FEBS J.* **2009**, *276* (8), 2213–2226. https://doi.org/10.1111/j.1742-4658.2009.06950.x.

(91) Goswami, R.; Mukherjee, S.; Ghadiyaram, C.; Wohlfahrt, G.; Sistla, R. K.; Nagaraj, J.; Satyam, L. K.; Subbarao, K.; Palakurthy, R. K.; Gopinath, S.; Krishnamurthy, N. R.; Ikonen, T.; Moilanen, A.; Subramanya, H. S.; Kallio, P.; Ramachandra, M. Structure-Guided Discovery of 1,3,5 Tri-Substituted Benzenes as Potent and Selective Matriptase Inhibitors Exhibiting in Vivo Antitumor Efficacy. *Bioorg. Med. Chem.* **2014**, *22* (12), 3187–3203. https://doi.org/10.1016/j.bmc.2014.04.013.

(92) Dramsi, S.; Trieu-Cuot, P.; Bierne, H. Sorting Sortases: A Nomenclature Proposal for the Various Sortases of Gram-Positive Bacteria. *Res. Microbiol.* **2005**, *156* (3), 289–297. https://doi.org/10.1016/j.resmic.2004.10.011.

(93) van 't Hof, W.; Maňásková, S. H.; Veerman, E. C. I.; Bolscher, J. G. M. Sortase-Mediated Backbone Cyclization of Proteins and Peptides. *Biol. Chem.* **2015**, *396* (4), 283–293. https://doi.org/10.1515/hsz-2014-0260.

(94) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. Staphylococcus Aureus Sortase, an Enzyme That Anchors Surface Proteins to the Cell Wall. *Science (80-.).* **1999**, *285* (5428), 760–763. https://doi.org/10.1126/science.285.5428.760.

(95) Comfort, D.; Clubb, R. T. A Comparative Genome Analysis Identifies Distinct Sorting Pathways in Gram-Positive Bacteria. *Infect. Immun.* **2004**, *72* (5), 2710–2722. https://doi.org/10.1128/IAI.72.5.2710-2722.2004.

(96) Navarre, W. W.; Schneewind, O. Surface Proteins of Gram-Positive Bacteria and Mechanisms of Their Targeting to the Cell Wall Envelope. *Microbiol. Mol. Biol. Rev.* **1999**, *63* (1), 174–229. https://doi.org/10.1128/MMBR.63.1.174-229.1999.

(97) Finn, R. D.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Mistry, J.; Mitchell, A. L.; Potter, S. C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; Salazar, G. A.; Tate, J.; Bateman, A. The Pfam Protein Families Database: Towards a More Sustainable Future. *Nucleic Acids Res.* **2016**, *44* (D1), D279–D285. https://doi.org/10.1093/nar/gkv1344.

(98) Barthels, F. Tools for the Design and Characterization of Covalent Cysteine Protease Inhibitors, Johannes Gutenberg-University, 2022.

(99) Kuk, A. C. Y.; Hao, A.; Guan, Z.; Lee, S.-Y. Visualizing Conformation Transitions of the Lipid II Flippase

MurJ. *Nat. Commun.* **2019**, *10* (1), 1736. https://doi.org/10.1038/s41467-019-09658-0.

(100) Deivanayagam, C. C. S. A Novel Variant of the Immunoglobulin Fold in Surface Adhesins of Staphylococcus Aureus: Crystal Structure of the Fibrinogen-Binding MSCRAMM, Clumping Factor A. *EMBO J.* **2002**, *21* (24), 6660–6672. https://doi.org/10.1093/emboj/cdf619.

(101) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. Structure of Sortase, the Transpeptidase That Anchors Proteins to the Cell Wall of Staphylococcus Aureus. *Proc. Natl. Acad. Sci.* **2001**, *98* (11), 6056–6061. https://doi.org/10.1073/pnas.101064198.

(102) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. The Structure of the Staphylococcus Aureus Sortase-Substrate Complex Reveals How the Universally Conserved LPXTG Sorting Signal Is Recognized. *J. Biol. Chem.* **2009**, *284* (36), 24465–24477. https://doi.org/10.1074/jbc.M109.022624.

(103) Tarini, M.; Cignoni, P.; Montani, C. Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12* (5), 1237–1244. https://doi.org/10.1109/TVCG.2006.115.

(104) Bentley, M. L.; Lamb, E. C.; McCafferty, D. G. Mutagenesis Studies of Substrate Recognition and Catalysis in the Sortase A Transpeptidase from Staphylococcus Aureus. *J. Biol. Chem.* **2008**, *283* (21), 14762–14771. https://doi.org/10.1074/jbc.M800974200.

(105) Marraffini, L. A.; Ton-That, H.; Zong, Y.; Narayana, S. V. L.; Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of Staphylococcus Aureus. *J. Biol. Chem.* **2004**, *279* (36), 37763–37770. https://doi.org/10.1074/jbc.M405282200.

(106) Perry, A. M.; Ton-That, H.; Mazmanian, S. K.; Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of Staphylococcus Aureus. *J. Biol. Chem.* **2002**, *277* (18), 16241–16248. https://doi.org/10.1074/jbc.M109194200.

(107) Ruzin, A.; Severin, A.; Ritacco, F.; Tabei, K.; Singh, G.; Bradford, P. A.; Siegel, M. M.; Projan, S. J.; Shlaes, D. M. Further Evidence That a Cell Wall Precursor [C 55 -MurNAc-(Peptide)-GlcNAc] Serves as an Acceptor in a Sorting Reaction. *J. Bacteriol.* **2002**, *184* (8), 2141–2147. https://doi.org/10.1128/JB.184.8.2141-2147.2002.

(108) Schneewind, O.; Model, P.; Fischetti, V. A. Sorting of Protein a to the Staphylococcal Cell Wall. *Cell* **1992**, *70* (2), 267–281. https://doi.org/10.1016/0092-8674(92)90101-H.

(109) Jacobitz, A. W.; Kattke, M. D.; Wereszczynski, J.; Clubb, R. T. Sortase Transpeptidases: Structural Biology and Catalytic Mechanism. In *Physiology & behavior*; 2017; Vol. 176, pp 223–264. https://doi.org/10.1016/bs.apcsb.2017.04.008.

(110) Schneewind, O.; Missiakas, D. Sortases, Surface Proteins, and Their Roles in Staphylococcus Aureus Disease and Vaccine Development. In *Protein Secretion in Bacteria*; ASM Press: Washington, DC, USA, 2019; Vol. 7, pp 173–188. https://doi.org/10.1128/9781683670285.ch15.

(111) Montanaro, L.; Poggi, A.; Visai, L.; Ravaioli, S.; Campoccia, D.; Speziale, P.; Arciola, C. R. Extracellular DNA in Biofilms. *Int. J. Artif. Organs* **2011**, *34* (9), 824–831. https://doi.org/10.5301/ijao.5000051.

(112) Mandakhalikar, K. D. Medical Biofilms. In *ACS Symposium Series*; 2019; Vol. 1323, pp 83–99. https://doi.org/10.1021/bk-2019-1323.ch004.

(113) Koo, H.; Allan, R. N.; Howlin, R. P.; Stoodley, P.; Hall-Stoodley, L. Targeting Microbial Biofilms: Current and Prospective Therapeutic Strategies. *Nat. Rev. Microbiol.* **2017**, *15* (12), 740–755. https://doi.org/10.1038/nrmicro.2017.99.

(114) Lister, J. L.; Horswill, A. R. Staphylococcus Aureus Biofilms: Recent Developments in Biofilm Dispersal. *Front. Cell. Infect. Microbiol.* **2014**, *4* (DEC), 1–9. https://doi.org/10.3389/fcimb.2014.00178.

(115) Rumbaugh, K. P. Convergence of Hormones and Autoinducers at the Host/Pathogen Interface. *Anal. Bioanal. Chem.* **2007**, *387* (2), 425–435. https://doi.org/10.1007/s00216-006-0694-9.

(116) Hughes, D. T.; Sperandio, V. Inter-Kingdom Signalling: Communication between Bacteria and Their Hosts. *Nat. Rev. Microbiol.* **2008**, *6* (2), 111–120. https://doi.org/10.1038/nrmicro1836.

(117) Novick, R. P.; Geisinger, E. Quorum Sensing in Staphylococci. *Annu. Rev. Genet.* **2008**, *42* (1), 541–564. https://doi.org/10.1146/annurev.genet.42.110807.091640.

(118) Zong, Y.; Xu, Y.; Liang, X.; Keene, D. R.; Höök, A.; Gurusiddappa, S.; Höök, M.; Narayana, S. V. L. A 'Collagen Hug' Model for Staphylococcus Aureus CNA Binding to Collagen. *EMBO J.* **2005**, *24* (24), 4224–4236. https://doi.org/10.1038/sj.emboj.7600888.

(119) Foster, T. J. The Remarkably Multifunctional Fibronectin Binding Proteins of Staphylococcus Aureus. *Eur. J. Clin. Microbiol. Infect. Dis.* **2016**, *35* (12), 1923–1931. https://doi.org/10.1007/s10096-016-2763-0.

(120) Tsompanidou, E.; Denham, E. L.; Sibbald, M. J. J. B.; Yang, X.; Seinen, J.; Friedrich, A. W.; Buist, G.; van Dijl, J. M. The Sortase A Substrates FnbpA, FnbpB, ClfA and ClfB Antagonize Colony Spreading of Staphylococcus Aureus. *PLoS One* **2012**, *7* (9), e44646. https://doi.org/10.1371/journal.pone.0044646.

(121) O´Brien, L.; Kerrigan, S. W.; Kaw, G.; Hogan, M.; Penadés, J.; Litt, D.; Fitzgerald, D. J.; Foster, T. J.; Cox, D. Multiple Mechanisms for the Activation of Human Platelet Aggregation by Staphylococcus Aureus: Roles for the Clumping Factors ClfA and ClfB, the Serine-Aspartate Repeat Protein SdrE and Protein A. *Mol. Microbiol.* **2002**, *44* (4), 1033–1044. https://doi.org/10.1046/j.1365-2958.2002.02935.x.

(122) Wardenburg, J. B.; Patel, R. J.; Schneewind, O. Surface Proteins and Exotoxins Are Required for the Pathogenesis of Staphylococcus Aureus Pneumonia. *Infect. Immun.* **2007**, *75* (2), 1040–1044. https://doi.org/10.1128/IAI.01313-06.

(123) Pishchany, G.; Sheldon, J. R.; Dickson, C. F.; Alam, M. T.; Read, T. D.; Gell, D. A.; Heinrichs, D. E.; Skaar, E. P. IsdB-Dependent Hemoglobin Binding Is Required for Acquisition of Heme by Staphylococcus Aureus. *J. Infect. Dis.* **2014**, *209* (11), 1764–1772. https://doi.org/10.1093/infdis/jit817.

(124) Muryoi, N.; Tiedemann, M. T.; Pluym, M.; Cheung, J.; Heinrichs, D. E.; Stillman, M. J. Demonstration of the Iron-Regulated Surface Determinant (Isd) Heme Transfer Pathway in Staphylococcus Aureus. *J. Biol. Chem.* **2008**, *283* (42), 28125–28136. https://doi.org/10.1074/jbc.M802171200.

(125) Falugi, F.; Kim, H. K.; Missiakas, D. M.; Schneewind, O. Role of Protein A in the Evasion of Host Adaptive Immune Responses by Staphylococcus Aureus. *MBio* **2013**, *4* (5), e00575-13. https://doi.org/10.1128/mBio.00575-13.

(126) Flemming, H.-C.; Wingender, J.; Szewzyk, U.; Steinberg, P.; Rice, S. A.; Kjelleberg, S. Biofilms: An Emergent Form of Bacterial Life. *Nat. Rev. Microbiol.* **2016**, *14* (9), 563–575. https://doi.org/10.1038/nrmicro.2016.94.

(127) Harrison, J. J.; Turner, R. J.; Ceri, H. Persister Cells, the Biofilm Matrix and Tolerance to Metal Cations in Biofilm and Planktonic Pseudomonas Aeruginosa. *Environ. Microbiol.* **2005**, *7* (7), 981–994. https://doi.org/10.1111/j.1462-2920.2005.00777.x.

(128) Mazmanian, S. K.; Liu, G.; Jensen, E. R.; Lenoy, E.; Schneewind, O. Staphylococcus Aureus Sortase Mutants Defective in the Display of Surface Proteins and in the Pathogenesis of Animal Infections. *Proc. Natl. Acad. Sci.* **2000**, *97* (10), 5510–5515. https://doi.org/10.1073/pnas.080520697.

(129) Mu, D.; Xiang, H.; Dong, H.; Wang, D.; Wang, T. Isovitexin, a Potential Candidate Inhibitor of Sortase A of Staphylococcus Aureus USA300. *J. Microbiol. Biotechnol.* **2018**, *28* (9), 1426–1432. https://doi.org/10.4014/jmb.1802.02014.

(130) Cascioferro, S.; Raffa, D.; Maggio, B.; Raimondi, M. V.; Schillaci, D.; Daidone, G. Sortase A Inhibitors: Recent Advances and Future Perspectives. *J. Med. Chem.* **2015**, *58* (23), 9108–9123. https://doi.org/10.1021/acs.jmedchem.5b00779.

(131) Rasko, D. A.; Sperandio, V. Anti-Virulence Strategies to Combat Bacteria-Mediated Disease. *Nat. Rev. Drug Discov.* **2010**, *9* (2), 117–128. https://doi.org/10.1038/nrd3013.

(132) Cascioferro, S.; Cusimano, M. G.; Schillaci, D. Antiadhesion Agents against Gram-Positive Pathogens. *Future Microbiol.* **2014**, *9* (10), 1209–1220. https://doi.org/10.2217/fmb.14.56.

(133) Cascioferro, S.; Totsika, M.; Schillaci, D. Sortase A: An Ideal Target for Anti-Virulence Drug Development. *Microb. Pathog.* **2014**, *77*, 105–112. https://doi.org/10.1016/j.micpath.2014.10.007.

(134) Frankel, B. A.; Kruger, R. G.; Robinson, D. E.; Kelleher, N. L.; McCafferty, D. G. Staphylococcus Aureus Sortase Transpeptidase SrtA: Insight into the Kinetic Mechanism and Evidence for a Reverse Protonation Catalytic Mechanism. *Biochemistry* **2005**, *44* (33), 11188–11200. https://doi.org/10.1021/bi050141j.

(135) Connolly, K. M.; Smith, B. T.; Pilpa, R.; Ilangovan, U.; Jung, M. E.; Clubb, R. T. Sortase from Staphylococcus Aureus Does Not Contain a Thiolate-Imidazolium Ion Pair in Its Active Site. *J. Biol. Chem.* **2003**, *278* (36), 34061–34065. https://doi.org/10.1074/jbc.M305245200.

(136) Zong, Y.; Bice, T. W.; Ton-That, H.; Schneewind, O.; Narayana, S. V. L. Crystal Structures of Staphylococcus Aureus Sortase A and Its Substrate Complex. *J. Biol. Chem.* **2004**, *279* (30), 31383–31389. https://doi.org/10.1074/jbc.M401374200.

(137) Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M pro by QM/MM Computational Methods. *Chem. Sci.* **2020**, *11* (39), 10626–10630. https://doi.org/10.1039/D0SC02823A.

(138) Sulpizi, M.; Laio, A.; VandeVondele, J.; Cattaneo, A.; Rothlisberger, U.; Carloni, P. Reaction Mechanism of Caspases: Insights from QM/MM Car-Parrinello Simulations. *Proteins Struct. Funct. Genet.* **2003**, *52* (2), 212–224. https://doi.org/10.1002/prot.10275.

(139) Elsässer, B.; Zauner, F. B.; Messner, J.; Soh, W. T.; Dall, E.; Brandstetter, H. Distinct Roles of Catalytic Cysteine and Histidine in the Protease and Ligase Mechanisms of Human Legumain As Revealed by DFT-Based QM/MM Simulations. *ACS Catal.* **2017**, *7* (9), 5585–5593. https://doi.org/10.1021/acscatal.7b01505.

(140) Mock, W. L. Theory of Enzymatic Reverse-Protonation Catalysis. *Bioorg. Chem.* **1992**, *20* (4), 377–381. https://doi.org/10.1016/0045-2068(92)90047-7.

(141) Ke, Z.; Zhou, Y.; Hu, P.; Wang, S.; Xie, D.; Zhang, Y. Active Site Cysteine Is Protonated in the PAD4 Michaelis Complex: Evidence from Born−Oppenheimer Ab Initio QM/MM Molecular Dynamics Simulations. *J. Phys. Chem. B* **2009**, *113* (38), 12750–12758. https://doi.org/10.1021/jp903173c.

(142) Sims, P. A.; Larsen, T. M.; Poyner, R. R.; Cleland, W. W.; Reed, G. H. Reverse Protonation Is the Key to General Acid−Base Catalysis in Enolase „. *Biochemistry* **2003**, *42* (27), 8298–8306. https://doi.org/10.1021/bi0346345.

(143) Huang, X.; Aulabaugh, A.; Ding, W.; Kapoor, B.; Alksne, L.; Tabei, K.; Ellestad, G. Kinetic Mechanism of Staphylococcus Aureus Sortase SrtA. *Biochemistry* **2003**, *42* (38), 11307–11315. https://doi.org/10.1021/bi034391g.

(144) Clancy, K. W.; Melvin, J. A.; McCafferty, D. G. Sortase Transpeptidases: Insights into Mechanism, Substrate Specificity, and Inhibition. *Biopolymers* **2010**, *94* (4), 385–396. https://doi.org/10.1002/bip.21472.

(145) Tian, B.-X.; Eriksson, L. A. Catalytic Mechanism and Roles of Arg197 and Thr183 in the Staphylococcus Aureus Sortase A Enzyme. *J. Phys. Chem. B* **2011**, *115* (44), 13003–13011. https://doi.org/10.1021/jp2058113.

(146) Shrestha, P.; Wereszczynski, J. Discerning the Catalytic Mechanism of Staphylococcus Aureus Sortase A with QM/MM Free Energy Calculations. *J. Mol. Graph. Model.* **2016**, *67* (1), 33–43. https://doi.org/10.1016/j.jmgm.2016.04.006.

(147) Ton-That, H.; Liu, G.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. Purification and Characterization of Sortase, the Transpeptidase That Cleaves Surface Proteins of Staphylococcus Aureus at the LPXTG Motif. *Proc. Natl. Acad. Sci.* **1999**, *96* (22), 12424–12429. https://doi.org/10.1073/pnas.96.22.12424.

(148) Jaudzems, K.; Kurbatska, V.; Jēkabsons, A.; Bobrovs, R.; Rudevica, Z.; Leonchiks, A. Targeting Bacterial Sortase A with Covalent Inhibitors: 27 New Starting Points for Structure-Based Hit-to-Lead Optimization. *ACS Infect. Dis.* **2020**, *6* (2), 186–194. https://doi.org/10.1021/acsinfecdis.9b00265.

(149) Ugur, I.; Schatte, M.; Marion, A.; Glaser, M.; Boenitz-Dulat, M.; Antes, I. Ca2+ Binding Induced Sequential Allosteric Activation of Sortase A: An Example for Ion-Triggered Conformational Selection. *PLoS One* **2018**, *13* (10), e0205057. https://doi.org/10.1371/journal.pone.0205057.

(150) Tonlolo, C.; Benedetti, E. The Polypeptide 310-Helix. *Trends Biochem. Sci.* **1991**, *16* (C), 350–353. https://doi.org/10.1016/0968-0004(91)90142-I.

(151) Naik, M. T.; Suree, N.; Ilangovan, U.; Liew, C. K.; Thieu, W.; Campbell, D. O.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. Staphylococcus Aureus Sortase A Transpeptidase. *J. Biol. Chem.* **2006**, *281* (3), 1817–1826. https://doi.org/10.1074/jbc.M506123200.

(152) Zhou, H.-X.; Pang, X.; Lu, C. Rate Constants and Mechanisms of Intrinsically Disordered Proteins Binding to Structured Targets. *Phys. Chem. Chem. Phys.* **2012**, *14* (30), 10466. https://doi.org/10.1039/c2cp41196b.

(153) Wright, P. E.; Dyson, H. J. Intrinsically Disordered Proteins in Cellular Signalling and Regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (1), 18–29. https://doi.org/10.1038/nrm3920.

(154) Ferreon, A. C. M.; Ferreon, J. C.; Wright, P. E.; Deniz, A. A. Modulation of Allostery by Protein Intrinsic Disorder. *Nature* **2013**, *498* (7454), 390–394. https://doi.org/10.1038/nature12294.

(155) Pang, X.; Zhou, H.-X. Disorder-to-Order Transition of an Active-Site Loop Mediates the Allosteric Activation of Sortase A. *Biophys. J.* **2015**, *109* (8), 1706–1715. https://doi.org/10.1016/j.bpj.2015.08.039.

(156) Wang, X.; Chen, J.-L.; Otting, G.; Su, X.-C. Conversion of an Amide to a High-Energy Thioester by Staphylococcus Aureus Sortase A Is Powered by Variable Binding Affinity for Calcium. *Sci. Rep.* **2018**, *8* (1), 16371. https://doi.org/10.1038/s41598-018-34752-6.

(157) Zong, Y.; Mazmanian, S. K.; Schneewind, O.; Narayana, S. V. L. The Structure of Sortase B, a Cysteine Transpeptidase That Tethers Surface Protein to the Staphylococcus Aureus Cell Wall. *Structure* **2004**, *12* (1), 105–112. https://doi.org/10.1016/j.str.2003.11.021.

(158) Chitnis, A. C. The University of Edinburgh's Natural History Museum and the Huttonian-Wernerian Debate. *Ann. Sci.* **1970**, *26* (2), 85–94. https://doi.org/10.1080/00033797000203437.

(159) Christensen, J. J.; Johnston, H. D.; Izatt, R. M. An Isothermal Titration Calorimeter. *Rev. Sci. Instrum.* **1968**, *39* (9), 1356–1359. https://doi.org/10.1063/1.1683671.

(160) Kanbour, F.; Joncich, M. J. Solution Calorimeter with Peltier Cooling for Operation at Constant Temperature. *Rev. Sci. Instrum.* **1967**, *38* (7), 913–916. https://doi.org/10.1063/1.1720921.

(161) Johnston, H. D. The Thermodynamics (Log K, ΔH°, ΔS°, ΔCP°) of Metal Ligand Interaction in Aqueous Solution.|nI.|p Design and Construction of an Isothermal Titration Calorimeter.|nII.|pThe Interaction of Cyanide Ion with Bivalent Nickel, Zinc, Cadmium and Mercury.|nIII.|p, 1968.

(162) Freire, E.; Mayorga, O. L.; Straume, M. Isothermal Titration Calorimetry. *Anal. Chem.* **1990**, *62* (18), 950A-959A. https://doi.org/10.1021/ac00217a002.

(163) Falconer, R. J. Applications of Isothermal Titration Calorimetry - the Research and Technical Developments from 2011 to 2015. *J. Mol. Recognit.* **2016**, *29* (10), 504–515. https://doi.org/10.1002/jmr.2550.

(164) Perozzo, R.; Folkers, G.; Scapozza, L. Thermodynamics of Protein–Ligand Interactions: History, Presence, and Future Aspects. *J. Recept. Signal Transduct.* **2004**, *24* (1–2), 1–52. https://doi.org/10.1081/RRS-120037896.

(165) Zhang, Q.; Liu, J.; Nai, X.; Bao, H.; Sun, D.; Liu, M. Examination and Improvement of Undergraduate Laboratory Experiment: Thermodynamics of a Surfactant Micellization. *J. Chem. Educ.* **2020**, *97* (12), 4490–4498. https://doi.org/10.1021/acs.jchemed.0c00886.

(166) Burnouf, D.; Ennifar, E.; Guedich, S.; Puffer, B.; Hoffmann, G.; Bec, G.; Disdier, F.; Baltzinger, M.; Dumas, P. KinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J. Am. Chem. Soc.* **2012**, *134* (1), 559–565. https://doi.org/10.1021/ja209057d.

(167) Ferenczy, G. G.; Keserű, G. M. Thermodynamics Guided Lead Discovery and Optimization. *Drug Discov. Today* **2010**, *15* (21–22), 919–932. https://doi.org/10.1016/j.drudis.2010.08.013.

(168) O'Brien, L. C.; Root, H. B.; Wei, C.-C.; Jensen, D.; Shabestary, N.; De Meo, C.; Eder, D. J. M 2+ •EDTA Binding Affinities: A Modern Experiment in Thermodynamics for the Physical Chemistry Laboratory. *J. Chem. Educ.* **2015**, *92* (9), 1547–1551. https://doi.org/10.1021/acs.jchemed.5b00159.

(169) Wei, C.-C.; Jensen, D.; Boyle, T.; O'Brien, L. C.; De Meo, C.; Shabestary, N.; Eder, D. J. Isothermal Titration Calorimetry and Macromolecular Visualization for the Interaction of Lysozyme and Its Inhibitors. *J. Chem. Educ.* **2015**, *92* (9), 1552–1556. https://doi.org/10.1021/ed5002569.

(170) Gibbs, J. W. A Method of Geometrical Representation of the Thermodynamic Properties of Substances By Means of Surfaces. *Trans. Connect. Acad.* **1873**, *II*, 382–404.

(171) Cooper, A.; Johnson, C. M.; Lakey, J. H.; Nöllmann, M. Heat Does Not Come in Different Colours: Entropy–Enthalpy Compensation, Free Energy Windows, Quantum Confinement, Pressure Perturbation Calorimetry, Solvation and the Multiple Causes of Heat Capacity Effects in Biomolecular Interactions. *Biophys. Chem.* **2001**, *93* (2–3), 215–230. https://doi.org/10.1016/S0301-4622(01)00222-8.

(172) Cameron, D. L.; Jakus, J.; Pauleta, S. R.; Pettigrew, G. W.; Cooper, A. Pressure Perturbation Calorimetry and the Thermodynamics of Noncovalent Interactions in Water: Comparison of Protein−Protein, Protein−Ligand, and Cyclodextrin−Adamantane Complexes. *J. Phys. Chem. B* **2010**, *114* (49), 16228–16235. https://doi.org/10.1021/jp107110t.

(173) Cooper, A. Protein Heat Capacity: An Anomaly That Maybe Never Was. *J. Phys. Chem. Lett.* **2010**, *1* (22), 3298–3304. https://doi.org/10.1021/jz1012142.

(174) Cooper, A. Heat Capacity Effects in Protein Folding and Ligand Binding: A Re-Evaluation of the Role of Water in Biomolecular Thermodynamics. *Biophys. Chem.* **2005**, *115* (2–3), 89–97. https://doi.org/10.1016/j.bpc.2004.12.011.

(175) Ward, W. H. J.; Holdgate, G. A. 7 Isothermal Titration Calorimetry in Drug Discovery. In *Progress in Medicinal Chemistry*; 2001; Vol. 38, pp 309–376. https://doi.org/10.1016/S0079-6468(08)70097-3.

(176) Deng, H.; Jia, Y.; Zhang, Y. Protein Structure Prediction. *Int. J. Mod. Phys. B* **2018**, *32* (18), 1840009. https://doi.org/10.1142/S021797921840009X.

(177) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(178) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.

(179) Ferenczy, G. G.; Keserű, G. M. Enthalpic Efficiency of Ligand Binding. *J. Chem. Inf. Model.* **2010**, *50* (9), 1536–1541. https://doi.org/10.1021/ci100125a.

(180) Leeson, P. D.; Springthorpe, B. The Influence of Drug-like Concepts on Decision-Making in Medicinal Chemistry. *Nat. Rev. Drug Discov.* **2007**, *6* (11), 881–890. https://doi.org/10.1038/nrd2445.

(181) Kenny, P. W. The Nature of Ligand Efficiency. *J. Cheminform.* **2019**, *11* (1), 8. https://doi.org/10.1186/s13321-019-0330-2.

(182) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *23* (1–3), 3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.

(183) Johnson, T. W.; Gallego, R. A.; Edwards, M. P. Lipophilic Efficiency as an Important Metric in Drug Design. *J. Med. Chem.* **2018**, *61* (15), 6401–6420. https://doi.org/10.1021/acs.jmedchem.8b00077.

(184) Yang, H.; Chen, X.; Ji, X.; Xiong, Y.; Yang, K. *RCSB PDB - 5GXJ: Zika Virus NS2B-NS3 protease*. https://doi.org/10.2210/pdb5GXJ/pdb.

(185) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. *Science (80-.).* **2016**, *353* (6298), 503–505. https://doi.org/10.1126/science.aag2419.

(186) Lee, H.; Ren, J.; Nocadello, S.; Rice, A. J.; Ojeda, I.; Light, S.; Minasov, G.; Vargas, J.; Nagarathnam, D.; Anderson, W. F.; Johnson, M. E. Identification of Novel Small Molecule Inhibitors against NS2B/NS3 Serine Protease from Zika Virus. *Antiviral Res.* **2017**, *139* (March 2017), 49–58. https://doi.org/10.1016/j.antiviral.2016.12.016.

(187) Aleshin, A. E.; Bankston, L.; Liddington, R. C. *RCSB PDB - 5TFN: CRYSTAL STRUCTURE OF THE ZIKA*

VIRUS NS2B-NS3 PROTEASE in super-open conformation. A novel conformation for the Zika virus NS2B-NS3 protease offers new insights into biological regulation and inhibitor design. https://doi.org/10.2210/pdb5TFN/pdb.

(188) Yao, Y.; Huo, T.; Lin, Y.-L. L.; Nie, S.; Wu, F.; Hua, Y.; Wu, J.; Kneubehl, A. R.; Vogt, M. B.; Rico-Hesse, R.; Song, Y. Discovery, X-Ray Crystallography and Antiviral Activity of Allosteric Inhibitors of Flavivirus NS2B-NS3 Protease. *J. Am. Chem. Soc.* **2019**, *141* (17), jacs.9b02505. https://doi.org/10.1021/jacs.9b02505.

(189) Grinter, R. *[ccp4bb] Questionable Ligand Density*. https://www.mail-archive.com/ccp4bb@jiscmail.ac.uk/msg47072.html (accessed 2023-03-13).

(190) Li, Q.; Kang, C. Insights into Structures and Dynamics of Flavivirus Proteases from NMR Studies. *Int. J. Mol. Sci.* **2020**, *21* (7), 2527. https://doi.org/10.3390/ijms21072527.

(191) Li, Y.; Zhang, Z.; Phoo, W. W.; Loh, Y. R.; Wang, W.; Liu, S.; Chen, M. W.; Hung, A. W.; Keller, T. H.; Luo, D.; Kang, C. B. Structural Dynamics of Zika Virus NS2B-NS3 Protease Binding to Dipeptide Inhibitors. *Structure* **2017**, *25* (8), 1242-1250.e3. https://doi.org/10.1016/j.str.2017.06.006.

(192) Li, Y.; Zhang, Z.; Phoo, W. W.; Loh, Y. R.; Li, R.; Yang, H. Y.; Jansson, A. E.; Hill, J.; Keller, T. H.; Nacro, K.; Luo, D.; Kang, C. Structural Insights into the Inhibition of Zika Virus NS2B-NS3 Protease by a Small-Molecule Inhibitor. *Structure* **2018**, *26* (4), 555-564.e3. https://doi.org/10.1016/j.str.2018.02.005.

(193) Mahawaththa, M. C.; Pearce, B. J. G.; Szabo, M.; Graham, B.; Klein, C. D.; Nitsche, C.; Otting, G. Solution Conformations of a Linked Construct of the Zika Virus NS2B-NS3 Protease. *Antiviral Res.* **2017**, *142*, 141–147. https://doi.org/10.1016/j.antiviral.2017.03.011.

(194) Li, Y.; Phoo, W. W.; Loh, Y. R.; Zhang, Z.; Ng, E. Y.; Wang, W.; Keller, T. H.; Luo, D.; Kang, C. Structural Characterization of the Linked <scp>NS</Scp> 2B- <scp>NS</Scp> 3 Protease of Zika Virus. *FEBS Lett.* **2017**, *591* (15), 2338–2347. https://doi.org/10.1002/1873-3468.12741.

(195) Wu, H.; Bock, S.; Snitko, M.; Berger, T.; Weidner, T.; Holloway, S.; Kanitz, M.; Diederich, W. E.; Steuber, H.; Walter, C.; Hofmann, D.; Weißbrich, B.; Spannaus, R.; Acosta, E. G.; Bartenschlager, R.; Engels, B.; Schirmeister, T.; Bodem, J. Novel Dengue Virus NS2B/NS3 Protease Inhibitors. *Antimicrob. Agents Chemother.* **2015**, *59* (2), 1100–1109. https://doi.org/10.1128/AAC.03543-14.

(196) Maus, H.; Barthels, F.; Hammerschmidt, S. J.; Kopp, K.; Millies, B.; Gellert, A.; Ruggieri, A.; Schirmeister, T. SAR of Novel Benzothiazoles Targeting an Allosteric Pocket of DENV and ZIKV NS2B/NS3 Proteases. *Bioorganic Med. Chem.* **2021**, *47* (August), 116392. https://doi.org/10.1016/j.bmc.2021.116392.

(197) Li, Z.; Zhang, J.; Li, H. Flavivirus NS2B/NS3 Protease: Structure, Function, and Inhibition. In *Viral Proteases and Their Inhibitors*; Elsevier, 2017; pp 163–188. https://doi.org/10.1016/B978-0-12-809712-0.00007-1.

(198) Leung, D.; Schroder, K.; White, H.; Fang, N.-X.; Stoermer, M. J.; Abbenante, G.; Martin, J. L.; Young, P. R.; Fairlie, D. P. Activity of Recombinant Dengue 2 Virus NS3 Protease in the Presence of a Truncated NS2B Co-Factor, Small Peptide Substrates, and Inhibitors. *J. Biol. Chem.* **2001**, *276* (49), 45762–45771. https://doi.org/10.1074/jbc.M107360200.

(199) 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. https://doi.org/10.1016/j.antiviral.2020.104900.

(200) Kuiper, B. D.; Slater, K.; Spellmon, N.; Holcomb, J.; Medapureddy, P.; Muzzarelli, K. M.; Yang, Z.; Ovadia, R.; Amblard, F.; Kovari, I. A.; Schinazi, R. F.; Kovari, L. C. Increased Activity of Unlinked Zika Virus NS2B/NS3 Protease Compared to Linked Zika Virus Protease. *Biochem. Biophys. Res. Commun.* **2017**, *492* (4), 668–673. https://doi.org/10.1016/j.bbrc.2017.03.108.

(201) Rentero Rebollo, I.; McCallin, S.; Bertoldo, D.; Entenza, J. M.; Moreillon, P.; Heinis, C. Development of Potent and Selective S. Aureus Sortase A Inhibitors Based on Peptide Macrocycles. *ACS Med. Chem. Lett.* **2016**, *7* (6), 606–611. https://doi.org/10.1021/acsmedchemlett.6b00045.

(202) Srinivasan, R.; Santhakumari, S.; Poonguzhali, P.; Geetha, M.; Dyavaiah, M.; Xiangmin, L. Bacterial Biofilm Inhibition: A Focused Review on Recent Therapeutic Strategies for Combating the Biofilm Mediated Infections. *Front. Microbiol.* **2021**, *12* (May), 1–19. https://doi.org/10.3389/fmicb.2021.676458.

(203) Dickey, S. W.; Cheung, G. Y. C.; Otto, M. Different Drugs for Bad Bugs: Antivirulence Strategies in the Age of Antibiotic Resistance. *Nat. Rev. Drug Discov.* **2017**, *16* (7), 457–471. https://doi.org/10.1038/nrd.2017.23.

(204) Kruger, R. G.; Barkallah, S.; Frankel, B. A.; McCafferty, D. G. Inhibition of the Staphylococcus Aureus Sortase Transpeptidase SrtA by Phosphinic Peptidomimetics. *Bioorg. Med. Chem.* **2004**, *12* (13), 3723–3729. https://doi.org/10.1016/j.bmc.2004.03.066.

(205) SCOTT, C. J.; McDOWELL, A.; MARTIN, S. L.; LYNAS, J. F.; VANDENBROECK, K.; WALKER, B. Irreversible Inhibition of the Bacterial Cysteine Protease-Transpeptidase Sortase (SrtA) by Substrate-Derived Affinity Labels. *Biochem. J.* **2002**, *366* (3), 953–958. https://doi.org/10.1042/bj20020602.

(206) Kruger, R. G.; Dostal, P.; McCafferty, D. G. Development of a High-Performance Liquid Chromatography Assay and Revision of Kinetic Parameters for the Staphylococcus Aureus Sortase Transpeptidase SrtA. *Anal. Biochem.* **2004**, *326* (1), 42–48. https://doi.org/10.1016/j.ab.2003.10.023.

(207) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. Sortagging: A Versatile Method for Protein Labeling. *Nat. Chem. Biol.* **2007**, *3* (11), 707–708. https://doi.org/10.1038/nchembio.2007.31.

(208) Popp, M. W.-L.; Ploegh, H. L. Making and Breaking Peptide Bonds: Protein Engineering Using Sortase. *Angew. Chemie Int. Ed.* **2011**, *50* (22), 5024–5032. https://doi.org/10.1002/anie.201008267.

(209) Wu, Z.; Hong, H.; Zhao, X.; Wang, X. Efficient Expression of Sortase A from Staphylococcus Aureus in Escherichia Coli and Its Enzymatic Characterizations. *Bioresour. Bioprocess.* **2017**, *4* (1), 13. https://doi.org/10.1186/s40643-017-0143-y.

(210) Witte, M. D.; Cragnolini, J. J.; Dougan, S. K.; Yoder, N. C.; Popp, M. W.; Ploegh, H. L. Preparation of Unnatural N-to-N and C-to-C Protein Fusions. *Proc. Natl. Acad. Sci.* **2012**, *109* (30), 11993–11998. https://doi.org/10.1073/pnas.1205427109.

(211) Wu, Z.; Guo, X.; Guo, Z. Sortase A-Catalyzed Peptide Cyclization for the Synthesis of Macrocyclic Peptides and Glycopeptides. *Chem. Commun.* **2011**, *47* (32), 9218. https://doi.org/10.1039/c1cc13322e.

(212) Chan, L.; Cross, H. F.; She, J. K.; Cavalli, G.; Martins, H. F. P.; Neylon, C. Covalent Attachment of Proteins to Solid Supports and Surfaces via Sortase-Mediated Ligation. *PLoS One* **2007**, *2* (11), e1164. https://doi.org/10.1371/journal.pone.0001164.

(213) Guo, X.; Wang, Q.; Swarts, B. M.; Guo, Z. Sortase-Catalyzed Peptide−Glycosylphosphatidylinositol Analogue Ligation. *J. Am. Chem. Soc.* **2009**, *131* (29), 9878–9879. https://doi.org/10.1021/ja903231v.

(214) Beerli, R. R.; Hell, T.; Merkel, A. S.; Grawunder, U. Sortase Enzyme-Mediated Generation of Site-Specifically Conjugated Antibody Drug Conjugates with High In Vitro and In Vivo Potency. *PLoS One* **2015**, *10* (7), e0131177. https://doi.org/10.1371/journal.pone.0131177.

(215) Voloshchuk, N.; Liang, D.; Liang, J. Sortase A Mediated Protein Modifications and Peptide Conjugations. *Curr. Drug Discov. Technol.* **2016**, *12* (4), 205–213. https://doi.org/10.2174/1570163812666150903115601.

(216) Glasgow, J. E.; Salit, M. L.; Cochran, J. R. In Vivo Site-Specific Protein Tagging with Diverse Amines Using an Engineered Sortase Variant. *J. Am. Chem. Soc.* **2016**, *138* (24), 7496–7499. https://doi.org/10.1021/jacs.6b03836.

(217) Antos, J. M.; Popp, M. W.-L.; Ernst, R.; Chew, G.-L.; Spooner, E.; Ploegh, H. L. A Straight Path to Circular Proteins. *J. Biol. Chem.* **2009**, *284* (23), 16028–16036. https://doi.org/10.1074/jbc.M901752200.

(218) Clark, R. J.; Craik, D. J. Invited Reviewnative Chemical Ligation Applied to the Synthesis and Bioengineering of Circular Peptides and Proteins. *Biopolymers* **2010**, *94* (4), 414–422. https://doi.org/10.1002/bip.21372.

(219) Popp, M. W.; Dougan, S. K.; Chuang, T.-Y.; Spooner, E.; Ploegh, H. L. Sortase-Catalyzed Transformations That Improve the Properties of Cytokines. *Proc. Natl. Acad. Sci.* **2011**, *108* (8), 3169–3174. https://doi.org/10.1073/pnas.1016863108.

(220) Wong, C. T. T.; Rowlands, D. K.; Wong, C.-H.; Lo, T. W. C.; Nguyen, G. K. T.; Li, H.-Y.; Tam, J. P. Orally Active Peptidic Bradykinin B1 Receptor Antagonists Engineered from a Cyclotide Scaffold for Inflammatory Pain Treatment. *Angew. Chemie Int. Ed.* **2012**, *51* (23), 5620–5624. https://doi.org/10.1002/anie.201200984.

(221) Tam, J. P.; Wong, C. T. T. Chemical Synthesis of Circular Proteins. *J. Biol. Chem.* **2012**, *287* (32), 27020–27025. https://doi.org/10.1074/jbc.R111.323568.

(222) Hirakawa, H.; Ishikawa, S.; Nagamune, T. Design of Ca 2+ -Independent Staphylococcus Aureus Sortase A Mutants. *Biotechnol. Bioeng.* **2012**, *109* (12), 2955–2961. https://doi.org/10.1002/bit.24585.

(223) Jeong, H.-J.; Abhiraman, G. C.; Story, C. M.; Ingram, J. R.; Dougan, S. K. Generation of Ca2+-Independent Sortase A Mutants with Enhanced Activity for Protein and Cell Surface Labeling. *PLoS One* **2017**, *12* (12), e0189068. https://doi.org/10.1371/journal.pone.0189068.

(224) Hammamy, M. Z.; Haase, C.; Hammami, M.; Hilgenfeld, R.; Steinmetzer, T. Development and Characterization of New Peptidomimetic Inhibitors of the West Nile Virus NS2B-NS3 Protease. *ChemMedChem* **2013**, *8* (2), 231–241. https://doi.org/10.1002/cmdc.201200497.

(225) DeLorbe, J. E.; Clements, J. H.; Teresk, M. G.; Benfield, A. P.; Plake, H. R.; Millspaugh, L. E.; Martin, S. F. Thermodynamic and Structural Effects of Conformational Constraints in Protein−Ligand Interactions. Entropic Paradoxy Associated with Ligand Preorganization. *J. Am. Chem. Soc.* **2009**, *131* (46), 16758–16770. https://doi.org/10.1021/ja904698q.

(226) McCauley, J. A.; Rudd, M. T. Hepatitis C Virus NS3/4a Protease Inhibitors. *Curr. Opin. Pharmacol.* **2016**, *30*, 84–92. https://doi.org/10.1016/j.coph.2016.07.015.

(227) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* **2004**, *3* (11), 935–949. https://doi.org/10.1038/nrd1549.

(228) Halgren, T. A. 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.

(229) Fu, H.; Chen, H.; Cai, W.; Shao, X.; Chipot, C. BFEE2: Automated, Streamlined, and Accurate Absolute Binding Free-Energy Calculations. *J. Chem. Inf. Model.* **2021**, *61* (5), 2116–2123. https://doi.org/10.1021/acs.jcim.1c00269.

(230) Furtmann, N.; Häußler, D.; Scheidt, T.; Stirnberg, M.; Steinmetzer, T.; Bajorath, J.; Gütschow, M. Limiting the Number of Potential Binding Modes by Introducing Symmetry into Ligands: Structure-Based Design of Inhibitors for Trypsin-Like Serine Proteases. *Chem. - A Eur. J.* **2016**, *22* (2), 610–625. https://doi.org/10.1002/chem.201503534.

(231) Béliveau, F.; Tarkar, A.; Dion, S. P.; Désilets, A.; Ghinet, M. G.; Boudreault, P.-L.; St-Georges, C.; Marsault, É.; Paone, D.; Collins, J.; Macphee, C. H.; Campobasso, N.; Groy, A.; Cottom, J.; Ouellette, M.; Pope, A. J.; Leduc, R. Discovery and Development of TMPRSS6 Inhibitors Modulating Hepcidin Levels in Human Hepatocytes. *Cell Chem. Biol.* **2019**, *26* (11), 1559-1572.e9. https://doi.org/10.1016/j.chembiol.2019.09.004.

(232) Pilgram, O.; Keils, A.; Benary, G. E.; Müller, J.; Merkl, S.; Ngaha, S.; Huber, S.; Chevillard, F.; Harbig, A.; Magdolen, V.; Heine, A.; Böttcher-Friebertshäuser, E.; Steinmetzer, T. Improving the Selectivity of 3-Amidinophenylalanine-Derived Matriptase Inhibitors. *Eur. J. Med. Chem.* **2022**, *238*, 114437. https://doi.org/10.1016/j.ejmech.2022.114437.

(233) Hofelich, T.; Wadsö, L.; Smith, A. L.; Shirazi, H.; Mulligan, S. R. The Isothermal Heat Conduction Calorimeter: A Versatile Instrument for Studying Processes in Physics, Chemistry, and Biology. *J. Chem. Educ.* **2001**, *78* (8), 1080. https://doi.org/10.1021/ed078p1080.

(234) Bartle, K. D.; Osborn, P. M. A Simplified Undergraduate Calorimetry Experiment. *J. Chem. Educ.* **1973**, *50* (9), 637. https://doi.org/10.1021/ed050p637.

(235) Wadsö, L.; Li, X. A Simple Rate Law Experiment Using a Custom-Built Isothermal Heat Conduction Calorimeter. *J. Chem. Educ.* **2008**, *85* (1), 112. https://doi.org/10.1021/ed085p112.

(236) Wadsö, L.; Li, Y.; Li, X. Isothermal Titration Calorimetry in the Student Laboratory. *J. Chem. Educ.* **2011**, *88* (1), 101–105. https://doi.org/10.1021/ed100649e.

(237) Zhang, Q.; Liu, J.; Bao, H.; Nai, X.; Sun, D.; Ma, B. Determining the Ternary Phase Diagram of Benzene–Acetic Acid–Water Using Isothermal Titration Microcalorimetry to Train Upper-Level Undergraduates in Advanced Calorimetry Methods. *J. Chem. Educ.* **2020**, *97* (5), 1470–1475. https://doi.org/10.1021/acs.jchemed.0c00002.

(238) Moore, D. E.; Goode, D. R.; Seney, C. S.; Boatwright, J. M. Isothermal Titration Calorimetry Can Provide Critical Thinking Opportunities. *J. Chem. Educ.* **2016**, *93* (2), 304–310. https://doi.org/10.1021/acs.jchemed.5b00575.

(239) Millies, B.; von Hammerstein, F.; Gellert, A.; Hammerschmidt, S.; Barthels, F.; Göppel, U.; Immerheiser, M.; Elgner, F.; Jung, N.; Basic, M.; Kersten, C.; Kiefer, W.; Bodem, J.; Hildt, E.; Windbergs, M.; Hellmich, U. A.; Schirmeister, T. 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.

(240) 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* (9), 1037–1042. https://doi.org/10.1021/ml500245v.

(241) Wunnicke, D.; Hänelt, I. The Synergetic Effects of Combining Structural Biology and EPR Spectroscopy on Membrane Proteins. *Crystals* **2017**, *7* (4), 117. https://doi.org/10.3390/cryst7040117.

(242) GraphPad Prism. GraphPad Software: San Diego, California USA 2014. www.graphpad.com.

(243) Zhang, Z.; Li, Y.; Loh, Y. R.; Phoo, W. W.; Hung, A. W.; Kang, C.; Luo, D. Crystal Structure of Unlinked NS2B-NS3 Protease from Zika Virus. *Science (80-.).* **2016**, *354* (6319), 1597–1600. https://doi.org/10.1126/science.aai9309.

(244) Hammerschmidt, S. J.; Huber, S.; Braun, N. J.; Lander, M.; Steinmetzer, T.; Kersten, C. Thermodynamic Characterization of a Macrocyclic Zika Virus NS2B/NS3 Protease Inhibitor and Its Acyclic Analogs. *Arch. Pharm. (Weinheim).* **2023**, *356* (4), e2200518. https://doi.org/10.1002/ardp.202200518.

(245) Diskowski, M.; Mehdipour, A. R.; Wunnicke, D.; Mills, D. J.; Mikusevic, V.; Bärland, N.; Hoffmann, J.; Morgner, N.; Steinhoff, H. J.; Hummer, G.; Vonck, J.; Hänelt, I. Helical Jackknives Control the Gates of the Double-Pore K+ Uptake System KtrAB. *Elife* **2017**, *6*, 1–21. https://doi.org/10.7554/eLife.24303.

(246) Pannier, M.; Veit, S.; Godt, A.; Jeschke, G.; Spiess, H. W. Dead-Time Free Measurement of Dipole–Dipole Interactions between Electron Spins. *J. Magn. Reson.* **2011**, *213* (2), 316–325. https://doi.org/10.1016/j.jmr.2011.08.035.

(247) Hänelt, I.; Wunnicke, D.; Müller-Trimbusch, M.; Vor der Brüggen, M.; Kraus, I.; Bakker, E. P.; Steinhoff, H.-J. Membrane Region M2C2 in Subunit KtrB of the K+ Uptake System KtrAB from Vibrio Alginolyticus Forms a Flexible Gate Controlling K+ Flux. *J. Biol. Chem.* **2010**, *285* (36), 28210–28219. https://doi.org/10.1074/jbc.M110.139311.

(248) Jeschke, G.; Chechik, V.; Ionita, P.; Godt, A.; Zimmermann, H.; Banham, J.; Timmel, C. R.; Hilger, D.; Jung, H. DeerAnalysis2006—a Comprehensive Software Package for Analyzing Pulsed ELDOR Data. *Appl. Magn. Reson.* **2006**, *30* (3–4), 473–498. https://doi.org/10.1007/BF03166213.

(249) Polyhach, Y.; Bordignon, E.; Jeschke, G. Rotamer Libraries of Spin Labelled Cysteines for Protein Studies. *Phys. Chem. Chem. Phys.* **2011**, *13* (6), 2356–2366. https://doi.org/10.1039/C0CP01865A.

(250) von Hammerstein, F.; Lauth, L. M.; Hammerschmidt, S.; Wagner, A.; Schirmeister, T.; Hellmich, U. A. 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.

(251) de Sousa, L. R. F.; Wu, H.; Nebo, L.; Fernandes, J. B.; da Silva, M. F. das G. F.; Kiefer, W.; Kanitz, M.; Bodem, J.; Diederich, W. E.; Schirmeister, T.; Vieira, P. C. Flavonoids as Noncompetitive Inhibitors of Dengue Virus NS2B-NS3 Protease: Inhibition Kinetics and Docking Studies. *Bioorg. Med. Chem.* **2015**, *23* (3), 466–470. https://doi.org/10.1016/j.bmc.2014.12.015.

(252) 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* (12), 2744–2752. https://doi.org/10.1021/cb400612h.

(253) Maier, G. P.; Bernt, C. M.; 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.

(254) Gibbs, A. C.; Steele, R.; Liu, G.; Tounge, B. A.; Montelione, G. T. Inhibitor Bound Dengue NS2B-NS3pro Reveals Multiple Dynamic Binding Modes. *Biochemistry* **2018**, *57* (10), 1591–1602. https://doi.org/10.1021/acs.biochem.7b01127.

(255) Zheng, J. *Biomedical Applications of Biophysics*; Jue, T., Ed.; Humana Press: Totowa, NJ, 2010. https://doi.org/10.1007/978-1-60327-233-9.

(256) Stryer, L. Fluorescence Energy Transfer as a Spectroscopic Ruler. *Annu. Rev. Biochem.* **1978**, *47* (1), 819–846. https://doi.org/10.1146/annurev.bi.47.070178.004131.

(257) Maus, H.; Hinze, G.; Hammerschmidt, S. J.; Basché, T.; Schirmeister, T. A Competition <scp>smFRET</Scp> Assay to Study Ligand-induced Conformational Changes of the Dengue Virus Protease. *Protein Sci.* **2023**, *32* (1), e4526. https://doi.org/10.1002/pro.4526.

(258) 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), 677–680. https://doi.org/10.1016/j.bbrc.2015.04.090.

(259) Kumar, A.; Liang, B.; Aarthy, M.; Singh, S. K.; Garg, N.; Mysorekar, I. U.; Giri, R. Hydroxychloroquine Inhibits Zika Virus NS2B-NS3 Protease. *ACS Omega* **2018**, *3* (12), 18132–18141. https://doi.org/10.1021/acsomega.8b01002.

(260) Nitsche, C.; Zhang, L.; Weigel, L. F.; Schilz, J.; Graf, D.; Bartenschlager, R.; Hilgenfeld, R.; Klein, C. D. 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.

(261) Yang, Y.; Cao, L.; Gao, H.; Wu, Y.; Wang, Y.; Fang, F.; Lan, T.; Lou, Z.; Rao, Y. Discovery, Optimization, and Target Identification of Novel Potent Broad-Spectrum Antiviral Inhibitors. *J. Med. Chem.* **2019**, *62* (8), 4056–4073. https://doi.org/10.1021/acs.jmedchem.9b00091.

(262) Lei, J.; Hansen, G.; Nitsche, C.; Klein, C. D.; Zhang, L.; Hilgenfeld, R. Crystal Structure of Zika Virus NS2B-NS3 Protease in Complex with a Boronate Inhibitor. *Science (80-.).* **2016**, *353* (6298), 503–505. https://doi.org/10.1126/science.aag2419.

(263) Preugschat, F.; Yao, C. W.; Strauss, J. H. 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.

(264) Bera, A. K.; Kuhn, R. J.; Smith, J. L. Functional Characterization of Cis and Trans Activity of the Flavivirus NS2B-NS3 Protease. *J. Biol. Chem.* **2007**, *282* (17), 12883–12892. https://doi.org/10.1074/jbc.M611318200.

(265) Chitambar, C. R.; Narasimhan, J. Targeting Iron-Dependent DNA Synthesis with Gallium and Transferrin-Gallium. *Pathobiology* **1991**, *59* (1), 3–10. https://doi.org/10.1159/000163609.

(266) Gugala, N.; Lemire, J. A.; Turner, R. J. The Efficacy of Different Anti-Microbial Metals at Preventing the Formation of, and Eradicating Bacterial Biofilms of Pathogenic Indicator Strains. *J. Antibiot. (Tokyo).* **2017**, *70* (6), 775–780. https://doi.org/10.1038/ja.2017.10.

(267) Wu, C.; Labrie, J.; Tremblay, Y. D. N.; Haine, D.; Mourez, M.; Jacques, M. Zinc as an Agent for the Prevention of Biofilm Formation by Pathogenic Bacteria. *J. Appl. Microbiol.* **2013**, *115* (1), 30–40. https://doi.org/10.1111/jam.12197.

(268) Dumas, P.; Ennifar, E.; Bec, G.; Piñero, A.; Sabín, J.; Muñoz, E.; Rial, J. Implementation of KinITC into AFFINImeter. *Appl. Note* **2015**.

(269) El-Araby, A. M.; Feltzer, R.; Kim, C.; Mobashery, S. Application of 2D-ITC to the Elucidation of the Enzymatic Mechanism of N -Acetylmuramic Acid/ N -Acetylglucosamine Kinase (AmgK) from Pseudomonas Aeruginosa. *Biochemistry* **2023**. https://doi.org/10.1021/acs.biochem.3c00090.

(270) Wang, Y.; Mittermaier, A. K. Characterizing Bi-Substrate Enzyme Kinetics at High Resolution by 2D-ITC. *Anal. Chem.* **2021**, *93* (37), 12723–12732. https://doi.org/10.1021/acs.analchem.1c02705.

(271) Drescher, M. EPR in Protein Science: Intrinsically Disordered Proteins. In *Topics in Current Chemistry*; 2012; Vol. 321, pp 91–120. https://doi.org/10.1007/128_2011_235.

(272) Jäger, H.; Koch, A.; Maus, V.; Spiess, H. W.; Jeschke, G. Relaxation-Based Distance Measurements

between a Nitroxide and a Lanthanide Spin Label. *J. Magn. Reson.* **2008**, *194* (2), 254–263. https://doi.org/10.1016/j.jmr.2008.07.012.

(273) Stryer, L.; Haugland, R. P. Energy Transfer: A Spectroscopic Ruler. *Proc. Natl. Acad. Sci.* **1967**, *58* (2), 719–726. https://doi.org/10.1073/pnas.58.2.719.

Poster presentations

[redacted]

Mainz, 17.04.2023 Stefan Josef Hammerschmidt